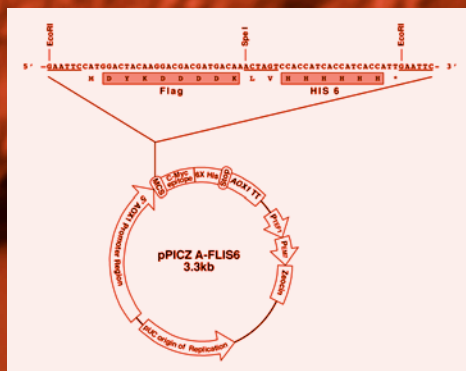


RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



RNA Interference

Historical Overview and Significance

Mary K. Montgomery

Summary

In the early 1990s, attempts to manipulate gene expression by researchers working in three different fields resulted in unanticipated gene silencing. Rather than ignoring such results, these researchers went on to document and further investigate the nature of such silencing, which was named “co-suppression” in plants, “quelling” in fungi, and “RNA interference” (RNAi) in nematodes. By the late 1990s, it was discovered that silencing could be initiated in this diverse set of organisms by exposing cells to double-stranded RNA (dsRNA), which directed the destruction of mRNAs containing similar sequences. Soon afterward, such dsRNA-mediated silencing was employed as a reverse genetic technique to analyze the functions of specific genes in a broad variety of organisms. Biochemical and genetic studies designed to uncover the components of the RNA silencing machinery identified a common core of proteins that serve to amplify the interfering RNA signal and direct endonucleolytic cleavage of target RNAs. A subset of silencing events may also direct DNA methylation of targeted genes. RNA silencing is thought to have evolved as a defense mechanism to suppress viral replication and transposon mobilization. However, additional functions involving the RNAi machinery have been uncovered, including posttranscriptional regulation of endogenous genes, and maintenance of structure and function of heterochromatin. Whereas many researchers have focused on understanding the natural biological functions of RNA silencing, others are testing its utility in antiviral and cancer therapies and in other biotechnological and biomedical applications.

Key Words: RNA silencing; RNA interference; cosuppression; quelling; posttranscriptional gene silencing; antisense RNA; double-stranded RNA; dicer; RNA-induced silencing complex; RNA-directed RNA polymerase.

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

1. RNA Interference: A Historical Overview

1.1. The Dawn of Reverse Genetics

Mutagenesis screens begun in 1908 in Thomas Hunt Morgan's laboratory led to the isolation of dozens of mutations in the fruit fly *Drosophila* (1). In addition to construction of the first genetic map, the analyses of genes identified as a result of such "forward" genetic screens (phenotype → gene) have made outstanding contributions to our understanding of how cells and organisms function and develop. However, genetic screens of this kind can only be realistically performed on organisms with certain lifestyle characteristics (ability to breed under laboratory conditions, short life cycle, abundant offspring), and it often requires months or years of dedicated work to identify the mutated gene responsible for a specific phenotype. With the advent of molecular biological techniques, researchers began looking for a "back door entrance" that could quickly reveal the biological functions of genes identified on the basis of sequence homology or biochemical function, namely a reverse genetic approach (gene → phenotype).

In 1984, a significant step in this direction was reported by Izant and Weintraub (2), who transformed tissue culture cells with a DNA construct engineered to express antisense RNA complementary to thymidine kinase mRNA; the accumulation of thymidine kinase protein was "dramatically and specifically" inhibited in these transgenic cells. This initial study was followed by several more (2–6), in which the efficacy of using an antisense strategy to inhibit the activity of specific genes was demonstrated. The technical advances pioneered by Weintraub, Melton, and others would prove to be of particular importance to studies involving cell lines and whole organisms such as *Xenopus* that are not particularly amenable to classic genetic analysis, but exhibit many other attributes making them attractive models for the study of gene function.

1.2. The Discovery of RNA Interference in *Caenorhabditis elegans*

In antisense studies using RNAs synthesized in vitro, sense RNAs have been typically introduced as negative controls for specificity. Thus, it was surprising when Guo and Kemphues (7) found that both injected control sense and antisense RNAs gave similar phenocopies. This proved a general observation in *C. elegans*: injection of sense RNA as well as antisense RNA generated a specific and reproducible phenotype. Two keys, opposite but complementary in design, appeared to open the same back door. Craig Mello and colleagues (8) coined this phenomenon "RNA interference," or RNAi, to distinguish it from classic antisense inhibition. Adding to the mystery, Mello and coworkers noticed that silencing spread to cells beyond the site of injection, suggesting that the interfering RNA could be transported from the site of initial delivery to

most cells and tissues in the worm, eliciting a systemic response (8,9). Thrilled with having a novel reverse genetic method to target their favorite genes, many *C. elegans* researchers embraced RNAi, even in the absence of an understanding of how the knockdown in expression was being achieved. Several investigators, however, were motivated to investigate the molecular mechanism of the RNAi response in worms.

One aspect of the mystery was solved by Andrew Fire at the Carnegie Institution of Washington's Department of Embryology when he first observed that double-stranded RNA (dsRNA), rather than single-stranded antisense RNA, was responsible for triggering the sequence-specific degradation of targeted endogenous mRNAs in *C. elegans*. Fire reasoned that sense and antisense RNA preps generated by in vitro transcription reactions using plasmid templates might be contaminated with small amounts of RNA of the opposite polarity, owing to the infidelity of viral RNA polymerases used in the in vitro synthesis reactions. This would lead to contamination with small amounts of dsRNA. Fire proceeded to test the hypothesis using purified preps of single-stranded RNAs (ssRNAs) and dsRNAs. When ssRNAs were gel purified, their abilities to generate phenotypes when injected were markedly reduced, whereas the deliberate introduction of dsRNA (sense and antisense strands annealed in vitro and subsequently injected) caused a relatively rapid and specific degradation of mRNAs containing sequences similar to those of the sense strand of the dsRNA (9).

One of the many interesting features of RNAi is not only that the "trigger" proved to be dsRNA rather than ssRNA, but that the dsRNA functions substoichiometrically (9) with one molecule triggering the degradation of dozens, perhaps hundreds, of individual mRNAs, again distinguishing the phenomenon from antisense inhibition (which typically requires an excess of asRNA molecules to target messages). But how was the dsRNA able to target cognate sequences? Any mechanistic model would also have to explain dsRNA's potency. There were at least three possible explanations, none of which were mutually exclusive: (1) the interfering RNA was amplified, (2) the dsRNA acted catalytically, or (3) the dsRNA targeted the gene directly.

We observed by *in situ* hybridization that mRNAs corresponding in sequence to the dsRNA trigger failed to accumulate (9). We initially entertained the idea that the dsRNA might somehow directly target the gene by interfering with either initiation or elongation of transcription; this would explain how one or two dsRNA molecules per cell could prevent mRNA accumulation. However, we were unable to find any experimental evidence to support this hypothesis; instead, we found that targeted genes continued to be transcriptionally active in the presence of dsRNA, but the mRNAs they encoded were rapidly degraded before they could be translated (10). Evidence supporting a posttranscriptional mechanism also included an inability to target introns or promoters (suggesting

a cytoplasmic mechanism) (9). Additionally, we observed no sequence changes corresponding to target mRNAs (10). This was not surprising because RNAi effects are typically fully reversible. We did not examine the possibility that targeted genes might become methylated, because *C. elegans* apparently lacks a DNA methylation system (11).

1.3. RNAi as a Form of Posttranscriptional Gene Silencing First Documented in Plants and Fungi

A form of posttranscriptional gene silencing (PTGS) in response to transgene sequences had previously been documented in plants and fungi (12,13). First, Napoli et al. (12) introduced a transgene designed to overexpress chalcone synthase in an attempt to increase pigment production in petunias; unexpectedly, in >40% of the transgenic plants, the flowers appeared white or variegated rather than purple. This phenotype was owing to silencing or suppression of not only the transgene but also endogenous copies of chalcone synthase and was thus termed *cosuppression*. In 1992, Romano and Macino (13) reported a similar phenomenon in the fungus *Neurospora crassa*, which they coined “quelling” (in an interesting coincidence these researchers were also manipulating genes involved in pigment production). In both studies, reversion to parental or wild-type phenotypes in some of the progeny of affected cells was observed. Thus, attempts to overexpress a gene could instead lead to silencing, but the effect was often transient, leading these investigators to hypothesize that either DNA methylation or an RNA intermediate was mediating the response.

Cogoni et al. (14) obtained strong evidence that quelling in fungi resulted from targeting of posttranscriptional events; moreover, they established that the effector molecule was cytoplasmic, and thus most likely RNA. By contrast, some cases of cosuppression observed in plants could be correlated to methylation of the gene itself, resulting in downregulation of expression (15). Cosuppression in plants could also arise from targeting posttranscriptional stages of gene expression (termed *PTGS* and, more recently, simply *RNA silencing*). Researchers investigating plant RNA silencing also concluded that RNA was the effector molecule, but the nature and/or structure of the RNA that triggered the silencing response was not immediately apparent.

The finding that dsRNA could trigger PTGS in *C. elegans* led to the hypothesis that the inadvertent production of dsRNA in transgenic plants (e.g., owing to unintended antisense production via cryptic promoters) might explain at least some cases of plant PTGS (16). Waterhouse et al. (17) soon demonstrated that tobacco and rice plants engineered to deliberately synthesize dsRNA did indeed exhibit gene silencing. This was an exciting result—two evolutionarily divergent organisms (plants and nematodes) appeared to respond to the presence of

dsRNA in the same manner. Another common feature was that the interfering RNA could spread to other tissues and activate systemic silencing. Nonetheless, this result did not mean that all of the previously documented cases of transgene-induced PTGS were owing to production of dsRNA from the *in vivo* annealing of sense with unintended antisense transcripts; one model proposed that high levels of sense transcripts could trigger the formation of copy RNA (cRNA) via an RNA-dependent RNA polymerase (reviewed in **ref. 18**); either dsRNA composed of the sense plus RNA or the cRNA itself might be the source of interfering RNA. Furthermore, it was hypothesized that such RNA might feed back to the gene and mediate DNA methylation in plants. Several early, as well as more recent, findings lend support to this model and are further discussed in **Subheading 1.6**.

Following the discovery of RNAi in *C. elegans*, researchers began to test whether dsRNA could elicit gene silencing in other organisms as well. Soon RNAi was shown to shut down or reduce gene expression in trypanosomes (a protist), the fruit fly *Drosophila*, and many other animal as well as plant species. Initially, it was thought that RNAi would have limited utility in mammals and other vertebrates owing to the activity of a dsRNA-dependent protein kinase called PKR that responds to viral or exogenous dsRNA by setting off a global panic response that can ultimately lead to cell death (**19**). But by targeting oocytes and early embryos in mice prior to development of the immune system and the onset of PKR expression, Wianny and Zernicka-Goetz (**20**) were able to demonstrate dsRNA-induced sequence-specific silencing. Similar studies in zebrafish, however, led to less than satisfactory results, with more instances of nonspecific effects (**21**). Thus, although the introduction of dsRNA was shown to result in sequence-specific silencing in an extraordinarily diverse set of organisms, it does not necessarily elicit the same response in all eukaryotes.

Further work has shown that the mechanisms underlying RNAi, PTGS, and quelling are related and appear to represent an evolutionarily conserved mechanism for thwarting the expression of nonself or “selfish” genetic information, such as that encoded by viruses and transposable elements. The presence of aberrant RNA structure, such as extended dsRNA or replicative intermediates expressed by some RNA viruses and the inverted repeat structure of some transposable elements, alerts the host cell to the presence of foreign invaders and triggers silencing. Supporting this interpretation of the function of RNAi and related cosuppression/PTGS responses is that some mutations in RNA-silencing pathway components result in increased susceptibility to viral infection (**22**). Not surprisingly, given the escalating arms race between host and pathogen, mechanisms for circumventing host PTGS responses have also been documented (**23**). So, how does RNAi/PTGS work?

1.4. A Mechanistic Model Emerges

RNAi has drawn as much attention for the mystery of how it works as for its potential utility in analyzing gene function. The potency of the dsRNA trigger to directly target multiple copies of an mRNA suggested that it acted catalytically and/or was amplified. A mechanistic model emerged relatively rapidly from a flurry of biochemical and genetic studies on flies, worms, fungi, plants, and mammalian cells (reviewed in **ref. 24**). The “core” RNA-silencing response involves processing of the trigger dsRNA into smaller 21- to 25-bp fragments with dinucleotide 3′ overhangs by an adenosine triphosphate (ATP)-dependent enzyme named Dicer (*see Fig. 1*). Dicer encodes a multidomain protein containing an ATP-dependent RNA helicase, PAZ domain, two tandem RNase III domains, and a dsRNA-binding domain (**25**). The 21- to 25-bp products of Dicer activity are referred to as short interfering RNAs (siRNAs) (**26**), which are thought to serve as “guides” to bring nuclease machinery to the target mRNA: each siRNA associates with a protein complex called RNA-induced silencing complex (RISC), and presumably the siRNA is unwound by a helicase component of RISC to allow base pairing between the antisense strand and the target mRNA (**27**). Such base pairing leads to endonucleolytic cleavage of the target mRNA, producing one fragment missing its polyA tail and the other missing its 5′ 7-methylguanosine cap, leaving each segment vulnerable to further degradation by RNA surveillance machinery. Following cleavage, it is thought that the siRNA/RISC complex becomes available to target another messenger molecule (*see Fig. 1*).

Thus, the initial trigger dsRNA generates several siRNAs (e.g., approx 20 siRNAs could be generated from a single 500-bp dsRNA), each of which recruits and activates a RISC, which together may function catalytically to target multiple messages. This model nicely explains the substoichiometric potency of the trigger dsRNA and encompasses both an amplification step and a catalytic component. Dicer and RISC may represent some of the most evolutionarily conserved components of the RNA silencing machinery, and the processes that catalyze the “core” RNAi response.

More recently, evidence has been accumulating that, at least in some species, additional amplification may occur. Secondary siRNAs may be generated by a mechanism termed *transitive RNAi*, in which an siRNA (**28**) or even a short antisense RNA (**29**) may function to prime specific RNA-dependent RNA polymerases (RdRPs) to synthesize cRNA of the target, producing dsRNAs *de novo* that are cleaved by Dicer into additional siRNAs (**Fig. 2**). There is some evidence that some RdRPs may also produce cRNA in the absence of priming in response to exceptionally high levels of a particular transcript or detection of other forms of aberrant RNA (perhaps interpreted by the cell as a potential

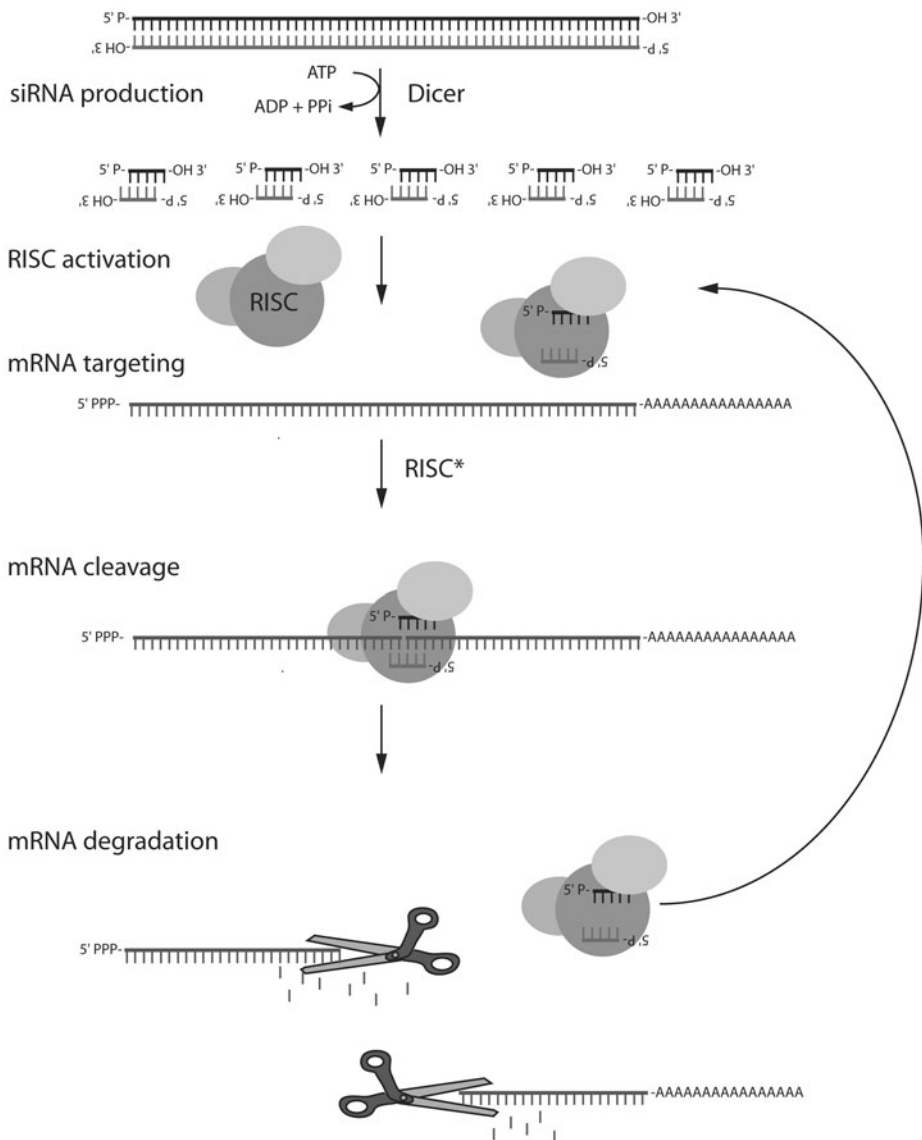


Fig. 1. Mechanistic model for RNAi. dsRNA, whether endogenously produced or exogenously provided, is cleaved by the ATP-dependent RNase III-like enzyme Dicer, producing 21- to 25-bp siRNAs with 2-nt 3' overhangs. Each siRNA recruits a RISC, which unwinds the siRNA, exposing each strand for potential binding to complementary target sequences. When the now-activated RISC (RISC*) binds to a target mRNA, it cleaves it at a site approx 10 bp from the 5' end of the siRNA. This results in two cleavage products, one missing its polyA tail and the other its 5' 7-methyl guanosine cap, making each vulnerable to additional degradation by mRNA surveillance machinery. Meanwhile, the activated RISC is free to target additional messages.

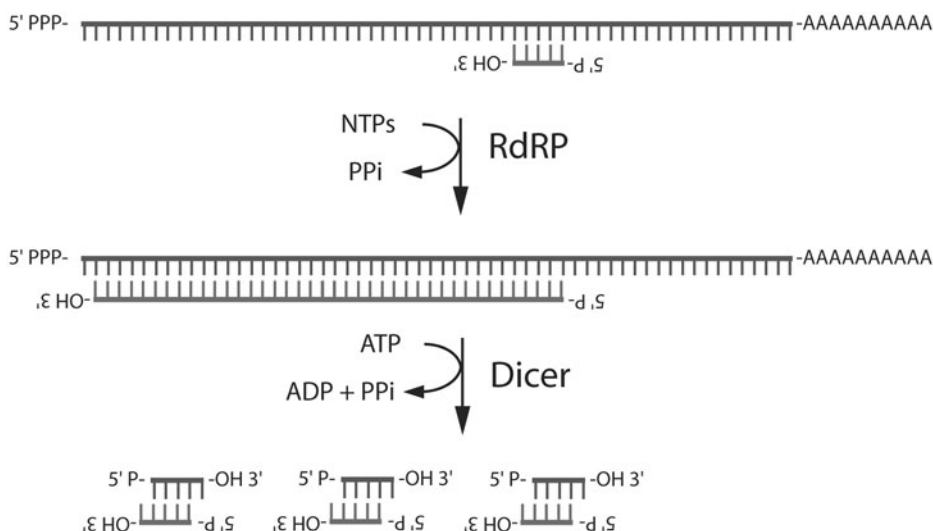


Fig. 2. "Transitive" RNAi. Binding of siRNAs or short antisense RNAs to a target mRNA may prime RNA synthesis by RdRPs. The original mRNA that served as template plus the complementary RNA together form dsRNA that is recognized by Dicer and cleaved to form secondary siRNAs. RdRPs in some organisms may be able to synthesize RNA in the absence of priming.

viral invasion), an idea first proposed to explain transgene-induced PTGS in plants (*see* **ref. 18**). Presumably, the advantage to the cell would be that more triggers would be generated to inactivate more targets enabling the cell to efficiently clear itself of viral infection.

Mutations in RdRPs in *C. elegans*, *Arabidopsis*, and *N. crassa* indicate that these proteins play a conserved and essential role in RNAi/PTGS; interestingly, some of these mutants also result in increased viral susceptibility, and/or developmental defects (reviewed in **ref. 24**). Curiously, no RdRP homolog has been identified by sequence similarity in the fly or human genomes, leaving open the question, Is this form of trigger amplification a universal feature of RNAi/PTGS? Although siRNA-primed RNA synthesis has been observed in *Drosophila* cell-free extracts (**30**), siRNAs with modified 3'-OH, which would presumably be unable to serve as primers, are nonetheless capable of triggering RNAi in both *Drosophila* and human HeLa cells (**31**); in addition, RNA synthesis is not required for efficient RNAi in mouse oocytes (**32**). These data taken together indicate that even if such amplification occurs in flies or humans, it is not essential for the RNAi response in these organisms. A potential advantage of lacking transitive RNAi is increased specificity, such that isoforms of alternatively spliced mRNAs could be targeted separately (**33**).

1.5. RNA Silencing Performs a Variety of Biological Functions

If one of the main functions of RNA silencing is to prevent expression of viruses and transposable elements that utilize an RNA intermediate, the advantages to the host of a mechanism that involves cleavage of the interfering dsRNA seem obvious. It would further benefit the cell to amplify the means by which to target as many unwanted RNAs by increasing the number of “guide” RNAs at the same time it prevents both the trigger and target mRNAs from being translated into functioning proteins. Of particular interest to those working with mammalian systems is that siRNAs do not trigger PKR activity but do mediate sequence-specific silencing; the nonspecific global panic response to the presence of dsRNA can thus be circumvented by exposing mammalian cells, either in culture or in vivo, to siRNAs. This and similar siRNA-based RNAi strategies are now being applied to a much broader range of experiments involving mammalian models.

Screens for RNA-silencing mutants were first performed in *Neurospora* (34), and more recently in *C. elegans* and *Arabidopsis*, and have led to the identification of several classes of genes essential for the RNA-silencing response, including Dicer homologs, RdRPs, RISC components, and various helicases (reviewed in ref. 24). Continued genetic dissection of the silencing pathway is leading to the ordering of some components as well as revealing additional biological functions. Some RNAi-defective mutants have additional phenotypes, including increased susceptibility to viral infection, increased transposon mobilization, defects in developmental timing, defects in mitosis and meiosis, sterility, and other developmental defects. These phenotypes provide important insight into the natural functions of RNA silencing.

One of the most exciting discoveries has been that RNA silencing plays a role in regulating the expression of genes “native” to the organism. RNA-silencing components are involved in processing a class of small regulatory RNAs into microRNAs (miRNAs) that bind to complementary sequences in the 3′ untranslated region (UTR) of target mRNAs; some miRNAs may function as siRNAs, recruiting RISC and resulting in cleavage of the target message. However, many miRNAs (e.g., *let-7*) work through an alternative pathway, in which binding does not destabilize the target mRNA but, rather, prevents accumulation of the encoded protein product by interfering with translation (see Chapter 8 for more details). How does the cell distinguish between these two classes of miRNAs? Recent work indicates that the number and placement of mismatches between the interfering RNA and its complementary target are most likely important, with miRNAs affecting translation typically containing multiple mismatches in the middle of the sequence resulting in a “bulge” during hybridization with their targets (35). A single siRNA may be sufficient to direct cleavage

of a target RNA provided it has access for base pairing (presumably it would not be able to target sequences masked by protein binding or secondary structure). Whereas there is some experimental evidence that a single binding site can also mediate translational downregulation, endogenous targets of *let-7*-type miRNA activity typically contain multiple binding sites within their 3' UTRs, suggesting that in some cases a single binding event is not sufficient to repress target expression.

1.6. RNA Silencing May Trigger Covalent Modification of DNA

Our current understanding of how RNA silencing works in the cell encompasses both amplification steps and a catalytic component. In plants, it has been shown that RNA can also trigger *de novo* DNA methylation leading to transcriptional silencing (36,37). Now, there is experimental evidence that siRNAs can inactivate transcription through direct DNA methylation and other types of covalent modification in the genomes of certain species (38). This activity in plants may be directed by a separate class of siRNAs (39).

One of the most remarkable recent findings has been that RNAi machinery in the fission yeast *Schizosaccharomyces pombe* plays a critical role in formation and maintenance of higher-order chromatin structure and function. Deletions in critical RNAi pathway genes lead to loss of epigenetic silencing of centromeric DNA and other types of heterochromatin; these changes are accompanied by changes in the methylation state of these regions, and subsequent loss of centromere cohesion, resulting in chromosome missegregation during nuclear division (40,41). It is hypothesized that in wild-type cells expression of centromeric repeats results in the formation of a dsRNA that is cleaved by Dicer into siRNAs that direct DNA methylation of heterochromatic sites. siRNAs have been reported to mediate/direct even more radical changes in chromosome structure, namely the elimination of specific DNA sequences that occurs in programmed genome arrangement in the protist *Tetrahymena* (42).

2. Significance

2.1. Impact of RNAi on Functional Analyses of Genes

The discovery of RNAi and its application as a reverse genetic technology seem almost perfectly timed to provide a reasonable strategy for analyses of the thousands of predicted genes emerging from various genome projects. RNAi is being used widely in functional studies of genes in *C. elegans* and *Drosophila*, both by individual researchers interested in dissecting specific genetic pathways and by others interested in systematically analyzing every gene in the genome (43). RNAi is also being used in genomewide screens to identify genes that result in a specific phenotype or that alter transgene expression (44). Perhaps one

of the most ambitious projects to date is a joint initiative by researchers in the United Kingdom and the Netherlands to design siRNAs against every gene in the human genome and then use them to systematically target each gene in a cancerous cell line in hopes of identifying those genes required for malignant transformation of cells (45). A more targeted approach is to couple RNAi with microarray analysis, selectively targeting genes that are upregulated or overexpressed in certain cell or tissue types; such an approach has led to the identification of genes controlling proliferation in colon cancer (46).

Perhaps one of RNAi's most significant legacies will be the ability to carry out functional studies of genes in organisms for which "the awesome power of genetics" (i.e., forward genetics) cannot be readily applied. The reverse genetic approach that RNAi can offer is giving new life to animal models that have been around for quite some time but that have proved genetically intractable, such as planaria (47), hydra (48), leech (49), and *Xenopus* (50). RNAi has also facilitated comparative studies between *Drosophila* and other insects (51–53) and between *C. elegans* and congeneric species (54,55).

2.1.1. Systemic vs Cell-Autonomous RNA Silencing and the Need for Alternative Delivery Methods

As noted earlier, dsRNA does not elicit an RNAi response in all species: zebrafish tend to respond to injected dsRNA with nonspecific lethality (21), whereas some other species, including sea urchins (A. Cameron, personal communication) and many nematode species outside the *Caenorhabditis* clade (ref. 56 and personal observation), appear RNAi resistant. Whether the use of siRNAs rather than long dsRNAs might alleviate the nonspecific response in zebrafish has yet to be tested. While the ability to execute RNAi may have been lost in multiple, independent evolutionary lines, it is also possible that some species that have been labeled RNAi resistant lack only the systemic as opposed to cell-autonomous response. Indeed, delivery may be the biggest block to more widespread use of the technology. In vascular plants the interfering RNA can spread to other tissues via the phloem and plasmodesmatal macromolecular trafficking system (see ref. 18). *C. elegans* has the ability to take up and transport exogenously supplied dsRNA (see Chapter 2), although this ability may not be expressed by all cell types, particularly those lacking one of the genetically identified transporters, *sid-1* (57).

Strategies for effective delivery of dsRNA or siRNAs into mammalian cells include transfection, electroporation, and transgenic approaches involving improved viral vectors (58,59); stable expression of siRNAs from Pol III-driven snap-back constructs (60) and related strategies are further discussed in Chapter 5. In an effort to explore the potential therapeutic capabilities offered by RNAi, several laboratories have been exploring methods for efficient *in vivo*

delivery of siRNAs into mammals. Both viral and nonviral delivery methods are being developed. One study found that direct injection of siRNAs into the tail vein of mice caused reduced mRNA and protein levels of the targeted gene in hepatocytes (**61**); these results suggest that delivery to at least certain tissues or cells may not be as high a potential hurdle as previously thought. Nonetheless, safe, efficient, and tissue-specific delivery of siRNAs may be one of the most significant obstacles to development of siRNA-based drugs in humans.

2.1.2. RNAi: Faster, Cheaper, Better?

One intriguing question is, can RNAi offer any unique advantages over traditional genetic approaches? In many cases, use of RNAi allows researchers to gain insight into the function of a gene years or decades more quickly than they could otherwise, particularly when a forward genetic approach is impractical. Moreover, genetic screens are inherently biased, since only mutants with expected phenotypes are isolated; a relevant gene with multiple roles, or expressed at an earlier stage, might be missed. RNAi can overcome this bias to some extent, revealing the functions of genes with unexpected phenotypes. Furthermore, the ease of generating double and triple “mutants” is greatly enhanced using RNAi, which is especially useful if two or more of the genes of interest are tightly linked genetically and creating the double mutant using more traditional approaches would rely on rare crossover events. RNAi is also a relatively cheap method for analyzing gene function, depending on how one weighs the cost of enzymes and reagents against that of research hours spent tracking down a gene via screening, mapping, cloning, rescuing, and sequencing. And with thousands of sequenced genes in the databases whose functions are unknown, faster and cheaper may be a valid enough reason for resorting to such strategies as RNAi for preliminary assignment/description of a gene’s role. But, can RNAi really offer anything new? Can RNAi provide any insight into a gene’s function that could not be revealed by a clever geneticist and a lucky mutation?

The ability to selectively eliminate gene function at various stages of development is one of RNAi’s main advantages, as is the potential to eliminate maternal and zygotic pools of mRNA simultaneously. However, researchers for decades have been able to manipulate loss of gene activity using conditional or inducible mutants. There is one type of conditional situation in which RNAi may prove truly advantageous. In some cases it is possible to selectively eliminate or reduce the activity of a specific gene in only a subset of cells given the appropriate promoters (**33**). For example, it may be possible to drive expression of a dsRNA from a transgene such that expression of the targeted gene is affected in only a specific subset of its total range. This might be thought of as a more tactical approach to targeted gene disruption, using the right promoters to drop isolated “smart bombs,” in essence producing a type of targeted genetic

mosaic. The use of such an RNAi strategy allows the researcher to more precisely direct where loss of expression will occur.

In fact, the use of transgenes to drive dsRNA expression has eliminated many of the potential disadvantages of RNAi, depending on how one wishes to apply the technique. The response to dsRNA is a transient effect, lasting a few hours to several days. If a more stable “mutant” line is desired, the organism or cell line can be transformed with transgenes that will generate dsRNA (either long or miRNA sized), such as by a “snap back” or inverted repeat construct driven by ubiquitous, inducible, or tissue-specific promoters. In certain applications, the transient nature of RNAi can be advantageous; for example, RNAi can be used to generate a phenotypic series, often in a single experiment, ranging from slight reduction of function to virtual nulls.

RNAi now joins the expanding repertoire of reverse genetic methods for interfering with or disrupting the activity of a specific gene in order to better understand its function. Depending on the organism, these other methods may include homologous recombination and targeted deletions, transposon-mediated gene disruption, traditional antisense approaches, and most recently morpholino-modified antisense oligonucleotides. RNA-silencing methods may be more versatile than many of these other techniques, but, in general, a set of tools should be chosen for thorough analysis of gene function.

2.2. Promises and Potential Pitfalls in Biotechnological and Biomedical Applications

RNAi is a powerful tool for inhibiting gene function in a sequence-specific manner. As discussed earlier, it is a remarkably useful method for generating loss-of-function or reduction-of-function phenotypes that can provide useful insights into the function of a gene. Applications of RNA-induced silencing are being explored in both the biotechnological and biomedical fields as well. The use of siRNAs targeting male fertility genes to generate sterile males is being tested as a new method for pest control. Transgene-induced RNA silencing is being used to repress expression of genes in plants with potential commercial value in efforts to either control expression of endogenous genes or induce resistance to viral or bacterial disease. The current approach to regulation of genetically modified crops in the United States seems to be one of “safe until proven otherwise”; stricter guidelines and adequate testing need to be applied to accurately determine the risks of such approaches to consumers and the environment. For example, at this point it is unknown what, if any, effect ingestion of dsRNA-containing food might have on consumers.

The therapeutic potential of siRNAs in humans is also being explored. Current applications include antiviral and cancer therapies in mammalian cell lines and animal models. RNAi-based strategies have successfully inhibited viral

infection from human papilloma virus (62), hepatitis B virus (63), hepatitis C virus (64), murine gammaherpesvirus (65), and human immunodeficiency virus type 1 (66). Despite these initial successes, one potential stumbling block may be RNAi-evading strategies that exist in some viruses (67).

Several research groups and biopharmaceutical companies are using RNAi in target discovery and target validation studies as part of their overall drug development strategy. One of the most encouraging studies to date reported that RNAi targeting of a gene upregulated in colon tumor cell lines (named *survivin*) resulted in reduced tumor growth (46). Efforts are also being directed at the development of RNAi-based medicines. As with current gene therapy trials, all appropriate measures to protect patients (e.g., independent review, voluntary informed consent) would need to be taken if siRNAs or their derivatives ever reach the clinical trial stage. Improvements in viral vector delivery methods are encouraging (58,59), but the development of effective and safe delivery strategies will most likely remain a primary area of interest.

Specificity is another issue that is being carefully examined. A single mismatch (on the antisense strand) between an siRNA and its target can be enough to substantially reduce the RNAi effect (68). Still, the connection between the number and placement of mismatches and the efficiency with which a particular siRNA can disrupt its target remain unclear. One concern is that an siRNA might target multiple gene products that share a stretch of similar sequence, leading to disruption of unintended as well as intended targets. Some siRNAs designed against one region of a target mRNA can be more effective than siRNAs designed against a different region of the same target; the basis underlying such differences in performance is not understood but may be related to the ability of a particular siRNA to access the complementary region on the target mRNA. Coupling siRNA testing with sophisticated expression analyses (e.g., microarrays) should allow researchers to identify siRNAs with the appropriate level of specificity and effectiveness. RNAi has become a hot and exciting field and RNAi's potential applications are tremendous. However, whether RNAi-based strategies eventually lead to effective therapies remains to be seen. Rather than fanning the flames of hyperbolic news accounts, perhaps researchers should remain cautiously optimistic as they continue to experiment.

Acknowledgments

I thank Veronica Descotte, Paul Overvoorde, Ann Rougvie, Jan Serie, and Lisa Timmons for useful discussions and comments on the manuscript.

References

1. Kohler, R. (1994) *Lords of the Fly: Drosophila Genetics and the Experimental Life*. University of Chicago Press, Chicago.

- 2 Izant, J. G. and Weintraub, H. (1984) Inhibition of thymidine kinase gene expression by anti-sense RNA: a molecular approach to genetic analysis. *Cell* **36**, 1007–1015.
- 3 Rosenberg, U. B., Preiss, A., Seifert, E., Jackle, H., and Knipple, D. C. (1985) Production of phenocopies by Kruppel antisense RNA injection into *Drosophila* embryos. *Nature* **313**, 703–704.
- 4 Harland, R. and Weintraub, H. (1985) Translation of mRNA injected into *Xenopus* oocytes is specifically inhibited by antisense RNA. *J. Cell Biol.* **101**, 1094–1099.
- 5 Melton, D. A. (1985) Injected anti-sense RNAs specifically block messenger RNA translation in vivo. *Proc. Natl. Acad. Sci. USA* **82**, 144–148.
- 6 Fire, A., Albertson, D., Harrison, S.W., and Moerman, D. G. (1991) Production of antisense RNA leads to effective and specific inhibition of gene expression in *C. elegans* muscle. *Development* **113**, 503–514.
- 7 Guo, S. and Kemphues, K. J. (1995) *par-1*, a gene required for establishing polarity in *C. elegans* embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. *Cell* **81**, 611–620.
- 8 Rocheleau, C. E., Downs, W. D., Lin, R., Wittmann, C., Bei, Y., Cha, Y. H., Ali, M., Priess, J. R., and Mello, C. C. (1997) Wnt signaling and an APC-related gene specify endoderm in early *C. elegans* embryos. *Cell* **90**, 707–716.
- 10 Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811.
- 10 Montgomery, M. K., Xu, S., and Fire, A. (1998) RNA as a target of double-stranded RNA-mediated genetic interference in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* **95**, 15,502–15,507.
- 11 Simpson, V. J., Johnson, T. E., and Hammen, R. F. (1986) *C. elegans* DNA does not contain 5-methylcytosine at any time during development or aging. *Nucl. Acids Res.* **14**, 6711–6717.
- 12 Napoli, C., Lemieux, C., and Jorgensen, R. (1990) Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. *Plant Cell* **2**, 279–289.
- 13 Romano, N. and Macino, G. (1992) Quelling: transient inactivation of gene expression in *Neurospora crassa* by transformation with homologous sequences. *Mol. Microbiol.* **6**, 3343–3353.
- 14 Cogoni, C., Irelan, J. T., Schumacher, M., Schmidhauser, T. J., Selker, E. U., and Macino, G. (1996) Transgene silencing of the *al-1* gene in vegetative cells of *Neurospora* is mediated by a cytoplasmic effector and does not depend on DNA-DNA interactions or DNA methylation. *EMBO J.* **15**, 3153–3163.
- 15 Matzke, M. A. and Matzke, A. (1995) How and why do plants inactivate homologous (trans)genes? *Plant Physiol.* **107**, 679–685.
- 16 Montgomery, M. K. and Fire, A. (1998) Double-stranded RNA as a mediator in sequence-specific genetic silencing and co-suppression. *Trends Genet.* **14**, 255–258.
- 17 Waterhouse, P. M., Graham, M. W., and Wang, M. B. (1998) Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proc. Natl. Acad. Sci. USA* **95**, 13,959–13,964.

- 18 Jorgensen, R. A., Que, Q., and Stam, M. (1999) Do unintended anti-sense transcripts contribute to sense cosuppression in plants? *Trends Genet.* **15**, 11, 12.
- 19 Williams, B. R. (1999) PKR; a sentinel kinase for cellular stress. *Oncogene* **18**, 6112–6120.
- 20 Wianny, F. and Zernicka-Goetz, M. (2000) Specific interference with gene function by double-stranded RNA in early mouse development. *Nat. Cell Biol.* **2**, 70–75.
- 21 Oates, A. C., Bruce, A. E., and Ho, R. K. (2000) Too much interference: injection of double-stranded RNA has nonspecific effects in the zebrafish embryo. *Dev. Biol.* **224**, 20–28.
- 22 Morel, J. B., Godon, C., Mourrain, P., Beclin, C., Boutet, S., Feuerbach, F., Proux, F., and Vaucheret, H. (2002) Fertile hypomorphic ARGONAUTE (*ago1*) mutants impaired in post-transcriptional gene silencing and virus resistance. *Plant Cell* **14**, 629–639.
- 22 Marathe, R., Anandalakshmi, R., Smith, T. H., Pruss, G. J., and Vance, V. B. (2000) RNA viruses as inducers, suppressors and targets of post-transcriptional gene silencing. *Plant Mol. Biol.* **43**, 295–306.
- 27 Hutvagner, G. and Zamore, P. D. (2002) RNAi: nature abhors a double-strand. *Curr. Opin. Genet. Dev.* **12**, 225–232.
- 25 Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363–366.
- 26 Hammond, S. M., Bernstein, E., Beach, D., and Hannon, G. J. (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404**, 293–296.
- 27 Zamore, P., Tuschl, T., Sharp, P., and Bartel, D. (2000) RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**, 25–33.
- 28 Sijen, T., Fleenor, J., Simmer, F., Thijssen, K. L., Parrish, S., Timmons, L., Plasterk, R. H., and Fire, A. (2001) On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* **107**, 465–476.
- 29 Tijsterman, M., Ketting, R. F., Okihara, K. L., Sijen, T., and Plasterk, R. H. (2002) RNA helicase MUT-14-dependent gene silencing triggered in *C. elegans* by short antisense RNAs. *Science* **295**, 694–697.
- 30 Lipardi, C., Wei, Q., and Paterson, B. M. (2001) RNAi as random degradative PCR: siRNA primers convert mRNA into dsRNAs that are degraded to generate new siRNAs. *Cell* **107**, 297–307.
- 31 Schwarz, D. S., Hutvagner, G., Haley, B., and Zamore, P. D. (2002) Evidence that siRNAs function as guides, not primers, in the *Drosophila* and human RNAi pathways. *Mol. Cell* **10**, 537–548.
- 32 Stein, P., Svoboda, P., Anger, M., and Schultz, R. M. (2003) RNAi: mammalian oocytes do it without RNA-dependent RNA polymerase. *RNA* **9**, 187–192.

- 33 Roignant, J.-Y., Carre, C., Mugat, B., Szymczak, D., Lepesant, J.-A., and Antoinewski, C. (2003) Absence of transitive and systemic pathways allows cell-specific and isoform-specific RNAi in *Drosophila*. *RNA* **9**, 299–308.
- 34 Cogoni, C. and Macino, G. (1997) Isolation of *quelling-defective (qde)* mutants impaired in posttranscriptional transgene-induced gene silencing in *Neurospora crassa*. *Proc. Natl. Acad. Sci. USA* **94**, 10,233–10,238.
- 35 Doench, J. G., Petersen C. P., and Sharp, P. A. (2003) siRNAs can function as miRNAs. *Genes Dev.* **17**, 438–442.
- 36 Wassenegger, M. (2000) RNA-directed DNA methylation. *Plant Mol. Biol.* **43**, 203–220.
- 37 Matzke, M. A., Matzke, A. J. M., Pruss, G., and Vance, V. B. (2001) RNA-based silencing strategies in plants. *Curr. Opin. Genet. Dev.* **11**, 221–227.
- 38 Zilberman, D., Cao, X., and Jacobsen, S. E. (2003) ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* **299**, 716–719.
- 39 Tang, G., Reinhart, B. J., Bartel, D. P., and Zamore, P. D. (2003) A biochemical framework for RNA silencing in plants. *Genes Dev.* **17**, 49–63.
- 40 Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I., and Martienssen, R. A. (2002) Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**, 1833–1837.
- 41 Hall, I. M., Noma, K., and Grewal, S. I. S. (2003) RNA interference machinery regulates chromosome dynamics during mitosis and meiosis in fission yeast. *Proc. Natl. Acad. Sci. USA* **100**, 193–198.
- 42 Mochizuki, K., Fine, N. A., Fujisawa, T., and Gorovsky, M. A. (2002) Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in tetrahymena. *Cell* **110**, 689–699.
- 43 Kamath, R. S., Fraser, A. G., Dong, Y., et al. (2003) Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* **421**, 231–237.
- 44 Ashrafi, K., Chang, F. Y., Watts, J. L., Fraser, A. G., Kamath, R. S., Ahringer, J., and Ruvkun, G. (2003) Genome-wide RNAi analysis of *Caenorhabditis elegans* fat regulatory genes. *Nature* **421**, 268–272.
- 45 Frankish, H. (2003) Consortium uses RNAi to uncover genes' function. *Lancet* **361**, 584.
- 46 Williams, N. S., Gaynor, R. B., Scoggin, S., Verma, U., Gokaslan, T., Simmang, C., Fleming, J., Tavana, D., Frenkel, E., and Becerra, C. (2003) Identification and validation of genes involved in the pathogenesis of colorectal cancer using cDNA microarrays and RNA interference. *Clin. Cancer Res.* **9**, 931–946.
- 47 Sanchez Alvarado, A. and Newmark, P.A. (1999) Double-stranded RNA specifically disrupts gene expression during planarian regeneration. *Proc. Natl. Acad. Sci. USA* **96**, 5049–5054.
- 48 Lohmann, J. U., Endl, I., and Bosch, T. C. (1999) Silencing of developmental genes in Hydra. *Dev. Biol.* **214**, 211–214.

- 49 Baker, M. W. and Macagno, E. R. (2000) RNAi of the receptor tyrosine phosphatase HmLAR2 in a single cell of an intact leech embryo leads to growth-cone collapse. *Curr. Biol.* **10**, 1071–1074.
- 50 Nakano, H., Amemiya, S., Shiokawa, K., and Taira, M. (2000) RNA interference for the organizer-specific gene *Xlim-1* in *Xenopus* embryos. *Biochem. Biophys. Res. Commun.* **274**, 434–439.
- 51 Brown, S. J., Mahaffey, J. P., Lorenzen, M. D., Denell, R. E., and Mahaffey, J. W. (1999) Using RNAi to investigate orthologous homeotic gene function during development of distantly related insects. *Evol. Dev.* **1**, 11–15.
- 52 Hughes, C. L. and Kaufman, T. C. (2000) RNAi analysis of *Deformed*, *proboscipedia* and *Sex combs reduced* in the milkweed bug *Oncopeltus fasciatus*: novel roles for *Hox* genes in the hemipteran head. *Development* **127**, 3683–3694.
- 53 Schroder, R. (2003) The genes *orthodenticle* and *hunchback* substitute for *bicoid* in the beetle *Tribolium*. *Nature* **422**, 621–625.
- 54 Haag, E. S. and Kimble, J. (2000) Regulatory elements required for development of *Caenorhabditis elegans* hermaphrodites are conserved in the *tra-2* homologue of *C. remanei*, a male/female sister species. *Genetics* **155**, 105–116.
- 55 Rudel, D. and Kimble, J. (2001) Conservation of *glp-1* regulation and function in nematodes. *Genetics* **157**, 639–654.
- 56 Louvet-Vallee, S., Kolotuev, I., Podbilewicz, B., and Felix, M. A. (2003) Control of vulval competence and centering in the nematode *Oscheius* sp. 1 CEW1. *Genetics* **163**, 133–146.
- 57 Winston, W. M., Molodowitch, C., and Hunter, C. P. (2002) Systemic RNAi in *C. elegans* requires the putative transmembrane protein SID-1. *Science* **295**, 2456–2459.
- 58 Robinson, D. A., Dillon, C. P., Kwiatkowski, A. V., et al. (2003) A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference. *Nat. Genet.* **33**, 401–406.
- 59 Stewart, S. A., Dykxhoorn, D. M., Palliser, D., et al. (2003) Lentivirus-delivered stable gene silencing by RNAi in primary cells. *RNA* **9**, 493–501.
- 60 Kawasaki, H. and Taira, K. (2003) Short hairpin type of dsRNAs that are controlled by tRNA^{Val} promoter significantly induce RNAi-mediated gene silencing in the cytoplasm of human cells. *Nucleic Acids Res.* **31**, 700–707.
- 61 Song, E., Lee, S. K., Wang, J., Ince, N., Ouyang, N., Min, J., Chen, J., Shankar, P., and Lieberman, J. (2003) RNA interference targeting Fas protects mice from fulminant hepatitis. *Nat. Med.* **9**, 347–351.
- 62 Jiang, M. and Milner, J. (2002) Selective silencing of viral gene expression in HPV-positive human cervical carcinoma cells treated with siRNA, a primer of RNA interference. *Oncogene* **21**, 6041–6048.
- 63 Shlomai, A. and Shaul, Y. (2003) Inhibition of hepatitis B virus expression and replication by RNA interference. *Hepatology* **37**, 764–770.
- 64 Kapadia, S. B., Brideau-Andersen, A., and Chisari, F. V. (2003) Interference of hepatitis C virus RNA replication by short interfering RNAs. *Proc. Natl. Acad. Sci. USA* **100**, 2014–2018.

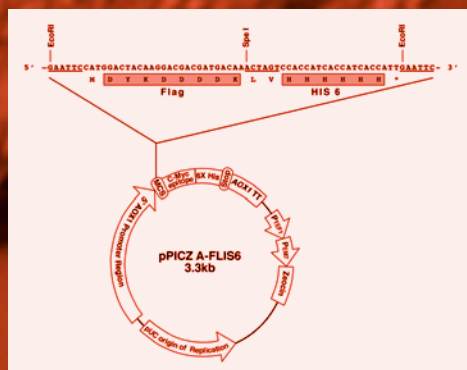
- 65 Jia, Q. and Sun, R. (2003) Inhibition of gamma herpesvirus replication by RNA interference. *J. Virol.* **77**, 3301–3306.
- 66 Yamamoto, T., Omoto, S., Mizuguchi, M., Mizukami, H., Okuyama, H., Okada, N., Saksena, N. K., Brisibe, E. A., Otake, K., and Fuji, Y. R. (2002) Double-stranded nef RNA interferes with human immunodeficiency virus type 1 replication. *Microbiol. Immunol.* **46**, 809–817.
- 67 Lichner, Z., Silhavy, D., and Burgyan, J. (2003) Double-stranded RNA-binding proteins could suppress RNA interference-mediated antiviral defences. *J. Gen. Virol.* **84**, 975–980.
- 68. Hamada, M., Ohtsuka, T., Kawaida, R., Koizumi, M., Morita, K., Furukawa, H., Imanishi, T., Miyagishi, M., and Taira, K. (2002) Effects on RNA interference in gene expression (RNAi) in cultured mammalian cells of mismatches and the introduction of chemical modifications at the 3'-ends of siRNAs. *Antisense Nucleic Acid Drug Dev.* **12**, 301–309.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



Methods for Delivery of Double-Stranded RNA into *Caenorhabditis elegans*

Dawn Hull and Lisa Timmons

Summary

The nematode *Caenorhabditis elegans* is often employed in investigations of diverse aspects of biology, including behavior, development, basic cellular processes, and disease states. The ability to utilize double-stranded RNA (dsRNA) to inhibit specific gene function in this organism has dramatically increased its value for these kinds of studies and has provided more flexibility in experimental design that include procedures. Here, we have collected a set of protocols from the *C. elegans* community for propagation of *C. elegans*, techniques for dsRNA preparation, four basic methods for delivery of dsRNA to *C. elegans* (injection, soaking, feeding, and in vivo delivery), and we suggest schemes that should facilitate detection of specific gene silencing.

Key Words: Double-stranded RNA; single-stranded RNA; RNA silencing; feeding; soaking; injection; HT115; HT115(DE3); transgene; transcription; RNAi.

1. Introduction

Caenorhabditis elegans is a small (1 mm long), nonparasitic nematode easily managed in the laboratory. The worm is cultured in Petri dishes using bacteria as a food source; has a short life-span (2 to 3 wk); and is transparent, with all 959 somatic cells visible under a microscope. The completed DNA sequence and the ability to generate mutants have made this an attractive system for analysis of gene function, and RNA silencing technology (RNA interference [RNAi]) has opened up amazing possibilities for genetic manipulations. RNAi in *C. elegans* affords particular advantages: First, long double-stranded RNAs (dsRNA) can be utilized as trigger molecules, and these are synthesized easily and inexpensively by in vitro transcription. Second, *C. elegans* apparently does not exhibit nonsequence-specific responses, such as the interferon/

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

protein kinase R responses, that are prevalent in mammalian systems. Third, RNA silencing molecules are readily imported into the cells of most tissues. Fourth, several methods are available for delivering dsRNA to *C. elegans*, as individuals or a population. In addition, RNAi has proven particularly valuable in uncovering roles for genes in the developing germline—roles that may not be observable in genetic mutants owing to maternal deposition of gene product (1).

RNAi in *C. elegans* entails manufacture of dsRNA, delivery of dsRNA to worms, and extensive phenotypic analysis of affected animals. We present four protocols that should allow efficient delivery of dsRNA molecules to the animal: *injection* involves preparing dsRNA by in vitro transcription followed by injection of the dsRNA into animals; *soaking* is a forced ingestion or absorption of dsRNA by suspension of animals in concentrated dsRNA solutions; *feeding* relies on ingestion of bacteria engineered to express dsRNA; and *in vivo delivery* is accomplished by creating DNA constructs designed to express dsRNA from *C. elegans* promoters and transforming such constructs into worms to generate transgenic lines.

1.1. Considerations for the mRNA Target

Many genes, when mutated, elicit no obvious phenotypes in the laboratory setting; thus, the corresponding RNAi experiments may not result in a phenotype. Additionally, some sequences are more efficiently targeted for RNA silencing than others. The intrinsic stability of an mRNA, its accessibility to the RNAi machinery, and additional factors may determine susceptibility. In some instances, the preexisting protein, as well as the mRNA, must be diminished before a phenotype is observed. A comparative analysis of dsRNA-treated animals vs mock-treated animals by *in situ* hybridization or protein immunolocalization ensures proper execution of the protocol. Another indication that the gene of interest is not amenable to RNA silencing can be revealed by a failure to reduce green fluorescent protein (GFP) fluorescence in a transgenic animal expressing a *gene of interest::gfp* fusion following treatment with dsRNA corresponding to the gene target. Some genes expressed in the nervous system, in the pharynx, and in males can prove to be difficult targets for silencing (2), but the use of animals defective for *rrf-3* has enabled some of these limitations to be overcome (3). *rrf-3* encodes a protein with some homology to RNA-dependent RNA polymerases (4) and is one of several genes that function to suppress RNA-silencing mechanisms.

1.2. Considerations for Choosing an Appropriate Region from the Gene of Interest for dsRNA Trigger

It is not necessary, nor is it always advisable, to use the entire coding region of a gene as a template for dsRNA synthesis. The particular sequence chosen

for use as a dsRNA trigger can have important consequences because distinct sequences derived from the same gene may elicit dissimilar RNAi results. When selecting a DNA sequence for use as a template, one should consider the following:

1. *Region*: dsRNA trigger sequences should be derived from exons and not introns or promoter regions. Since many *C. elegans* genes are interspersed with small introns, it may be necessary to obtain a cDNA clone or to generate one using reverse transcriptase followed by DNA amplification (RT-PCR) in order to obtain a contiguous stretch of exonic sequence sufficient to trigger robust RNAi. cDNAs are available from several sources, most notably from the expressed sequence tag database of Yuji Kohara: <http://nematode.lab.njg.ac.jp/dbest/keysrch.html>.
2. *Length*: Longer dsRNAs may be more effective in eliciting RNAi since they are fragmented into a greater number of effector siRNA molecules than shorter dsRNAs. However, when longer sense and antisense strands are synthesized in separate tubes, efficient annealing of the two strands may prove difficult. Two hundred to 300 bp of sequence is often sufficient for RNAi, and in most cases, strands of this length anneal fairly readily.
3. *Specificity*: It is possible to target multiple RNAs for destruction using only one dsRNA trigger. If this is not the intent, BLAST or more sophisticated algorithms should be used under low-stringency search conditions to compare the chosen sequence to both *C. elegans* genomic DNA and expressed sequences. It may prove advisable to avoid any region of the trigger sequence that has more than 15–20 bp of perfect homology to a nontarget sequence.

1.3. Considerations for Choosing Delivery Method

It can be difficult to predict the best method for silencing a given gene target. Each of the delivery methods has particular advantages and limitations, and when the null phenotype is not known, all methods should be tried. The relative dosage of dsRNA may account for some limitations. For injection and soaking, a wide range of dsRNA concentrations can be utilized; however, the dosage of dsRNA delivered by soaking, as well as feeding, relies on an adequate ingestion or absorption by the animal. Feeding also relies on bacterial accumulation of dsRNA. The feeding protocol can be less effective for silencing in older embryos and L1 larvae because eggshell deposition prevents uptake of maternally deposited dsRNA, and animals are thus shielded from exposure to dsRNA until they feed again upon hatching. Conversely, the feeding protocol can elicit phenotypes that are not generally observed in injected animals, possibly owing to long-term or continuous exposure of worms to dsRNA. For in vivo delivery of dsRNAs from transgenes, the timing and location of silenced cells is determined by the transcriptional properties of the transgene as well as the ability of RNA-silencing signals to spread from the tissue of origin.

1.4. Considerations for Experimental Reproducibility

The penetrance and expressivity of RNAi phenotypes can vary between experiments. Growth temperature, developmental stage at the time of dsRNA delivery, or the presence of contaminating microorganisms can contribute to efficacy, and the rate of protein clearance may be affected by outside variables that may be manipulable. To fully assess an RNAi phenotype, it is advisable to administer dsRNA to animals of the same developmental stage, clone treated animals on individual culture plates, group plates into sets, and place sets under different growth conditions. Each animal and its progeny should be monitored for phenotype(s).

The methods we have provided include detailed descriptions of techniques that can be considered worm-specific or method-of-delivery-specific, or that have demonstrated advantages to the task at hand. Since space is limited, we have not provided protocols for many basic cloning and molecular techniques—especially those that might be gene specific (such as PCR conditions)—because these are available elsewhere (5). RNAi results can vary not only among laboratories but also among experiments performed by the same investigator; thus, it is important to repeat experiments and use different delivery methods for the same gene target. False negative results abound, especially from the large genomewide screens, and phenotypes that require more detailed analysis to decipher are often missed. Finally, remember that although RNAi can reliably produce a null phenocopy for some genes, it is by no means a substitute for genetic mutants.

2. Materials

2.1. *C. elegans* Husbandry

Wild-type *C. elegans* (N2 is a commonly used wild-type strain) and mutant worm stocks, as well as *Escherichia coli* OP50 bacteria, are available from the Caenorhabditis Genetics Center (CGC), a central repository for strains under contract from the NIH National Center for Research Resources (<http://biosci.umn.edu/CGC/CGChomepage.htm>). (A streptomycin-resistant OP50 strain is also available and can be useful for minimizing unwanted bacterial contamination when streptomycin is included in the culture plates.) A common means of rearing worms is to culture them on nematode growth medium (NGM) plates seeded with OP50 bacteria. OP50 is a uracil auxotroph with a slower growth rate than wild-type *E. coli* (6) and is not resistant to antibiotics. More complete guides for worm husbandry that include freezing and thawing protocols (for long-term storage) and decontamination protocols are available elsewhere (7). Protocols for bacterial maintenance and so on are also available (5).

2.1.1. Bacterial Culture

1. Erlenmeyer flask (2 L), Bacto-tryptone, Bacto-yeast extract, NaCl, agar, Teflon-coated stir bar, aluminum foil, stir plate, appropriate antibiotics.
2. *E. coli* OP50 strain.

2.1.2. Preparation of NGM Plates

1. Erlenmeyer flask (2 L), Bacto-tryptone, agar, NaCl, 10 mg/mL of cholesterol in ethanol, 2 M Tris base (sterilize by autoclaving), 3.5 M Tris-HCl (sterilize by autoclaving).
2. Sterile Petri plates (60 mm diameter).

2.1.3. Maintenance of *C. elegans* Stocks

1. Dissecting stereomicroscope with transmitted light source ($\times 1$ to $\times 65$ magnification).
2. Thin metal spatula.
3. Worm pick: This can be fashioned from a 1- to 2-in piece of 30- to 32-gage platinum wire mounted into a bacteriological loop holder or onto a glass Pasteur pipet (by melting the tip of the pipet around the base of the wire). The tip of the wire can be hammered flat and shaped into a tapered point by cutting with scissors. Alternatively, the tip can be shaped into a hook or loop using dissection forceps.
4. Refrigerated incubator(s).

2.2. Generation of DNA Templates Used in dsRNA Production

The dsRNA trigger used in soaking or injection protocols is synthesized using a DNA template that consists of a bacteriophage promoter sequence properly oriented with respect to the gene of interest. PCR products, as well as plasmids, can be used as templates for in vitro transcription reactions. In vivo delivery of dsRNA requires a DNA template with a *C. elegans* promoter to drive expression of dsRNA. dsRNA delivery by bacterial feeding requires either a plasmid with two T7 promoter sites in inverted configuration flanking the gene of interest, or a plasmid with a single T7 promoter site preceding two copies of the gene of interest, in inverted repeat configuration. The protocols in this chapter outline some required elements for transcription from plasmid templates, strategies for transcription from PCR templates, points to consider when choosing the target region for the gene of interest, and assembly of the final DNA construct to be used as a template for in vitro transcription.

2.2.1. Preparation of Plasmid DNA Templates for dsRNA Production

Plasmid vectors are available from several suppliers of molecular biology products. These should contain bacteriophage promoters flanking a restriction bank. Transcription from these promoters proceeds toward the restriction

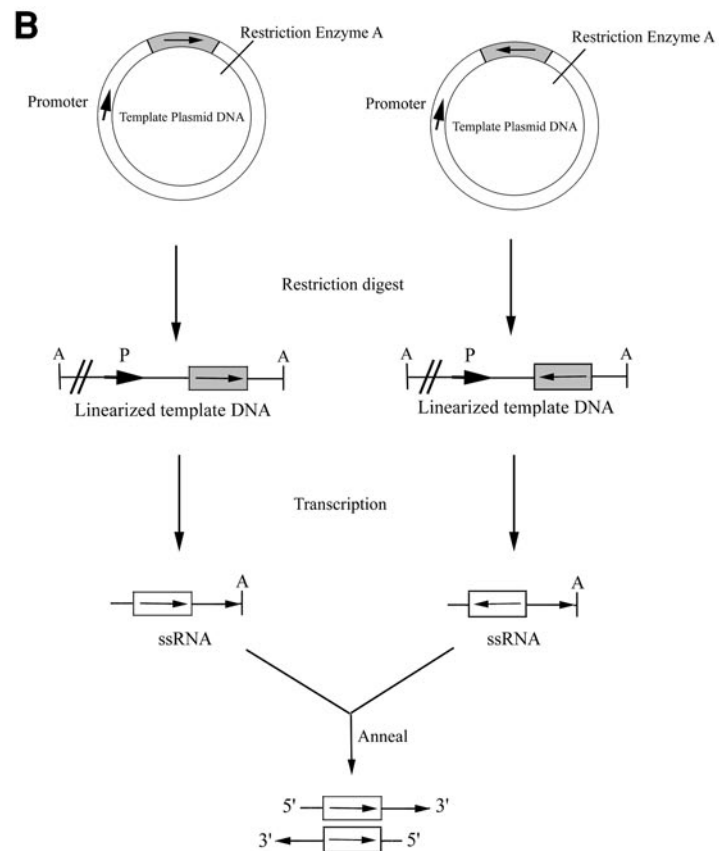
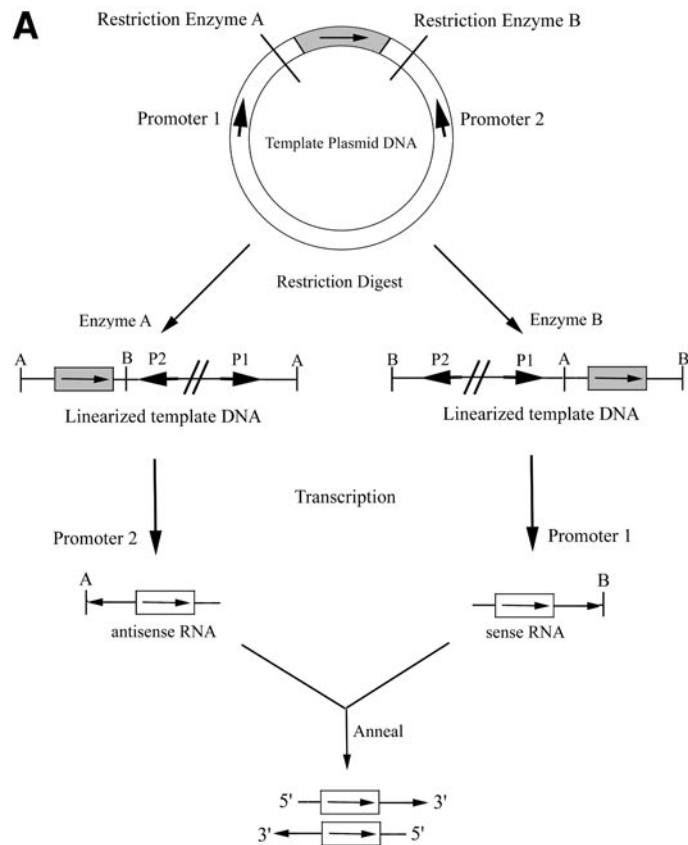
bank—where the target gene sequence (RNAi trigger) will be inserted (**Figs. 1–3**; gray boxes represent the target gene sequence). The promoter sequences generally employed in these vectors are derived from a defined promoter region from a gene in T7, T3, or SP6 bacteriophage; however, the precise promoter sequence may not be identical in all vectors. The vector may have a combination of two distinct promoters (T3 and T7, or T7 and SP6; *see Fig. 1A*), or two identical promoters in inverted repeat orientation flanking a restriction bank (double T7; *see Fig. 2*), or a single promoter (generally T7; *see Figs. 1B and 3*). A double-T7 plasmid (L4440) is available from the Caenorhabditis Genetics Center (<http://biosci.umn.edu/CGC/CGChomepage.htm>).

1. Plasmid vector, plasmid with DNA corresponding to target gene, restriction enzymes, supplies for agarose gel electrophoresis and DNA analysis.
2. DNA affinity purification kit, phenol/chloroform/isoamyl alcohol (25:24:1), 100% ethanol, 3 M sodium acetate (pH 5.2), 70% ethanol.

2.2.2. Preparation of PCR-Amplified Templates for dsRNA Production

1. N2 worms cultured on OP50-seeded plates with a thin layer of 2% agarose overlay, 15-mL conical centrifuge tubes, sterile dH₂O; Pasteur pipets, clinical centrifuge, benchtop microfuge, vortexer.

Fig. 1. (*see facing page*) Plasmid configurations for in vitro dsRNA synthesis. (**A**) Dual-promoter plasmids. The represented plasmid contains two oppositely positioned bacteriophage promoters (bold arrows) and a section of the gene of interest (gray box) pasted between restriction sites A and B. Linearized plasmid can be used to transcribe sense and antisense RNA strands (white boxes)—two separate RNA transcription reactions are required for this. An antisense strand is generated by linearizing the plasmid DNA at the position of enzyme A and using the RNA polymerase corresponding to promoter 2. A sense strand can be generated in a separate tube by linearizing the plasmid with enzyme B and using the RNA polymerase corresponding to promoter 1. Annealing of sense and antisense strands produces a dsRNA trigger for injection and soaking. (**B**) Single-promoter plasmids. A typical single-promoter plasmid with one copy of the *C. elegans* gene of interest (gray boxes) pasted behind a bacteriophage promoter (bold arrow is shown). This configuration yields both sense and antisense strands (white boxes) provided that two versions of the plasmid—with the target gene in both orientations—are available. The two plasmids can be linearized with the same restriction enzyme (A), but in separate tubes so that each reaction can be monitored for completeness. The in vitro transcription reactions can be performed in separate tubes that are then mixed for annealing into dsRNA, or by combining both linearized plasmids (if both plasmids have the same promoter). Efficient annealing of strands is generally best when the strands are cotranscribed; however, it is sometimes desirable to have ssRNA preps as controls.



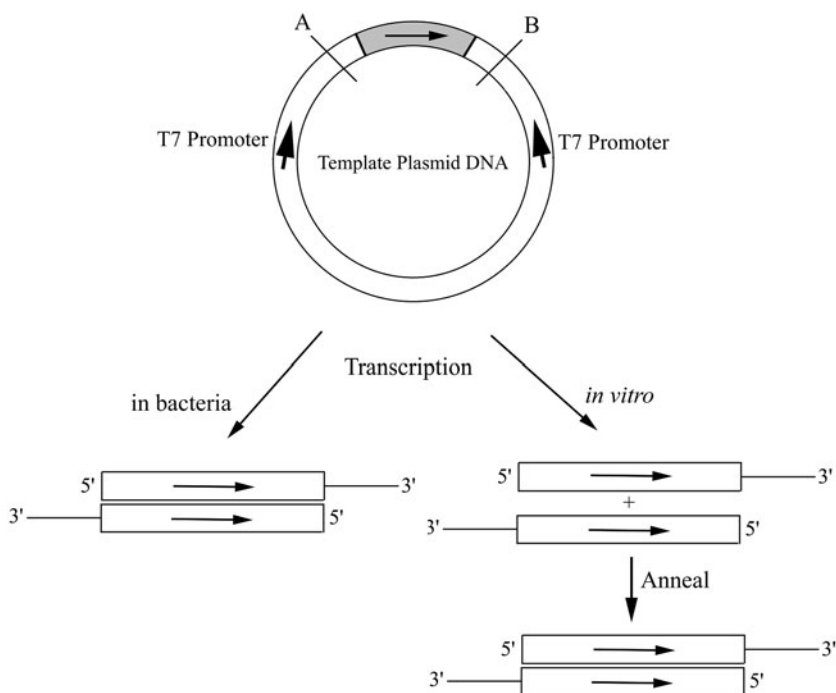


Fig. 2. Double-promoter configurations for *in vitro* dsRNA synthesis or for dsRNA feeding experiments. Plasmids contain two identical, inverted promoters (e.g., two T7 promoters, bold arrows) flanking a DNA insert (gray box). When utilized in the bacterial feeding method (left), T7 RNA polymerase is manufactured by the bacteria and utilizes the T7 promoter to synthesize dsRNA by cotranscription. This plasmid may also be used to generate dsRNA *in vitro* (right) (*see Subheadings 2.3. and 3.3.*; also *see Fig. 1A*). Two versions of linearized plasmid should be generated using restriction enzymes A and B in separate tubes. Since the same promoter is present at each end of the template, the linearized plasmids can be combined and one *in vitro* transcription reaction can be performed. Efficient annealing of strands is generally best when strands are cotranscribed.

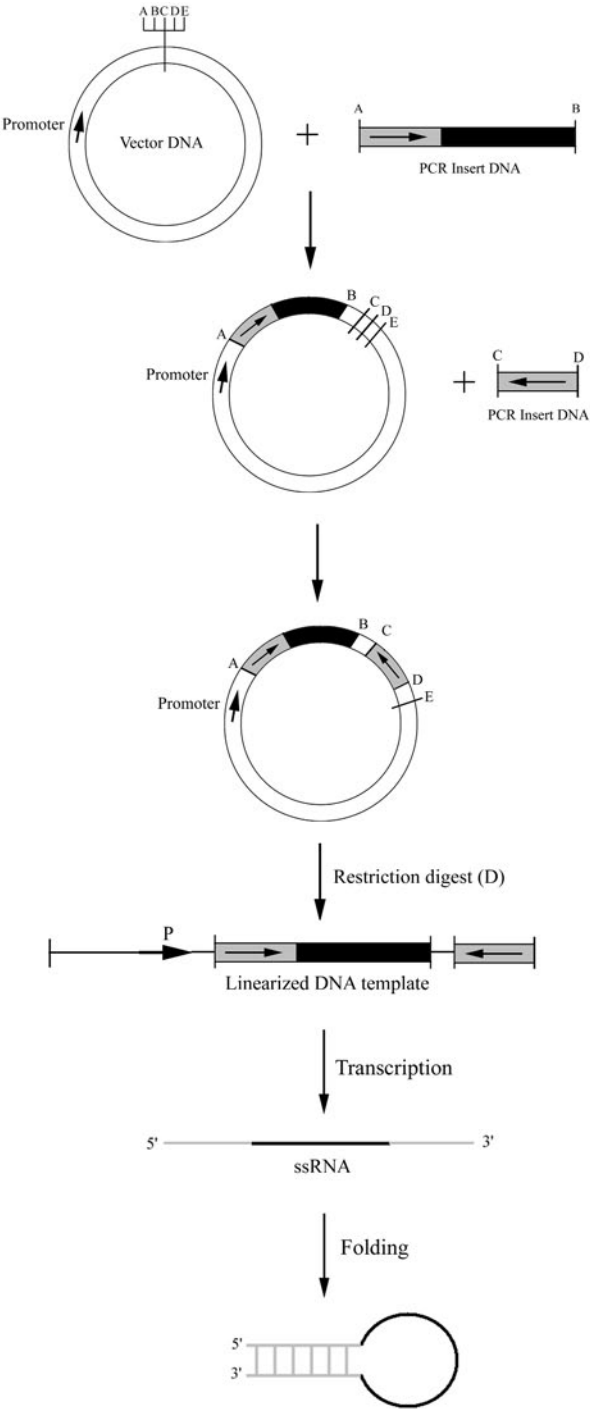
2. NTE: 100 mM NaCl; 50 mM Tris; 20 mM EDTA, proteinase K, 10% sodium dodecyl sulfate (SDS).
3. Water baths, incubators, or PCR machine for various incubations.
4. Phenol/chloroform (1:1), 100% ethanol, 3 M sodium acetate (pH 5.2), 70% ethanol, DNase-free RNase A.
5. Trizol reagent (cat. no. 15596-026; Life Technologies, Gaithersburg, MD), chloroform, RNA affinity purification kit, RNase-free DNase I.
6. Ultraviolet (UV) spectrophotometer and supplies.
7. Materials for reverse transcription reactions (commercial kits are available).

8. PCR primers with promoter sites at 5' ends, thermostable DNA polymerase and reagents for PCR amplification, PCR machine and supplies.
9. Agarose gel electrophoresis supplies for DNA analysis, DNA affinity purification kit.
10. Gloves.
11. PCR primers: Each primer should have a bacteriophage promoter sequence at its 5' end followed by sequences corresponding to the targeted gene. The melting temperature of the primer should be calculated based on the *C. elegans*-derived sequences alone because the promoter sequences will not hybridize to *C. elegans* DNA. This melting temperature should be higher than standard annealing conditions in PCR reactions. (Web-based sources such as <http://alces.med.umn.edu/rawtm.html> can aid in calculating the T_m for each primer.) The same promoter site can be included in both primers to generate a double-promoter construct. This strategy has the advantage that dsRNA can be synthesized in one tube. Two separate in vitro transcription reactions may be required if two different promoter sites are incorporated into the PCR primers.

2.3. Production of dsRNA by In Vitro Transcription

dsRNA of sufficient quantities (several micrograms or more) for most soaking and injection experiments can be obtained relatively easily using in vitro transcription protocols. These reactions employ a simple bacteriophage RNA polymerase and a DNA template with promoter sequences corresponding to the RNA polymerase. The RNA polymerase binds to the promoter and synthesizes a copy of single-stranded RNA (ssRNA) in a directional and strand-specific manner (relative to the DNA template) that is specified by the orientation of the promoter sequence. PCR products as well as plasmids can be used as templates. In vitro transcription kits are available from several commercial sources. These kits are supplied with detailed instructions, and since other protocols are readily available (5), we provide here only a general outline for performing these reactions. In vitro transcription involves preparation of the template DNA (purification and restriction digestion); synthesis of RNA; annealing of sense and antisense strands; and, finally, analysis of the ssRNA and dsRNA products.

1. Purified template DNA (plasmid or PCR product; see Figs. 1, 3, and 4), restriction enzymes, agarose gel electrophoresis apparatus, UV transilluminator.
2. 0.5X TAE: 20 mM Tris-acetate, 0.5 mM EDTA, 0.3 µg/mL of ethidium bromide [EtBr]; filtered; 1% agarose gel for RNA analysis (gel should be made and run in 0.5X TAE).
3. DNA affinity purification kit, phenol/chloroform/isoamyl alcohol (25:24:1).
4. 100% ethanol.
5. 3 M sodium acetate (pH 5.2).
6. 70% ethanol.



7. Reagents for transcription reactions: 200 mM dithiothreitol (DTT), ribonucleotide solution (GTP, ATP, UTP, and CTP, each at 5 mM), nuclease-free H₂O, 10X transcription buffer (standard buffer: 400 mM Tris, pH 7.5; 100 mM NaCl; 60 mM MgCl₂; 20 mM spermidine [5]), RNase inhibitor, bacteriophage RNA polymerases (e.g., T7, T3, and SP6). Alternatively, a commercially available in vitro transcription kit can be used.
8. TE: 25 mM Tris-HCl; 10 mM EDTA, pH 8.0; 3X injection buffer (20 mM phosphate buffer, pH 7.5; 3 mM potassium citrate, pH 7.5; 2% PEG 6000).
9. Incubators, water baths, or PCR machine for various incubation temperatures.
10. Gloves.

2.4. Delivery of dsRNA into *C. elegans* by Injection

RNAi by injection is the “classic” method for dsRNA delivery (8), and it is effective (dilute solutions of dsRNA can induce RNAi) and multipurpose (several distinct mRNAs can be targeted simultaneously using an injection mix comprising multiple dsRNAs) (9,10). (However, the RNAi machinery can be saturated—several groups have had difficulty targeting more than three separate genes at once.) Injection can also result in transmission of RNA-silencing signals to progeny, resulting in a larger subject population (8,11). With a workable injection system, even novice injectors can achieve RNAi successfully, because targeting the delivery needle to a specific tissue is not required for phenocopy production. Injected dsRNA can elicit systemic RNA silencing in the injected animal and its progeny even when the dsRNA is delivered to a body cavity (8). Injection allows some flexibility regarding the study of gene function at particular stages. For example, it is possible to analyze gene function in later-developing tissue for genes that are essential earlier in development. In these instances, dsRNA may be injected into young larvae (L1/L2 larvae), allowing the requirements for gene function to be examined at a later stage (e.g., L4 larvae or adults).

Fig. 3. (see opposite page) Plasmid used as template for transcribing hairpin dsRNAs in vitro. A plasmid containing the gene of interest configured as an inverted repeat (gray boxes) with a stuffer fragment (black boxes) flanked by a single promoter (bold arrows) can be generated in two steps. First, a fragment of the coding region corresponding to the gene of interest is inserted behind the promoter using restriction sites A and B. A second DNA fragment is then inserted behind the first using sites C and D. (Sites C and D can be created by appropriately designed PCR primers.) This suggested design for an inverted repeat makes use of sequences from the gene of interest as the stuffer region (black boxes). The purified plasmid is linearized at site D or E, and runoff transcription is performed using RNA polymerase. This plasmid configuration can be used in the feeding protocol if a bacteriophage T7 promoter is present.

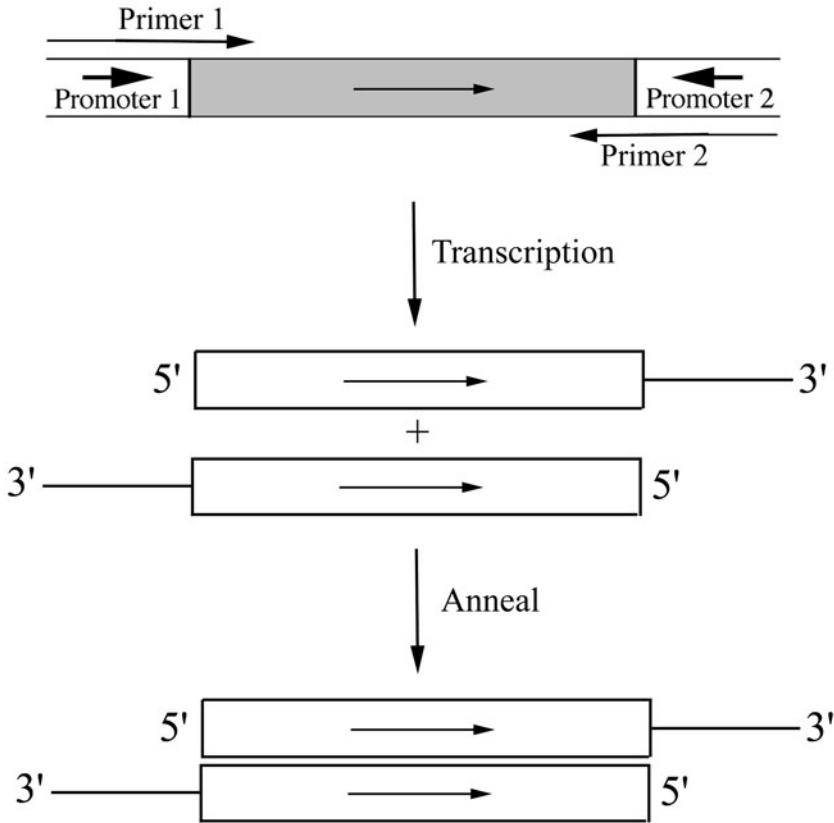


Fig. 4. PCR-amplified template DNA for RNA soaking and injection experiments. PCR product contains a DNA insert (gray box) flanked by two promoters. The promoter sites can be incorporated into the primer sequence, as indicated. The PCR fragment can be used directly (after cleanup) in an in vitro transcription reaction. A runoff transcript will be produced with either RNA polymerase. If the two promoters are identical, only one in vitro transcription reaction is required to generate dsRNA.

Delivery of dsRNA to *C. elegans* involves preparing a dsRNA injection mix, injection needles, and an “injection pad” of agarose; loading the mix into an injection needle; injecting the dsRNA into an animal; recovering the animal; and assessing any phenotype(s). More complete descriptions of microinjection are available (*12,13*). References for DNA injections (*14,15*) may provide additional information.

1. dsRNA (0.1–3 $\mu\text{g}/\mu\text{L}$) prepared by in vitro transcription (see **Subheadings 2.3** and **3.3**.) or other means.

2. Microinjection equipment: inverted microscope with attached micromanipulator, needle puller, Pasteur pipets, forceps, glass slides, standard borosilicate thin-wall filamented capillary tubes with outer diameter/inner diameter within the range of 1.0/0.58–1.5/0.84 mm. The microinjection needles will be made from the capillary tubes and thus the choice in size will depend on the design of the needle holder.
3. Freshly prepared 2% agarose solution in water, 60°C water bath or incubator, 37°C incubator, microfuge.
4. Large Petri dishes or other covered containers for holding needles.
5. Clay fashioned into a pencil shape and inserted into the needle container. Pulled needles can be mounted horizontally into the clay, preventing accidental breakage of tips. A similar setup can be used to hold needles loaded with dsRNA solution; a small paper towel moistened with water should be included to prevent dehydration of the injection solution.
6. Binocular dissection microscope with transmitted light source, mineral oil (heavy white oil; viscosity at 100°C: 340–360), M9 medium (for 1 L, 3 g of KH_2PO_4 , 6 g of Na_2HPO_4 , 5 g of NaCl, 1 mL of 1 M MgSO_4), recovery buffer (M9 media + 4% glucose).
7. OP50-seeded NGM plates (*see Subheadings 2.1.2. and 3.1.2.*).

2.5. Delivery of dsRNA by Soaking

An RNAi phenotype can be induced in *C. elegans* by soaking the worms in a concentrated solution of dsRNA made by in vitro transcription (**16,17**).

1. dsRNA (0.2–5 $\mu\text{g}/\mu\text{L}$), 1.5-mL microfuge tubes, sterile dH_2O and M9 medium (*see Subheading 2.4.*), appropriate strain of *C. elegans*.
2. 15°C Incubator, additional incubators at appropriate temperatures.
3. OP50-seeded NGM plates (*see Subheadings 2.1.2. and 3.1.2.*), mineral oil.
4. Appropriate microscope for phenotypic analysis.

2.6. Delivery of dsRNA by Feeding Worms dsRNA-Expressing Bacteria

The laboratory food source for *C. elegans* is bacteria, and bacterial strains have been established that can transcribe specific RNAs from engineered plasmids. Sufficient quantities of dsRNA accumulate within the bacterial cell such that these strains can be used as a food source for *C. elegans* and can also induce RNAi (**2**). The feeding protocol involves generating a plasmid with the gene of interest following a T7 bacteriophage promoter (production of these plasmids is described in **Subheadings 2.2.1. and 3.2.1.**; *see Figs. 2 and 3*), transforming an HT115(DE3) bacterial strain with such a plasmid, plating transformed cells on NGM plates under induction conditions, applying worms to plates, and monitoring phenotype(s).

1. DNA plasmid containing the gene of interest inserted between two T7 promoter sites (**Fig. 2**). A common double-T7 plasmid (Amp^r) is L4440, available from the CGC (<http://biosci.umn.edu/CGC/CGChomepage.htm>). Alternatively, DNA plas-

mid containing a single T7 promoter site followed by the gene of interest in inverted repeat orientation can be used (**Fig. 3**).

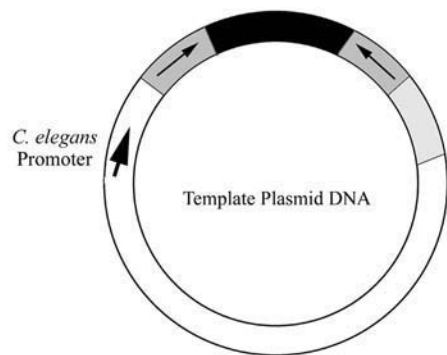
2. HT115(DE3) bacterial strain (**2**)—a tetracycline (Tet)-resistant, RNaseIII (–) strain available from the CGC; LB broth and agar plates; 12.5 mg/mL of Tet; 50–100 mg/mL of ampicillin (Amp); 37°C shaking incubator.
3. Cold 50 mM CaCl₂, filter sterilized; clinical centrifuge housed in a cold room or a refrigerated centrifuge; 15- and 50-mL sterile polypropylene centrifuge tubes.
4. 80% Glycerol, autoclaved; isopropyl-β-D-thiogalactopyranoside (IPTG): stock solution = 4 mM (1000X).
5. Plates (60 × 15 mm) containing NGM agar supplemented with 12.5 μg/mL of Tet, the antibiotic appropriate for plasmid selection (e.g., 50–100 μg/mL of Amp), and 0.4 mM IPTG; appropriate strain of *C. elegans*.

2.7. Production of dsRNA by In Vivo Transcription

Transcription of specific dsRNA within worm cells can be achieved by generating worm strains transformed with DNA constructs designed to express dsRNA. First, a plasmid is configured with a worm promoter and the gene of interest. The plasmid is then injected into the germline of a suitable worm strain. The DNA enters the nuclei of the syncytial germline, where the plasmids recombine/ligate into a “minichromosomalike” structure that is repetitive in nature and is maintained in cells as an extrachromosomal array (**15**). Multiple plasmids from the same injection mix are often incorporated into the array, and most arrays are marked with a dominant selectable marker that is derived from a plasmid in the injection mix. It is possible to observe RNAi when a transgene array is formed from a mix of two different plasmids with the same promoter: one plasmid capable of expressing the sense strand of an RNAi trigger and a second plasmid expressing the antisense strand (**18**). Alternatively, a single worm promoter can be placed in front of two copies of the gene of interest arranged as inverted repeats (**Fig. 5**).

1. Plasmid containing inverted DNA repeats (gray boxes in **Fig. 3**) and a stuffer fragment (black boxes in **Fig. 3**) flanked by one promoter and various restriction endonuclease sites (see **Subheadings 2.2.1.** and **3.2.1.2.**).
2. Plasmid containing a *C. elegans* promoter and 3′ untranslated region (UTR).

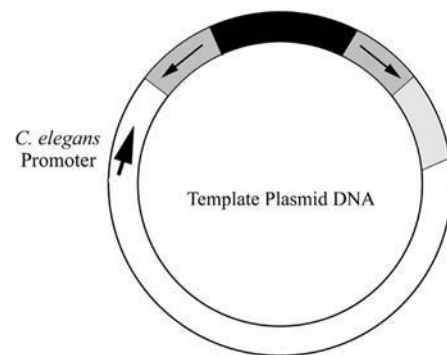
Fig. 5. (see facing page) Plasmid configuration for in vivo transcription of RNA hairpins. A hairpin-generating plasmid that contains inverted DNA repeats (gray boxes) and a stuffer fragment (black boxes) as in **Fig. 3** can be used for expression in *C. elegans* cells. A *C. elegans* promoter (bold arrows) should be inserted in front of the inverted repeat and a 3′ UTR (light gray) from a stably expressing gene should be inserted at the end of the inverted repeat. No differences in effectiveness have been reported to date between the configurations of inverted repeats depicted here. Such plasmids are used to generate transgenic worms that transcribe dsRNA in vivo.



Transcription



Folding



Transcription



Folding



3. Microinjection equipment and associated supplies (*see Subheading 2.4.*).
4. Injection mix composed of transformation marker plasmid (e.g., *rol-6* [22]) and dsRNA-expressing plasmid (*see Subheading 3.2.1.2.*).

3. Methods

3.1. *C. elegans* Husbandry

3.1.1. Bacterial Culture

1. Luria-Bertani (LB) medium, pH 7.5: Dissolve in a 2-L Erlenmeyer flask: 10 g/L of Bactotryptone, 5 g/L of Bacto yeast extract, 5 g/L of NaCl. Dispense into 100-mL bottles. Sterilize by autoclaving.
2. LB plates: Follow instructions for LB medium (do not dispense) and add 15 g/L of agar before autoclaving. Carefully drop a Teflon-coated stir bar into the flask, cover the flask with aluminum foil, and autoclave. (Agar will not dissolve unless heated. Use of the stir bar will speed cooling of the solution when removed from the autoclave and will prevent solidification at the bottom of the flask.) Stir the autoclaved solution until the temperature lowers to ~60°C. Add antibiotics as necessary. Pour into sterile 100-mm Petri dishes and allow to solidify.
3. Using sterile technique, streak the starter OP50 bacterial culture onto sterile LB agar plates and incubate, inverted, overnight at 37°C. The plate can be stored for routine use for several months at 4°C when sealed with Parafilm to prevent desiccation.
4. Inoculate a bottle of LB with a single colony from the OP50 culture plate. Incubate overnight at 37°C.
5. Store the OP50 liquid stock at 4°C. The stock can be used for several months to seed NGM plates, barring contamination (*see Subheadings 2.1.2. and 3.1.2.*).

3.1.2. Preparation of NGM Plates

1. Add the following to a 2-L Erlenmeyer flask: 1 mL of 2 M Tris base, 1 mL of 3.5 M Tris-Cl, 0.5 mL of 10 mg/mL cholesterol, 3 g of tryptone, 2 g of NaCl, 17 g of agar. Bring the volume to 1 L with dH₂O and carefully drop a Teflon-coated stir bar into the flask. Cover the flask with aluminum foil and autoclave.
2. Stir the autoclaved solution on a stir plate until the temperature lowers to ~60°C. Dispense into 60 × 15 mm Petri dishes under sterile conditions (*see Note 1*).
3. Allow time (overnight) for the plates to evaporate excess moisture. The plates can then be stored at 4°C or at room temperature in airtight containers.
4. Using a sterile technique, apply the OP50 bacterial culture dropwise onto the surface of the plates (~100 µL). Tilt the plates to spread out the lawn, but do not seed with so much bacteria that the lawn reaches the edge of the plate because worms may crawl up the plastic sides and dehydrate.
5. Allow the bacterial lawn to grow overnight at room temperature. Seeded plates can be stored in an airtight container for 2 to 3 wk at room temperature or at 15°C.

3.1.3. Maintenance of *C. elegans* Stocks

Worms can be transferred from plates lacking food (starved plates) to freshly seeded plates using a worm pick. The platinum pick is first flamed for sterility

before each attempt at transfer—this also prevents cross-contamination of stocks. Mounting the worms onto the pick is achieved using a gentle swiping motion to lift the worm off the plate. The animals are removed from the pick using a reverse motion into the bacterial lawn of the fresh plate. Bacteria can facilitate adherence of the worm onto the pick, acting as a sticky surface bridge between the pick and the worm. A successful transfer will gouge neither the worm nor the surface of the agar plate. Worms tend to burrow into the agar using these imperfections, and monitoring and collection of burrowed worms can prove difficult. Novice worm pickers may require a few days of practice to master the art of transfer.

Worms can also be transferred to a fresh plate by transferring a “chunk” of agar. This method quickly moves many worms to a fresh plate and is particularly useful when plates are starved, when the genotype of each worm is identical, or when a mating is not required to maintain the stock. To chunk worms, a metal spatula is flame-sterilized, then used to cut a cube of agar from the old plate. The spatula is then used as a spoon or shovel to scoop the worm-laden cube onto a new plate, preferably just adjacent to, but not directly on, the bacterial lawn.

C. elegans is generally reared at temperatures between 15 and 25°C (typically 20°C). Higher temperatures produce a faster growth rate: worms grow twice as fast at 25°C than at 16°C. Transferring worms every 1 to 2 d ensures a good supply of worms at every developmental stage.

3.2. Generation of DNA Templates Used in dsRNA Production

Subheading 3.2.1. describes methods that can be utilized when the DNA sequence corresponding to the dsRNA trigger is in hand (e.g., a cDNA). **Subheading 3.2.2.** describes methods to generate a trigger DNA by PCR or RT-PCR if the appropriate sequence is not readily available.

3.2.1. Assembly of Plasmid DNA Templates for dsRNA Production

3.2.1.1. ASSEMBLY OF DOUBLE- AND SINGLE-PROMOTER PLASMIDS HARBORING TRIGGER DNA

The template sequence (the *C. elegans* gene of interest) is inserted using restriction enzymes corresponding to sites in the restriction bank and standard cloning techniques (see **Figs. 1** and **2**; also see **Note 2**). For double-promoter templates, the final construct will contain a DNA insert (gray boxes in **Figs. 1A** and **2**) flanked by two promoters with some restriction endonuclease sites remaining. For single-promoter templates, the final construct will contain a DNA insert (gray boxes in **Fig. 1B**) flanked at one end by a promoter. For dsRNA production from single-promoter plasmids, it will be necessary to con-

struct two plasmids that differ with respect to gene orientation (**Fig. 1B**). dsRNA transcribed in vitro from these templates can be used for injection (*see Subheadings 2.4. and 3.4.*) and soaking (*see Subheadings 2.5. and 3.5.*) experiments; double-T7 plasmids can be utilized in the bacterial feeding protocol (*see Subheadings 2.6. and 3.6.*).

3.2.1.2. ASSEMBLY OF SINGLE-PROMOTER PLASMIDS WITH INVERTED-REPEAT CONFIGURATION OF TRIGGER DNA

The final construct (**Fig. 3**) will contain two oppositely oriented copies of the same DNA segment (gray boxes in **Fig. 3**) flanking a stuffer fragment (black box in **Fig. 3**). The promoter will be located at one end of the inverted repeat sequence, while the other end should contain at least one restriction endonuclease site (**Fig. 3**). A promoter sequence for RNA polymerases T7, T3, or SP6 can be utilized.

Use standard cloning techniques to insert a *C. elegans* coding region into the restriction bank of the plasmid. Choose a set of enzymes that leaves at least two restriction sites at one end of the restriction bank. If the DNA segment is cloned into a singly cut vector, it will be necessary to determine the orientation of the insert before proceeding.

Figure 3 depicts a cloning strategy in which the left-hand end of the inverted repeat is the first to be cloned into the vector. (A mirror image cloning strategy is also possible where the right portion of the inverted repeat is cloned into the vector first.) For the second fragment to be oriented properly, it may be necessary to synthesize PCR primers with restriction sites on the 5' ends that correspond to insertion sites in the vector (**Fig. 3**, sites C and D). It is often convenient and advantageous to generate the stuffer fragment from sequences derived from the target gene itself; this eliminates a cloning step and alleviates concern that a nonrelated stuffer fragment might elicit RNAi for another gene. This is accomplished when one of the sections is longer than the other (black boxes in **Fig. 3**). For the repeat to be properly maintained by bacterial cells, the stuffer should consist of at least 100 bp. Cloning of DNA with inverted repeats can be challenging, as they are often not faithfully maintained in bacteria. Transformation of ligations into a bacterial strain that is defective in multiple recombination mechanisms (such as SURE cells from Stratagene) can improve the cloning success rate.

Once the construct has been generated, a large, clean batch of DNA should be prepared. A variety of DNA purification kits that utilize affinity resins are available for this purpose. The final purified construct can be digested with a restriction enzyme at the end of the inverted repeat (not the promoter end; **Fig. 3**, site E) so that a runoff transcription product can be generated. dsRNA generated from such templates is suitable for soaking (*see Subheadings 2.5.*

and 3.5.) and injection (*see Subheadings 2.4. and 3.4.*) experiments and can be in bacterial feeding experiments if a T7 promoter site is present (*see Subheadings 2.6. and 3.6.*).

3.2.2. Assembly of PCR Templates for dsRNA Production

A PCR fragment can be generated (**Fig. 4**) that is composed of the coding region from the target gene of interest (gray box in **Fig. 4**) flanked by two promoters. The PCR reaction can be performed on purified genomic DNA (*see Subheading 3.2.2.1.*), directly on DNA released from worms (*see Subheading 3.2.2.2.*), or on cDNA synthesized from purified worm RNA (*see Subheading 3.2.2.3.*) The latter protocol is preferred when the target sequence is broken up into small exons; an RT-PCR fragment allows a more contiguous stretch of homology between trigger dsRNA and target mRNA. Sequences corresponding to T7, T3, and SP6 bacteriophage promoters are small enough that they can be incorporated into a PCR primer (**5,19**). However, SP6 RNA polymerase does not efficiently transcribe RNA from a PCR-amplified DNA template (**20**); therefore, primers with T7 or T3 promoter sequences are preferred.

3.2.2.1. PREPARATION OF N2 WORM GENOMIC DNA

1. Pick (not chunk) N2 worms onto 10 NGM/OP50 plates (100 × 15 mm) with a thin layer of 2% agarose overlayed and grow until worms are gravid. (The agarose prevents worms from digging into the plates and provides a barrier to the agar, which might contain impurities that could prevent subsequent molecular procedures.)
2. Collect worms from all stock plates into a 15-mL conical centrifuge tube by dislodging the worms from the plates with sterile-filtered dH₂O and pooling washes into the 15-mL tube using a Pasteur pipet.
3. Centrifuge the worms at low speed in a clinical centrifuge for 10–20 s to pellet the worms.
4. Remove the dH₂O by aspiration, and gently wash the worms with an additional 2 mL of dH₂O.
5. Repellet the worms at low speed in a clinical centrifuge for 10–20 s and aspirate the dH₂O.
6. Freeze the worm pellet at –80°C for at least 1 h (this helps crack open the worms).
7. Thaw the pellet on ice and then add 2 mL of NTE, 20 µg of proteinase K, and 100 µL of 10% SDS. Mix.
8. Incubate at 65°C for 2 h.
9. Transfer the worm mixture to microfuge tubes (500 µL/tube) and add an equal volume of phenol/chloroform, preequilibrated to room temperature.
10. Vortex the mixture well, and then centrifuge for 5 min in a room temperature microfuge at maximum speed.
11. Transfer the top aqueous phase to a new microfuge tube and repeat the phenol/chloroform extraction an additional two times (do not transfer more than 400 µL of aqueous solution to each new tube after the final extraction).

12. Add 1 mL of ice-cold 100% ethanol to each tube of 400 μ L and mix gently by inverting.
13. Centrifuge for 30 min at 4°C in a microfuge at maximum speed to pellet the DNA.
14. Carefully remove the ethanol and gently add 500 μ L of 70% ethanol to the pellet.
15. Centrifuge for 5 min at 4°C in a microfuge at maximum speed.
16. Carefully remove the ethanol and dry the pellet at room temperature (approx 1 h).
17. Dissolve the DNA in 400 μ L of sterile dH₂O.
18. Add 8 μ g of RNase A and incubate for 30 min at 37°C.
19. Phenol/chloroform extract the DNA three times (transfer 400 μ L of the aqueous DNA solution to each fresh tube after the final extraction).
20. Add 1/10 vol of 3 M sodium acetate, pH 5.2, and 1 mL of ice-cold 100% ethanol to each tube. Mix gently by inverting the tubes.
21. Centrifuge for 30 min at 4°C in a microfuge at maximum speed to pellet the DNA.
22. Carefully remove the ethanol and add 1 mL of ice-cold 70% ethanol, being careful not to dislodge the pellet.
23. Centrifuge for 5 min at 4°C in a microfuge at maximum speed and carefully remove the ethanol.
24. Air-dry the DNA pellet at room temperature (do not completely dry the pellet because this can cause the DNA to fragment).
25. Resuspend the DNA in 200 μ L of sterile dH₂O. Estimate the yield on a 0.7% agarose gel by comparing the EtBr staining intensity with that of a known concentration of marker DNA.
26. Use 0.1–1 μ g of this genomic DNA in a 50- μ L PCR reaction. To generate a PCR fragment for use in an in vitro transcription reaction (**Fig. 4**), use hybrid primers with bacteriophage promoter sequence at the 5' end (T7, T3, or Sp6) and target gene sequence at the 3' end (see **Note 3**) (23).

3.2.2.2. SINGLE WORM PCR OF TARGET SEQUENCE

1. In a PCR tube, add 9 μ L of sterile dH₂O, 1 μ L of 10X PCR polymerase buffer (with 1.0–2.5 mM Mg⁺²; 1.5 mM Mg⁺² is standard), and one to three gravid N2 worms (10- μ L total volume per tube).
2. Place at –80°C for at least 20 min (this helps crack open the worms).
3. Remove the tubes from the freezer and thaw on ice.
4. Add 0.5 μ L (~5–10 μ g) of proteinase K.
5. Incubate at 65°C for 1 h, then 95°C for 15 min. Proteinase K will help lyse the worm cuticle and also degrade proteins, especially DNases. A 95°C inactivation step is required so that the DNA polymerase will not be degraded when added later during the amplification step.
6. Add the remaining PCR reaction components (more buffer [with Mg⁺²], dH₂O, dNTPs, primers, and thermostable polymerase) to generate a PCR fragment for use in an in vitro transcription reaction (**Fig. 4**). Some parameters, such as Mg⁺² concentration and cycle conditions, will be sequence- or primer-specific. A number of sources are available for advice on these matters (5).

3.2.2.3. TOTAL RNA PREPARATION FROM N2 WORMS AND RT-PCR AMPLIFICATION OF TRIGGER DNA

1. Grow N2 worms on a minimum of 20 NGM/OP50 plates (60×15 mm) until gravid.
2. Collect the worms from all stock plates into one 15-mL conical tube using sterile-filtered dH₂O and a Pasteur pipet.
3. Centrifuge the worms at low speed in a clinical centrifuge for 10–20 s to pellet the worms.
4. Remove the dH₂O by aspiration and wash the worms with an additional 2 mL of dH₂O.
5. Repellet the worms at low speed in a clinical centrifuge for 10–20 s and aspirate the dH₂O.
6. Resuspend the worm pellet in the residual dH₂O and transfer to a microfuge tube.
7. Pellet the worms one final time by centrifuging for 10–20 s in a microfuge at low speed. Remove the remaining dH₂O.
8. For every 50 μ L of worms, add 400 μ L of Trizol Reagent (Life Technologies) (*see* **Notes 4** and **5**).
9. Shake the mixture by hand frequently over a 10-min incubation period at room temperature.
10. Add 80 μ L of chloroform, mix by inverting for 15 s, and incubate for 2 to 3 min at room temperature.
11. Centrifuge the sample at 8000 rpm for 15 min at 4°C in a microfuge.
12. Transfer the top aqueous layer to a fresh microfuge tube.
13. Purify the RNA by an affinity method to concentrate; commercial kits are available.
14. Add 10 U of RNase-free DNase I, 25 μ L of 10X DNase buffer, and sterile dH₂O to purified RNA (final reaction volume of 250 μ L) and incubate at 37°C for 20 min to remove the DNA.
15. Add an equal volume of phenol/chloroform, mix well, centrifuge at 4°C for 5 min at maximum speed in a microfuge, and transfer the top aqueous phase to a new microfuge tube.
16. Add 1/10 vol of 3 M sodium acetate, pH 5.2, and 1 mL of ice-cold 100% ethanol to the tube and mix by gently inverting.
17. Centrifuge for 30 min at 4°C in a microfuge at maximum speed to pellet the RNA.
18. Carefully remove the ethanol and add 1 mL of ice-cold 70% ethanol, being careful not to dislodge the pellet.
19. Centrifuge for 5 min at 4°C in a microfuge at maximum speed. Carefully remove the ethanol.
20. Air-dry the RNA pellet at room temperature (approx 1 h). Do not dry completely—the RNA pellet can be difficult to resuspend.
21. Resuspend the RNA in 50 μ L or less of sterile dH₂O.
22. Determine the concentration of the purified RNA by spectrophotometric measurement. Generally, 5 μ g of total RNA is utilized for a reverse transcription reac-

tion (*see* **Note 6**). A number of commercially available kits are available for performing these reactions. The reverse transcription product is used as a template for PCR (*see* **Note 7**). The amplified DNA can be ethanol precipitated or purified using a number of commercially available kits. The purified DNA can then be subcloned into a plasmid or used directly as a template for in vitro transcription (*see* **Subheadings 2.3.** and **3.3.**).

3.3. Production of dsRNA by In Vitro Transcription

The methods in **Subheading 3.2.** provide guidelines to generating plasmid DNA or PCR-amplified DNA that can be used as a template for dsRNA production. The following protocols describe how dsRNA is produced from such templates. The resulting dsRNA can be delivered to worms by injection or soaking, as described in **Subheadings 3.4.** and **3.5.** **Subheading 3.3.1.** describes the preparation necessary for the template DNA, **Subheading 3.3.2.** describes the in vitro transcription reaction, and **Subheading 3.3.3.** describes how to anneal sense and antisense strands to generate dsRNA.

3.3.1. Preparation of Template DNA

3.3.1.1. dsRNA PREPARATION FROM A PLASMID TEMPLATE

Plasmid DNA should be of high quality and can be purified by an affinity method that minimizes salt concentrations (a number of commercial kits are available).

1. Linearize the template DNA (5–10 μ g) with the appropriate restriction enzymes (*see* **Figs. 1–3**). For dual-promoter plasmids, two digestions in separate tubes will be required: one tube will be used to generate the sense strand and the other for the antisense strand. For double promoter (e.g., double T7) plasmids, the two digested plasmids can be combined into one tube and transcribed simultaneously using one RNA polymerase. For single-promoter plasmids harboring an inverted repeat DNA sequence, only one site (at the end of the inverted repeat) should be cut. Enzymes that leave a 5' overhang or blunt end should be chosen since RNA polymerases may initiate transcription from a 3' overhang.
2. Analyze a small aliquot of the digestions by agarose gel electrophoresis to confirm that the plasmid DNA was completely linearized.
3. Clean up the digestions to remove enzymes and salts by purifying over a DNA affinity column (commercial kits are available) or by utilizing phenol/chloroform extraction and ethanol precipitation.

3.3.1.2. dsRNA PREPARATION FROM A PCR TEMPLATE

DNA generated by PCR (**Fig. 4**) can be used directly as a template in in vitro transcription reactions following a cleanup step. Purification of the DNA

from reaction components can be easily accomplished by using commercial kits or by phenol/chloroform extraction followed by ethanol precipitation.

3.3.2. *In Vitro Transcription* (see **Notes 8–10**)

1. “Homemade mixes”: Mix the following reaction components in the order listed to prevent precipitation of DNA with the spermidine in the buffer (equilibrate components to room temperature unless otherwise noted):
 - a. 0.5–1 μg of linearized template DNA or PCR fragment.
 - b. 1 μL of 200 mM DTT.
 - c. 2 μL of 5 mM NTPs.
 - d. H_2O to 16 μL .
 - e. 2 μL of 10X transcription buffer.
 - f. 24 U of RNase inhibitor (stored on ice).
 - g. 15–20 U of RNA polymerase (stored on ice).The total reaction volume is 20 μL .
2. Commercially available kits: Mix the reaction components and template DNA per the manufacturer’s instructions.
3. Incubate the reaction for 1–4 h at 37°C, or per the manufacturer’s instructions.

3.3.3. *Annealing RNA Strands* (see **Note 11**)

When sense and antisense strands are synthesized separately, it is necessary to perform an annealing step. An annealing step may also optimize the yield of fully double-stranded RNA when the strands are synthesized in the same tube. If loss of ssRNA is a concern or if the transcription product is a hairpin, **steps 2a–e** can be omitted.

1. Remove a 0.5- μL sample from each single-stranded reaction for analysis by agarose gel electrophoresis (**Fig. 6**; see **Notes 12** and **13**). Proceed to **step 2** if clean ssRNA preparations are required; otherwise proceed to **step 3**.
2. The following protocol, adapted from Andrew Fire, results in dsRNA, as well as clean preparations of sense and antisense RNA for use as controls:
 - a. Bring the volume of single-stranded reactions to 400 μL by adding 40 μL of 3 M NaOAc (pH 5.2) and 360 μL of H_2O .
 - b. Add 200 μL of phenol/chloroform (1:1), mix by inverting, and centrifuge at maximum speed for 5 min in a microfuge at 4°C.
 - c. Transfer the upper aqueous phase to a new tube, add 200 μL of chloroform, mix by inverting, and centrifuge as in **step 2b**.
 - d. Transfer the upper aqueous phase to a new tube, add 1 mL of ethanol, mix by inverting, and centrifuge for 20 min at maximum speed in a microfuge at 4°C.
 - e. Dry the ssRNA pellets at room temperature and resuspend in 10 μL of TE.
 - f. Add 2 U of RNase-free DNase I, 2 μL of 10X DNase buffer, and sterile dH_2O to each ssRNA sample (20- μL final reaction volume) and incubate at 37°C for 15 min. (This step is important when the ssRNA will be injected but is not required for soaking experiments.)

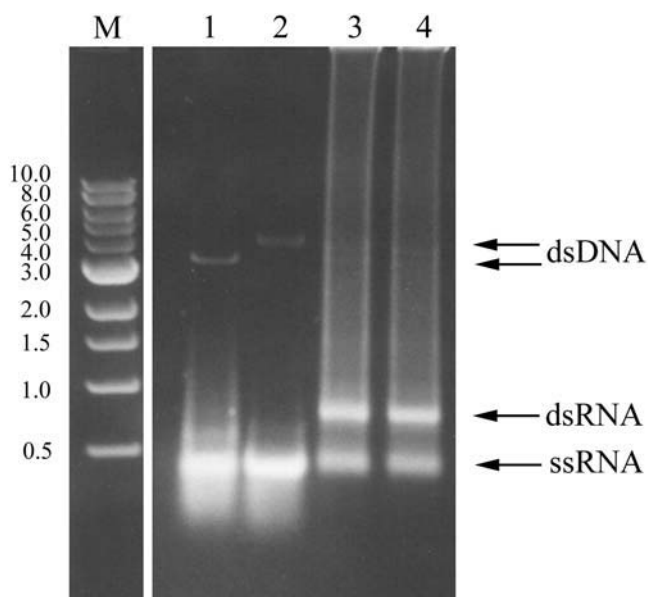


Fig. 6. Agarose gel analysis of ssRNA and dsRNA synthesized by in vitro transcription. A plasmid with a configuration similar to **Fig. 1A** was used for in vitro transcription to generate dsRNA corresponding to the *C. elegans unc-22* gene. T7 and T3 polymerases were used to generate the single strands in separate transcription reactions that were incubated for 4 h at 37°C to produce sense and antisense RNA. The reactions were then mixed and incubated an additional 2 h at 37°C. Further annealing of the strands was accomplished by incubating the RNA mixture at 75°C for 5 min, then cooling to room temperature for 30 min. M, DNA marker; lane 1, sense RNA; lane 2, antisense-RNA; lane 3, mixed ssRNA; lane 4, mixed and annealed ssRNA. Arrows indicate dsDNA template, dsRNA, and ssRNA. Note the shift in size between the ssRNA and the dsRNA and that this protocol resulted in partial annealing of strands (see **Note 14**).

- g. Phenol/chloroform extract, chloroform extract, and ethanol precipitate as described in **steps 2a–d**. Dry the ssRNA pellets at room temperature and resuspend in 10–50 µL of TE.
- h. Reserve the necessary amount of clean ssRNA for use as controls and analyze yield by agarose gel electrophoresis (**Fig. 6**; see **Notes 12** and **13**). Combine the remaining volume of the single-stranded samples into one tube and anneal by incubating at 68°C for 10 min, then 37°C for 30 min in 1X injection buffer.
3. Alternatively, to maximize the yield of dsRNA (and when clean ssRNA is not required), combine the transcription reactions. Then try one or more of the following methods followed by phenol/chloroform extraction and ethanol precipitation to achieve complete annealing:

- a. Continue incubating the reactions at 37°C for an additional 2 h, heat to 75°C for 5 min, and slowly cool to room temperature over a 30-min period.
- b. Heat to 75°C for 5 min and slowly cool to room temperature.
- c. Heat to 65°C for 15 min, then 37°C for 30 min.
- d. Heat to 65°C for 15 min, decrease to 37°C at a rate of $-1^{\circ}\text{C}/\text{min}$, incubate at 37°C for an additional 30 min, then decrease to room temperature at $-1^{\circ}\text{C}/\text{min}$.

3.4. Delivery of dsRNA into *C. elegans* by Injection

3.4.1. Preparation of Injection Pads (see *Note 15*)

1. Using a Pasteur pipet (preheated in an oven or heated by repeated pipetting of water maintained at 60°C), place a drop of 1 to 2% agarose at 60°C onto a glass slide or cover slip that is suitable for your microinjection system.
2. Quickly place another glass slide on top of the agarose drop, orienting the top slide perpendicular to the bottom slide.
3. Allow a few minutes for the agarose to solidify, and then peel apart the slides. A quick sliding motion generally accomplishes this; some practice may be required. A square of agarose will be left on the surface of one of the slides. Trim the square with a razor to remove bulges at the edges that might interfere with needle movement later.
4. Allow the slides to dry overnight at room temperature or dry in a 60°C oven for 1 h.

3.4.2. Preparation of Injection Needles

Several needles may be necessary for each injection mix, because some may become plugged during injection. Most clogged needles are useless and must be disposed. Several different needle-pulling devices are available from Sutter, Narishige, and other companies that will shape a needle for injection.

3.4.3. Loading Injection Needles With dsRNA Solutions

Several methods can be used to load the needles with the solutions. With filamented needles, the bottom of the needle can be sterilized by brief passage through a flame. Then do the following:

1. Warm the needles in a 37°C incubator for 10 min.
2. First, centrifuge the dsRNA solution briefly in a microfuge to remove sediment that might clog the injection needle. Then, insert the bottom end of the prewarmed needle into dsRNA-containing solution using forceps to hold the needle. As the air inside the needle cools, the liquid will rise by capillary action. (Using your fingers to hold the needle prevents efficient cooling.)
3. When some liquid has risen into the injection needle, remove the needle from the solution and gently lodge, tip down, into the clay in the hydrated needle chamber (do not let the tip touch the bottom of the chamber). The liquid should continue to move while the needle is stored in the chamber. Once the liquid has reached the

tip, the needle is ready for injecting; liquid movement can be monitored using a dissection microscope and generally takes 5–20 min.

4. Mount the loaded needle into the needle holder of a microscope set up for micro-injection and focus on the tip of the needle under low power. Place a slide with an agarose pad topped with mineral oil onto the microscope, lower the needle into the oil, focus on the tip, and check the flow rate. It may be necessary to break the tip of the needle to get the liquid to flow out. This can be accomplished by using the tip to pick at imperfections on the pad while the needle is under pressure. Replace this slide with one mounted with worms for injections.

3.4.4. Mounting *C. elegans* onto Injection Pads and Injecting (see **Note 16**)

1. Place a drop of mineral oil onto an injection pad (see **Note 17**).
2. Transfer well-maintained worms from an OP50 bacterial lawn onto an unseeded plate (this allows the worms to shed bacteria from their cuticles).
3. Using a platinum worm pick fashioned into a point (see **Subheading 2.1.3.**), gently transfer worms from the unseeded plate onto the agarose pad. Monitor transfer under a dissection microscope. Here it is important to avoid a swiping motion of the pick. When a part of the worm has touched the agarose surface, it should immediately adhere, and any further motion of the pick will rip the worm apart. The rest of the worm is brought to the pad as the worm struggles to free itself. For novice injectors, limit one to five worms per pad.
4. Mount the slide onto the injection microscope already loaded with a working needle.
5. Inject the worms with liquid. Injection into the body cavity or gut is sufficient to elicit an RNAi response. However, if the phenotype is to be monitored in the progeny of injected animals, a greater number of affected progeny may result by targeting the gonad (**Fig. 7**).

3.4.5. Recovery of Injected Animals

1. Place a large drop of recovery buffer or M9 medium over the injected animals. The animals should be released from the pad immediately. Allow 15 min for recovery.
2. Place a drop of M9 and a drop of mineral oil onto an unseeded area of an OP50-seeded plate. Remove the worms from the agarose pad using a 200- μ L pipetman set at 20 μ L (higher volume settings will provide more pipet tip surface area for the worms to stick).
3. Place the worms into the M9 drop on the NGM plate—use the same pipet tip to transfer all worms. Worms should be counted before and after transfer. If worms have stuck to the pipet tip, they can be dislodged by pipetting mineral oil from the plate slowly up and down the pipet tip, expelling the mineral oil onto the plate, followed by pipetting similarly with more M9. The mineral oil will release the worms from the plastic tip surface, and the M9 will help wash the worms and mineral oil from the tip.
4. Allow the worms to recover on the plates overnight.

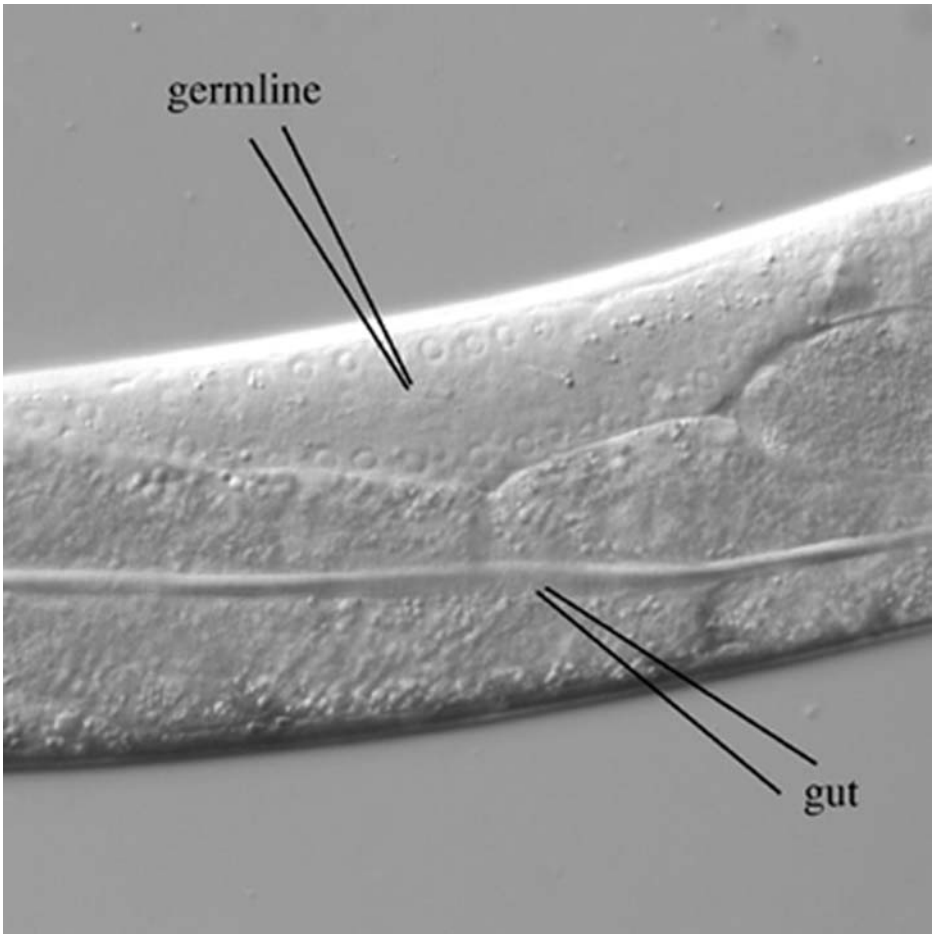


Fig. 7. Image of adult hermaphrodite highlighting the gut and gonad ($\times 40$). Lines represent an injection needle. dsRNA may be injected into the gut, body cavity, or germline of the animal; DNA must be injected into the gonad for germline transformation. The gut cells have a darker, grainy appearance, are much larger than other cells of the animal; and the central lumen is often visible. The gonad consists of two U-shaped structures that are enclosed in basement membrane and is generally observed as a clear region of the animal. At one end lies a region of mitotically dividing nuclei. The mitotic gonad is a syncytium of nuclei that are arranged similar to the “kernels on a corn cob.” In this image, outermost gonadal nuclei are in focus and are colinear with the central point of the lumen where DNA is injected.

3.4.6. Analysis of RNAi Phenotypes

1. Transfer individual worms from the recovery plate onto separate OP50-seeded NGM plates.
2. Transfer each injected worm onto a fresh plate on a daily (or more frequent) basis. Label the plates so that all progeny of an injected individual can be monitored as a group.
3. Monitor each batch of progeny for phenotypes. The first batch of progeny may contain unaffected animals, because, they may have been too developmentally progressed at the time of dsRNA delivery. Later batches of progeny may not be affected because the dsRNA may have become degraded or may become limiting.
4. It is wise to perform several sets of injections, culturing each set of worms post-injection under different temperature or other growth conditions. Compare the phenotypic distributions among the sets of progeny and among sets of injections.

3.5. Delivery of dsRNA by Soaking

1. Set up varying dilutions of dsRNA (0.2–5 $\mu\text{g}/\mu\text{L}$) in 1.5-mL microfuge tubes (*see Note 18*) in a 5- μL volume (minimum). Dilutions can be made using a 1:1 mixture of sterile M9 in water.
2. Add 10–20 worms of the appropriate strain and developmental stage to the diluted dsRNA.
3. Incubate overnight at 15°C. The length and temperature of incubation may be varied.
4. After incubation, carefully transfer the worms from the tube to a seeded NGM plate using a 200- μL pipet tip set at 20 μL . Carefully rinse the pipet tip and the tube with a small amount of sterile dH_2O and mineral oil to ensure that all worms have been transferred (*see Subheading 3.4.5.*).
5. After a few hours, transfer soaked worms individually to an OP50-seeded NGM plates.
6. Once F_1 embryos are observed, transfer the soaked worms to a fresh plate. This is done to flush out the first F_1 progeny that may have been present within the adult animal at the time of soaking, and therefore may not show a phenotype. Subsequent transfers of soaked animals on following days is also advisable because the RNAi effects do wear off and phenotypes may not be observed in later progeny.
7. Periodically monitor the soaked animals and F_1 progeny for phenotype (*see Fig. 8* for results of soaking experiments using ds *unc-22* RNA; *also see Note 19*).

3.6. Feeding *C. elegans* dsRNA-Expressing Bacteria

The feeding protocol requires a plasmid with T7 promoter sequences and DNA corresponding to the dsRNA trigger. The plasmid can be configured in a “double T7” configuration with two T7 promoter sites (plasmid L4440, available from the CGC, is one example (*see Subheading 3.2.1.1.* and **Fig. 2**) or the insert can be configured as an inverted repeat behind a single T7 promoter

Worm #	Experiment A		Experiment B		Experiment C	
	Twitching P0	Twitching F1	Twitching P0	Twitching F1	Twitching P0	Twitching F1
1	+	++	-	+++	+	-
2	+	++	-	+++	+	+
3	+	++	-	+++	+	+
4	+	++	-	+++	+	-
5	+	-	-	+++	+	-
6	-	-	-	+++	+	-
7	-	++	-	+++	+	-

Fig. 8. Experimental variability among soaking experiments. Experiments A and B: N2 worms were incubated in *unc-22* dsRNA overnight at 15°C, recovered, cloned, and kept at 20°C during phenotypic analysis of F₁ progeny. Experiment C: N2 worms were incubated in *unc-22* dsRNA overnight at 15°C, recovered, cloned, kept at 15°C for 48 h, then shifted to 20°C before phenotypic analysis of F₁ progeny. These experiments were performed using the same batch and concentration of dsRNA. Animals were scored for the corresponding loss-of-function twitching phenocopy: +, 25–50% of progeny exhibited the twitching phenotype; ++, 50–75% of progeny exhibited the twitching phenotype; +++, 75–100% of progeny exhibited the twitching phenotype. This set of experiments highlights the variability among individual experiments, and this may be influenced by factors such as the developmental stage of the soaked worm and incubation temperature after soaking (cf. experiment A with C). Although the penetrance varied, the expressivity of the phenotype was strong in all cases.

site (see **Subheading 3.2.1.2.** and **Fig. 3**). The plasmid is transformed into HT115(DE3) host cells, and dsRNA production is maintained on NGM plates supplemented with antibiotics and IPTG. Generally, feeding experiments are performed using one bacterial strain harboring one plasmid at a time (see **Note 20**). Worms are placed directly on such plates and phenotypes are monitored in the presence of food. The following protocols require Amp selection for maintenance of plasmids in the bacterial strain. If another plasmid is used that contains a different antibiotic resistance gene, replace the Amp with the appropriate antibiotic. Use sterile techniques for all the protocols.

3.6.1. Producing CaCl_2 -Competent HT115(DE3) Cells

1. Inoculate 2 mL of LB+Tet (12.5 µg/mL) medium with HT115(DE3) host cells. Incubate overnight at 37°C with shaking (150–225 rpm).
2. Inoculate 0.5–1 mL of the overnight culture into 20 mL of LB+Tet (a 50-mL sterile centrifuge tube is a convenient vessel). Incubate at 37°C with shaking (150–225 rpm) until an OD₆₀₀ of 0.4–0.8 is attained (this usually requires 1–4 h).
3. Pellet the cells by centrifuging the culture tube in a clinical centrifuge for 15 min at maximum speed at 4°C.
4. Decant the medium. Resuspend the bacterial pellet by gently vortexing in the residual medium.

5. Incubate the tube on ice for 5 min. *In subsequent steps, the cells should be kept on ice.*
6. Add 10 mL of sterile, cold 50 mM CaCl₂. Swirl the tube gently to mix, and incubate on ice for 1 h.
7. Pellet the cells by centrifuging the tube at 4°C in a clinical centrifuge for 15 min at maximum speed.
8. Decant the medium. Resuspend the cells in the residual solution by gently flicking the bottom of the tube.
9. Add 0.1–0.5 mL of cold 50 mM CaCl₂, and swirl the tube gently to mix. The cells are now ready for transformation. Cells stored at 4°C may be used for 72 h with little loss of competency.

3.6.2. Transforming HT115(DE3) Cells With Plasmids

Use 50–200 µL of competent cells prepared in **Subheading 3.6.1.** for transformation of supercoiled plasmids (5).

1. Using a sterile technique, add 10–100 ng of plasmid to the cells and incubate on ice for 1 h.
2. Heat-pulse the cells in a 42°C water bath for 30 s, and then incubate the tube on ice for 2 min.
3. Add 500 µL of LB medium without antibiotics to the tube and incubate at 37°C for 1 h.
4. Plate the cells on LB-agar plates containing 12.5 µg/mL of Tet and 50–100 µg/mL of Amp. *Do not add IPTG to the plates.*

3.6.3. Freezing HT115(DE3) Expression Strains

1. Using a sterile technique, inoculate a colony of freshly transformed HT115(DE3)+plasmid into 1 mL of LB broth+Tet and 50–100 µg/mL of Amp. Incubate at 37°C with shaking (225 rpm) until an OD₆₀₀ of 0.4–0.8 is attained (this usually requires 1–4 h).
2. Add 750 µL of cells and 250 µL of sterile 80% glycerol to a labeled, sterile freezer vial and mix by gently inverting. Quick-freeze in a dry ice/ethanol bath. Place immediately in a –80°C freezer.

3.6.4. Coculture of *C. elegans* With dsRNA-Expressing Bacteria (see **Notes 21 and 22**)

1. Inoculate a 20-mL 2X YT culture containing 12.5 µg/mL of Tet and 50–100 µg/mL of Amp with a single colony of HT115(DE3)+plasmid and incubate overnight at 37°C with shaking at 225 rpm (see **Note 23**).
2. Dilute the culture more than 100-fold, and continue to grow until the culture reaches OD₆₀₀ = 0.4–0.8.
3. Add IPTG to the culture to a final concentration of 0.4 mM, and incubate with shaking (225 rpm) for an additional 1 h at 37°C. This induces transcription of T7 RNA polymerase within the cells.

4. Supplement the culture with an additional 50 µg/mL of Amp, 12.5 µg/mL of Tet, and 0.4 mM IPTG.
5. Directly seed the cells onto NGM plates containing 50–100 µg/mL of Amp, 12.5 µg/mL of Tet, and 0.4 mM IPTG. Allow the cells to incubate and the plates to dry at room temperature overnight (see **Note 24**).
6. Transfer the worms to plates using a worm pick to transfer individuals or a metal spatula to transfer a small agar chunk with more worms.
7. Monitor phenotypes in the transferred animals and in their progeny (see **Notes 25–28**).

3.7. *In Vivo* Transcription of dsRNA

1. Subclone the inverted DNA repeats (gray boxes in **Fig. 5**) and stuffer fragment (black boxes in **Fig. 5**) behind a *C. elegans* promoter and in front of a *C. elegans* 3' UTR (light gray) using standard cloning methods (**Fig. 5**) (5).
2. Refer to **Subheading 3.4.** for instructions on *C. elegans* injections (see **Notes 29** and **30**); the injections of dsRNA and plasmid DNA are similar (see **ref. 12** for more information).
3. Recover each set of animals onto NGM/OP50 plates. Clone the progeny that exhibit a phenotype corresponding to the transformation marker onto separate plates. Monitor these F₁ progeny for the presence of the marker phenotype—not all F₁s will give rise to lines. Maintain those plates producing F₂ animals with the marker phenotype as separate lines.
4. Examine each transformed line for an RNAi phenotype using different culture conditions (e.g., different temperatures).
5. Monitor the efficacy of RNA silencing by performing *in situ* RNA hybridizations or protein immunolocalization if an antibody to the target protein is available.

4. Notes

1. The plate protocols can be scaled up, and an automatic dispenser such as Wheaton Omnisense may be used to facilitate pouring. If pouring plates by hand, it may be more convenient to make the medium in an autoclavable container with a spout and handle.
2. It is possible to produce a hybrid dsRNA molecule by inserting two trigger sequences into an RNA expression plasmid. Efficient RNAi for both gene targets can be observed. If one of the sequences corresponds to *rrf-3* or other endogenous inhibitor of RNAi, the efficiency of RNAi can be enhanced.
3. Bacterial DNA is likely to be present in this genomic DNA prep.
4. Wear gloves when preparing RNA to protect against the introduction of RNases. All solutions used during an RNA preparation should be autoclaved or filter sterilized to remove RNases and other contaminants.
5. Care should be taken when using the Trizol reagent because it contains phenol and can cause burns.
6. When a contiguous coding region of sufficient length is not available (i.e., the gene is disrupted by too many introns), RT-PCR can be utilized to generate a trig-

ger molecule of adequate length (≥ 200 bp). The resulting DNA can be inserted into a plasmid (**Fig. 1**), or, alternatively, if hybrid primers composed of promoter and target gene sequences were utilized in the RT-PCR reaction, the DNA can be used directly (after cleanup) in in vitro transcription reactions (**Fig. 4**).

7. When performing RT-PCR reactions, a control tube lacking the RT enzyme should be included with each primer set. This control should help detect the presence of contaminating genomic DNA in the RNA preparation (no amplification product should be produced). If DNA is present, another round of DNase treatment, followed by phenol/chloroform extraction and ethanol precipitation, can be performed on the RNA sample.
8. Wear gloves whenever working with RNA or reagents used to synthesize RNA to protect against the introduction of RNases.
9. All solutions used during an RNA preparation should be autoclaved or filter sterilized to remove RNases and other contaminants.
10. To obtain the highest yield of ssRNA, it may be necessary to optimize the reaction conditions. The suggested buffer is typical, but buffers that are supplied with commercially available RNA polymerases can vary: for SP6 reactions, NaCl may be omitted and DTT lowered to 1 mM; for T3 and T7 reactions, concentrations for NaCl, MgCl₂, and DTT can range from 10 to 25, 6 to 8, and 5 to 10 mM, respectively. The amount of template DNA and incubation time can also be varied to increase the yield of RNA.
11. To obtain the highest yield of dsRNA, it may be necessary to optimize the annealing conditions. This can be influenced by factors such as length and complexity of the ssRNA transcripts. For RNAs that are difficult to anneal, try using a smaller fragment of the gene of interest as template.
12. EtBr is a mutagen and a carcinogen, and gloves should be worn during its use.
13. Wear safety glasses to protect eyes from UV light when viewing agarose gel.
14. A linearized plasmid will migrate much more slowly than the resulting dsRNA or ssRNA when resolved on an agarose gel (**Fig. 6**). The yield of dsRNA can be estimated by comparison with the fluorescence intensity of a known quantity of DNA in the marker lane. However, a PCR template used for in vitro transcription may be similar in size to the resulting dsRNA, possibly skewing quantitation of the dsRNA. To avoid this problem, the dsRNA preparation should be treated with DNase.
15. The agarose pad provides a sticky surface for mounting worms and prevents them from moving during injection. Worms mounted on this surface slowly dehydrate, allowing them to accept the injected fluid. The pad composition can be manipulated to suit work style (or injection speed). A faster-dehydrating pad may be preferred by experienced users while a slower-dehydrating pad may be preferred by novice injectors or by injectors using younger (smaller) animals. The rate of dehydration depends on the following:
 - a. The thickness of the pad: We have found that the weight of a glass slide is sufficient to spread out the drop of agarose. Furthermore, since glass slides have a relatively consistent weight, the resulting pads are also relatively

- consistent in thickness. (We have not found it necessary to measure the precise amount of agarose dropped onto the first slide; a large drop from a Pasteur pipet is spread completely underneath the top slide.)
- b. The temperature of the agarose: Hot agarose will spread faster and produce a thinner (slower-dehydrating) pad. Slightly thicker pads will result when using cooler agarose. The agarose solution can be maintained at 60°C in covered glass vials in a heat block for about 1 wk with good results.
 - c. The concentration of agarose: Some batch-to-batch variability in performance has been observed. We test each new bottle of agarose using concentrations within the range of 1–4% and determine the best working concentration for that bottle, which is then reserved exclusively for injections. Generally, 1% agarose works best.
16. For novice injectors, it is advisable to master mounting and recovery of worms before attempting injections. To practice, place worms onto pads (in mineral oil), wait 20 min, add recovery buffer, wait 20 min, and then transfer the worms to seeded plates. When this can be accomplished with full survival, proceed with mastering the injection. This strategy also allows the quality of the reagents to be checked.
 17. It is not a bad idea to test a new batch of mineral oil for toxic effects: mount some worms onto pads and overlay with mineral oil, allow to sit for 10 min, drop M9 medium onto the worms, and transfer to a seeded NGM plate. Assay for viability. If worms float off the agarose pads into the mineral oil, it may need to be replaced.
 18. Some proprietary buffers in commercial in vitro transcription kits may kill worms. Dilution or ethanol precipitation of in vitro transcribed dsRNA may be necessary. Worms should be soaked in various concentrations (e.g., 1X, 0.5X) of a commercial buffer in the absence of dsRNA to determine toxicity.
 19. Since it is not possible to control how much dsRNA a worm will ingest during soaking, the penetrance and expressivity of the resulting loss-of-function phenotypes may vary greatly from worm to worm, especially for dilute dsRNA solutions.
 20. “Multiplex” RNAi—in which more than one dsRNA trigger is delivered to the worm simultaneously—is possible using the feeding method. In general, bacterial strains harbor only one double-T7 plasmid, and two bacterial strains each expressing a different dsRNA can be mixed and fed to worms. However, the success rate for achieving RNAi for two targets simultaneously is very low. Better success is achieved when one bacterial strain expresses two distinct dsRNA sequences simultaneously. This is most easily accomplished by juxtaposing two DNA sequences into the double-T7 vector such that a hybrid dsRNA can form.
 21. This protocol is particularly useful when a few dsRNA-expressing bacterial strains (i.e., a few triggers) are used to study phenotypic effects on multiple strains of worms. However, another variation of this protocol may facilitate analyses of one strain of worms (wild-type) with multiple strains of dsRNA-expressing bacteria (i.e., multiple RNA triggers) (21). In this protocol, bacterial cells are grown in liquid culture to log phase or saturation in the absence of IPTG and Tet, seeded

onto plates containing 1 mM IPTG and 75 µg/mL of carbenicillin without Tet, and allowed to induce expression of trigger RNA overnight at room temperature. This protocol reduces the need to handle many tubes of different bacterial strains but also adds the hazard and expense of higher concentrations of IPTG in the plates.

22. dsRNA can also be extracted from the bacterial cells (2). The extracted dsRNA can elicit RNAi when injected into worms.
23. We have found that the feeding protocol works best when the HT115(DE3) cells are freshly transformed. Storage of expression strains at 4°C on LB plates often results in a loss of competency for dsRNA production. We make freezer stocks from newly transformed colonies and generally inoculate from frozen stocks. Additionally, we regenerate our frozen stocks of commonly used dsRNA-expressing bacterial strains at least every 3 mo, again from freshly transformed colonies.
24. Freshly seeded plates or plates stored for as long as 3 wk at 15°C can produce RNAi phenotypes; however, it is always best to use freshly seeded plates.
25. RNAi phenotypes are usually observable within 16 h to 3 d, depending on the target gene and quality of food.
26. Plates contain sufficient quantities of bacteria to support growth of the worms for one generation; subsequent generations can be transferred to fresh plates.
27. At no time during dsRNA administration should the animals be depleted of food because the RNAi phenotype can diminish. Best results are achieved when animals are transferred frequently onto freshly seeded plates—allowing more food for fewer F₁ animals.
28. For long-term maintenance of animals on dsRNA-expressing food, worms from successive generations should be transferred to fresh feeding plates.
29. Plasmid DNA *must* be injected into the gonad of the worm in order to obtain stable lines that will express the trigger RNA in vivo.
30. The dsRNA hairpin may not elicit a visibly discernable phenotype, so a transformation marker must be injected along with the dsRNA to confirm that the worms have incorporated the injected DNA. When plasmid pRF4 is injected along with the inverted repeat plasmid, worms with pRF4-containing arrays will exhibit a “roller” phenotype. (pRF4 harbors the *rol-6* gene with a dominant mutation that produces this phenotype.) Other dominant markers, including GFP-expressing plasmids, can be used.

Acknowledgments

Most of the presented methods are an accumulated set of protocols used by members of the worm community. We owe a debt of gratitude to the original inventors and apologize that space limitations restricted proper referencing of their contributions. We wish to thank Guy Caldwell, Shelli Williams, Erik Lundquist, and Mary Montgomery for reviewing the manuscript. This work was supported in part by funds from the National Institutes of Health (P20 RR015563 from the COBRE Program of the National Center for Research

Resources) as well as funds from the National Science Foundation (EPS-0236913) that include matching support from the State of Kansas and the University of Kansas.

References

- 1 Hubbard, E. J. and Greenstein, D. (2000) The *Caenorhabditis elegans* gonad: a test tube for cell and developmental biology. *Dev. Dyn.* **218**, 2–22.
- 2 Timmons, L., Court, D. L., and Fire, A. (2001) Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. *Gene* **263**, 103–112.
- 3 Simmer, F., Tijsterman, M., Parrish, S., Koushika, S. P., Nonet, M. L., Fire, A., Ahringer, J., and Plasterk, R. H. (2002) Loss of the putative RNA-directed RNA polymerase RRF-3 makes *C. elegans* hypersensitive to RNAi. *Curr. Biol.* **12**, 1317–1319.
- 4 Sijen, T., Fleenor, J., Simmer, F., Thijssen, K. L., Parrish, S., Timmons, L., Plasterk, R. H., and Fire, A. (2001) On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* **107**, 465–476.
- 5 Sambrook, J. and Russell, D. W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- 6 Brenner, S. (1974) The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94.
- 7 Stiernagle, T. (1999) Maintenance of *C. elegans*, in *C. elegans: A Practical Approach* (Hope, I. A., ed.), Oxford University Press, Oxford, pp. 51–67.
- 8 Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811.
- 9 Gonczy, P., Echeverri, C., Oegema, K., et al. (2000) Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. *Nature* **408**, 331–336.
- 10 Dudley, N. R., Labbe, J. C., and Goldstein, B. (2002) Using RNA interference to identify genes required for RNA interference. *Proc. Natl. Acad. Sci. USA* **99**, 4191–4196.
- 11 Grishok, A., Tabara, H., and Mello, C. C. (2000) Genetic requirements for inheritance of RNAi in *C. elegans*. *Science* **287**, 2494–2497.
- 12 Jin, Y. (1999) Transformation, in *C. elegans: A Practical Approach* (Hope, I. A., ed.), Oxford University Press, Oxford, pp. 69–96.
- 13 Epstein, H. F. and Shakes, D. C. (1995) *Caenorhabditis elegans: Modern Biological Analysis of an Organism*, vol. 48. Academic, San Diego.
- 14 Kimble, J., Hodgkin, J., Smith, T., and Smith, J. (1982) Suppression of an amber mutation by microinjection of suppressor tRNA in *C. elegans*. *Nature* **299**, 456–458.
- 15 Stinchcomb, D. T., Shaw, J. E., Carr, S. H., and Hirsh, D. (1985) Extrachromosomal DNA transformation of *Caenorhabditis elegans*. *Mol. Cell. Biol.* **5**, 3484–3496.
- 16 Tabara, H., Grishok, A., and Mello, C. C. (1998) RNAi in *C. elegans*: soaking in the genome sequence. *Science* **282**, 430–431.

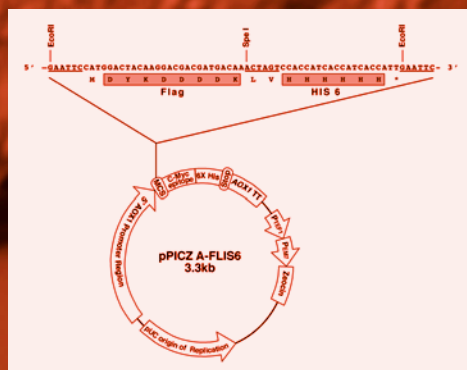
- 17 Maeda, I., Kohara, Y., Yamamoto, M., and Sugimoto, A. (2001) Large-scale analysis of gene function in *Caenorhabditis elegans* by high-throughput RNAi. *Curr. Biol.* **11**, 171–176.
- 18 Tabara, H., Sarkissian, M., Kelly, W. G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C. C. (1999) The *rde-1* gene, RNA interference, and transposon silencing in *C. elegans*. *Cell* **99**, 123–132.
- 19 Jorgensen, E. D., Durbin, R. K., Risman, S. S., and McAllister, W. T. (1991) Specific contacts between the bacteriophage T3, T7, and SP6 RNA polymerases and their promoters. *J. Biol. Chem.* **266**, 645–651.
- 20 Logel, J., Dill, D., and Leonard, S. (1992) Synthesis of cRNA probes from PCR-generated DNA. *Biotechniques* **13**, 604–610.
21. Kamath, R. S., Martinez-Campos, M., Zipperlen, P., Fraser, A. G., and Ahringer, J. (2001) Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*. *Genome Biol.* **2**, RESEARCH0002.
- 22 Williams, B. D., Schrank, B., Huynh, C., Shownkeen, R., and Waterston, R. H. (1992) A genetic mapping system in *Caenorhabditis elegans* based on polymorphic sequence-tagged sites. *Genetics* **131**, 609–624.
23. Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991) Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



Induction and Biochemical Purification of RNA-Induced Silencing Complex From *Drosophila* S2 Cells

Amy A. Caudy and Gregory J. Hannon

Summary

The discovery of RNA interference (RNAi) has greatly simplified the process of suppressing genes in many experimental systems, including *Caenorhabditis elegans*, *Drosophila*, and mammalian cells. A sequence-specific nuclease complex, called the RNA-induced silencing complex (RISC), can be purified from cells undergoing RNAi. RISC shows RNase activity when exposed to RNAs homologous to the input double-stranded RNA (dsRNAs) but lacks activity in the presence of nontargeted RNAs. We describe the induction of RNAi by dsRNA in cultured *Drosophila* Schneider-2 (S2) cells and detail procedures for RISC purification from these cells. This purification approach has allowed us to identify several RISC components, including siRNAs, Argonaute 2 (Ago-2), *Drosophila* Fragile X related protein (dFXR), *Vasa* intronic gene (VIG), and the micrococcal nuclease family member Tudor-SN (*Drosophila* CG7008). RNAi is carried out by an endogenous pathway important for normal development in many organisms. In fact, organisms express hundreds of different microRNAs (miRNAs), small hairpin RNAs that function through the RNAi pathway to suppress expression of endogenous genes. The function of miRNAs is poorly understood, and most of their targets are unknown. Purified RISC complexes contain short interfering RNAs and endogenously expressed miRNAs and will be useful for studying many aspects of the RNAi machinery.

Key Words: RNA interference; Dicer; Argonaute-2; Ago-2; Tudor-SN; let-7; *Drosophila*.

1. Introduction

The discovery of RNA interference and related posttranscriptional gene-silencing phenomena in plants now allows scientists working in systems including *Caenorhabditis elegans*, *Drosophila*, mammalian cell culture, and plants

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

to easily and specifically silence virtually any gene of interest (reviewed in **ref. 1**). RNA interference (RNAi) and related phenomena are not just a tool for reverse genetics—research on RNAi is revealing new pathways of gene expression directed by the expression of endogenously encoded small RNAs (reviewed in **ref. 2**).

Several biochemical systems that recapitulate aspects of RNA have been developed using *Drosophila* (**3,4**), *C. elegans* (**5**), mammalian cells (**6–10**), and plants (**11**). Two *Drosophila* systems have been developed: one using cultured S2 cells (**3**) and the other a method originally developed by the Sharp laboratory using 0- to 4-h embryos (**4**). In both systems, input double-stranded RNAs (dsRNAs) are cleaved to ~22 nt short interfering RNAs (siRNAs) by Dicer (**12,13**), and a sequence-specific RNase activity is formed. This chapter describes the preparation of extracts from S2 cells, in which the sequence-specific nuclease activity has been named RNA-induced silencing complex, or RISC. The purification of RISC activity from these extracts has been used to identify five components: small RNAs homologous to the triggering dsRNA, Argonaute 2 (Ago-2), dFXR, *Vasa* intronic gene (VIG), and Tudor-SN (**1,3,14,15**). The discovery of Ago-2 linked the biochemistry of the *Drosophila* system with the Argonaute family members that had been linked to RNAi by genetics in *C. elegans*, *Neurospora*, and *Arabidopsis*. The *Drosophila* system also has proven predictive value—the *C. elegans* homolog of VIG is important for targeting genes via the let-7 miRNA pathway (**15**). In addition to siRNAs derived from the triggering dsRNA, RISC complexes contain endogenously encoded small RNAs, called microRNAs (miRNAs) (**16**). These miRNAs function to control the expression of endogenous protein-coding genes (**5,17–25**). *C. elegans* homologs of *Drosophila* RISC components are required for miRNA function (**15**), which predicts that further analysis of RISC will provide insight into miRNA biology.

In our *Drosophila* S2 cell RNAi model, sequence-specific mRNA degradation activity is induced by introducing dsRNA into cells and allowing the cells to grow for several days before preparing a RISC extract. This is different from the *Drosophila* embryo system, in which dsRNA or siRNAs are introduced into prepared extract to induce nuclease activity. The S2 system completely degrades targeted RNAs, while the embryo system degrades the targeted message at specific intervals corresponding to the processing of the input dsRNA into siRNAs (**13,26**). These both resemble in vivo outcomes since in some cases siRNAs and miRNAs direct the complete degradation of the message (**27–30**), whereas in others site-specific cleavage is detected. We describe herein our method for inducing RISC activity and purifying it from *Drosophila* S2 cells.

2. Materials

1. *Drosophila* S2 cell line.
2. Schneider's medium for S2 cells (Sigma, St. Louis, MO; or Invitrogen/Gibco-BRL, Gaithersburg, MD).
3. Laminar flow hood for tissue culture.
4. Pipets, flasks, hemocytometer, and other standard equipment for tissue culture. If transfecting cells, autoclaved pasteur pipets are necessary.
5. HBS: 16 mg/mL of NaCl, 0.7 mg/mL of KCl, 0.4 mg/mL of Na₂HPO₄ (anhydrous), 2 mg/mL of dextrose, 10 mg/mL of HEPES. Prepare five times the desired quantity (generally, we prepare 2.5 L) and adjust the pH to 6.95 using 10 N NaOH. Remove 500 mL, adjust the pH to 7.00, and remove another 500 mL. Repeat for pH 7.05, 7.10, and 7.15. Filter sterilize and store at 4°C. This reagent will inexplicably lose transfection activity, so it is important to monitor transfection efficiency with a marker such as green fluorescent protein (GFP) or lacZ and prepare new reagent when transfection efficiencies decline.
6. Hypotonic lysis buffer: 20 mM HEPES, pH 7.0; 2 mM MgCl₂, 0.2 mM CaCl₂, 1 mM dithiothreitol (DTT). Add an EDTA-free protease inhibitor tablet (Roche) for each 25 mL of buffer.
7. Buffer A for protein purification: 20 mM HEPES, pH 7.0; 2 mM MgCl₂, 0.2 mM CaCl₂, 1 mM DTT, 0.5% *n*-octyl β -D-glucopyranoside. Add KCl to 0.4 M for size fractionation or to 1 M for ion-exchange methods. Some lots of *n*-octyl β -D-glucopyranoside form a precipitate in low salt at 4°C. Thus, it is best to prepare this buffer 1 d in advance and filter the material through a 0.2- μ m filter.
8. 5X buffer F for RISC assays: 550 mM KOAc (not pH adjusted); 5 mM MgCl₂; 10 mM CaCl₂, 15 mM EGTA, 100 mM HEPES, pH 7.0, 5 mM DTT.
9. Tris/EDTA (TE): 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, pH 8.0.
10. Formamide loading dye: 95% formamide, 0.01% bromophenol blue, 0.01% xylene cyanol, 0.1% sodium dodecyl sulfate.
11. Materials for denaturing polyacrylamide gel electrophoresis.
12. Reagents for polymerase chain reaction (PCR).
13. Megascript (Ambion) or Ribomax (Promega, Madison, WI) or other kit for large-scale in vitro transcription of RNA.
14. Riboprobe (Promega) or other kit for radiolabeled RNA probe preparation.
15. Dounce homogenizers of appropriate volumes for preparation of extract. For the amounts described here, a 40- and a 15-mL Dounce are required.
16. Trizol (Invitrogen), glycogen, chloroform, and isopropanol.

3. Methods

3.1. S2 Cell Culture

We initiated our studies by obtaining cell lines from several different laboratories and the American Type Culture Collection. Some lines seemed to be

more effective as substrates for RISC purification than others. In large part, this may have been because the S2 line that we selected is contaminated with the Flockhouse virus, a nodavirus. Furthermore, we suspect that it was contaminated with *Drosophila* virus X as well. The consequence of these infections is that the RNAi machinery is preprimed and RISC complexes are already abundant. By contrast, in human embryonic kidney 293 cells, we observe that complexes between the RNAi components Tudor-SN, EIF2C2, and FMRP are undetectable until cells are transfected with siRNAs (15).

Aeration is very important for large-scale Schneider-2 cell culture. If the culture volume exceeds 20% of the spinner flask volume, cell growth stalls. When cells are continually aerated by bubbling, spinner flasks can be filled to capacity. We have aerated cultures with both aquarium pumps and automated pipettors (we simply tape the pipettor button so that it constantly expels air). Replace one of the side-arm caps of the spinner flask with flame-sterilized aluminum foil, and insert a 2- or 5-mL pipet so that it will extend into the culture medium. Insert a 0.45- or 0.22- μ m syringe filter into the top end of the pipet. Then connect the top of the filter to the aquarium pump or automatic pipettor via autoclaved latex tubing.

Cells are cultured in Schneider's medium. We have successfully used both Invitrogen's pre-prepared medium and Sigma's dry medium. The dry medium is more cost-effective for large-scale cultures used for purification. We reconstitute the medium according to the manufacturer's instructions and adjust the final pH to 6.4. The medium is filtered via positive pressure in a Millipore 20 L apparatus. The medium is supplemented with 10% fetal bovine serum (FBS) and 1X penicillin/streptomycin. For suspension culture, the surfactant Pluronic (Invitrogen) must be added to 1X; otherwise, the cells will be damaged by shear forces.

3.1.1. Passage of S2 Cells

Cells should not be split below 1×10^6 , because low culture densities grow poorly. This is probably partially owing to the cells' need for secreted insect factors not present in FBS. For routine culture we split the cells to 2×10^6 /mL. Cells will saturate in 3 to 4 d. Older or nonaerated cultures typically grow more slowly than fresh, aerated cultures. Cells generally arrest at a density of $8\text{--}10 \times 10^6$ cells/mL.

3.1.2. Freezing S2 Cells

It is important to maintain adequate frozen culture stocks, because the large-scale cultures will periodically arrest and have to be discarded. It can take several weeks for cultures to grow to significant scale from frozen stocks, because there is always a slow phase of growth after the thaw.

This freezing protocol uses conditioned medium, in which cells have already been growing. Because S2 cells presumably release autocrine growth factors into the medium, the use of conditioned medium in the freezing buffer will promote recovery. This protocol has been slightly modified from that provided by Invitrogen with its *Drosophila* expression system.

1. When cells are growing well and at a density of $5\text{--}10 \times 10^6$ cells/mL, count the cells. Calculate the amount of culture necessary to achieve a final concentration of 1.1×10^7 cells/mL. Centrifuge the cells for 5 min at 3000g.
2. Remove the supernatant to a new, sterile tube. This will be the conditioned medium used for freezing.
3. Prepare freezing medium: 45% conditioned medium, 45% fresh medium (with 10% FBS), and 10% dimethylsulfoxide. Filter the freezing medium through a 0.45- μ m filter.
4. Resuspend the cell pellet for a final cell concentration of 1.1×10^7 cells/mL.
5. Aliquot the cells into vials for freezing.
6. Place the vials inside a styrofoam container, and allow to freeze overnight at -80°C before transferring to liquid nitrogen.

3.1.3. Thawing S2 Cells

1. Place a vial of frozen cells in a 37°C water bath.
2. When cells are just thawed, remove from the water bath and sanitize the vial with 70% ethanol.
3. Transfer the cells to a fresh flask containing 5 mL of fresh medium/mL of frozen cells.
4. The following day, change the medium. Some cells will adhere to the flask, but others will be in suspension. Remove the culture medium containing suspended cells to a sterile tube, and pellet the cells by centrifuging at 3000g for 5 min.
5. Resuspend the cells in an equal volume of fresh medium and return to the culture flask with the adherent cells.
6. Maintain the culture by splitting at appropriate intervals. It is normal for the initial growth after freezing to be somewhat slow.

3.2. Induction of an RNAi Response by dsRNA

Introducing a dsRNA allows the assay of specific RNase activity. Generally, we use exogenous genes such as luciferase to induce RNase activity (14,16), although we have also observed similar activity targeted against endogenous genes (3).

3.2.1. Preparation of dsRNA

We prepare dsRNA using large-scale kits such as Ambion Megascript or Promega RiboMax. We generally prepare transcription templates by PCR, using primers that add T7 RNA polymerase promoters on both ends of the PCR

product. The PCR products are gel purified and used for transcription reactions in accordance with the manufacturer's instructions. The reaction volumes are scaled up to achieve the desired quantity of RNA. Following the reactions, RNA is purified by sequential phenol/chloroform extraction, chloroform extraction, and ethanol/sodium acetate precipitation. No further annealing steps are necessary.

3.2.2. Introducing dsRNA: Soaking and Transfection

One of the tremendous conveniences of *Drosophila* S2 cells is that they will take up dsRNA introduced into the medium, a process called soaking. The majority of our experiments are carried out using either of the soaking protocols described in **Subheadings 3.2.2.1.** or **3.2.2.2.** However, when it is necessary to simultaneously introduce a DNA expression construct and a dsRNA, transfection can be used.

3.2.2.1. SOAKING DSRNA FOR SMALL-SCALE CULTURE

For RNAi in a small volume of cells (typically <100 mL final), it is effective and convenient to follow the protocol initially developed in the Dixon laboratory (31,32).

1. Count the cells using a hemocytometer. Calculate the number of cells needed for a final cell count of 0.66×10^6 cells/mL. Centrifuge the needed volume of cells at 3000g for 5 min.
2. Aseptically remove and discard the supernatant. Resuspend the cells in serum-free medium to one third of the desired final culture volume.
3. Add dsRNA to 15 μ g/mL.
4. Incubate the cells at room temperature for 30 min.
5. Add normally supplemented medium (10% FBS, 1X antibiotic, and 1X pluronic if a spinner culture) to reach the final culture volume.
6. The amount of time needed for maximal suppression depends on the targeted gene product. Generally, this occurs between 3 and 5 d postsoaking. If necessary, the soaking process can be repeated. If the target is a nonendogenous gene such as luciferase, we can detect nuclease activity for approx 7 d following soaking.

3.2.2.2. SOAKING DSRNA FOR LARGE-SCALE CULTURE

For larger-scale culture, the quantity of dsRNA required by the protocol in **Subheading 3.2.2.1.** would be expensive to prepare. This protocol is a slight adaptation of that one. It is somewhat less efficient at suppressing endogenous genes, but it will induce robust RISC activity, allowing for excellent purifications.

1. Count the cells using a hemocytometer. Calculate the number of cells needed for a final cell count of 1×10^6 cells/mL. Centrifuge the needed volume of cells at 3000g for 10 min.

2. Aseptically remove and discard the supernatant. Resuspend the cells in 1/10 culture vol of serum-free medium and repeat centrifugation as in **step 1**.
3. Resuspend the cells in 1/20 culture vol of serum-free medium.
4. Add dsRNA to 30 $\mu\text{g/mL}$.
5. Incubate for 45 min on a rocking platform at room temperature.
6. Add the cells to normally supplemented medium, resulting in a cell concentration of 1×10^6 cells/mL.
7. Incubate the cells for 5–7 d before harvesting and preparation of extract.

3.2.2.3. TRANSFECTION

Although liposome-based transfection methods can be used successfully in S2 cells, the calcium phosphate method is much less expensive and therefore well suited to larger-scale biochemical experiments.

As discussed in the preparation of HBS for transfection, it is critical to determine the optimum pH by testing HBS at different pH values and monitoring the expression of a reporter construct. We generally use either a lacZ or a GFP reporter.

The transfection described next is for a 10-cm dish of cells containing a 12-mL cell volume. It can be scaled proportionally for 6-well plates and 15-cm dishes, and it is possible to scale up to prepare transfection reagent for multiple plates at a time.

1. In a sterile 5-mL polypropylene culture tube, mix 100 μL of 2.5 *M* CaCl₂, 24 μg of nucleic acid (any relative proportion of DNA and dsRNA), and add sterile water to a total volume of 1 mL.
2. Attach a glass Pasteur pipet to an automatic pipetting aid and bubble the CaCl₂/nucleic acid solution with one hand.
3. Using the other hand, slowly add 1 mL of HBS in individual drops while continuing to bubble the mixture. The solution will gradually become cloudy.
4. Wait 3–5 min before evenly dropping the solution onto the cells.
5. Place the cells on a rocking platform for 10–20 min, and then return to the incubator.
6. The following day, remove the culture medium from the cells by aspiration. The cells are generally well adhered after transfection, and it is not necessary to worry about slight losses.
7. Replace the culture medium with fresh medium. If using a copper-inducible promoter, include 1 *mM* CuSO₄ in the medium.
8. Incubate the cells for 3–5 d before harvesting.

3.3. Preparation of Extract

Extracts are prepared by hypotonically lysing cells, and then pelleting and discarding the nuclei and cell debris. Ribosomes are pelleted from the cytoplasmic supernatant, and RISC is extracted from the ribosomes using high salt.

The RISC activity can be recovered by diluting the high-salt extract, which creates a low-salt precipitate containing RISC activity. The hypotonic lysate, the high-salt extraction, and the low-salt precipitate can all be used for analysis of RISC activity (*see* **Note 1**).

3.3.1. Hypotonic Lysis

1. Spin down the cells for 10 min at 3000g in a centrifuge of appropriate capacity for the culture volume. If cells are growing in plates or flasks, they can be dislodged by gently pipetting the medium. Discard the supernatant.
2. Resuspend the cells in ice-cold phosphate-buffered saline (PBS) at approx 1/5 to 1/10 of the culture volume. Spin down the cells as in **step 1**. Discard the supernatant.
3. Repeat the PBS wash, and discard the supernatant. Note the volume of the cell pellet—the volumes used in the subsequent steps are adjusted for the pellet volume.
4. Wash the cells in 5 pellet vol of hypotonic lysis buffer supplemented with 60 mM KCl and protease inhibitors. Spin down the cells as in **step 1**. Discard the supernatant with care—the pellet will be loose as a consequence of cell swelling in the slightly hypotonic conditions.
5. Resuspend the cells in 0.5 pellet vol of hypotonic lysis buffer supplemented with protease inhibitors. Transfer to a Dounce of appropriate volume.
6. Reserve a few microliters of resuspended cells for microscopic analysis. Homogenize the cells in a Dounce firmly at a moderate speed using an A pestle until 90% of the cells are lysed (releasing the nuclei). Generally this requires 8–10 strokes, but it is better to stop, observe the cells, and continue homogenizing as necessary. To confirm lysis, pipet 1 to 2 μ L of cells in 15 μ L of hypotonic lysis buffer. Place the cells on a slide with no cover slip (pressure from a cover slip will rupture cells), and observe the cells using phase contrast microscopy. Compare the homogenized material to the reserved cells in order to be certain that there is a difference between the intact and lysed cells. Lysed cells are less refractile. In addition, most cells release nuclei, which are slightly smaller than the cells.
7. Centrifuge the extract at 30,000g for 20 min.
8. Reserve the supernatant in another tube.
9. Resuspend the nuclear pellet using 0.5 pellet vol of hypotonic lysis buffer (the same amount of buffer as in **step 5**). Transfer the suspension to the Dounce and homogenize using two strokes.
10. Centrifuge the extract at 30,000g for 20 min.
11. Combine the supernatant with the supernatant in **step 8**.

3.3.2. Extraction of RISC From Ribosomes

1. Pellet the supernatant from **Subheading 3.3.1., step 11** at 200,000g for 2.5 h.
2. Discard the supernatant, and retain the sticky, dense pellet, which contains the ribosomes.

3. Using a flamed metal spatula, scrape the ribosomal pellet into a Dounce on ice. Resuspend the pellet in 1/10 of the volume of the cytosolic extract using high-salt extraction buffer. Use some of this buffer to remove traces of the pellet remaining in the centrifuge tubes.
4. Resuspend the pellet using the A pestle. Incubate the extract on ice for 30 min.
5. Return the high-salt extract to fresh centrifuge tubes and centrifuge at 200,000g for 2.5 h.
6. Remove the supernatant to a fresh tube, leaving the pellet behind. Also remove 1 to 2 μL of the ribosomal pellet and mix with the supernatant (*see Note 2*). Dilute the supernatant with 5 vol of hypotonic lysis buffer. Incubate at 4°C overnight.
7. The following day, centrifuge the low-salt precipitate at 20,000g for 10 min. Decant the supernatant, and resuspend the pellet in 1 mL of buffer A amended to contain 0.4 M KCl. Transfer the resuspended pellet to a microfuge tube.
8. Put the tube containing the pellet suspension on a rotator, and allow the pellet to resuspend for 30 min.
9. Pellet the insoluble material for 5 min at 13,000g at 4°C in a standard microfuge.
10. Filter extract through a 0.45- μm syringe-tip filter. The sample is ready for column chromatography.

3.4. Purification of RISC by Column Chromatography

We have successfully used a variety of fractionation schemes, but we always begin with size fractionation. Without size fractionation, the material will precipitate when diluted to low-salt conditions for ion-exchange chromatography. In particular, overloading of the sizing column (by loading too large a volume) can result in fractions that precipitate when diluted for ion exchange. The purification scheme described contains size fractionation (Superose 6), anion-exchange (Source Q), and heparin affinity steps. We also used hydroxylapatite and cation exchange (Source S) in some purifications.

The activity is fairly unstable; therefore, it is best to minimize the purification time. With the use of a phosphorimager, it is possible to analyze RISC assays after a 1-h exposure. This brief exposure is sufficient to determine active fractions and proceed to the next purification step. The activity becomes progressively less stable with increasing degrees of purification.

We use an AKTA fast protein liquid chromatography (FPLC) system for purification, but any medium-pressure system should work well.

3.4.1. Size Fractionation

1. Equilibrate sizing columns with buffer A amended to contain 0.4 M KCl. We generally use two Superose 6 HR 10/30 columns connected in series. We operate these at a flow rate of 0.3 mL/min.
2. Load a 2-mL sample loop with no more than 1 mL of extract, either the resuspended low-salt precipitate described in **Subheading 3.3.2., step 7** or as described in **Note 3**.

3. Run a program with a 0.3 mL/min flow rate. Inject 4 mL of buffer through the sample loop, using 4 mL, followed by 1.2 column volumes of elution buffer. Collect 1-mL fractions.
4. Assay the fractions for RISC activity as in **Subheading 3.5**. The activity is generally observed at 30–32 mL from the beginning of sample injection.

3.4.2. Anion Exchange

1. Prepare FPLC with buffer A in pump A and buffer A containing 1 M KCl in pump B.
2. Preequilibrate 0.5 mL of Source Q HR 5/5 column with buffer A containing 1 M KCl, and then equilibrate the column with buffer A. This column is homemade from Source Q resin. To achieve such a small column, put adjustable beds on both ends to make as small a column as possible.
3. Dilute active size fractions with 5 vol of buffer A.
4. Load a superloop with the diluted extract. Run a program with a 0.5 mL/min flow rate that loads the entirety of the sample onto the column, washes with 2 column vol of buffer A, and runs a 20-column-vol gradient from pure buffer A to buffer A with 1 M KCl. Collect 0.5-mL fractions.
5. Assay the fractions for RISC activity as in **Subheading 3.5**. The activity generally elutes very early, just when the conductivity is beginning to increase—around fraction five or six.

3.4.3. Heparin

1. Equilibrate 1 mL of heparin HiPrep column with buffer A.
2. Dilute active fractions to approx 50 mM KCl (generally 1 to 2 vol of buffer A).
3. Load a sample loop or superloop with the diluted extract. Run a program with a 100 μ L/min flow rate that loads the sample onto the column, washes with 2 column vol of buffer A, and runs a 10-column-vol gradient to 100% buffer B. Collect 0.5-mL fractions.
4. Assay the fractions for RISC activity as in **Subheading 3.5**. The activity generally elutes around fraction 11 or 12.

3.5. RISC Assays

3.5.1. Preparation of Substrate

Radiolabeled RNAs are prepared using the Promega Riboprobe T7 kit and α -labeled rNTPs. The RNA templates are gel-purified PCR products with a single T7 promoter introduced on either the 5' (to produce a sense transcript) or the 3' terminus (to produce an antisense transcript). The RNA is transcribed and gel purified from low-melt agarose. To avoid RNase contamination, we use a dedicated gel box that has been carefully cleaned with hydrogen peroxide and washed extensively with water.

1. Set up transcription reactions according to the manufacturer's recommendations, except increase the concentration of the cold nucleotide counterpart to the labeled

nucleotide from 25 μM to 0.1 mM (the other nucleotides are at 0.5 mM in the reaction).

2. Incubate transcription reactions for 1 to 2 h at 37°C.
3. Following the reaction, add TE buffer to 200 μL and extract the sample with phenol/chloroform. Remove the top phase to a fresh tube.
4. Add 2 μL of 20 mg/mL glycogen, 200 μL of 5 M NH_4OAc , and 1 mL of ethanol. Centrifuge for 15 min.
5. Remove the supernatant and wash the pellet with 0.5 mL of 70% ethanol.
6. Resuspend the pellet in formamide loading dye.
7. Heat at 95°C for 2 min and place immediately on ice.
8. Load the samples on a 1X TAE 1% low-melt agarose gel.
9. Wrap the gel in Saran Wrap and apply a glow-in-the-dark marker to facilitate alignment of the developed film with the gel.
10. In a darkroom, place a piece of X-ray film directly on the gel and expose for 5 min. Develop the film.
11. Using the marker, align the film with the gel and use a marker to draw a box around the band to be cut.
12. Cutting directly through the Saran Wrap surrounding the gel, excise the band with a scalpel and transfer to a fresh tube.
13. Add 400 μL of water and 500 μL of phenol (not phenol/chloroform).
14. Heat the tube at 65°C until the gel slice is dissolved. The phenol and aqueous phases may merge; this is not a problem.
15. Centrifuge for 5 min to separate the phases.
16. Remove the top layer to a fresh tube, leaving behind the white interface.
17. Repeat extraction of the upper aqueous layer with phenol, and then with chloroform.
18. Transfer 400 μL of the chloroform-extracted material to a fresh tube. Add 2 μL of glycogen; 40 μL of 3 M sodium acetate, pH 4.8; and 1 mL of ethanol.
19. Centrifuge for 15 min and remove the supernatant.
20. Wash the pellet with 0.5 mL of 70% ethanol. Remove the supernatant and allow the pellet to dry.
21. Resuspend the pellet in 50 μL of water and count 1 μL in a scintillation counter.
22. Dilute the substrate to 1000 cpm/ μL .
23. Confirm RNA quality by running a 5% acrylamide/8 M urea/1X TBE gel. RNA should form a single band.

3.5.2. Activity Assay

1. Set up reactions for a 10- μL final volume containing 1X buffer F, 1 mM ATP, 1 mM MgCl_2 , 1 U of RNasin (Promega), 1000 cpm of RNA, and 2 μL of extract (see **Note 4**). Set up two reactions for each fraction, one with the targeted (homologous) substrate and the other with a control (nonhomologous) substrate. Be sure to include a control with no extract.
2. Incubate for 1 h at 30°C.
3. Terminate the reactions by adding 250 μL of Trizol amended to contain 10 μL of glycogen/ mL of Trizol.

4. Add 50 μ L of chloroform, vortex, and centrifuge at 16,000g for 3 min.
5. Remove the supernatant to a fresh tube containing 150 μ L of isopropanol.
6. Centrifuge at 16,000g for 15 min. Aspirate the supernatant.
7. Resuspend the pellet in formamide loading dye. Analyze the transcripts on a 5% acrylamide/8 M urea/1X TBE gel.

4. Notes

1. The hypotonic lysate, high-salt extract, and low-salt precipitate all contain RISC activity. These preparations retain robust activity after snap freezing on dry ice/ethanol (unlike column fractions, which retain little or no activity after freezing), and they can be useful both as controls and for characterizing RISC activity.
2. The addition of a small amount of ribosomes facilitates RISC precipitate formation. The low-salt precipitation solution should be somewhat cloudy.
3. For smaller-scale analytic purifications, we commonly load the initial high-salt-extracted ribosome pellet on the sizing column and fractionate without doing low-salt precipitation. Low-salt precipitation is most effective when working with 4 L or more of starting culture, although many experiments can be effectively done with smaller amounts of material.
4. RISC activity assays should be adjusted to accommodate high salt in column fractions. For instance, when assaying sizing fractions in 0.4 M KCl, reduce buffer F to 0.5X (55 mM KCl) and use 2 μ L of the column fraction to achieve approx 180 mM salt. Similar adjustments can be made for assays of heparin fractions.

Acknowledgments

G. J. Hannon is a Rita Allen Foundation Scholar and is supported by an Innovator Award from the U.S. Army Breast Cancer Research Program. A. A. Caudy is a George A. and Marjorie H. Anderson Fellow of the Watson School of Biological Sciences and an HHMI Predoctoral Fellow. This work was also supported by a grant from the National Institutes of Health.

References

- 1 Hannon, G. J. (2002) RNA interference. *Nature* **418**, 244–251.
- 2 Ambros, V. (2001) microRNAs: tiny regulators with great potential. *Cell* **107**, 823–826.
- 3 Hammond, S. M., Bernstein, E., Beach, D., and Hannon, G. J. (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404**, 293–296.
- 4 Tuschl, T., Zamore, P. D., Lehmann, R., Bartel, D. P., and Sharp, P. A. (1999) Targeted mRNA degradation by double-stranded RNA in vitro. *Genes Dev.* **13**, 3191–3197.
- 5 Ketting, R. F., Fischer, S. E., Bernstein, E., Sijen, T., Hannon, G. J., and Plasterk, R. H. (2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.* **15**, 2654–2659.

- 6 Hutvagner, G. and Zamore, P. D. (2002) A microRNA in a multiple-turnover RNAi enzyme complex. *Science* **297**, 2056–2060.
- 7 Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R., and Tuschl, T. (2002) Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* **110**, 563–574.
- 10 Yang, D., Buchholz, F., Huang, Z., Goga, A., Chen, C. Y., Brodsky, F. M., and Bishop, J. M. (2002) Short RNA duplexes produced by hydrolysis with *Escherichia coli* RNase III mediate effective RNA interference in mammalian cells. *Proc. Natl. Acad. Sci. USA* **99**, 9942–9947.
- 9 Billy, E., Brondani, V., Zhang, H., Muller, U., and Filipowicz, W. (2001) Specific interference with gene expression induced by long, double-stranded RNA in mouse embryonal teratocarcinoma cell lines. *Proc. Natl. Acad. Sci. USA* **98**, 14,428–14,433.
- 10 Yang, S., Tutton, S., Pierce, E., and Yoon, K. (2001) Specific double-stranded RNA interference in undifferentiated mouse embryonic stem cells. *Mol. Cell. Biol.* **21**, 7807–7816.
- 11 Tang, G., Reinhart, B. J., Bartel, D. P., and Zamore, P. D. (2003) A biochemical framework for RNA silencing in plants. *Genes Dev.* **17**, 49–63.
- 12 Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363–366.
- 13 Zamore, P. D., Tuschl, T., Sharp, P. A., Bartel, D. P. (2000) RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**, 25–33.
- 14 Hammond, S. M., Boettcher, S., Caudy, A. A., Kobayashi, R., and Hannon, G. J. (2001) Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **293**, 1146–1150.
- 15 Caudy, A. A., Ketting, R. F., Hammond, S. M., et al. (2003) A staphylococcal nuclease homolog as a component of RNAi effector complexes in *Drosophila*, *C. elegans* and mammals. *Nature*. **425(6956)**, 411–414.
- 16 Caudy, A. A., Myers, M., Hannon, G. J., and Hammond, S. M. (2002) Fragile X-related protein and VIG associate with the RNA interference machinery. *Genes Dev.* **16**, 2491–2496.
- 17 Lim, L. P., Lau, N. C., Weinstein, E. G., Abdelhakim, A., Yekta, S., Rhoades, M. W., Burge, C. B., and Bartel, D. P. (2003) The microRNAs of *Caenorhabditis elegans*. *Genes Dev.* **17**, 991–1008.
- 18 Brennecke, J., Hipfner, D. R., Stark, A., Russell, R. B., and Cohen, S. M. (2003) Bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene *hid* in *Drosophila*. *Cell* **113**, 25–36.
- 19 Abrahante, J. E., Daul, A. L., Li, M., et al. (2003) The *Caenorhabditis elegans* hunchback-like gene *lin-57/hbl-1* controls developmental time and is regulated by microRNAs. *Dev. Cell* **4**, 625–637.
- 21 Llave, C., Xie, Z., Kasschau, K. D., and Carrington, J. C. (2002) Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. *Science* **297**, 2053–2056.

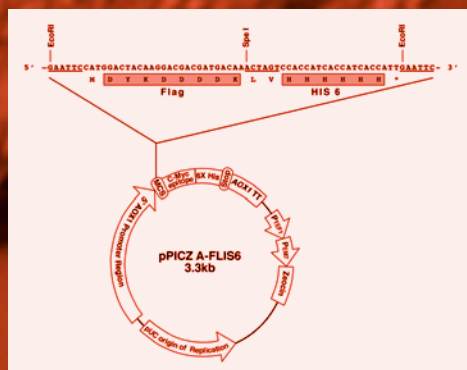
- 21 Olsen, P. H. and Ambros, V. (1999) The *lin-4* regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation. *Dev. Biol.* **216**, 671–680.
- 22 Grishok, A., Pasquinelli, A. E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D. L., Fire, A., Ruvkun, G., and Mello, C. C. (2001) Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* **106**, 23–34.
- 27 Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., and Ruvkun, G. (2000) The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **403**, 901–906.
- 24 Seggerson, K., Tang, L., and Moss, E. G. (2002) Two genetic circuits repress the *Caenorhabditis elegans* heterochronic gene *lin-28* after translation initiation. *Dev. Biol.* **243**, 215–225.
- 25 Slack, F. J., Basson, M., Liu, Z., Ambros, V., Horvitz, H. R., and Ruvkun, G. (2000) The *lin-41* RBCC gene acts in the *C. elegans* heterochronic pathway between the *let-7* regulatory RNA and the LIN-29 transcription factor. *Mol. Cell* **5**, 659–669.
- 26 Elbashir, S. M., Lendeckel, W., and Tuschl, T. (2001) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* **15**, 188–200.
- 27 Spankuch-Schmitt, B., Bereiter-Hahn, J., Kaufmann, M., and Strebhardt, K. (2002) Effect of RNA silencing of polo-like kinase-1 (PLK1) on apoptosis and spindle formation in human cancer cells. *J. Natl. Cancer Inst.* **94**, 1863–1877.
- 28 Capodici, J., Kariko, K., and Weissman, D. (2002) Inhibition of HIV-1 infection by small interfering RNA-mediated RNA interference. *J. Immunol.* **169**, 5196–5201.
- 29 Wilson, J. A., Jayasena, S., Khvorova, A., Sabatinos, S., Rodrigue-Gervais, I. G., Arya, S., Sarangi, F., Harris-Brandts, M., Beaulieu, S., and Richardson, C. D. (2003) RNA interference blocks gene expression and RNA synthesis from hepatitis C replicons propagated in human liver cells. *Proc. Natl. Acad. Sci. USA* **100**, 2783–2788.
- 30 McCaffrey, A. P., Nakai, H., Pandey, K., Huang, Z., Salazar, F. H., Xu, H., Wieland, S. F., Marion, P. L., and Kay, M. A. (2003) Inhibition of hepatitis B virus in mice by RNA interference. *Nat. Biotechnol.* **21**, 639–644.
- 31 Clemens, J. C., Worby, C. A., Simonson-Leff, N., Muda, M., Maehama, T., Hemmings, B. A., and Dixon, J. E. (2000) Use of double-stranded RNA interference in *Drosophila* cell lines to dissect signal transduction pathways. *Proc. Natl. Acad. Sci. USA* **97**, 6499–6503.
- 32 Worby, C. A., Simonson-Leff, N., and Dixon, J. E. (2001) RNA interference of gene expression (RNAi) in cultured *Drosophila* cells. *SciSTKE* **95**, PL1.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



Analysis of Gene Function in *Trypanosoma brucei* Using RNA Interference

Appolinaire Djikeng, Shuiyuan Shen, Christian Tschudi,
and Elisabetta Ullu

Summary

Trypanosoma brucei, a flagellate protozoa of the family Trypanosomatidae, has become one of the model systems for unicellular pathogens to study fundamentally important biological phenomena. Currently, the method of choice to examine gene function in these organisms is RNA interference (RNAi). mRNA degradation is triggered by double-stranded RNA (dsRNA) produced in vivo from transgenes transcribed from opposing tetracycline (tet)-inducible T7 RNA polymerase promoters, or hairpin RNA transcribed from the tet-inducible procyclic acidic repetitive protein promoter. In this chapter, we describe some of the methods we employ for ablation of gene expression by RNAi in *T. brucei* with particular emphasis on transfection and cloning of procyclic cells, induction of dsRNA expression, isolation of RNA and analysis of dsRNA, and target mRNA.

Key Words: DNA transfection; hairpin construct; *Trypanosoma brucei*; clonal cell line; RNA dot blot; inducible promoter.

1. Introduction

Over the last 5 yr, doubled-stranded RNA (dsRNA)-mediated genetic interference or RNA interference (RNAi) has been shown to exist in several organisms representing different levels of the evolutionary tree (*1*). RNAi is a posttranscriptional mechanism in which expression of dsRNA or delivery of synthetic dsRNA into cells causes specific degradation of the target mRNA. Within a relatively short time of its discovery, RNAi has rapidly been developed as a valuable tool for “reverse” genetic studies. The specific degradation of an

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

mRNA often results in a detectable phenotype, and in some cases RNAi phenotypes have been shown to be comparable with loss-of-function phenotypes.

In *Trypanosoma brucei*, RNAi has been used to study several biological processes including, but not limited to, kinetoplast DNA replication, glycosyl phosphatidylinositol biosynthesis, glycosome biosynthesis, RNA editing, and flagellum biogenesis (2–5). Here we describe methods and reagents available for inducible and heritable RNAi in procyclic *T. brucei* cells.

2. Materials

2.1. Generation of RNAi Constructs

1. pLew79- and pLew100-based vectors for in vivo expression of dsRNA.
2. Oligonucleotide primers.
3. Expand High Fidelity polymerase chain reaction (PCR) system (cat. no. 1732641; Roche).
4. Qiaquick PCR purification kit (cat. no. 28104; Qiagen).
5. Qiaprep Spin Miniprep Kit (cat. no. 27106; Qiagen).
6. Rapid DNA Ligation kit (cat. no. 1635379; Roche).
7. GeneClean II Kit (cat. no. 1001-400; Qbiogene).
8. HiSpeed Plasmid Midi Kit (cat. no. 12643; Qiagen).

2.2. Transfection of Procyclic *T. brucei* Cells by Electroporation

1. Procyclic form of *T. brucei* cells, named 29.13.6, that stably express the T7 RNA polymerase and the tetracycline repressor (ref. 6 and see Note 1).
2. Cunningham medium (ref. 7 and see Note 2) supplemented with 20% heat-inactivated fetal bovine serum (FBS) (cat. no. 100-106; Gemini), 2 mM L-glutamine (cat. no. 25030-081; Gibco, Gaithersburg, MD), 100 U/mL of penicillin and 100 µg/mL of streptomycin (cat. no. 15140-122; Gibco), and 0.01 mg/mL of gentamycin (cat. no. 100-106; Gemini).
3. CO₂ incubator at 28°C (5% CO₂ in air).
4. Cytomix: 120 mM KCl, 0.15 mM CaCl₂, 10 mM K₂HPO₄, 25 mM HEPES, 2 mM EGTA, 5 mM MgCl₂. Dissolve all the ingredients in 450 mL of high-performance liquid chromatography (HPLC)-grade water, adjust to pH 7.6 with KOH, adjust the volume to 500 mL, and filter sterilize. Store at room temperature (stable for several months).
5. Bio-Rad gene Pulser II (Bio-Rad, Hercules, CA).
6. Electroporation cuvet (0.4-cm gap) (cat. no. 165-2088; Bio-Rad).
7. Neomycin or G418 (cat. no. 1181-031; Gibco) used at 15 µg/mL for maintenance of stable cell lines or at 60 µg/mL for the initial selection of stable transfectants immediately after transfection, Hygromycin B (cat. no. 843555; Roche) used at 50 µg/mL and Phleomycin (cat. no. P9564; Sigma, St. Louis, MO) used at 2.5 µg/mL.

2.3. Cloning of *T. brucei* Cells

1. Multichamber pipet: Finnpiptet Varichannel 200-1000 (cat. no. 4347030; Labsystems).
2. 96-Well microtiter plates (cat. no. 3595; Corning).
3. 48-Well microtiter plates (cat. no. 35-3078; Falcon).

2.4. Induction of dsRNA Expression

1. Tetracycline hydrochloride (cat. no. T3383; Sigma): Dissolve the powder in 70% ethanol to make a stock solution of 10 mg/mL. Store in the dark at -20°C . Replenish every week.

2.5. Extraction of Total RNA

1. Cell wash: 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 3 mM MgCl_2 . Filter sterilize and store at room temperature (stable for several months).
2. RNeasy Tissue Kit (cat. no. 69506; Qiagen).
3. RNase-free DNase I (cat. no. 776-785; Roche).
4. HPLC-grade water (cat. no. 7732-18-5; J. T. Baker).

2.6. RNA Dot Blots

1. Dot-blot minifold apparatus (cat. no. 44-27510; Schleicher & Schuell).
2. Protran nitrocellulose membrane (cat. no. 10402594; Schleicher & Schuell).

2.7. Northern Blot Analysis

1. Seakem Gold Agarose (cat. no. 50150; BMA).
2. Electrophoresis apparatus with connections to a circulating pump.
3. Hybridization chamber with variable temperature settings from 38 to 60°C .
4. 20X SET: 3 M NaCl, 20 mM EDTA, 0.6 M Tris-HCl. Adjust the pH to 8.0, filter sterilize, and store at room temperature (stable for several months).
5. Prehybridization solution: 10X Denhardt's, 50% formamide (cat. no. AB00600; American Bioanalytical), 5X SET, 100 $\mu\text{g/mL}$ of yeast total RNA, 1% sodium dodecyl sulfate (SDS).
6. 10X MOPS buffer: 0.4 M MOPS, pH 7.0, 0.1 M NaOAc, 0.01 M EDTA.
7. Formaldehyde (37% solution) (cat. no. 50-00-0; J. T. Baker).
8. 20X saline sodium citrate (SSC): 3 M NaCl, 0.3 M Na citrate.
9. RNA gel-loading solution (cat. no. 8552; Ambion).
10. HybondTM-N membrane (cat. no. RPN303N; Amersham Pharmacia Biotech).
11. Methylene blue solution (cat. no. 193998; ICN): Make a working solution of 0.04% in 0.5 M NaCH_3COO (pH 5.2) and store at room temperature (stable for several months).
12. DNA fragment to be labeled.
13. $[\alpha\text{-}^{32}\text{P}]\text{dCTP}$ (cat. no. BLU513H; NEN).
14. Random Prime Labeling System (RediprimeTM II; cat. no. RPN1633; Amersham).
15. Bio-spin 6 (P6) chromatography columns (cat. no. 732-6002; Bio-Rad).

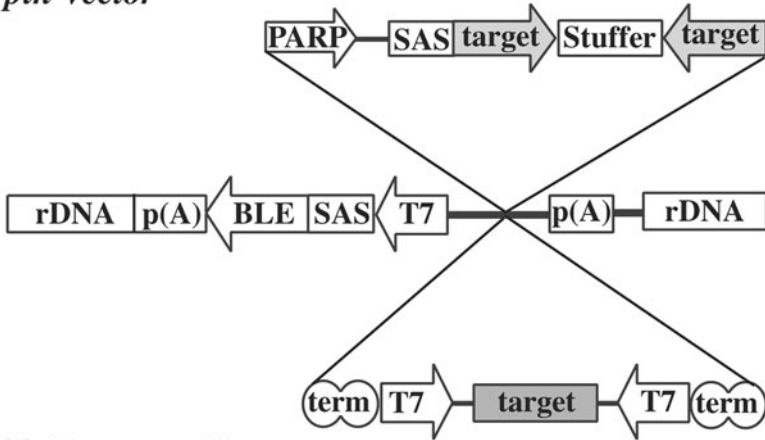
Hairpin Vector***Double Promoter Vector***

Fig. 1. Vectors for RNAi in *T. brucei*. PARP arrow, tetracycline-inducible PARP promoter; SAS, splice acceptor site; rDNA, ribosomal DNA nontranscribed spacer; p(A), poly(A) addition site; BLE, phleomycin resistance gene; T7 arrow, tetracycline-inducible T7 promoter; term, T7 transcription terminators.

3. Methods

3.1. Generation of RNAi Constructs

To obtain stable and inducible expression of dsRNA, an appropriate construct needs to be inserted into the genome and a stable cell line expressing the dsRNA needs to be selected. At the present time, there are two types of vectors that we and others have constructed for eliciting an RNAi response in *T. brucei* (**Fig. 1**). The first vector uses the tetracycline (tet)-inducible promoter from the procyclic acidic repetitive protein (PARP) genes (**4,8**) to drive expression of hairpin RNAs (*see Notes 3 and 4*). The generation of a construct containing two inverted repeats of the gene of interest (*see Note 5*) separated by a stuffer fragment requires a two-step cloning procedure (**2**). The second vector is equipped with opposing tet-inducible T7 RNA polymerase promoters (**2,9**). In this case, a simple cloning step is required to insert a portion of the gene of interest in between two T7 promoters of pZJM (**2**) for producing dsRNA in vivo (*see Note 3*). The hairpin and double-T7 constructs are linearized with *EcoRV* and *NotI*, respectively, for integration at the rDNA nontranscribed spacer region of a special recipient strain, named 29.13.6, expressing the tet repressor and T7 RNA polymerase (**ref. 6** and *see Note 1*).

3.2. Transfection of Procyclic *T. brucei* Cells by Electroporation

Procyclic *T. brucei* cultures are maintained in Cunningham's medium (7) containing 20% heat-inactivated FBS at 28°C in a standard incubator (no CO₂ is required, except for cloning cells in microtiter plates). Each transfection for establishing stable cells lines requires 5×10^7 cells and about 20 µg of DNA.

1. The day before the transfection, dilute the cells in such a way that they reach a density of not more than 10^7 cells/mL the next day.
2. The day of the experiment, first do the following:
 - a. Aliquot 10 mL of complete medium in the required number of T flasks.
 - b. Aliquot 20 µg of linearized DNA into Eppendorf tubes and adjust the volume to 100 µL with H₂O.
 - c. Include a no-DNA control tube with 100 µL of H₂O.
 - d. Prepare and label the required number of 0.4-cm gap cuvetts.
 - e. Turn on the Bio-Rad gene Pulser II and set the capacitance at 25 µF and the voltage at 1.5 kV.
3. Spin down the cells (5×10^7 cells/transfection) at 2000g for 5 min.
4. Discard the supernatant and resuspend the cell pellet in cytomix in $\frac{1}{2}$ of the original culture volume.
5. Spin again and resuspend the cell pellet in cytomix in $\frac{1}{20}$ of the original culture volume.
6. Aliquot 0.5 mL of cells each into the Eppendorf tubes containing the plasmid DNA, mix by pipetting up and down, and transfer the mixture into a 0.4-cm gap cuvet.
7. Close the cuvet cap loosely and zap twice, waiting 10 s in-between zaps (time constant is normally between 0.4 and 0.5).
8. Transfer the electroporated cells immediately into the flasks prepared in **step 2a** and rinse the cuvet with some medium.
9. Incubate the cells at 28°C for 24 h without drugs. This allows the transfected cells to recover and begin to express the selectable marker.
10. After 24 h, add the appropriate drug. At this point, cells can be either cloned (as described in **Subheading 3.3.**) or selected as a population (go to **step 11**).
11. Monitor the cells, and if they grow to approx 10^7 cells/mL, dilute them to 10^6 cells/mL and continue selecting.
12. Once the control cells die, the selected cells are considered stable. Now test the cells for expression of dsRNA and also for downregulation of the targeted mRNA.
13. Prepare a frozen stock of the population (*see Note 6*) and clone the cells (*see Subheading 3.3.*).

3.3. Cloning of Procyclic *T. brucei* Cells

Although certain RNAi phenotypes are going to be evident in a cell population, others will only become apparent in a clonal cell line (*see Note 7*). We

clone cells by limiting dilution, which can be applied to a stable cell population or to cells 2 d posttransfection (*see Note 8*).

1. Prewarm medium at 28°C for at least 30 min and add the appropriate drug(s) required for selection.
2. Count the cells and prepare four sets of dilutions using the medium prepared in **step 1**, so as to have one set at 0.3 cell/50 μ L, a second set at 2 cells/50 μ L, a third set at 4 cells/50 μ L, and a fourth set at 10 cells/50 μ L. Also dilute one set of wild-type cells so as to have 100 cells/50 μ L (the wild-type cells will condition the medium). Plan to use 30 wells of a 96-well microtiter plate for one set of dilutions. Add 50 μ L of each dilution and 50 μ L of wild-type cells.
3. Cover the microtiter plate, seal with Parafilm® and incubate in a humidified incubator at 28°C with 5% CO₂.
4. Monitor the cells microscopically every other day; by d 7–11 there should be a noticeable cell growth. For expansion, choose wells from the highest dilution. In a successful cloning, only one of three wells of a dilution should display growth.
5. Add 100 μ L of fresh complete medium to each of the chosen wells and continue the incubation. The following day, transfer the 200 μ L of cell culture into a 48-well dish and add 500 μ L of fresh medium.
6. When the cells are dense enough ($\sim 10^7$ cells/mL), transfer them into a T flask and bring the volume to 3 mL with medium. Cells can now be diluted to large volumes (10–20 mL), and at this point, aliquots should be frozen for long-term storage (*see Note 6*).

3.4. Induction of dsRNA Expression and Analysis of RNAi Cells

It is important to start with freshly diluted cells. If cells have been cloned, always plan to analyze several clones, since the level of expression of dsRNA can vary (**Fig. 2**) and that also correlates with differences in the efficiency of degradation of the target mRNA. The window for maximum degradation of the target mRNA should be determined for each experiment. This allows a better execution of a downstream assay for functional studies.

1. Dilute the cells to 1×10^6 cells/mL and add tetracycline (dissolved in 70% ethanol) to a final concentration of 10 μ g/mL.
2. Collect a total of 10^8 cells for isolation of total RNA at desired time points (then go to **Subheading 3.5.** for isolation of total RNA). The first time point for RNA analysis can be as early as 5 h. Keep the cell concentration at 1×10^6 /mL, and continue to add tetracycline to the culture medium every day. Also count the cells every day to monitor whether RNAi induction affects cell growth.

3.5. Isolation of Total RNA

1. Spin down 10^8 cells in a Falcon tube for 5 min at 2000g.
2. Wash the cells three times with cell wash buffer, and follow the RNeasy Tissue Kit for the isolation of total RNA.

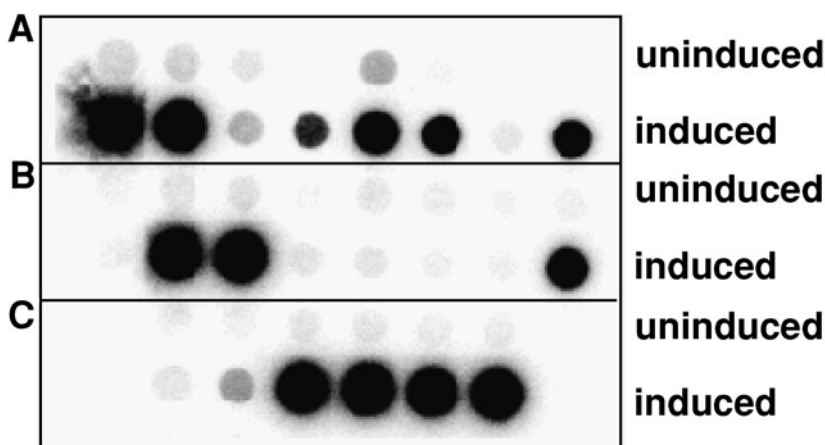


Fig. 2. Induction of dsRNA expression in clonal cells. RNAi cells targeting mRNAs for a *T. brucei* methyltransferase (A), and capping enzyme I (B) and II (C) were established. Hairpin constructs expressing dsRNAs were inserted at the rDNA locus and stable cell lines selected and cloned. The expression of dsRNA was induced with tetracycline (at 10 $\mu\text{g}/\text{mL}$) for 24 h. Total RNA was prepared, dot blotted on a membrane, and hybridized with a radiolabeled DNA fragment derived from the stuffer sequence.

3. Use an ultraviolet (UV) spectrophotometer to estimate the RNA concentration, and perform dot-blot analysis (see **Subheading 3.6.**) or Northern blot analysis (see **Subheading 3.7.**).

3.6. RNA Dot-Blot Analysis

Total RNA prepared from induced and noninduced cloned cells can be used to rapidly determine the expression level of dsRNA by dot-blot analysis. The probe can be a DNA fragment either homologous to the dsRNA or homologous to the stuffer fragment in the case of hairpin constructs.

3.6.1. Assembling Minifold Dot-Blot Apparatus

1. Wash the apparatus with 0.1 *N* NaOH and rinse well with water.
2. Assemble the apparatus according to the manufacturer's instructions, and use Whatman 3MM paper soaked in 20X SSC for at least 1 h.
3. Rinse the chambers with 10X SSC.
4. Refill the chambers with 10X SSC and turn off the vacuum.

3.6.2. Preparation and Processing of RNA Samples

1. Aliquot 2–5 μg of total RNA, in H_2O , in a total volume of 10 μL .
2. Add 20 μL of formamide, 7 μL of formaldehyde, and 2 μL of 20X SSC, and incubate at 68°C for 15 min.

3. Cool the samples on ice and add 2 vol of 20X SSC.
4. Turn on the vacuum to remove the liquid and then turn off the vacuum.
5. Load the samples into the chambers, turn on the vacuum, and wash twice with 1 mL of 10X SSC.
6. Keep the vacuum on for an additional 5 min.
7. Remove the nitrocellulose filter, dry at room temperature and bake the filter at 80°C for 2 h.
8. Set the filter for prehybridization and hybridization as described in **Subheading 3.7.3**.

3.7. Northern Blot Analysis

A successful study of gene function using RNAi depends on the degradation of the corresponding mRNA. It is therefore essential to check the expression of the dsRNA and the degradation of mRNA by Northern blot analysis.

3.7.1. Gel Electrophoresis and Transfer

1. Thoroughly clean the gel electrophoresis equipment and rinse it several times with Millipore water.
2. Make a 1.2% agarose slurry in 1X MOPS buffer: for a 100-mL gel slurry, weigh 1.2 g of agarose, and add MOPS and Millipore water to 83 mL. Melt in a microwave oven on medium heat and cool to 60°C.
3. In a chemical hood, add 17 mL of formaldehyde to the gel slurry, mix, and pour in the gel-casting tray. Let the gel solidify for 30 min.
4. In the meantime, prepare the RNA sample by mixing the following: 1 part of RNA in water and 3 parts of the RNA gel-loading solution. Incubate the mixture at 65°C for 15 min and spin briefly.
5. Transfer the gel to the electrophoresis tank containing 1X MOPS buffer and connect to a circulating pump.
6. Load the gel and run at 20–60 V for 6–12 h depending on the desired resolution.
7. After the run, rinse the gel several times with Millipore water for 20 min, changing the water every 5 min, and then soak the gel in 20X SSC for 20 min.
8. Set the transfer (*see Note 9*) with 10X SSC using a Hybond-N membrane.
9. After the transfer, dry the membrane and crosslink the RNA onto the filter by UV exposure. The length of UV exposure depends on the equipment used and should be determined experimentally.
10. To stain the RNA immobilized on the filter, transfer the filter into a solution of 0.04% methylene blue in 0.5 M Na acetate at room temperature for 30–120 s.
11. Wash the filter with water to destain, and scan or photocopy the filter for documentation.

3.7.2. Probe Labeling

1. Prepare DNA to be used as probe (*see Note 10*) by heating 20–25 ng in 45 µL of water to 100°C for 5 min to denature.

2. Place the tube on ice for 5 min, and then centrifuge briefly to collect the condensate.
3. Add the denatured DNA to a Rediprime labeling mix, add 5 μ L of [α - ^{32} P]dCTP, and incubate at 37°C for 60 min.
4. In the meantime, prepare one P-6 column per sample. Spin at 1000g for 2 min and discard the flowthrough.
5. Load the entire labeling reaction onto a P-6 column and spin at 1000g for 5 min. Collect the eluate and determine the specific activity by scintillation counting.

3.7.3. Hybridization

1. Place the filter in a clean hybridization container, add enough prehybridization solution to cover the entire filter, and prehybridize at 50°C for at least 30 min.
2. Denature the probe by heating to 100°C for 5 min and then place it on ice for 3 min. Add the probe to the prehybridized filter at a concentration of at least 10^6 cpm/mL of hybridization solution.
3. Allow the hybridization to go overnight and wash the filter with 2X SSC, 0.1% SDS at 60°C, until the background is acceptable. Expose the filter to a phosphorimager screen or X-ray film.

4. Notes

1. The 29.13.6 cells are a procyclic form of *T. brucei* that have been genetically modified by insertion into the genome of two cassettes for expression of the T7 RNA polymerase and the tetracycline repressor (6). They should be grown in the presence of G418 (at 15 μ g/mL) and Hygromycin B (at 50 μ g/mL), since expression of T7 RNA polymerase and tet repressor were selected with these markers.
2. We routinely culture the procyclic form of *T. brucei* cells in suspension at 28°C. We use Cunningham medium (7), which can be prepared in the laboratory, filter sterilized, and stored in a refrigerator for several months. Alternatively, we have the medium custom-made by HyClone (cat. no. SH3A1286.01). Just before use, add FBS, penicillin/streptomycin, glutamine, and gentamycin.
3. In the absence of tetracycline, expression from the tet-induced PARP promoter is almost negligible, allowing cloning of toxic products, whereas under full-induction conditions high levels of expression can be achieved. On the other hand, the double-T7 promoter vector has the problem that a considerable level of dsRNA is produced even in the absence of tetracycline (2), because binding of the tet repressor to the tet operators does not completely shut off the T7 promoters (6). Thus, in certain instances one should use caution in using this particular vector. In addition, the low-level expression can lead to the selection of cells that express dsRNA that target the corresponding mRNA, but that have adapted to live normally with low levels of the targeted mRNA. Thus, we use tetracycline-screened serum (cat. no. SH30070.03; HyClone) whenever we target an essential gene with RNAi. In this instance, we begin to use the tet-free medium immediately after transfection.
4. We previously used a spliced leader RNA gene fragment, but later we observed that a different stuffer fragment, a piece of the pex11 gene (2), worked better. In

fact, RNAi cells generated with constructs containing the pex11 stuffer fragment acquired the phenotype much more rapidly than those generated with constructs made with the SL stuffer (unpublished findings).

5. We generally use 100- to 1000-bp DNA fragments for RNAi. If possible, we use only a portion of the gene for RNAi, leaving another region to be used as a probe, since the detection of the dsRNA can, in some cases, lead to hybridization background.
6. To prepare glycerol stocks for long-term storage of *T. brucei* cells, first prepare the freezing medium (20% glycerol in complete Cunningham medium) and filter sterilize. Then spin down approx 10^8 cells at 3500 rpm for 5 min. Next remove the supernatant and resuspend the pellet with 1 mL of freezing medium. Finally, transfer the cells into a 1.5-mL screw-cap tube, place the tube in between two styro-foam racks, and transfer directly to -80°C for long-term storage.
7. It is our experience that sometimes the RNAi phenotype is much stronger in clonal cell lines, as compared to populations of stable transfectants. Therefore, we prefer to perform the analysis of the RNAi phenotype in clonal cell lines.
8. Two methods are presently available for the cloning of trypanosome cells. The first is by plating cells on agarose plates (10) and the second is by limiting dilution of cells. The cloning method by plating cells onto agarose plates involves several steps and is sometimes problematic.
9. There is a significant variation in methods used for transferring RNA to membranes. The upward capillary method is widely used and requires at least 16 h for maximum transfer of total RNA. We routinely use the downward capillary method. In our hands, this method allows a maximum transfer within 5 to 6 h.
10. We generally use as probe a region of the targeted mRNA not included in the dsRNA construct. This allows the analysis of mRNA degradation after induction of RNAi. We have also successfully used probes homologous to the stuffer fragment in Northern blot analysis.

Acknowledgments

This study was supported by National Institutes of Health grant AI28798.

References

- 1 Hannon, G. J. (2002) RNA interference. *Nature* **418**(6894), 244–251.
- 2 Wang, Z., Morris, J. C., Drew, M. E., and Englund, P. T. (2000) Inhibition of *Trypanosoma brucei* gene expression by RNA interference using an integratable vector with opposing T7 promoters. *J. Biol. Chem.* **275**, 40,174–40,179.
- 3 Wang, Z. and Englund, P. T. (2001) RNA interference of a trypanosome topoisomerase II causes progressive loss of mitochondrial DNA. *EMBO J.* **20**, 4674–4683.
- 4 Bastin, P., Ellis, K., Kohl, L., and Gull, K. (2000) Flagellum ontogeny in trypanosomes studied via an inherited and regulated RNA interference system. *J. Cell Sci.* **113**, 3321–3328.

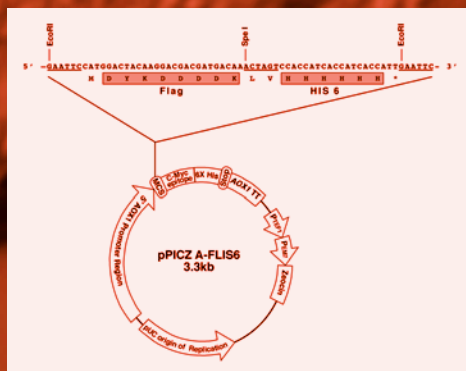
- 5 Drozd, M., Palazzo, S. S., Salavati, R., O'Rear, J., Clayton, C., and Stuart, K. (2002) TbMP81 is required for RNA editing in *Trypanosoma brucei*. *EMBO J.* **21**, 1791–1799.
- 6 Wirtz, E., Leal, S., Ochatt, C., and Cross, G. A. (1999) A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **99**, 89–101.
- 7 Cunningham, I. (1977) New culture medium for maintenance of tsetse tissues and growth of trypanosomatids. *J. Protozool.* **24**, 325–329.
- 10 Shi, H., Djikeng, A., Mark, T., Wirtz, E., Tschudi, C., and Ullu, E. (2000) Genetic interference in *Trypanosoma brucei* by heritable and inducible double-stranded RNA. *RNA* **6**, 1069–1076.
- 9 LaCount, D. J., Bruse, S., Hill, K. L., and Donelson, J. E. (2000) Double-stranded RNA interference in *Trypanosoma brucei* using head-to-head promoters. *Mol. Biochem. Parasitol.* **111**, 67–76.
10. Carruthers, V. B. and Cross, G. A. (1992) High-efficiency clonal growth of bloodstream- and insect-form *Trypanosoma brucei* on agarose plates. *Proc. Natl. Acad. Sci. USA* **89**, 8818–8821.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



Short Hairpin Activated Gene Silencing in Mammalian Cells

Patrick J. Paddison, Amy A. Caudy, Ravi Sachidanandam,
and Gregory J. Hannon

Summary

RNA interference (RNAi) is now a popular method for silencing gene expression in a variety of systems. RNAi methods use double-stranded RNAs (dsRNAs) to target complementary RNAs for destruction. In mammalian systems, very short dsRNAs (22–25 bp) such as short interfering RNAs (siRNAs) or short hairpin RNAs (shRNAs) are used to avoid endogenous nonspecific antiviral responses that target longer dsRNAs. siRNAs elicit a transient silencing response, while shRNAs can be expressed continuously to establish stable gene silencing. shRNAs can be introduced into cells and animals using a variety of standard vectors as well as retroviral or lentiviral expression systems. This chapter describes the design, construction, validation, and use of shRNAs for silencing genes. We report our results from testing a variety of shRNA design features and shRNA expression vectors. We also provide methods that use shRNAs to permit different levels of gene expression. Additionally, we discuss some aspects important for constructing an information pipeline to support development of a large shRNA library.

Key Words: RNA interference; U6; small hairpin; short interfering RNA; short hairpin activated gene silencing; library.

1. Introduction

RNA interference (RNAi) has now gained popularity in a variety of biological systems as a methodology of choice for knocking down gene expression (reviewed in **ref. 1**). In mammals, given the strong conservation of RNAi-related genes in vertebrates, including Dicer and Argonaute family members, the expectation was that RNAi would operate in some capacity. The use of conventional double-stranded RNA (dsRNA) triggers (such as the ~500-nt

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

RNAs used in plants and *Caenorhabditis elegans*) in somatic cells, however, is limited by antiviral/interferon responses, including the PKR and RNaseL pathways ([2]; reviewed in **ref. 3**), which trigger translational repression and apoptosis in response to dsRNA of >30 bp in length. One way around these nonspecific dsRNA responses is to simply create dsRNA triggers of RNAi <30 bp in length.

Over the past 2 yr, two short dsRNA structures have emerged that will evoke sequence-specific gene silencing in somatic cells without activating antiviral responses. These are the small interfering RNA (siRNA) and the short hairpin RNA (shRNA). Both are modeled after biologically active structures in the RNAi pathway: the Dicer cleavage product and small temporal RNAs or (microRNAs [miRNAs]), respectively.

Elbashir et al. (4) first demonstrated that small dsRNAs of ~21 nt resembling siRNAs from other systems can induce sequence-specific gene silencing when transiently transfected into mammalian cells. These siRNAs are chemically synthesized emulations of Dicer cleavage products, which are short RNA duplexes 21–22 nt in length containing 2-nt 3' overhangs. Thus, siRNAs presumably bypass the requirement for Dicer and enter the silencing pathway by incorporation into RNA-induced silencing complex (RISC) complexes. The use of siRNAs has been recently reviewed, in detail, and numerous resources for the design and use of siRNAs are available online from a number of vendors and investigators.

As an alternative strategy, we and others have developed in vivo expression constructs for small dsRNA triggers in mammalian cells, which resemble endogenously expressed hairpin RNAs (5–11). This approach uses small inverted repeats (19–29 nt) expressed from an RNA polymerase III promoter to create shRNAs, which can then be processed by Dicer and shunted into the RNAi pathway.

Both siRNAs and shRNAs have distinct advantages. Because this chapter focuses on the use of shRNAs, their positive characteristics are elaborated here. First, shRNAs are inexpensive, compared with the considerable cost of chemically synthesized siRNAs (~\$300.00 per duplex). Second, shRNAs provide the possibility of enforceable and stable expression of the silencing trigger. Third, shRNAs can be delivered using viral and other delivery vehicles into primary cell types and in whole organisms (e.g., mouse). Several methods have been used to drive expression of shRNAs, including RNA polymerase II and III promoters. Our preferred method is to drive expression of shRNAs by placing them behind the human RNA polymerase III U6 promoter. This expression system has now been demonstrated to be effective both in vitro (5,12,13) and in vivo: transiently in mouse (14), stably during hematopoiesis (13), and stably in the generation of transgenic mice (15). The methods presented herein focus

on the use of the U6 promoter. However, we note that similar results have also been obtained for the H1 RNA promoter, and in parallel experiments, we have not noted any specific advantage of either promoter for shRNA expression.

RNAi appears to be remarkably compatible with most of the commonly used techniques to introduce exogenous genes into mammals and mammalian cell systems. Early analyses of RNAi approaches have indicated that changes in gene expression are highly specific (**16,17**). However, some nonspecific effects have been observed, particularly when sequences show similarity to the targeted sequence (**18**). In general, most RNAi experiments seem to act in a specific manner. Improvements to current silencing strategies, such as optimal target sequences and introduction of inducible expression cassettes, will only serve to improve this technology in the future. The method described herein outlines how to design and clone shRNAs for expression in mammalian cells.

2. Materials

1. Cloning vector or polymerase chain reaction (PCR) template containing RNA polymerase III promoter, such as human or mouse U6, tRNA, or H1. Our laboratory commonly uses pGEM-zeo-U6 for PCR cloning (**12,13**).
2. Oligonucleotide primers containing shRNA sequences, designed as described in **Subheading 3.1.1**.
3. PCR reagents: *Taq* DNA polymerase, dNTPs, buffer, MgCl₂, dimethylsulfoxide (DMSO).
4. Restriction enzymes.
5. Agarose gel equipment and a kit to gel purify DNA fragments from agarose gels (e.g., Qiagen QiaexII).
6. Electro-competent or chemically competent *Escherichia coli* for cloning (e.g., DH5 α , DH10B).
7. Zeocin (Invitrogen); antibiotic for selection of inserts.
8. Appropriate reagents and apparatus for testing efficacy of shRNAs (e.g., Western blotting, Northern blotting, reverse transcriptase [RT-PCR]).
9. Viral expression vector (e.g., MSCV-puro [Clontech]).
10. Luria Bertani (LB)—low-salt medium for the use of Zeocin: 5 g/L of NaCl, 10 g/L of tryptone, 5 g/L of yeast extract, 20 g/L of agar (for plates). Adjust the pH to 7.5 using NaOH and autoclave. For plates, cool the medium to 55°C before adding Zeocin to 250 μ g/mL, and then pour the plates. Zeocin is not functional in standard LB medium—low-salt LB medium must be used, and the pH is critical.

3. Methods

We strongly recommend constructing three to six shRNAs per gene and carrying out a validation step before doing the actual biological experiment. For example, we often transiently transfect shRNA plasmids into any commonly used cell line that also expresses the gene of interest and assay efficacy by Western blotting, Northern blotting, or RT-PCR. Alternately, the target gene

can be introduced transiently along with the shRNA and assayed in a cell type lacking endogenous expression. Western blots or RT-PCR are good predictors of efficacy in these assays. We find a direct correlation between whether an shRNA works well in transient assays and when it is expressed stably (e.g., from a retrovirus), although very often the degree of suppression following stable expression from a retrovirus exceeds that seen in transient studies. In fact, we have been able to build “epiallelic” series by this method—a series of shRNAs with different efficacies that then give rise to phenotypes of a corresponding severity (*see Subheading 3.4.1.*). The correlation of knockdown and phenotype appears to hold true from transient and stable experiments in vitro to stable expression in vivo (*13*).

Concerning in vivo experiments in mouse, we have recently demonstrated that shRNAs can be used to make transgenic mice, and that the silencing effect is transmittable to the next generation (*15*). Important conclusions from this work are that RNAi is apparently functional in most tissues, and that mammalian development does not seem to present a barrier to the use of RNAi.

3.1. Designing and Cloning shRNAs

There are several methods for designing and cloning shRNAs, including annealed oligo- and PCR-based cloning. We have adopted the PCR-based method because it is cheaper for large-scale synthesis and is as reliable as the annealed oligo method. The PCR-based method is described in this chapter.

3.1.1. Designing shRNAs

In an attempt to optimize our shRNA strategy for triggering RNAi, we have first started by modifying variables that might affect the overall efficacy of hairpins containing the same target sequences ($n = 10$). These include promoter choice (human U6, human H1, mouse U6, mouse H1, cytomegalovirus); hairpin length (19, 21, 23, 25, 27, and 29 nt); loop structure (4–14 nt, mimicking miRNA loops, and so on); overall hairpin structure (e.g., miRNA mimic); and, finally, addition of the 27-nt U6 leader sequence. The results from these studies suggest the following (unpublished results). First, the RNA polymerase III promoters that we tested are largely functionally equivalent in commonly cultured cell lines. Second, simple hairpin structures containing 29 bp of non-bulged stem work more effectively than 19- and 21-nt hairpins. Third, for 29-nt hairpins, loop structure appears to be largely irrelevant. Fourth, adding the U6 leader sequence to shRNAs with mediocre efficacy can increase their potency.

These results suggest that 29-nt hairpins containing the U6 leader sequence and a simple loop structure will be the most effective (**Fig. 1**). Regarding promoter choice, we have decided to continue using the human U6 promoter in light of in vivo confirmation of U6-siRNA effectiveness in mouse (*13,14*).

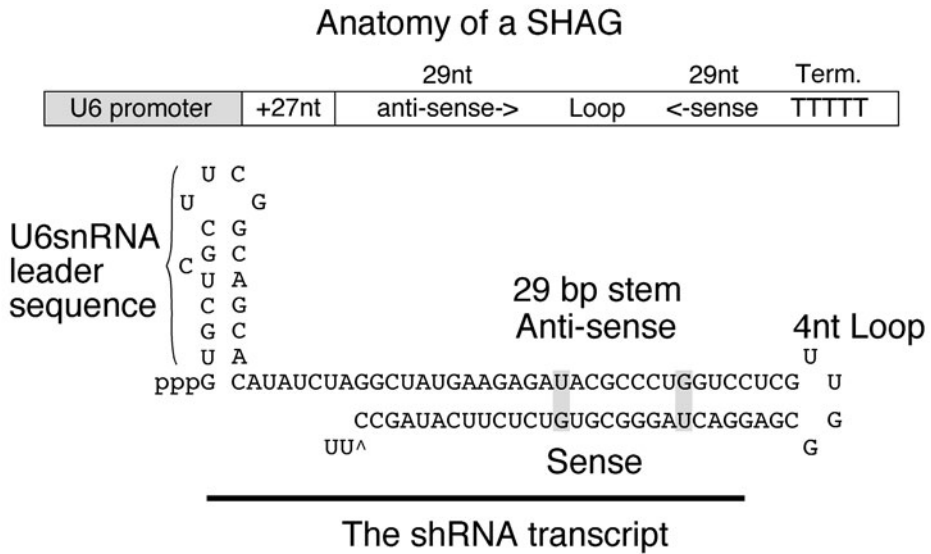


Fig. 1. Anatomy of short hairpin activated gene-silencing (SHAG) trigger. Our current design of the shRNAs contains the 27-nt U6 leader sequence, 29-nt dsRNA stem containing 29-nt of identity with the cognate gene target, several G-U pairs to promote stability in bacteria, and a small 4-nt loop.

The U6 snRNA leader transcript is a small RNA hairpin that directs the addition of a γ -monomethyl-phosphate guanosine cap (19). No work has yet suggested the mechanism that leads to improved shRNA efficacy, but increased stability, more effective transport, or a different localization are possibilities. Paul et al. (8) first suggested the use of the U6 leader sequence for 19-nt hairpins. For 29-nt hairpins that work well, the leader sequence is neutral (unpublished results).

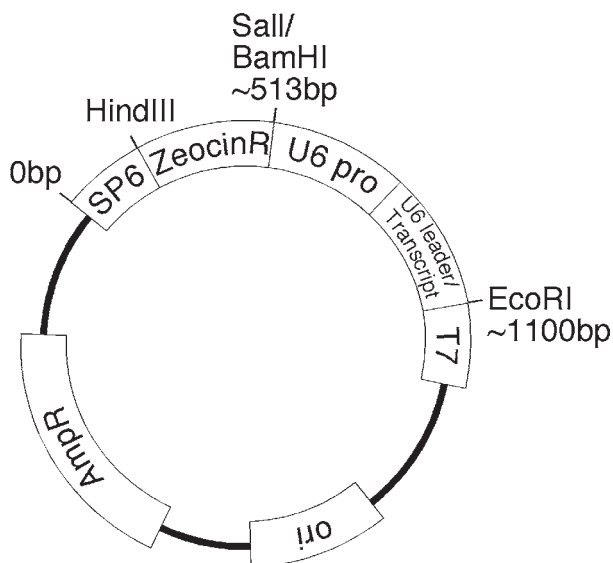
There is no conclusive evidence to explain why 29-nt hairpins are on average more effective than shorter hairpins. One possible explanation is that 29-nt hairpins may be a better substrate for Dicer than their 19- or 21-nt counterparts. Another is that this length increases the probability of obtaining a good targeting sequence. Twenty-nine-nucleotide hairpins indeed require Dicer for triggering RNAi (5), and human Dicer has been shown to efficiently cleave dsRNA as small as 30 bp (20).

3.1.2. Designing shRNA PCR Primers

To construct short hairpin expression vectors we use a PCR-based method (i.e., PCR-SHAG; Fig. 2A). The hairpin is added in its entirety to a 3' primer with identity to a U6 promoter PCR template (Fig. 2B). Using a universal

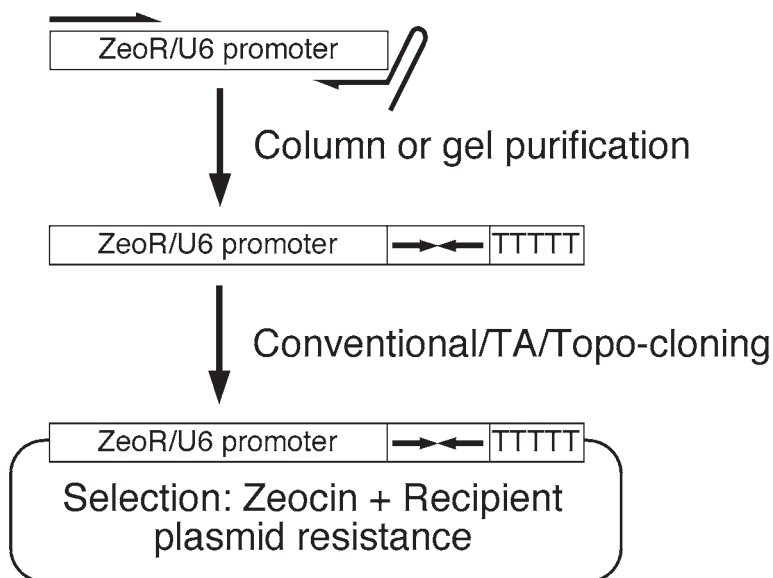
A

PCR-SHAG Template: pGEM-Zeo-HuU6



B

PCR-SHAG



5' primer, in this case SP6, for the PCR reaction results in a PCR product that contains the U6 promoter, the shRNA, and the RNA polymerase III termination sequence. To facilitate cloning and subsequent subcloning strategies, we have included a bacterial Zeocin resistance marker that is included within the U6 shRNA PCR product.

For conventional clonings, we generally add an *Xho*I site into the forward SP6 primer and an *Eco*RI site into the 3' hairpin primer. These enzymes work well and are compatible with many other vectors, such as pBluescript (Stratagene) and MSCV Puro (Clontech).

Following are the steps to design the primers for short hairpin expression vectors. Currently, we do not have any reliable rules to predict what part of the coding sequence will form the most effective hairpin. We do select regions with 40–60% GC content, avoid homopolymeric regions (which can be problematic in oligo synthesis), and confirm that sequences are unique. In the absence of a method for predicting the efficacy of any given hairpin, we generally construct three to six hairpins and then test them. We have automated this process to generate hairpins from large gene sets (described in detail in **Subheading 3.5.1.**).

1. To design the 3' hairpin primer, choose a 29-nt sequence from the sense strand of the gene. For example, a target sequence from Firefly luciferase 5'-ATCAGGTG GCTCCCGCTGAATTGGAATCC-3'.
2. Determine the reverse complement of the sense strand. In this case, it is 5'-GGATTCCAATTCAGCGGGAGCCACCTGAT-3'. When transcribed and cleaved by Dicer, this antisense RNA sequence will base pair with the mRNA to target it for degradation.
3. As previously described, we use a simple loop structure to separate the sense and antisense strands of the hairpin, 5'-UUGG-3'. Connect the sequence antisense-loop-sense, in this case 5'-GGATTCCAATTCAGCGGGAGCCACCTGATtgg ATCAGGTGGCTCCCGCTGAATTGGAATCC-3' (loop indicated in lowercase).
4. Create a few G-U base pairings in the stem of the predicted hairpin. GU-base pairs are permitted in dsRNA β -helices. The conversion of some G-C or A-U pairs to GU seems to stabilize the hairpins in bacteria (*see Note 1*) and may also facilitate DNA sequencing through the region. However, we only alter nucleotides in the "sense" strand of the predicted shRNA, so as not to affect pairings between the antisense strand and its mRNA target. In our example, we create the sequence 5'-GGATTCCAATTCAGCGGGAGCCACCTGATtggATCgGGTGGCTCtCGCT

Fig. 2. (*see opposite page*) (A) Map of pGEM-Zeo-U6 plasmid used as PCR template for short hairpin expression constructs; (B) a schematic drawing of PCR product produced using this vector and short hairpin primers.

GAgTTGGAATCC-3' (loop and altered bases to accommodate G-U pairings are indicated in lowercase).

5. Add the U6 termination sequence, TTTTT, to the transcript. The result is 5'-GGATTCCAATTCAGCGGGAGCCACCTGATttggATCgGGTGGCTCtCGC TGAgTTGGAATCC TTTTT-3'.
6. After the termination sequence, include a restriction site as desired for downstream cloning (we use *EcoRI*, GAATTC and add an extra base at the end to allow enzyme cleavage, giving GAATTCC). The result is 5'-GGATTCCAAT TCAGCGGGAGCCACCTGATttggATCgGGTGGCTCtCGCTGAgTTGGAATC CTTTTTgaattcc-3'.
7. Determine the reverse complement of the hairpin sequence. In this case, the result is 5'-ggaattcAAAAAGGATTCCAACtCAGCGaGAGCCACCCGATccaaATCA GGTGGCTCCCCTGAATTGGAATCC-3'.
8. Append a sequence that will recognize the 3' end of the U6 template sequence, TAGTATATGTGCTGCCGAAGC. (This sequence includes part of the 3' leader sequence, but the entirety of the leader is contained in the PCR template.) The result is 5'-ggaattcAAAAAGGATTCCAACtCAGCGaGAGCCACCCGATccaa ATCAGGTGGCTCCCCTGAATTGGAATCCTAGTATATGTGCTGCCGAAGC.
9. Design a forward primer for the U6 template. Since our vector has an Sp6 phage promoter upstream of the U6 promoter, we use an SP6 primer. The SP6 sequence is 5'-GATTTAGGTGACACTATAG-3'. At the 5' end we incorporate sequence for the restriction enzyme of our choice. We generally use *XhoI* (ctcgag) and add four extra bases to permit efficient restriction digestion (see **Note 2**). The result is 5'-GGCCctcgagGATTTAGGTGACACTATAG-3'.
10. Order the primers from your preferred oligonucleotide supplier. We order them at standard synthesis scale (usually 0.05 μ mol) and standard purification. We find gel purification to be costly and unnecessary.

When transcribed, the sequence will form a structure as shown in **Fig. 1**. The sequence of the entire U6 leader is 5'-GUGCUCGCUUCGGCAGCA CAUAUCUA-3', as shown in the diagram. The mFold program written by Michael Zucker can be used to predict the structure of your hairpin (**21**).

3.1.3. Carrying Out shRNA PCR

We have consistent success using PCR reactions with the following conditions (many variations would likely work):

1. Set up PCR reactions containing *Taq* polymerase, the vendor-supplied buffer, 4% DMSO, 200 μ M dNTPs, U6 template, and 50 pmol of each primer.
2. Program a thermocycler for 95°C for 3 min; 30 cycles of 95°C for 30 s, 55°C for 30 s, and 72°C for 1 min; followed by one cycle of 72°C for 10 min. The PCR product will be ~1.1 kb in length.

3.1.4. Cloning shRNA PCR Products

1. Digest PCR products and the vector of choice with the desired restriction enzymes (in our example *EcoRI* and *XhoI*).
2. Gel purify PCR products and vector using the preferred protocol. We generally use the QiaexII kit (Qiagen).
3. Set up a standard ligation reaction with vector and insert. Also prepare a vector-only control and an insert-only control.
4. Transform ligation reactions into competent bacteria. Select on LB–low salt medium containing Zeocin as well as antibiotic appropriate for the vector.
5. Observe the colonies. There should be many fewer (or zero) colonies on the vector-only and insert-only controls. The pGEM-zeo-U6 PCR template is an ampicillin-resistant plasmid. If you are cloning into a vector containing ampicillin resistance and are having background from the insert-only control, the vector template (pGEM-zeo-U6) can be cut with *XmnI* and *ScaI* (unique sites in the AmpR gene; see **Fig. 2A**) prior to the PCR reaction.
6. Pick clones for miniprep analysis. Our experience is that zeo-selected clones are correct about 50–75% (median frequency) of the time. This means that you should pick four to five clones to sequence per hairpin. If your efficiency is less than this, you should check the sequencing trace histograms. We use a sequencing primer at –42 relative to the U6 transcript initiation site, 5'-GTAACCTTGAAAGTATTTCG.

3.2. Expression Vectors

3.2.1. Standard Cloning Vectors

In mammals, RNAi seems to produce transient, cell-autonomous gene silencing. In other words, unlike RNAi in *C. elegans* and plants, the gene-silencing trigger is not amplified nor does it appear to be “transported” to nonaffected neighbors, at least in tissue culture. In mammalian cells, triggers of RNAi therefore have half-lives and must be continuously fed to the RNAi machinery to maintain silencing; thus, stable RNAi requires stable expression of the dsRNA trigger.

We have used small hairpins contained in basic cloning vectors such as pBSSK (Stratagene) to transiently silence genes of interest. However, cloning vectors generally lack markers for selection in tissue culture, so the generation of stable lines requires use of a vector such as pCDNA3 (Clontech). Vectors such as these are functional for gene silencing and allow generation of stably silenced lines. We prefer to use viral vectors, because infection is generally more efficient than transfection for many cell lines.

3.2.2. Retroviral Vectors

For retroviral vectors in general, we have had the most success putting the shRNA expression cassette between the 5' long tandem repeat (LTR) and

Retro-SHAG: Retroviruses expressing shRNAs

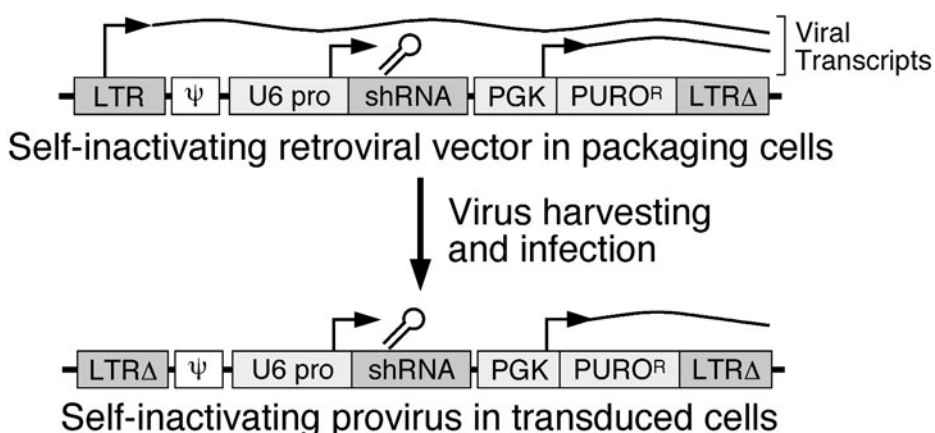


Fig. 3. Schematic of placement of shRNAs in retroviral vectors.

the drug selection marker. For example, the configuration of our MSCV-puro (Clontech) construct is shown in **Fig. 3**. We have carried out expression analysis of shRNA cassettes from the LTR and multiple cloning site positions (**Fig. 3**). The multiple cloning site configuration seems to generally work better when comparing the same shRNA expression cassettes. We are still examining the use of self-inactivating viruses (SINs), which do not produce proviral transcripts in transduced cells. In theory, proviral transcripts could act as a “dummy” target because it will contain the same targeting sequences as the gene of interest. However, in practice, we have found that non-SIN-based retroviruses such as MSCV and pBabe are effective both *in vitro* and *in vivo*. We find that no modifications are required for existing protocols for lenti- and MSCV-based viral production of short hairpin viruses.

3.3. Validation of shRNAs

We strongly recommend constructing three to six shRNAs per gene and carrying out a validation step before doing the actual biological experiment. For example, we often will transiently transfect shRNA plasmids into a commonly used cell line (293 or 3T3) that also expresses the gene of interest and assay efficacy by Northern blotting, RT-PCR, or Western blotting. If the target gene is not expressed in the cell line used for testing, it can be introduced transiently along with the shRNA and assayed. We find a direct correlation between whether an shRNA works well in transient assays and when it is expressed stably, though this may not be true for all test cases.

3.3.1. Monitoring mRNA Levels

Either standard RT-PCR or Northern analysis can be used to confirm the efficacy of silencing. In Northern analysis, there is sometimes a smear of apparently degraded RNA below the band corresponding to full-length transcript. In addition, it is important to remember that mRNA levels may not directly reflect changes in protein levels. This is of particular concern if hairpins are expressed in a transient setting but is less problematic for stably selected lines. However, when testing hairpins transiently as described in **Subheading 3.3.**; it may be useful to try longer time points, particularly for proteins that are expected to be stable.

3.3.2. Monitoring Protein Levels

The most direct method to monitor the efficacy of silencing is through a Western blot, assuming that appropriate antibodies are available.

3.3.3. Phenotype Analysis

When the degree of silencing has been confirmed, it is possible to use stably silenced cell lines in all types of experiments. The literature is already rich with applications for RNAi in many areas of biology. One advantage of RNAi can be the generation of intermediate silencing effects to create a range of phenotypes. This approach is described in **Subheading 3.4.1.**

3.4. Experimental Considerations

3.4.1. Epialleles

Frequently, expression analysis will reveal some short hairpins that induce intermediate levels of silencing. While using a hairpin that gives the highest possible level of knockdown may be of initial interest (in many cases, phenotypically very similar to a knockout allele), analysis of weaker hairpins may reveal interesting effects that reinforce the phenotypic analysis. Stable, short hairpin expression constructs behave very much like mutant alleles, and we have termed them *epialleles*.

We constructed such an epiallelic series directed against the p53 tumor suppressor (**I3**). A set of hairpins targeting p53 were tested transiently and assayed by Western blot to determine their relative efficacy. These hairpins were then stably expressed from a retroviral construct used to infect mouse hematopoietic stem cells. These cells are derived from Eμ-Myc transgenic mice, which are prone to lymphoma. The hematopoietic stem cells can be genetically modified and then injected to reconstitute the immune system in irradiated recipient mice. Previous studies had shown that p53 deficiency greatly accelerates lymphoma formation by these cells. Similarly, suppression of p53 by short hairpins

accelerated lymphoma development. Both the aggressiveness of the cancer and the time to cancer development correlated precisely with the reduction in p53 expression. The correlation of knockdown and phenotype appears to hold true from transient and stable experiments in vitro to stable expression in vivo.

3.4.2. Exon-Specific Silencing

One exciting application of RNAi is to determine the role of different splice variants of a gene. This approach has already been successfully applied in *Drosophila* and *C. elegans* and is now being used in mammalian systems with similar success (22). Short hairpins have even been used to silence a single allele of a gene that differed by a single missense mutation from the other allele (23).

Careful choice of target sequence can be used to silence either a single transcript or a group of transcripts. Additionally, closely related genes sometimes share sufficient homology to target a gene family with one or two short hairpins.

Because triggering RNAi using short RNAs of 19–29 nt in length is in theory highly specific, there exists the possibility of discriminating between “alleles” of genes that differ by as few as one nucleotide (e.g., RasV12). Indeed, this strategy has been employed for both siRNAs (<http://stke.sciencemag.org/cgi/reprint/sigtrans;2002/147/p113.pdf>) and shRNAs (27) with apparent success. However, such strategies may not work for all target sequences or biological assays. This is especially true when considering the efficacy of miRNAs (i.e., naturally expressed shRNAs); the antisense strands of most miRNAs are degenerate with respect to their cognate mRNA targets. Thus, researchers should be especially careful in designing these experiments.

3.4.3. “Knock In” of Altered Transcripts

Gene silencing by RNAi depends on a high degree of sequence homology between the short hairpin (or siRNA) and the mRNA target. Consequently, it is possible to introduce silent mutations in a cDNA and express this cDNA while simultaneously knocking down the endogenous mRNA. This technique has been used with siRNAs to validate phenotypes (25,26). Such validation is only one application of this approach—experiments could be designed to replace endogenous genes with specifically mutated proteins. Furthermore, this approach could be used to express an epitope-tagged construct in place of the endogenous protein. The purification of epitope-tagged protein complexes from higher eukaryotic cells is often complicated by competition from the endogenous protein. Removing the endogenous untagged protein population could allow better purification of complexes and reduce the effect of overexpression. This technique has been used in *Drosophila* to affinity purify a number of complexes, including exosomes (involved in RNA decay), nuclear export receptor complexes, and splicing complexes (27). In the cited work, long dsRNAs (which

can be used in *Drosophila* cells) were directed against an untranslated region of the gene, but a similar approach could be used for experiments in mammalian cells.

3.5. Bioinformatic Considerations for shRNA Library Construction

Our laboratory is in the process of building libraries of short hairpin expression constructs directed against 10,000 predicted human genes. This section describes our approach to some of the bioinformatics issues involved in creating and running an analysis and information pipeline for genomewide targeting by short hairpins. Two important requirements in the pipeline are web-based access to the data and a Lab Information Management System to track physical clones and other laboratory materials.

We are using a MySQL database and an Apache server for our Web server needs. Most of our scripts are in PERL or C. The system runs on a dual Athlon machine with 2 GB of RAM, and 80 GB of disk space. We regularly download copies of the Human Genome and the full set of human and mouse mRNAs from Genbank, since it is prohibitively time-consuming to work with the copies at NCBI. We have also downloaded all the information in the Mammalian Gene Collection (MGC) from the National Center for Biotechnology Information (NCBI), allowing us to access the MGC clones if one exists for a gene of interest. We built a name translation program, which lets us quickly annotate newly added accessions against the ones preexisting in the database. This program works by doing a high-similarity blast search against the oligos in the database, and if all the oligos from a given accession match, then it is safe to assume that the two accessions are either closely related or identical. In addition, we generated a mouse-human homolog map, because we try to design short hairpins that will be functional in both human and mouse cells.

3.5.1. Generation of Short Hairpins for Large Gene Sets

We began by creating a list of functionally annotated genes, which were then hand curated into groups such as kinases, phosphatases, and proteolysis-related genes. We built up incrementally a database of genes (the silencing database) that are targets of silencing experiments, always checking to make sure that a new entry was added to the database only if it was at least 40% (by length) different from other genes already in the database. Anything that failed this test we annotated as such in the silencing database. For each gene that survived this process, we selected 29-bp-long targets from across the coding region of the gene. We then mapped each of these targets against the genes in the silencing database, as well as a separate database of mouse mRNAs. We finally selected three target sequences per gene, using the following criteria for ranking them:

1. The three oligos should not map to more than one location in the silencing database.
2. The rank of an oligo is higher if a mouse hit is also found.
3. Oligos with GC content between 40 and 60% are ranked higher.
4. Oligos with homopolymer stretches of >4 ("TTTT," "GGGG," and so on) are eliminated.
5. The oligos are picked so that they are spread out as much as possible over the gene, while satisfying the previous criteria. The oligos must be separated by at least 60 bp.

The target sequences selected by these criteria are used to design short hairpins. We have created a program that carries out the steps described in **Subheading 3.1.4**.

This process of short hairpin design needs to be run in two separate rounds. In the first round, we ignore splice variants, as well as genes that are closely related to each other at the nucleotide level. We pick one member from each set of closely related genes and design oligos against them. In the second round of designs, we make splice-variant-specific hairpins. This will allow for large-scale screens of oligo sets in the initial stages, followed by experiments to resolve the fine structure of the effects of silencing specific splice variants.

3.5.2. Access to Data and Management of Samples

All the data are accessible through Web browsers, through CGI scripts on the Web server. These scripts can be used to query the database and present results in a variety of useful formats. This allows scientists to check the currently available short hairpin constructs for their gene of interest, and to confirm the location of targeted sequences on a transcript.

4. Notes

1. Inverted repeats such as the short hairpins described in this chapter are recognized by bacterial machinery and are frequently the targets of deletion or recombination. Introducing GU base pairing (GT at the DNA level) may evade these processes to some extent and does not affect the efficacy of silencing.
2. If you choose to use our pGEM-Zeo-U6 construct, the following enzymes are contained in this sequence and should not be used: *Bam*HI, *Hind*III, *Nde*I, *Sal*I, *Sma*I, and *Xma*I. Some usable enzymes include *Not*I, *Nhe*I, *Xho*I, *Eco*RI, and *Bgl*II.

References

- 1 Hannon, G. J. (2002) RNA interference. *Nature* **418**, 244–251.
- 2 Baglioni, C. and Nilsen, T. W. (1983) Mechanisms of antiviral action of interferon. *Interferon* **5**, 23–42.
- 3 Williams, B. R. (1997) Role of the double-stranded RNA-activated protein kinase (PKR) in cell regulation. *Biochem. Soc. Trans.* **25**, 509–513.

- 4 Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494–498.
- 5 Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J., and Conklin, D. S. (2002) Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev.* **16**, 948–958.
- 6 Paddison, P. J., Caudy, A. A., and Hannon, G. J. (2002) Stable suppression of gene expression by RNAi in mammalian cells. *Proc. Natl. Acad. Sci. USA* **99**, 1443–1448.
- 7 Brummelkamp, T. R., Bernards, R., and Agami, R. (2002) A system for stable expression of short interfering RNAs in mammalian cells. *Science* **296**, 550–553.
- 10 Paul, C. P., Good, P. D., Winer, I., and Engelke, D. R. (2002) Effective expression of small interfering RNA in human cells. *Nat. Biotechnol.* **20**, 505–508.
- 9 Sui, G., Soohoo, C., Affar el, B., Gay, F., Shi, Y., and Forrester, W. C. (2002) A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. *Proc. Natl. Acad. Sci. USA* **99**, 5515–5520.
- 10 Yu, J. Y., DeRuiter, S. L., and Turner, D. L. (2002) RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells. *Proc. Natl. Acad. Sci. USA* **99**, 6047–6052.
- 11 Zeng, Y., Wagner, E. J., and Cullen, B. R. (2002) Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol. Cell* **9**, 1–20.
- 12 Paddison, P. J. and Hannon, G. J. (2002) RNA interference: the new somatic cell genetics? *Cancer Cell* **2**, 17–23.
- 13 Hemann, M. T., Fridman, J. S., Zilfou, J. T., Hernando, E., Paddison, P. J., Cordon-Cardo, C., Hannon, G. J., and Lowe, S. W. (2003) An epi-allelic series of p53 hypomorphs created by stable RNAi produces distinct tumor phenotypes in vivo. *Nat. Genet.* **33**, 396–400.
- 14 McCaffrey, A. P., Nakai, H., Pandey, K., Huang, Z., Salazar, F. H., Xu, H., Wieland, S. F., Marion, P. L., and Kay, M. A. (2003) Inhibition of hepatitis B virus in mice by RNA interference. *Nat. Biotechnol.* **6**, 639–644.
- 15 Carmell, M. A., Zhang, L., Conklin, D. S., Hannon, G. J., and Rosenquist, T. A. (2003) Germline transmission of RNAi in mice. *Nat. Struct. Biol.* **10**, 91–92.
- 16 Semizarov, D., Frost, L., Sarthy, A., Kroeger, P., Halbert, D. N., and Fesik, S. W. (2003) Specificity of short interfering RNA determined through gene expression signatures. *Proc. Natl. Acad. Sci. USA* **100**, 6347–6352.
- 17 Chi, J. T., Chang, H. Y., Wang, N. N., Chang, D. S., Dunphy, N., and Brown, P. O. (2003) Genomewide view of gene silencing by small interfering RNAs. *Proc. Natl. Acad. Sci. USA* **100**, 6343–6346.
- 18 Jackson, A. L., Bartz, S. R., Schelter, J., Kobayashi, S. V., Burchard, J., Mao, M., Li, B., Cavet, G., and Linsley, P. S. (2003) Expression profiling reveals off-target gene regulation by RNAi. *Nat. Biotechnol.* **21**, 635–637.
- 19 Singh, R. and Reddy, R. (1989) Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. *Proc. Natl. Acad. Sci. USA* **86**, 8280–8283.

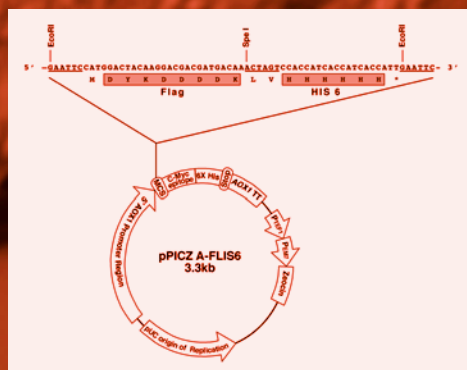
- 21 Zhang, H., Kolb, F. A., Brondani, V., Billy, E., and Filipowicz, W. (2002) Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *EMBO J.* **21**, 5875–5885.
- 21 Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415.
- 22 Kisielow, M., Kleiner, S., Nagasawa, M., Faisal, A., and Nagamine, Y. (2002) Isoform-specific knockdown and expression of adaptor protein ShcA using small interfering RNA. *Biochem. J.* **363**, 1–5.
- 27 Miller, V. M., Xia, H., Marrs, G. L., Gouvion, C. M., Lee, G., Davidson, B. L., and Paulson, H. L. (2003) Allele-specific silencing of dominant disease genes. *Proc. Natl. Acad. Sci. USA* **100**, 7195–7200.
- 24 Brummelkamp, T. R., Bernards, R., and Agami, R. (2002) Stable suppression of tumorigenicity by virus-mediated RNA interference. *Cancer Cell* **2**, 243–247.
- 25 Lassus, P., Rodriguez, J., and Lazebnik, Y. (2002) Confirming specificity of RNAi in mammalian cells. *Sci. STKE* **147**, PL13.
- 26 Lassus, P., Opitz-Araya, X., and Lazebnik, Y. (2002) Requirement for caspase-2 in stress-induced apoptosis before mitochondrial permeabilization. *Science* **297**, 1352–1354.
27. Forler, D., Kocher, T., Rode, M., Gentzel, M., Izaurralde, E., and Wilm, M. (2003) An efficient protein complex purification method for functional proteomics in higher eukaryotes. *Nat. Biotechnol.* **21**, 89–92.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



Geminivirus Vectors for Transient Gene Silencing in Plants

Nooduan Muangsang and Dominique Robertson

Summary

Both RNA and DNA viruses have been engineered to serve as vectors for transient silencing in intact plants. Host gene sequences carried by the virus are seen by the plant as “foreign,” and homologous gene-silencing machinery acts on both the viral vector RNA and the endogenous host gene mRNA. DNA viruses, such as geminiviruses, are advantageous for silencing because only their mRNAs are silenced and their DNA genomes continue to replicate and move. The conserved genome organization of geminiviruses and the fact that they can be cloned into *Escherichia coli* plasmids, propagated, and then inoculated into plants for infection simplifies the procedure for silencing specific chromosomal genes in intact plants. This chapter describes the development of a silencing vector from cabbage leaf curl virus for use in Arabidopsis and procedures for silencing two genes simultaneously.

Key Words: Geminivirus; transient gene silencing; Arabidopsis; functional genomics; DNA virus-induced gene silencing; cabbage leaf curl virus.

1. Introduction

The Arabidopsis genome has more than 25,000 putative genes, many with unknown function and others with one or more functions (**1**). Expression analyses, knockout libraries, and proteomics are making remarkable inroads into the identification of gene function. However, assigning function to genes that are required for growth can be difficult using these methods, and the lack of gene expression in seedlings may obscure identification of functions of the same gene later in development. We have developed a system that overcomes some of these limitations by using geminiviruses as silencing vectors. Geminiviruses are DNA viruses that replicate in plant nuclei without integrating into

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

chromatin. Their circular genomes contain an origin of replication that requires plant DNA replication machinery, and bidirectionally transcribed genes that use promoter and polyadenylation sequences recognized by plant RNA-processing enzymes (for review, *see* **refs. 2 and 3**). As a whole, geminiviruses infect a large number of plants, including many crop plants (www.danforthcenter.org/iltab/geminiviridae/). We describe the construction of a silencing vector from cabbage leaf curl virus (CbLCV), originally isolated in Florida and generously provided by Ernie Hiebert and James Strandberg (**4**). Using this vector, we have reliably verified gene function 14–21 d postinoculation (dpi).

Several other gene-silencing technologies use plant viruses, specifically vectors derived from RNA (**5–8**) (also *see* **ref. 9**) for methods). However, an efficient system using RNA viruses for gene silencing in *Arabidopsis* has not been developed. Indeed, even chromosomally integrated versions of geminivirus vectors have not given uniform silencing (**10**). A key component of episomal, geminivirus-induced silencing is high-level transcription of the silencing fragment (sense or antisense orientation) for effective spread of a silencing signal.

2. Materials

2.1. Construction of Silencing Vector

1. *Escherichia coli* DH5 α competent cells or equivalent.
2. Restriction enzymes and buffers.
3. Target DNA sequence (plant cDNA, clone, or genomic DNA).
4. Vector DNA (pCPCbLCV.007 or equivalent).
5. Agarose.
6. Ethidium bromide (EtBr): 0.5 mg/mL; use 1:1000.
7. Tris-borate-EDTA buffer: 10.8 g Tris, 5.5 g boric acid, 4 mL of 0.5 M EDTA for 1 L (pH 8.0).
8. T4 DNA ligase and buffer.
9. Luria Bertani (LB) medium: 5 g of tryptone, 5 g of salt, 10 g of yeast extract for 1 L. For plates, add 1.5% agar (w/v).
10. Ampicillin (100 mg/mL) or carbenicillin (50 mg/mL): Store at -20°C and use 1:1000 in LB.
11. Qiaquick polymerase chain reaction (PCR) purification kit (Qiagen, Valencia, CA).
12. Qiaquick Gel extraction kit (Qiagen).
13. Qiagen Plasmid maxiprep kit (Qiagen).
14. TE: 10 mM Tris-HCl (pH 8.0), 1 mM EDTA.

2.2. Plant Germination and Inoculation

1. Seeds of *Arabidopsis thaliana* ecotype Columbia (can be obtained from Lehle Seeds or the Arabidopsis Biological Resource Center, Ohio State University; www.arabidopsis.org).
2. Plastic tray pots (e.g., TLC plug flat) (Hummert, Earth City, MO).

3. Plastic pots (e.g., 2.5 × 2.5 in) (Hummert).
4. Soil mixture for *Arabidopsis* (e.g., Promix from Hummert).
5. Fertilizer.
6. Sterile distilled water.
7. Sterile 0.1% Type M agarose (Sigma, St. Louis, MO).
8. Microfuge tubes (1.5 mL).
9. Pipetmen and normal and aerosol tips.
10. Water bottle.
11. Forceps.
12. Growth chamber.
13. pCPCbLCVA.007 vector containing silencing fragment and pCPCbLCVB.002 (wt B component, necessary for viral movement).
14. 100, 95, and 70% ethanol.
15. Helium-driven particle gun (Bio-Rad PDS-1000/He™; Bio-Rad, Hercules, CA) or equivalent.
16. Benzylkonium chloride (5%).
17. Compressed helium cylinder and regulator.
18. Vortex with 6-in platform multitube head.
19. Gold or tungsten particles (1.0 μ in diameter; Bio-Rad).
20. Macrocarriers (Bio-Rad).
21. Rupture disks (1100 or 1300 psi) (Bio-Rad).
22. Stopping screens (Bio-Rad).
23. Spermidine (free-base, 0.1 M); filter sterilize.
24. CaCl₂ (2.5 M); filter sterilize.
25. Petri plates.

2.3. Verification of Silencing

1. Camera (preferably digital for cataloging pictures).
2. Oligonucleotide primers and reverse transcriptase (RT)-PCR kit.
3. RNAeasy isolation kit (Qiagen).
4. Oligonucleotide primers for checking vector insert size.
5. Digoxigenin DNA labeling kit.

3. Methods

The following sections describe how to clone a silencing fragment into a geminivirus vector, bombard plants with the recombinant geminivirus (*see Note 1*), and analyze plant material for silencing effectiveness.

3.1. Construction of Silencing Vector

The geminivirus CbLCV has been modified to serve as a silencing vector (*II*). The genome of this virus is bipartite, with genes on two circular 2.5-kb DNA molecules, the A and B components. The cloning of a silencing fragment into a vector derived from the A component of CbLCV is described (**Fig. 1**).

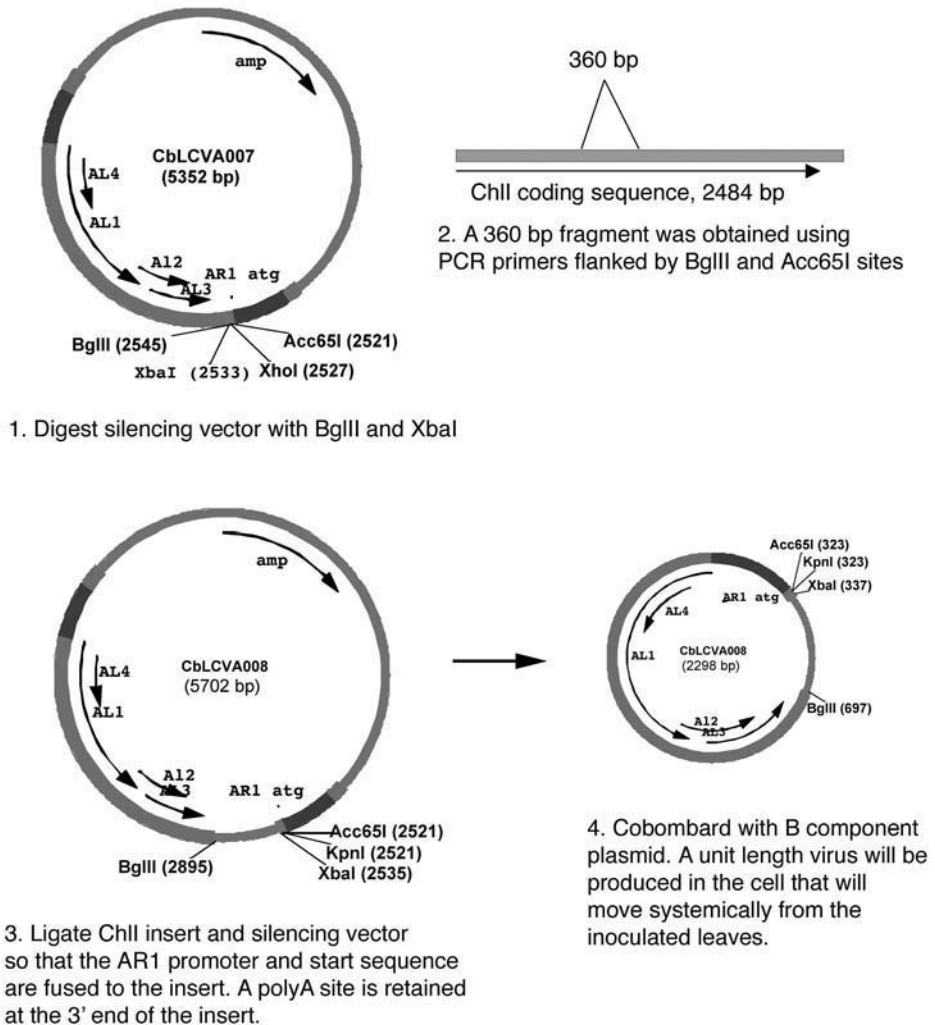


Fig. 1. Diagram of cloning of pMTCbLCVA.008. The common region is indicated by black boxes. The CR is conserved between the A and B genome but varies between viruses. To make pMTCbLCVA.008, the open reading frame of the coat protein gene, AR1, was replaced downstream of the AR1 ATG start codon with a fragment of the *ChII* gene. Release of viral episomes from plasmid DNA occurs in plant cells following microprojectile bombardment. AL1 is the Rep gene, which is needed for viral DNA replication; AL2 (Trap) is needed for transcription of the AR1 (coat protein) gene; and AL3 (Ren) enhances viral replication.

The silencing vector DNA is coinoculated with an equal amount of plasmid containing the wild-type B component, pCPCbLCVB.002 (*see Note 2*). Other geminivirus vectors that have been cloned as partial tandem direct repeats can also be used (*see Note 3*).

1. Digest the vector pCPCbLCVA.007 (**Fig. 1**) to completion with appropriate restriction enzymes (*Bgl*II, *Xba*I, *Xho*I, or *Acc*651).
2. Following electrophoresis in agarose, purify the fragment using a gel extraction kit. Quantify the DNA using a spectrophotometer to measure the absorbance at 260 nm (A_{260}).
3. Isolate one 600- to 800-bp gene fragment homologous to the gene targeted for silencing, or two approx 300–400 bp sequences for silencing two target genes (*see Note 3*). Isolation can be done by standard subcloning techniques or by PCR. Primers for PCR can contain embedded restriction sites for *Bgl*II, *Xba*I, *Xho*I, or *Acc*651 at the 5' ends of the oligonucleotides for directional cloning, or can be subcloned into a vector designed for PCR products (*see Note 4*).
4. Purify the fragment after enzyme digestion using a spin column (e.g., Qiaquick PCR purification kit).
5. Ligate vector and insert (approx 90 ng/ μ L of vector and a threefold molar excess of insert) using T4 DNA ligase and buffer.
6. Transform the ligation product into competent *E. coli* DH5 α or equivalent, and after 40–60 min, plate the cells onto LB plates containing ampicillin (100 μ g/L) or carbenicillin (50 μ g/L).
7. Miniprep the DNA from putative transformants and test for the presence of the insert. Verify by sequencing or restriction digests.
8. Perform large-scale plasmid DNA isolation (e.g., Qiagen). Resuspend the plasmid DNA in TE and quantify. Adjust the concentration to 1 μ g/ μ L with TE.
9. Perform a large-scale DNA isolation of plasmid carrying the B component of CbLCV (pCPCbLCVB.002); a negative control (e.g., pNMCbLCVA.Luc) carrying a 600- to 800-bp nonhomologous DNA, and a positive control, such as pMTCbLCVA.008, which carries a fragment of *ChI* (**Fig. 1**).

3.2. Plant Germination

Arabidopsis can be bombarded in soil or in sterile nutrient agar (*see Note 5*). One method of preparing plants for bombardment in pots is described. More information about *Arabidopsis* growth can be found in *Arabidopsis: A Laboratory Manual* (*see ref. 12*) and Boyes et al. (**13**).

1. Soak Promix with 1X Peter's 15:16:17 (or Miracle Gro) fertilizer in a large container. Mix well with a large scoop or your hands. The soil should stick together but not be soggy.
2. Place the soil in plastic flats with indented areas for the soil. The surface of the soil should be about even with the top. Level the soil surface by gently patting.

3. Place 500–1000 seeds (10–20 mg) into a microfuge tube or other tube containing 0.1% type M agar. Incubate the seeds and 0.1% agar at 4°C for 2 d to break dormancy.
4. Dispense the seeds onto the surface of the soil using a pipet. Place a plastic flat into a nested flat for support and cover it with a clear plastic cover.
5. Place the flats in a growth chamber under short-day conditions (8 h light/16 h dark) for vegetative growth or long-day conditions (16 h light/8 h dark) for rapid growth and earlier flowering.
6. Add water to a level of 1 to 2 cm above the base of the inner plastic tray. Cover with a plastic top and incubate at 22–24°C for 5–7 d.
7. Remove the plastic cover when the seedlings reach the four-leaf stage.
8. Water the plants only as needed, when the soil begins to dry. Overwatering should be avoided owing to the potential for algal or fungal growth on the soil surface. Add fertilizer at 2-wk intervals.
9. Transplant five- to six-leaf-stage seedlings into 6-cm² square plastic pots. Four healthy seedlings can be transplanted per pot and inoculated with an efficiency of 75–100%.
10. Insert closed forceps into the soil (prepared as above) for plant placement. Using the forceps, very gently remove the seedlings and transfer into the premade hole in the soil. Cover the base of the plant with soil, and spray the base of the seedlings with a few drops of water.
11. Place the plants in a large flat, cover with a plastic top, and place in a growth chamber. Cover the base of the pots with 1 to 2 cm of 1X fertilizer.
12. After 5 d, shake the plastic cover to remove moisture and return the cover slightly out of register to allow some air. At 6 d remove the cover. There should be no standing water at 7 d, and plants should not be watered 2 d before bombardment (see **Note 6**).

3.3. Microprojectile Bombardment

Microprojectile bombardment should include the combinations of plasmids shown in **Table 1**; an A component silencing vector and a wild-type B component. Five micrograms of each DNA (10 µg total) should be coated onto a 50-µL aliquot of microprojectiles shortly before bombardment (see **Note 7**). All plasmid coatings of microprojectiles can be done at the same time. Five bombardments should be done consecutively for each combination of plasmids. More test fragments can be added and 30–50 bombardments per session can be done in an afternoon. The chamber should be cleaned with water and then 95% ethanol after each five plants to prevent viral DNA from inoculating unintended target plants (carryover). For this reason, the positive control is shot after the test plants. The last bombardment checks for carryover silencing vector DNA.

Table 1
Four Combinations of A and B Components Recommended
for One Experiment

A component plasmid	B component plasmid
pNMCbLCVA.Luc (pCPCbLCVA.007 containing nonhomologous spacer DNA; negative control)	pCPCbLCVB.002
pCPCbLCVA.007 containing 300–800 bp of test silencing fragment	pCPCbLCVB.002
pCPCbLCVA.008 (positive control, silencing seen as loss of chlorophyll formation)	pCPCbLCVB.002
No A component; control for contaminating viral DNA and bombardment damage	pCPCbLCVB.002

3.3.1. Preparing Gold Particles

Gold microprojectile particles should be prepared and stored at -20°C in the dark (see **Note 7**).

1. Weigh 60 mg of microprojectiles and place in a microfuge tube.
2. Add 1 mL of absolute EtOH and vortex.
3. Centrifuge for 10 s at 13,000 rpm. Remove the ethanol, resuspend in 1 mL of sterile water, and vortex.
4. Centrifuge for 10 s at 13,000 rpm.
5. Discard the supernatant and resuspend in 1 mL of sterile water. Vortex frequently while transferring 50- μL aliquots into sterile microcentrifuge tubes being careful to keep the particles in suspension during the dispersal.

3.3.2. Coating Gold Particles With DNA

DNA-coated particles must be prepared fresh. The following procedure will yield particles for five bombardments and should be repeated for each combination of plasmids. The DNA-coated particles in ethanol can be stored at -20°C for several hours (overnight) before use. Longer periods of storage are not advised.

1. Add 5 μg of each plasmid DNA (A DNA and B DNA) to a tube containing 50 μL of gold particles and vortex. Add 50 μL of CaCl_2 (2.5 M) and immediately vortex. Pipet the solution in and out of the tip to help break up conglomerates if they form and vortex again.
2. Spermidine will cause precipitation, so it is important to disperse it uniformly and quickly. Add 20 μL of spermidine (0.1 M) to the side of the tube and vortex,

or add with the pipet tip submerged and immediately retrieve and expel the solution several times. Continue to vortex for 3–5 min using the vortex Genie equipped for multiple 1.5-mL microfuge tubes.

3. Spin the tubes at 10,000 rpm for 10 s. Carefully remove the supernatant and discard. Add 250 μ L of 100% ethanol and resuspend the pellet by vortexing.
4. Spin the tubes at 10,000 rpm for 10 s.
5. Carefully remove the supernatant and discard. Resuspend the DNA particles in 65 μ L of 100% ethanol (*see Note 7*) and keep the tubes on ice until use. This volume contains enough particles for five bombardments. The particles can be stored at -20°C for several hours before use.

3.3.3. Plant Inoculation

1. Safety glasses are required for all steps of the bombardments. Spray the bombardment area and the gene gun chamber with 5% benzyalkonium chloride followed by 70% ethanol.
2. Soak clean macrocarrier holders, macrocarriers, and stopping screens in 95% ethanol for 1 to 2 min, and allow them to air-dry in a laminar flow hood (*see Note 8*).
3. Vortex the tube containing the DNA-coated gold particles before removing aliquots to maximize uniform sampling. Break up conglomerates by pipetting or briefly sonicating.
4. Remove 10 μ L of the DNA particle mix and transfer to the center of a dry, sterile macrocarrier. Finger flick or vortex particle mix and add 10 μ L to the next macrocarrier. Repeat until all 5 have particles with as uniform distribution as possible.
5. While waiting for the carriers to dry, turn on the helium tank, the vacuum pump, and the power switch for the PDS1000He. The inner valve of the regulator should register about 2000 psi. Set the outer valve of the helium regulator to 1500 psi by turning the knob counterclockwise. Helium infiltration of the chamber is achieved later by pressing the Fire button. Place a nonsterile rupture disk (1100 psi) into the upper assembly and screw into place.
6. Assemble the microcarrier launch assembly (MCLA) by placing a sterile stopping screen into the orifice. With blunt forceps, transfer the macrocarrier into a stainless steel macrocarrier holder, particle side up. Using the supplied plastic tube, make sure the macrocarrier is flat. Invert the macrocarrier holder with DNA-coated gold particles so that the side containing DNA faces down. Screw in the retaining ring to secure the macrocarrier holder. Slide the MCLA into the top rack of the chamber (*see Note 9*).
7. Place a pot or Petri plate into the chamber and close the door (*see Note 10*).
8. Press the upper part of the Vac switch to evacuate air from the chamber.
9. When the vacuum gage registers 28 in (about 30 s), press the Fire button. Keep depressed (manually) until sufficient helium enters the chamber to burst the 1100-psi rupture disk, making a “poof” sound (*see Note 11*). Immediately press the Vac switch on the Vent position (intermediate position; do not press the Hold part of the three-way switch).

10. Remove the target plant when the vacuum gage reaches zero. Remove the upper bronze holder and discard the rupture disk. Remove the MCLA. Discard the macrocarrier disk that should now be fused to the stopping screen.
11. Repeat **steps 6–10** for the next four bombardments.
12. Clean the gene gun chamber with 70% ethanol and return the macrocarriers to 95% EtOH in Petri plates (*see Note 12*). Continue with the next set of plasmid-coated particles (**steps 3–11**).
13. At the end of the session, clean the MCLA as described, clean the chamber with water and 70% ethanol, and wipe out the laminar flow hood or bench. Close the main valve of the helium tank. Press the Fire button on the chamber to release helium in the line. The gage on the outer valve should drop. Turn off the power switch and the vacuum pump.

3.4. Analyzing Silencing Effectiveness

3.4.1. Silencing Onset

1. Return the plants to the growth chamber immediately after bombardment.
2. Cover the base of the pots with approx 1 to 2 cm of nutrient solution. For plants bombarded in soil, individual plants can be transplanted into separate pots 5–7 d after bombardment (*see Note 13*).
3. Place the pots containing the plants into a larger flat.
4. Add nutrient solution 1 to 2 cm above the base of the pots, cover the flats with a clear plastic lid for 5–7 d, and place the flats in a growth chamber (*see Note 14*).

Infected plants start to show symptoms including leaf curling and an uneven leaf surface about 8–10 dpi (**Fig. 2**). Plants infected with CbLCV vector carrying a fragment of *ChlI*, a subunit of magnesium chelatase required for chlorophyll biosynthesis (**14**), show yellow or white in new growth owing to the lack of chlorophyll (10–12 dpi) (**Fig. 2**). Silencing spreads evenly throughout new growth. At 20–25 dpi, most or all systemically infected leaves are yellow whereas mature leaves remain green (**Fig. 2**). Infected older seedlings (10- to 12-leaf stage) take 14 to 15 d to develop silencing. The positive control, *ChII*, can be monitored for silencing onset and spread. It is important to include this control because environmental conditions or DNA coating of microprojectiles can alter silencing efficiency.

3.4.2. Detection of Viral DNA Accumulation Levels in Infected Plants

Inoculated leaves and systemically infected leaves can be harvested 3 to 4 wk after bombardment. DNA extraction can be performed using the methods described in **refs. 15 and 16**), or other methods suitable for Arabidopsis (*see Note 15*). A phenol chloroform extraction of DNA from the Dellaporta method may be necessary. PCR can be performed on small leaf samples frozen in liquid nitrogen, ground, suspended in 10 μ L of TE, and centrifuged. This technique

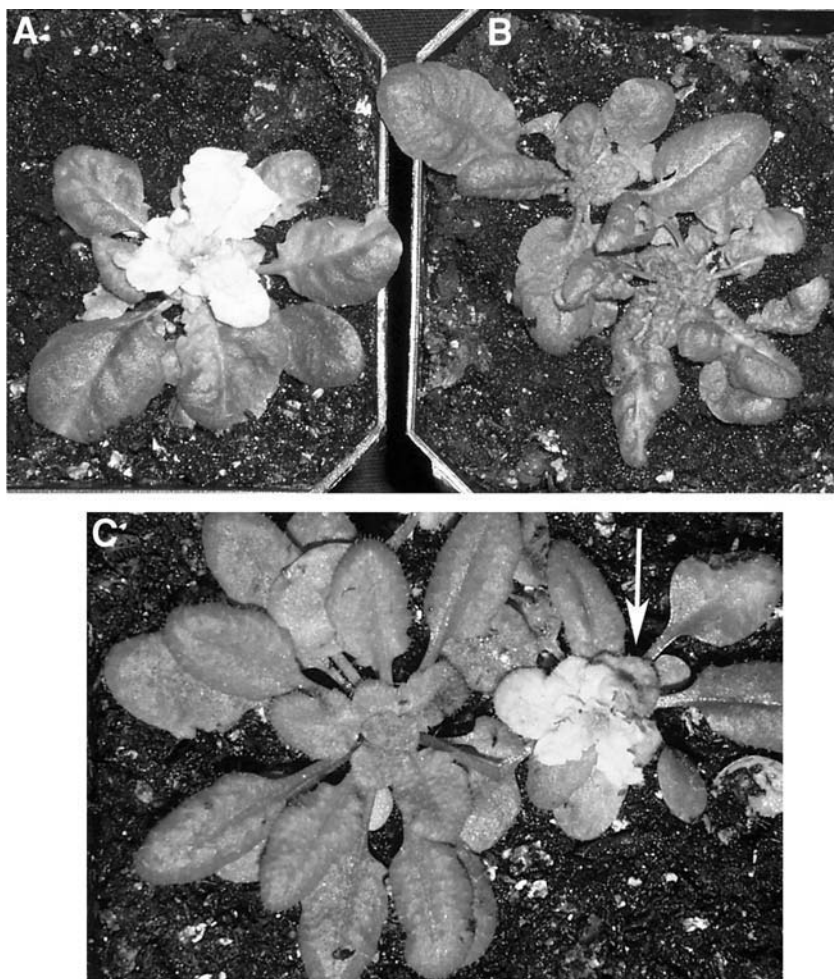


Fig. 2. Arabidopsis bombarded with pMTCbLCV.008 (A) or pNMCbLCV.LUC (B) carrying a nonhomologous insert. The B component was cobombarded with the plasmids. Symptoms of leaf curling but no chlorosis are evident in the control (right), while the *ChII*-silenced plant has lost chlorophyll in new rosette leaves (21 dpi). The arrow shows a leaf with some curling but no silencing. (C) Pot containing wild-type plant and *ChII*-silenced plant. Growth of the silenced plant is reduced.

should always include a no-template control (same reaction mix, no leaf sample).

To verify the size of the viral DNA vector and insert, perform a DNA gel blot hybridization (17). Digest 5 μ g of total DNA from each sample with enzyme(s) flanking the insert (see Fig. 1). Separate the fragments by electro-

phoresis on a 1% agarose–0.4 µg/L of EtBr gel, and photograph on an ultra-violet (UV) transilluminator. Transfer to a nylon membrane (e.g., Hybond, AmershamBiosciences.com) and crosslink using UV light as recommended (Stratagene.com). Prehybridize and hybridize using standard techniques. Digoxigenin-labeled probes (Roche-applied-science.com) or ³²P-labeled DNA probes corresponding to the viral DNA and insert DNA can be made by PCR. To detect the silencing vector (A component), a CbLCV probe derived from the AL1 (or Rep) can be labeled by PCR using the upper primer (5'AGAGAGGAA CATTGACACG G 3'), and lower primer (5'AGCACGATTGAGGGTATGCC 3') can be made. Double-stranded DNA (dsDNA) lacking an insert should be 1.7 kb while DNA with insert can be up to 2.5 kb (*see Note 16*). Because the vector lacks coat protein, single-stranded DNA (ssDNA) will be significantly less than dsDNA compared to a similar blot of wt CbLCV DNA (*4*). ssDNA migrates as a diffuse band of lower molecular weight when compared with dsDNA.

Both RNA gel blot hybridizations and RT-PCR can be used to assess endogenous gene expression. If RT-PCR is used, it is preferable to design primers to anneal to the target gene outside of the area of homology with the viral insert. If this is not possible, at least a 3-h incubation with RNase-free DNase is required for RT-PCR. It is important to include a no-RT control PCR reaction using the same reagents.

When two genes are silenced simultaneously from the same vector (*11*), both endogenous gene sequences need to be verified. Tandem sequences cloned into the polylinker of pCPCbLCV.007 should add up to at least 600 bp in size (*see Note 17*). The length of each sequence can vary but should be >90 bp for reliable silencing. In some cases it may be desirable to clone into pMTCbL CV.008 to use loss of chlorophyll as a visual marker for silenced tissue. A transgene, such as green fluorescent protein (GFP), should not be used as a marker for endogenous gene silencing because spread of silencing is much greater for transgenes than endogenous genes.

3.4.3. Interpreting Silencing Phenotypes

A transient silencing system can streamline the process of gene identification and biochemical pathway modifications by providing rapid silencing in wild-type and existing mutants. Because the virus also causes changes in gene regulation, silencing results should be validated by other methods. RNAi from chromosomal constructs containing inducible promoters or searches for existing mutants can be done to verify results. The Arabidopsis Information Resource is very useful for identifying insertion mutants in a given gene. Photodocumentation is useful for comparing phenotypes, especially across experiments. Using a systematic approach to Arabidopsis growth stages is essential for interpreting results (*13,18*).

We have uncovered phenotypic responses to essential genes including *ChlI* (DNA VIGS phenotype that alters senescence patterns [19], proliferating cell nuclear antigen (DNA VIGS phenotype that alters exit from the cell cycle [20], and pRb (DNA VIGS phenotype is cell death and developmental deformities (unpublished findings). We consider transient silencing using DNA VIGS to be only one part of a toolbox for identification of gene function in Arabidopsis, but one that if used properly (see **Note 18**) has unique advantages. For more information about the vectors used in this chapter, see **Note 19**.

4. Notes

1. *Agrobacterium* can also be used for inoculation (21). Pathogen and wounding effects on physiology should be considered in the interpretation of silencing results.
2. Bombardment with the silencing vector A component carrying a sequence homologous to an endogenous gene in the absence of the B component will cause local spots of silencing. The A component can replicate by itself but needs B component genes for movement. If the A component has sequence homologous to a transgene, spread of transgene silencing in the absence of the B component is much more extensive than for endogenous genes.
3. Each geminivirus/host combination must be optimized to produce effective silencing. The requirements for silencing from a geminivirus vector are that a sequence 90–800 bp long with homology to a target gene be cloned downstream of a geminivirus promoter or gene such that it is also transcribed. It is the mRNA, not the viral DNA, that activates silencing. Some geminiviruses require a coat protein for movement and can accept only 90- to 160-bp insertions in their genomes. In this case silencing fragments should be cloned downstream of the stop codon of gene (20). Geminiviruses that can move as DNA can contain up to 800 bp, which replace the coat protein gene. CbLCV does not require a coat protein for movement in Arabidopsis. Because the coat protein is required for insect transmission, the CbLCV vector is not infectious.
4. Exact homology between the silencing fragment and the target gene is not required, and PCR enzymes with average fidelity (*Taq* DNA polymerase) can be used. If oligonucleotide primers contain restriction sites, they must have 2 to 3 nt 5' of the restriction site or they will not be cut.
5. Seeds can also be sterilized and germinated on sterile medium. Ten to 15 seeds per plate should be prepared and deposited in a ring with a diameter of about 2.5 cm. Petri plates should be placed at the middle of the chamber. After bombardment, incubate 3 d in plates and then transfer to soil. In this case, 50–75% of plants get infected.
6. If plants have excess water in the soil, it will seep out during the vacuum stage of bombardment. If *Agrobacterium* or a handheld Helios gun (Bio-Rad) is used for inoculation, this step is not critical.
7. Theoretically 50 μ L of EtOH should suffice, but some evaporation occurs. We have found that some loss occurs and that 65 μ L ensures enough particles for five bombardments.

8. A laminar flow hood is not necessary for plants in soil. Drying macrocarrier holders is faster with air flow. Inverting on a paper towel will also help, and Kimwipes can be used to remove residual alcohol in the rim of the holder.
9. Tungsten particles can also be used. These should not be stored at -20°C for more than 1 mo and can oxidize, greatly reducing inoculation efficiency. Complete procedures for particle preparation and gene delivery using the PDS-1000/He Biolistic device (Bio-Rad) are described in the instruction manual.
10. The chamber door gasket must be kept in good shape. Apply a small amount of vacuum grease periodically to ensure a good seal.
11. The valve on the helium tank should rise until 1100 or 1350 and then fall. The gage will not be exact but should be within 100 psi of the rupture disk. If the rupture disk will not burst, two disks may have been inserted. Recheck to determine whether this is the problem.
12. If cross-contamination is found, it will be necessary to clean the MCLA each time a new combination of plasmids is to be bombarded. Clean the MCLA white platform with soapy water or 5% benzylkonium chloride, rinse seven times with water, and spray with 95% ethanol.
13. For plants bombarded on Petri plates, transplant into pots of soil 3 d after the bombardment to avoid fungal contamination.
14. Growth condition can affect silencing. For silencing vegetative tissues, 8 h of light/16 h of dark at 22°C (short days) will produce more rosette leaves, while 16 h of light/8 h of dark at 22°C (long days) will promote flowering. For silencing *ChII*, we use short-day conditions with 70% humidity and 120 μEin^{-2} lighting.
15. For protocols on DNA extraction from Arabidopsis, go to arabidopsis.org/comguide/table_of_contents.html.
16. We have found that fragments larger than 1 kb in the related geminivirus, tomato golden mosaic virus, are not propagated unless they have deletions that restore the size to approx 2.5 kb, or wild-type size (20).
17. Insert size is a constraint on systemic spread. CbLCV can increase to wild-type size (2.5 kb) if fragments smaller than about 650 bp are inserted. Because there is selection for wild-type (2.5 kb) size (21), deletion of foreign DNA is very rare.
18. CbLCV silencing vectors are compromised because they lack the coat protein necessary for insect transmission. It is nonetheless prudent to follow procedures for working with biohazardous materials. Infected plants must be autoclaved at 121°C , 15 psi for at least 1 h.
19. All vectors in this chapter are available by contacting Niki_Robertson@ncsu.edu. Three plasmids have accession nos: AY279345 (pCPCbLCVA.007), AY279344 (pCPCbLCVB.002), and AY279346 (pMTCbLCVA.008).

References

1. Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796–815.
- 2 Gutierrez, C. (2000) DNA replication and cell cycle in plants: learning from geminiviruses. *EMBO J.* **19**, 792–799.

- 3 Hanley-Bowdoin, L., Settlage, S. B., Orozco, B. M., Nagar, S., and Robertson, D. (1999) Geminiviruses: models for plant DNA replication, transcription and cell cycle regulation. *Crit. Rev. Plant Sci.* **18**, 71–106.
- 4 Hill, J. E., Strandberg, J. O., Hiebert, E., and Lazarowitz, S. G. (1998) Asymmetric infectivity of pseudorecombinants of cabbage leaf curl virus and squash leaf curl virus: implications for bipartite geminivirus evolution and movement. *Virology* **250**, 283–292.
- 5 Ratcliff, F., Martin-Hernandez, A. M., and Baulcombe, D. C. (2001) Technical advance: tobacco rattle virus as a vector for analysis of gene function by silencing. *Plant J.* **25**, 237–245.
- 6 Gossele, V., Fache, I., Meulewaeter, F., Cornelissen, M., and Metzclaff, M. (2002) SVISS—a novel transient gene silencing system for gene function discovery and validation in tobacco plants. *Plant J.* **32**, 859–866.
- 7 Angell, S. M. and Baulcombe, D. C. (1997) Consistent gene silencing in transgenic plants expressing a replicating potato virus X RNA. *EMBO J.* **16**, 3675–3684.
- 8 Kumagai, M. H., Donson, J., della-Cioppa, G., Harvey, D., Hanley, K., and Grill, L. K. (1995) Cytoplasmic inhibition of carotenoid biosynthesis with virus-derived RNA. *Proc. Natl. Acad. Sci. USA* **92**, 1679–1683.
- 10 Dinesh-Kumar, S. P., Anandalakshmi, R., Marathe, R., Schiff, M., and Liu, Y. (2003) Virus-induced gene silencing, in *Plant Functional Genomics Methods and Protocols*, vol. 236 (Grotewold, E., ed.), Humana, Totowa, NJ, p. 287–293.
- 10 Atkinson, R. G., Bielecki, G. R., Gleave, A. P., et al. (1998) Post-transcriptional silencing of chalcone synthase in petunia using ageminivirus-based episomal vector. *Plant J.* **15**, 593–604.
- 11 Turnage, M. A., Muangsan, N., Peele, C. G., and Robertson, D. (2002) Geminivirus-based vectors for gene silencing in Arabidopsis. *Plant J.* **30**, 107–114.
12. Weigel, D. and Glazebrook, J. (2002) *Arabidopsis: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- 13 Boyes, D., Zayed, A., Ascenzi, R., McCaskill, A., Hoffman, N., Davis, K., and Grolach, J. (2001) Growth stage-based phenotypic analysis of Arabidopsis: a model for high throughput functional genomics in plants. *Plant Cell* **13**, 1499–1510.
- 14 Koncz, C., Mayerhofer, R., Koncz-Kalman, Z., Nawrath, C., Reiss, B., Redei, G. P., and Schell, J. (1990) Isolation of a gene encoding a novel chloroplast protein by T-DNA tagging in Arabidopsis thaliana. *EMBO J.* **9**, 1337–1346.
- 15 Dellaporta, S. L., Wood, J., and Hicks, J. B. (1993) A plant DNA miniprep: version II. *Plant Mol. Biol. Rep.* **1**, 19–21.
- 16 Jose, J. and Usha, R. (2003) Bhendi yellow vein mosaic disease in India is caused by association of a DNA Beta satellite with a begomovirus. *Virology* **305**, 310–317.
17. Sambrook, J. and Russell, D. W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- 18 Kjemtrup, S., Boyes, D. C., Christensen, C., McCaskill, A. J., Hylton, M., and Davis, K. (2003) Growth stage-based phenotypic profiling of plants, in *Plant Functional Genomics Methods and Protocols*, vol. 236 (Grotewold, E., ed.), Humana, Totowa, NJ, pp. 427–441.

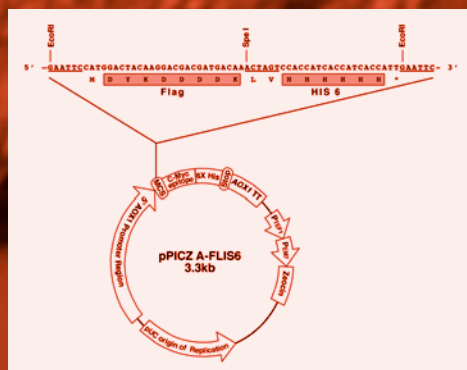
- 19 Kjemtrup, S., Sampson, K., Peele, C., Nguyen, L. V., Conkling, M. A., Thompson, W. F., and Robertson, D. (1998) Gene silencing from plant DNA carried by a Geminivirus. *Plant J.* **14**, 91–100.
- 20 Peele, C., Jordan, C. V., Muangsan, N., Turnage, M., Egelkrout, E., Eagle, P., Hanley-Bowdoin, L., and Robertson, D. (2001) Silencing of a meristematic gene using geminivirus-derived vectors. *Plant J.* **27**, 357–366.
21. Elmer, S. and Rogers, S. G. (1990) Selection for wild type size derivatives of tomato golden mosaic virus during systemic infection. *Nucleic Acids Res.* **18**, 2001–2006.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



Posttranscriptional Gene Silencing in Plants

Susan Varsha Wesley, Chris Helliwell, Ming-Bo Wang,
and Peter Waterhouse

Summary

Double-stranded RNA when introduced into cells results in severe reduction of the target mRNA. This phenomenon is known as posttranscriptional gene silencing in plants and RNA interference in animals. Hairpin RNA-mediated gene silencing exploits this cellular mechanism. A convenient way of generating hairpin constructs is to use generic vectors such as pHANNIBAL and pHELLSGATE, vectors based on the Gateway® technology. These vectors are suitable for high-throughput gene silencing, and the silencing effect is stably inherited over many generations.

Key Words: Gene silencing; RNA interference; knockout; hairpin RNA; pHANNIBAL; pHELLSGATE; Gateway; double-stranded RNA.

1. Introduction

Double-stranded RNA (dsRNA) is perceived as foreign and triggers the degradation of itself and homologous RNA within the cell. Two riboprotein complexes, DICER and RNA-induced silencing complex, are now implicated in cleaving the dsRNA into small RNAs (small interfering RNA or siRNA of approx 21 bases long) and using them as guides to recognize cognate mRNA for sequence-specific degradation. This process, called posttranscriptional gene silencing ([PTGS], also termed *RNA interference* in animals) (*1*), can be exploited as a functional genomics tool. Already it has been used to ascertain the function of several genes in *Drosophila* and *Caenorhabditis elegans* (*2*).

Gene silencing can be achieved by transformation of plants with constructs that express self-complementary (termed *hairpin*) RNA containing sequences homologous to the target genes. The DNA sequences encoding the self-

complementary regions of hairpin RNA (hpRNA) constructs form an inverted repeat (3). The complementary regions of the inverted repeat have to be separated with a spacer region or the plasmid is often unstable. The “spacer” region can be of any fragment and plays no role in directing PTGS. However, using an intron as a DNA spacer seems to improve the efficiency of the construct with up to 100% of the transformants generated with a particular gene construct showing some degree of silencing (4). There are at least three ways in which hpRNA constructs can be made. The construct may be generated from standard binary plant transformation vectors in which the hairpin-encoding region is generated *de novo* for each gene. Alternatively, generic gene-silencing vectors such as the pHANNIBAL and the pHELLSGATE series (5–7) can be used. Or the PCR products may be simply inserted into these vectors by conventional cloning or by using the Gateway® directed recombination system.

The features of intron interrupted hpRNA (ihpRNA)-mediated gene silencing make it particularly attractive in the production of silenced or knockdown plants for functional genomics applications. Because hpRNA targets specific genes, each gene in the group under study can be targeted with an hpRNA construct. The hpRNA construct is genetically dominant, and therefore phenotypes can be screened in primary transformed plants without the need to produce homozygous lines. Most plant transformation systems give rise to a number of transformation events that are propagated as separate transgenic lines. Thus, if a phenotype is replicated among the population of plants generated using a particular hpRNA transgene, it is highly likely that the phenotype is the result of silencing of the target gene rather than caused by a mutation introduced by the transformation procedure. Differing degrees of silencing are usually obtained in the lines produced from one transformation (5,6). Rather than being a disadvantage, the less severely silenced plants may allow survival of lines for genes for which a complete loss of function would be lethal. The sequence specificity of gene silencing allows the use of unique sequences to target specific genes and the potential to use conserved sequences to target multigene families. This enables researchers to custom make “knockdown” plants to suit their requirements.

1.1. Selecting a Target Gene Fragment

Gene fragments ranging from 50 bp to 1.6 kb have been successfully used as targets (*see Table 1*). Two factors can influence the choice of length of the fragment: (1) shorter fragments result in a lower frequency of silencing; (2) very long hairpins increase the chance of recombination in bacterial host strains. The effectiveness of silencing also appears to be gene dependent and could reflect accessibility of the target mRNA or the relative abundances of the target mRNA and the hpRNA in cells where the gene is active. We recom-

Table 1
hpRNA Silencing^a

Gene	Species	Construct type	Prom	Intron	Target	Stem (nt)	Silenced primary transformed (%)
PPO ⁵	Tobacco	ihp	35S	Pdk	ORF	572	70
GUS ⁵	Tobacco	hp	35S	NA	ORF	800	48
PVY-Nia ⁵	Tobacco	hp/ihp	35S	Pdk	ORF	730	58/96
EIN2 ⁵	Arabidopsis	ihp	35S	Pdk	ORF	600	65
FLC1 ⁵	Arabidopsis	ihp	35S	Pdk	ORF	650	100
FLC1 ⁵	Arabidopsis	ihp	35S	Pdk	ORF	400	100
CHS ⁵	Arabidopsis	ihp	35S	Pdk	ORF	741	91
Δ 12 ⁵	Arabidopsis	hp/ihp	Napin	Δ 12a	3' UTR	120	69/100
AG ¹⁷	Arabidopsis	hp	35S	NA	ORF	554	99
CLV3 ¹⁷	Arabidopsis	hp	35S	NA	ORF	288	88
AP1 ¹⁷	Arabidopsis	hp	35S	NA	ORF	409	96
PAN ¹⁷	Arabidopsis	hp	35S	NA	ORF	369	87
CBL ¹⁷	Arabidopsis	hp	35S	NA	ORF	1146	91
PDS ⁶	Arabidopsis	ihp	35S	Pdk	ORF	300	100
PhyB	Arabidopsis	ihp	35S	Pdk	3' UTR	300	70
Δ 12 ¹¹	Cotton	hp	Lectin	NA	ORF	853	58
Δ 12 ⁵	Cotton	ihp	Δ 12c	Δ 12c	5' UTR	98	100
Δ 9 ¹¹	Cotton	hp	Lectin	NA	ORF	514	57
BYDV-Pol ¹⁸	Barley	hp	Ubi	NA	ORF	1600	36
GUS ⁵	Rice	hp	Ubi	NA	ORF	560	85

^aAG, agamous; AP1, apetal; BYDV-Pol, barley yellow dwarf virus RNA-dependent RNA polymerase; CBL, cystathionine β -lyase; CHS, chalcone synthase; CLV3, clavata 3; Δ 9, Δ 9-desaturase; Δ 12, Δ 12-desaturase; EIN2, ethylene signaling gene; FLC1, flowering repression gene; GUS, β -glucuronidase; hp, hairpin; ihp, intron-interrupted hairpin; N/A, not applicable; ORF, open reading frame; PAN, Periantha; PdK, pyruvate orthophosphate dikinase; PDS, phytoene desaturase; PhyB, phytochrome B; PPO, polyphenol oxidase; PVY-NIa, potato virus Y Nia; Ubi, ubiquitin; UTR, untranslated region; 35S, 35S promoter from cauliflower mosaic virus.

mend a fragment length of between 300 and 1000 bp as a suitable size to maximize the efficiency of silencing obtained.

Both translated as well as untranslated regions (UTRs) have been used with equally good results (*see Table 1*). Because the mechanism of silencing depends on sequence homology, there is potential for cross-silencing of related mRNA sequences. Where this is not desirable, a region with low sequence similarity to other sequences, such as a 5' or 3' UTR, should be chosen. To reduce cross-silencing, blocks of sequence with identity over 20 bases between the construct and nontarget gene sequences should be avoided.

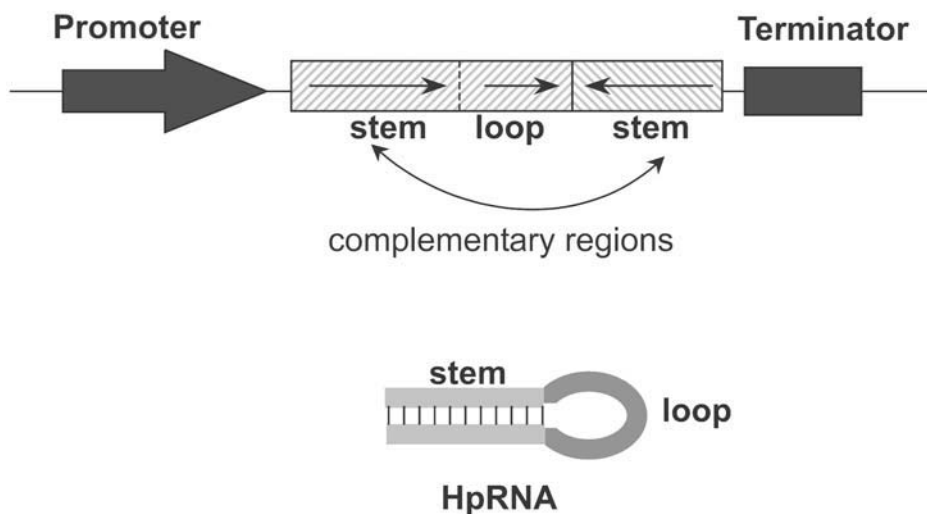


Fig. 1. Conventional hpRNA constructs are made by joining a 400- to 600-bp target gene sequence to an approx 300-bp fragment from the 3' end of the same sequence in an inverted orientation. Transcripts transcribed from such constructs will have regions of self-complementarity that have the potential to form hpRNA duplexes. hpRNA construct consists of a sense and an antisense arm separated by a spacer or loop DNA.

1.2. Conventional hpRNA Constructs

In their simplest form, hpRNA constructs can be made from either the whole or part of the target gene sequence, as illustrated in **Fig. 1**. These constructs usually give good silencing but the efficiency may not be as high as the hairpins containing an intron (4) (see **Table 1** for details regarding efficacy of such constructs in plants).

Constructs can be assembled in generic plant transgene cloning vectors such as pART7 (8) or pBluescript that contain a promoter for constitutive expression in both monocot and dicot plants. Once the assembly of the inverted repeat is complete, it can then be cloned into an appropriate binary vector such as pBIN19 (9) for transformation and expression in plants. A 400- to 600-bp sequence of the target gene is amplified in a polymerase chain reaction (PCR). This PCR fragment can then be ligated to an approx 300-bp fragment from the 3' end of the same target gene sequence in an inverted orientation. Transcripts transcribed from such constructs will have regions of self-complementarity that have the potential to form hpRNA duplexes.

1.3. *pHANNIBAL* and *pKANNIBAL* Vectors

The *pHANNIBAL* (with ampicillin resistance in bacteria)/*pKANNIBAL* (with kanamycin resistance in bacteria) system (**Fig. 2A**) has been found to work extremely efficiently and effectively for a number of genes (**Table 1**). PCR fragments are inserted into these vectors using conventional restriction enzyme digestion and DNA ligation techniques, in the sense orientation into the *XhoI.EcoRI.KpnI* polylinker and in the antisense orientation into the *Clal.HindIII.BamHI.XbaI* polylinker. These vectors are suitable for silencing a small number of genes. The construction of each hpRNA construct usually takes about 2 wk.

The *pKANNIBAL* vector is particularly useful because the PCR fragments from the target gene can be directly cloned, without prior restriction enzyme digestion, into a commercially prepared 3'-T-overhang, ampicillin-resistant vector such as pGEM®-Teasy (Promega, Madison, WI) and subsequently subcloned into *pKANNIBAL* using differential antibiotic selection. The *NotI* fragment from *pH/KANNIBAL* containing the hpRNA cassette can then be subcloned into a convenient binary vector such as pART27 (resistance to spectinomycin in bacteria and to kanamycin in plants) and used to transform plants. This approach bypasses the need for gel-purifying DNA fragments.

1.4. *pHELLSGATE* Vectors

The *pHELLSGATE* vectors were designed as high-throughput alternatives to the *pH/KANNIBAL* vectors. They are used in conjunction with the commercially available Gateway cloning system (www.invitrogen.com) which facilitates directional, recombinational in vitro cloning. The system incorporates a negative selection marker (*ccdB*) that selects against vectors that have not undergone a recombination reaction, resulting in a high frequency of recovery of recombined plasmids.

The *pHELLSGATE* vectors contain two recombination cassettes consisting of either attP1-*ccdB*-attP2 or attR1-*ccdB*-attR2 in an inverted repeat configuration such that when gene fragments flanked by the appropriate att sites are recombined with the vector, an ihpRNA-encoding construct is produced (**Fig. 3**). A series of *HELLSGATE* vectors have been generated (**Fig. 4**). Constructs in *pHELLSGATE4* are generated by a single recombination with an attB-flanked PCR product; however, more effective silencing is observed with constructs in *pHELLSGATE 8* or *12*. In these vectors, the gene fragment is recombined into an intermediate vector such as pDONR201 before a second recombination into *pHELLSGATE8/12*. *pHELLSGATE12* contains two introns

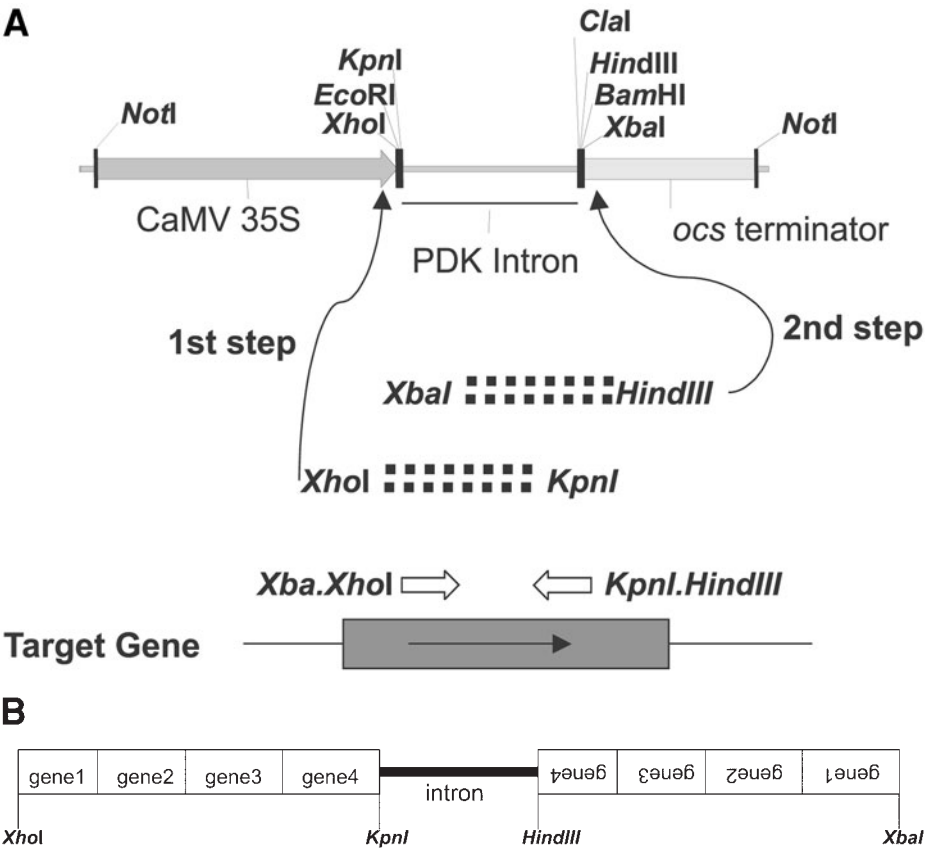


Fig. 2A. (A) The gene of interest is PCR amplified with the indicated restriction enzyme sites appended to the 5' end of the primers and sequentially cloned into similarly cut pHANNIBAL or pKANNIBAL vectors. Cloning into *XhoI.EcoRI.KpnI* poly-linker gives the sense arm of the hairpin and cloning into *ClaI.HindIII.BamHI.XbaI* polylinker the antisense arm. (B) When silencing multiple genes, fragments from various genes are PCR amplified, stitched together, and the whole cassette is cloned in the sense and antisense orientation into pH/KANNIBAL vectors.

in opposite orientations so that the final product of recombination will always contain one spliceable intron, thus reducing the number of recombinant plasmids that must be screened to obtain ihpRNA constructs.

2. Materials

2.1. Conventional hpRNA Constructs

1. PCR primers to amplify approx 800 bp of the target sequence with *EcoRI* and *XhoI* appended to the 5' end of the forward and reverse primers, respectively.

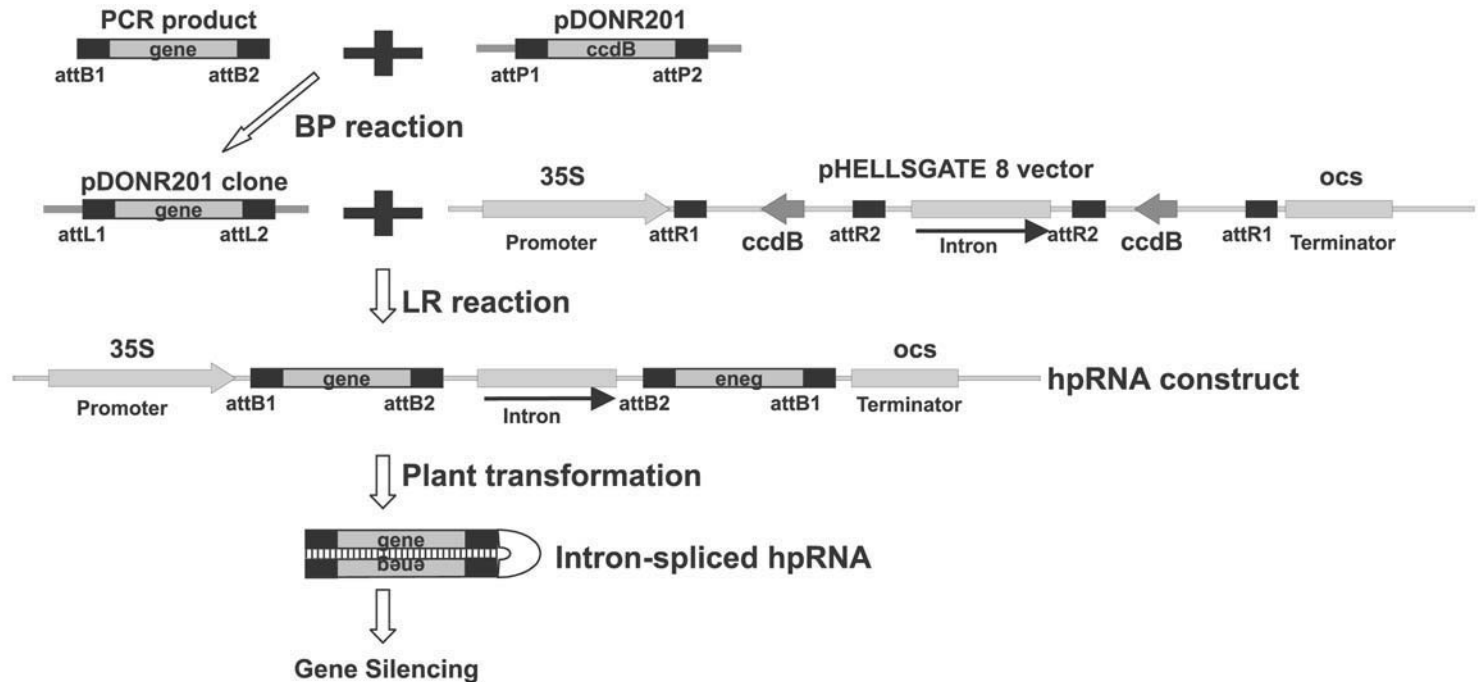
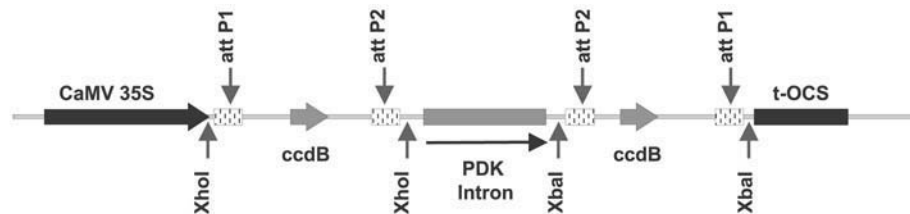
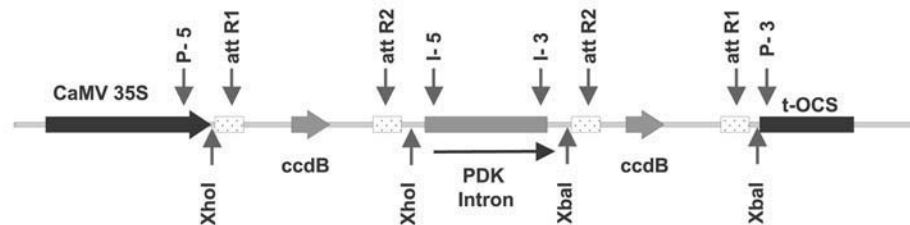


Fig. 3. To clone into pHELLSGATE8, the gene of interest is amplified with primers that have attB1 and attB2 sites appended to the 5' and 3' end, respectively. The PCR product is directionally recombined into pDONR201 vector through an in vitro recombination reaction using the enzyme BP clonase. The pDONR201 clones are then recombined into pHELLSGATE8 in a second recombination reaction using an enzyme LR clonase. The resultant plasmid is capable of producing hpRNA in plant cells transformed with it.

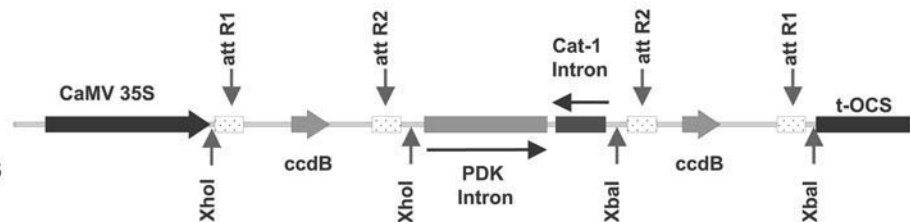
pHELLSGATE 4
P sites
BP reaction



pHELLSGATE 8
R sites
BP and LR reactions



pHELLSGATE 12
R sites
double intron
BP and LR reactions



2. PCR primers to amplify approx 300- to 1000-bp fragments from the 5' end of the target sequence with *Hind*III and *Sma*I appended to the 5' end of the forward and reverse primers, respectively.
3. pART7 (8) (modify restriction sites on the primers if using other vectors).
4. PCR purification kit or columns (Wizard® PCR Kit; Promega).
5. DNA template (20 ng).
6. 10 μ M of each primer.
7. 10 mM dNTP mix.
8. 25 mM MgCl₂.
9. *Taq* DNA polymerase (Perkin Elmer).
10. Buffer: 50 mM KCl, 10 mM Tris-HCl, pH 8.3.
11. Appropriate restriction enzymes from companies such as Promega or MBI Fermentas.
12. pART27 (8) or other binary vectors compatible with the plasmid containing 35S promoter.
13. Luria Bertani (LB) medium (liquid and solid) with appropriate antibiotics.

2.2. pHANNIBAL and pKANNIBAL Vectors

1. Forward primer: 5'-*Xba*I.*Xho*I + gene-specific sequence.
2. Reverse primer: 5'-*Cla*I.*Kpn*I + gene-specific sequence.
3. Vectors such as pGEM® Teasy from Promega (with resistance to ampicillin in bacteria) to clone the PCR product.
4. pHANNIBAL and pKANNIBAL vectors (www.pi.csiro.au).
5. pART27 vector (www.pi.csiro.au).
6. PCR purification kit or columns (Promega or others).
7. PCR amplification reagents (e.g., *Taq* polymerase, buffer, dNTPs).
8. Appropriate restriction enzymes.
9. LB medium (liquid and solid) with appropriate antibiotics.
10. Primers for sequence verification of hairpin constructs (P-5: 5'-GGGATGACG CACAATCC-3'; P-3: 5'-GAGCTACACATGCTCAGG-3'; I-5: 5'-ATAATCAT ACTAATTAACATCAC-3'; I-3: 5'-TGATAGATCATGTCATTGTG-3').
11. LB plates containing rifampicin (25 mg/L), gentamycin (25 mg/L), and spectinomycin (50 mg/L).
12. Plants to be transformed.
13. 5% Sucrose.
14. Silwet L-77.
15. MS agar plates containing kanamycin (100 mg/L).

Fig. 4. (see opposite page) Constructs in pHELLSGATE4 are generated by a single recombination with an attB-flanked PCR product. In pHELLSGATE8 and 12, the gene fragment is recombined into an intermediate vector, pDONR201, before a second recombination into pHELLSGATE8/12. pHELLSGATE12 contains two introns in opposite orientations so that the final product of recombination will always contain one spliceable intron.

2.3. *pHELLSGATE Vectors*

Materials are given for cloning into pHELLSGATE8 vector.

1. Forward primer: attB1-(5'-GGGGACAAGTTTGTACAAAAAAGCAGGCT) + gene sequence.
2. Reverse primer: attB2-(5'-GGGACCACTTTGTACAAGAAAGCTGGGT) + gene sequence.
3. AttP1 primer: 5'-GCTAGCATGGATCTCGG.
4. AttP2 primer: 5'-GAGCTGCAGCTGGATGG.
5. BP Clonase, buffer, and proteinase K (no. 11789013; Invitrogen, Carlsbad, CA).
6. LR Clonase, buffer, and proteinase K (no. 11791019; Invitrogen).
7. pHELLSGATE 8 (*see Note 1*) (www.pi.csiro.au).
8. pDONR201 (*see Note 1*) (no. 11798014; Invitrogen).
9. 30% PEG8000, 30 mM MgCl₂, TE.
10. PCR reagents (as in **Subheading 2.1.**).
11. Water bath at 25°C.
12. LB plates containing kanamycin (50 mg/L).
13. LB plates containing spectinomycin (100 mg/L).

3. Methods

3.1. *Conventional hpRNA Constructs*

1. Clone the target gene in a sense orientation in the *XhoI*/*EcoRI* sites of the pART7 vector.
2. Set up a standard PCR reaction using 20 ng of DNA template, 0.2 μM of each primer, 200 μM of each nucleotide, 1.5 mM MgCl₂, and 2.5 U of *Taq* DNA polymerase in 1X buffer. Adjust the reaction volume to 100 μL with water, and carry out 30 cycles of amplification using a PCR program consisting of denaturation at 94°C for 1 min, annealing at 55°C for 1 min, and extension at 72°C for 2 min, followed by a further extension at 72°C for 7 min.
3. Clean the PCR reaction with a purification column.
4. Digest approx 500 ng of the PCR product with *SmaI* and *HindIII*.
5. Clean the reaction, resuspend in 10 μL of water, and clone into *SmaI*/*HindIII*-restricted pART7 containing the target gene fragment from **step 1**.
6. Digest a positive clone from **step 5** with *NotI*, and clone into a *NotI*-digested pART27 binary vector.

3.2. *Cloning into pHANNIBAL and pKANNIBAL Vectors*

1. Set up the PCR as described in **Subheading 3.1., step 2**.
2. Clean the PCR product with a column, and digest with *XhoI* and *KpnI* for sense arm cloning and *XbaI* and *ClaI* for antisense arm cloning.
3. Ligate digested fragments sequentially to *XhoI*/*KpnI*- and *XbaI*/*ClaI*-digested pHANNIBAL or pKANNIBAL cloning vectors.

4. Clone the *NotI* fragment containing the ihpRNA cassette from pH/KANNIBAL into the *NotI* site of binary vector pART27.
5. For sequence verification (*see Note 2*), digest miniprep DNA with *BglII* (it cuts once in the *pdK* intron sequence found in pHANNIBAL, pKANNIBAL, and pHELLSGATE8).
6. Set up two separate PCR reactions using P-5 and I-5 primers to amplify the sense arm and I-3 and P-3 to amplify the antisense arm (the size of the product is 250 bases longer than the insert; *see Fig. 4*).
7. Purify the PCR product and sequence using the appropriate primers.
8. Transform the hpRNA construct into an *Agrobacterium tumefaciens* strain such as GV3101, and plate the cells on rifampicin, gentamycin, and spectinomycin plates (*see Note 3*).
9. Grow liquid cultures of *Agrobacterium* with antibiotics overnight; spin the cultures, and resuspend in a 2X vol of 5% sucrose and 0.05% Silwet.
10. Transform plants (any *Arabidopsis* ecotype) by the floral dip method (**10**): dip them twice 1 wk apart, collect the seed, and select the transformed plants on kanamycin (100 mg/L).
11. Screen at least 20 independent transformed lines, and measure the varying degrees of silencing either by the severity of the phenotype or by RNA levels (*see Note 4*).

3.3. Cloning into pHELLSGATE8 Vector

1. PCR amplify the gene of interest using the forward and reverse primers.
2. Check the PCR products by agarose gel electrophoresis for yield and product size.
3. Purify by diluting the PCR reaction with 3 vol of TE and precipitating with 2 vol of 30% PEG8000, 30 mM MgCl₂.
4. Collect the precipitate by centrifuging at >13,000g for 15 min and remove the supernatant using a pipet.
5. Resuspend the DNA pellet in 1 vol of TE.
6. Set up the BP reaction by mixing 2 μ L of BP Clonase buffer, 2 μ L of PCR product, 2 μ L (150 ng) of pDONR201, and 2 μ L of TE.
7. Incubate at room temperature (25°C) for 1 h.
8. Add 1 μ L of proteinase K mix (supplied with BP Clonase), and incubate for 10 min at 37°C.
9. Use 2 μ L to transform *Escherichia coli* DH5 α cells (the competent cells should have a transformation efficiency of at least 10⁷ colonies/mg of plasmid DNA).
10. Plate the transformation mixture on the kanamycin plates.
11. Screen the clones (typically six; *see Note 5*) for the insert by restriction digestion with enzymes such as *ApaI* and *PstI* that cut on either side of the insert. Alternatively, PCR amplify the fragment using AttP1 and AttP2 primers.
12. Set up the LR reaction by mixing 2 μ L of LR Clonase buffer, 2 μ L (100–200 ng) of pDONR clone (positive clone from **step 11**), 2 μ L (300 ng) of pHELLSGATE8 vector, and 2 μ L of TE.
13. Incubate for 1–16 h at room temperature (25°C), with longer incubations being better. Treat the reaction with 1 μ L of proteinase K for 10 min at 37°C.

14. Use 2 μ L of the reaction mix to transform DH5 α , and select colonies on the spectinomycin plates (the plates generally require 24 h of incubation at 37°C before colonies are visible).
15. Screen the clones (typically six; *see Note 4*) by digesting the miniprep DNA with *Xho*I (sense arm) and *Xba*I (antisense arm) separately (*see Note 6*). The size of the fragment should be the size of the insert plus 250 bp (**Fig. 4**).
16. Sequence verify the final clones as in **Subheading 3.2., steps 5–7**.
17. Transform plants as in **Subheading 3.2., steps 8–11**. *See refs. 11–13 and Note 7* for more applications.

4. Notes

1. Vectors containing the negative selectable marker *ccdB* (pHELLSGATE 8 and 12 and pDONR201) must be maintained in the DB3.1 *E. coli* strain. Competent cells can be purchased from Invitrogen; alternatively, electrocompetent cells can be prepared using standard methods.
2. It is difficult to sequence hpRNA constructs because the two arms of the hairpin anneal to each other before the primers can anneal to them.
3. Once the hairpin constructs are assembled, they can be stably integrated into plant genome by plant transformation (**10**) or delivered in a transient manner through bombardment or agroinfiltration (*see Note 5*; for review *see ref. 14*). HpRNA silencing is stably inherited up to five generations (**15**).
4. When designing probes for Northern hybridizations, or primers for real-time PCR, use a region in the gene that is not used in the hpRNA construction since some of the hpRNA can remain intact (**16**).
5. The percentage of positive clones obtained in pDONR201 and pHELLSGATE vectors sometimes depends on the gene sequence, which means more than six colonies may have to be screened.
6. The *Xho*I, *Xba*I digestion is not very good on DNA from *Agrobacterium*. Back transformation to DH5 α cells may be necessary.
7. Multiple genes: It has been possible to combine different hpRNA-mediated silenced traits through sexual crossing of relevant transgenic lines (**14**). However because the different hpRNA transgenes are inserted at different locations, they will segregate in subsequent generations, making the task of stacking modified traits through crossing laborious and time-consuming. This will limit the number of genes that can be combined.

References

- 1 Hannon, G. J. (2002) RNA interference. *Nature* **418**, 244–251.
- 2 Kamath, R. S., Fraser, A. G., Dong, Y., et al. (2003) Systematic functional genomic analysis of *Caenorhabditis elegans* genome using RNAi. *Nature* **421**, 231–237.
- 3 Waterhouse, P. M., Graham, M. W., and Wang, M.-B. (1998) Virus resistance and gene silencing in plants is induced by double-stranded RNA. *Proc. Nat. Acad. Sci. USA* **95**, 13,959–13,964.

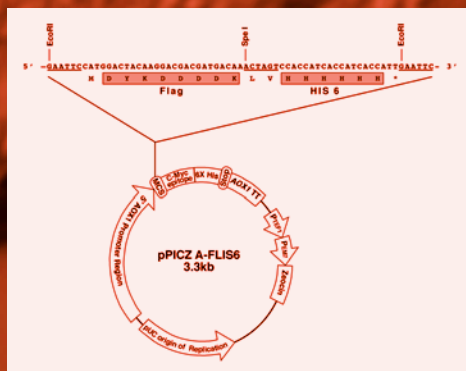
- 4 Smith, N. A., Singh, S. P., Wang, M.-B., Stoutjesdijk, P., Green, A. and Waterhouse, P. M. (2000) Total silencing by intron-spliced hairpin RNAs. *Nature* **407**, 319, 320.
- 5 Wesley, S. V., Helliwell, C. A., Smith, N., et al. (2001) Construct design for efficient, effective and high-throughput gene silencing in plants. *Plant J.* **27**, 581–590.
- 6 Helliwell, C. A., Wesley, S. V., Wielopolska, A. J., and Waterhouse, P. M. (2002) High throughput vectors for efficient gene silencing in plants. *Funct. Plant Biol.* **29**(10), 1217–1225.
- 7 Waterhouse, P. M. and Helliwell, C. A. (2003) Exploring plant genomes by RNA-induced gene silencing. *Nat. Rev. Genet.* **4**, 29–38.
- 8 Gleave, A. P. (1992) A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. *Plant Mol. Biol.* **20**, 1203–1207.
- 10 Hellens, R., Mullineaux, P., and Klee, H. (2000) A guide to *Agrobacterium* binary Ti vectors. *Trends Plant Sci.* **5**(10), 446–451.
- 10 Clough, S. J. and Bent, A. F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743.
- 11 Wang, M.-B. and Waterhouse, P. M. (2002) Application of gene silencing in plants. *Curr. Opin. Plant. Biol.* **5**, 146–150.
- 12 Klink, V. P. and Wolniak, S. M. (2000) The efficacy of RNAi in the study of the plant cytoskeleton. *J. Plant Growth. Regul.* **19**, 371–384.
- 13 Schweizer, P., Pokorny, J., Schulze-Lefert, P., and Dudler, R. (2000) Double-stranded RNA interferes with gene function at the single-cell level in cereals. *Plant J.* **24**, 895–903.
- 14 Liu, Q., Singh, S., and Green, A. (2002) High-stearic and oleic cottonseed oils produced by hairpin RNA-mediated post-transcriptional gene silencing. *Plant Physiol.* **129**, 1732–1743.
- 15 Stoutjesdijk, P., Singh, S. P., Liu, Q., Hurlstone, C. J., Waterhouse, P. M., and Green, A. G. (2002) HpRNA-mediated targeting of the *Arabidopsis* FAD2 gene gives highly efficient and stable silencing. *Plant Physiol.* **129**, 17.
- 16 Levin, J. Z., Framond, A. J., Tuttle, A., Bauer, M. W., and Heifetz, P. B. (2000) Methods of double-stranded RNA-mediated gene inactivation in *Arabidopsis* and their use to define an essential gene in methionine biosynthesis. *Plant Mol. Biol.* **44**, 759–775.
- 17 Chuang, C. F. and Meyerowitz, E. M. (2000) Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. *Proc. Nat. Acad. Sci. USA* **97**, 4985–4990.
- 18 Wang, M.-B., Abbott, D. C., and Waterhouse, P. M. (2000) A single copy of a virus-derived transgene encoding hairpin RNA gives immunity to barley yellow dwarf virus. *Mol. Plant Pathol.* **1**, 347–356.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



Identification of microRNAs and Other Tiny Noncoding RNAs by cDNA Cloning

Victor Ambros and Rosalind C. Lee

Summary

MicroRNAs (miRNAs) and other small RNAs can be identified by cloning and sequencing cDNAs prepared from the ~22-nt fraction of total RNA. Methods are described for the construction of cDNA libraries from small noncoding RNAs through the use of T4 RNA ligase, reverse transcriptase, and polymerase chain reaction. cDNAs are cloned in λ or plasmid vectors, and the sequences are compared to annotated genomic sequence databases, and analyzed by RNA folding programs to distinguish miRNA sequences from other small RNAs of similar size. Northern blot hybridization is used to confirm the expression of small RNAs in vivo.

Key Words: microRNA; noncoding RNA; RNA ligase; SMART; RNA folding; cDNAs; Northern blot.

1. Introduction

Noncoding RNAs come in many classes and perform diverse functions in eukaryotic and prokaryotic cells (*1*). The identification of noncoding RNAs requires special experimental strategies that take advantage of completely sequenced genomes and that are tailored to the unusual properties of the RNAs themselves. This chapter describes protocols for the cDNA cloning and detection in vivo of RNA transcripts in the ~22-nt-size class. Noncoding RNAs of this class include the microRNAs (miRNAs) implicated in the posttranscriptional regulation of genes during the development of plants and animals (*2–10*). The first miRNA genes to be characterized, *lin-4* and *let-7*, were identified in *Caenorhabditis elegans* by their mutant phenotypes and were cloned by conventional positional cloning (*11,12*). Many more miRNA genes have been identified in plants and animals by adapting methods developed for the cDNA

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

cloning of the ~22-nt short interfering RNAs (siRNAs) involved in RNA interference (RNAi) (**3–6,13**).

This chapter describes protocols for the production of cDNA libraries enriched for sequences of transcripts in the ~22-nt-size fraction of total RNA. Although the specific emphasis here is on the identification of miRNAs, these methods also yield cDNA clones corresponding to siRNAs, other miscellaneous ~22-nt noncoding RNAs, as well as less interesting sequences corresponding to degradation products of longer RNAs (**10**). For this reason, a complete annotated genomic sequence of the organism under study is essential for determining the genomic location and sequence context of each cDNA. miRNA sequences are distinguished from other kinds of ~22-nt cDNAs by three essential criteria: (1) a genomic location outside of protein coding sequence (either intergenic or intronic), (2) accumulation of detectable levels of a discrete ~22 nt single-stranded RNA in vivo, and (3) a predicted hairpin precursor transcript of about 70 nt in length that would yield the ~22 nt miRNA by Dicer processing of the helical portion of the hairpin (**14**). The mature miRNAs are implicated in a variety of regulatory phenomena in plants and animals (**15–17**).

miRNAs and other noncoding RNAs with evolutionarily conserved predicted secondary structures can also be discovered through the analysis of conserved genomic sequences (**3,9,10,18**). Computational methods are described in Chapter 21 of this volume, so the discussion here is restricted to identification of miRNAs by cDNA cloning. However, analysis of miRNA expression by Northern blots is covered here and can be used essentially as described to test for the expression of computationally predicted small RNAs.

2. Materials

2.1. Total RNA Extraction

1. RNase-free deionized water (*see Note 1*).
2. Gibco Trizol reagent.
3. Chloroform.
4. Phenol/chloroform (50/50).
5. 100% Isopropanol.
6. 70% Ethanol.
7. 100% Ethanol.
8. TES buffer: 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 1% sodium dodecyl sulfate (SDS).
9. TE buffer: 10 mM Tris-HCl, pH 8.0, 1 mM EDTA.
10. 3 M Na acetate.
11. Siliconized, RNase-free 1.5-mL and 0.5-mL microcentrifuge tubes (*see Note 2*).
12. Glycogen (20 µg/mL) (store at –20°C; *see Note 3*).

2.2. Preparation of Internal Size Marker for Size Selection

1. T7 top-strand DNA oligonucleotide (100 μ M): 5' AATTTAATACGACTCAC TATAG 3'.
2. Marker template DNA oligonucleotide (100 μ M): 5' TTTGGATTCCCTA CACGCTTCGCCTATAGTGAGTCGTATTAAATT 3'.
3. 10X T7 transcription buffer (400 mM Tris, pH 7.5; 100 mM NaCl; 60 mM MgCl_2).
4. 100 mM dithiothreitol (DTT).
5. 10 mM ATP.
6. 10 mM GTP.
7. 10 mM CTP.
8. 10 mM UTP.
9. α - ^{32}P UTP (3000 Ci/mmol) (10 $\mu\text{Ci}/\mu\text{L}$).
10. T7 RNA polymerase (10 U/ μL) (Ambion).
11. RNase-free DNase I (2 U/ μL) (Ambion).

2.3. Gel Purification of Marker RNA Oligonucleotide

1. Acrylamide gel system (National Diagnostics Sequagel; store these solutions at room temperature): gel diluent (8 M urea), gel concentrate (19% acrylamide, 1% bisacrylamide, 8 M urea), 10X gel buffer (890 mM Tris base; 890 mM boric acid; 20 mM EDTA, pH 8.3; 8 M urea).
2. TEMED (store at 4°C).
3. 10% Ammonium persulfate (store at 4°C).
4. 1X TBE electrophoresis buffer: 89 mM Tris base; 89 mM boric acid; 2 mM EDTA, pH 8.3.
5. Denaturing gel sample buffer: 95% formamide, 0.025% xylene cyanol, 0.025% bromophenol blue, 18 mM EDTA, 0.025% SDS.
6. Vertical slab gel electrophoresis apparatus, plates, and combs; Bio-Rad Protean II minigel system or equivalent (Bio-Rad, Hercules, CA) (*see Note 4*).
7. Bio-Rad Mini Spin Filters (cat. no. 732-027; Bio-Rad).
8. Gel elution buffer: 20 mM Tris-HCl, pH 7.5, 5 mM EDTA, 400 mM Na acetate.

2.4. Preparation of Size-Selected RNA From Total RNA

1. Millipore Microcon YM-100 columns (or equivalent) (0.5-mL).
2. Denaturing acrylamide gel-running and extraction solutions (*see Subheading 2.1.3.*).

2.5. Ligation of 3' Linker Oligonucleotide

1. 10X New England Biolabs restriction buffer 3: 1 M NaCl, 500 mM Tris-HCl, pH 7.9, 100 mM MgCl_2 , 10 mM DTT.
2. Calf intestinal phosphatase (CIP) (10 U/ μL) (New England Biolabs).

3. T4 RNA ligase (10 U/ μ L) (cat. no. E2050Y; Amersham, Piscataway, NJ).
4. 10X RNA ligase buffer: 500 mM Tris-HCl, pH 7.5, 100 mM MgCl_2 , 100 mM DTT, 10 mM ATP, 600 μ g/mL of bovine serum albumin (BSA).
5. Dimethylsulfoxide (DMSO).
6. 60 μ M ATP (store at -20°C).
7. 100 μ M Stock of RNA 3' linker oligonucleotide 5'-pAr₆A₂₄dd-3' (see **Note 5**).
8. 100 μ M Stock of DNA 3' linker oligonucleotide 5'-pA₃₀dd-3' (see **Note 5**).
9. 10 mL Stock of 1 M methylimidazole prepared as follows:
 - a. Mix 820 μ L of methylimidazole (Sigma, St. Louis, MO) and 7 mL of water.
 - b. Add concentrated HCl dropwise until the pH is 6.0.
 - c. Add water to make a final volume of 10 mL.
 - d. Aliquot into 0.5-mL portions and store at -80°C .
10. 1.0 mL Stock of 1 M adenosine monophosphate in water, pH 6.0. Aliquot into 100- μ L portions and store at -80°C .
11. 1.0 mL Stock of 1 M carbodimide in water; adjust the pH as needed to 6.0–6.5. Aliquot into 100- μ L portions and store at -80°C .
12. Ethidium bromide (EtBr) stock (10 mg/mL).

2.6. Reverse Transcription

1. Clontech SMART cDNA library construction kit (optional).
2. 10 μ M Stock of SMART template oligonucleotide (**Fig. 1**): 5'-AAGCAGTG GTATCAACGCAGAGTGGCCATTACGGCCGGG-3'.
3. 10 μ M Stock of reverse transcriptase (RT) primer oligonucleotide (**Fig. 1**): 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-d(T)₃₀(A, G, or C)(A, G, C, or T)-3' (The last two 3' bases are redundant, as indicated.)
4. 5X First-Strand RT Buffer: 250 mM Tris-HCl, pH 8.3, 30 mM MgCl_2 , 375 mM KCl.
5. RNase H-deficient MMLV Reverse Transcriptase (Clontech PowerScript[®]).
6. Deoxynucleotide triphosphate (dNTP) mix (10 mM each of the four dNTPs).

2.7. Polymerase Chain Reaction Amplification of cDNA

1. 10 μ M cDNA 5' polymerase chain reaction (PCR) primer: 5'-AAGCAGTG-GTATCAACGCAGAGT-3'.
2. 10 μ M cDNA 3' PCR primer: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-3'.
3. dNTP mix (2 mM each of the four dNTPs).
4. *Taq* polymerase (2.5 U/ μ L).
5. 10X *Taq* polymerase buffer: 200 mM Tris-HCl, pH 8.8; 100 mM KCl; 100 mM $(\text{NH}_4)_2\text{SO}_4$, 20 mM MgSO_4 ; 1% Triton X-100.
6. Proteinase K (20 mg/mL).
7. TE buffer: 20 mM Tris-HCl, pH 7.5, 1 mM EDTA.
8. Phenol extraction reagents (see **Subheading 2.1**).

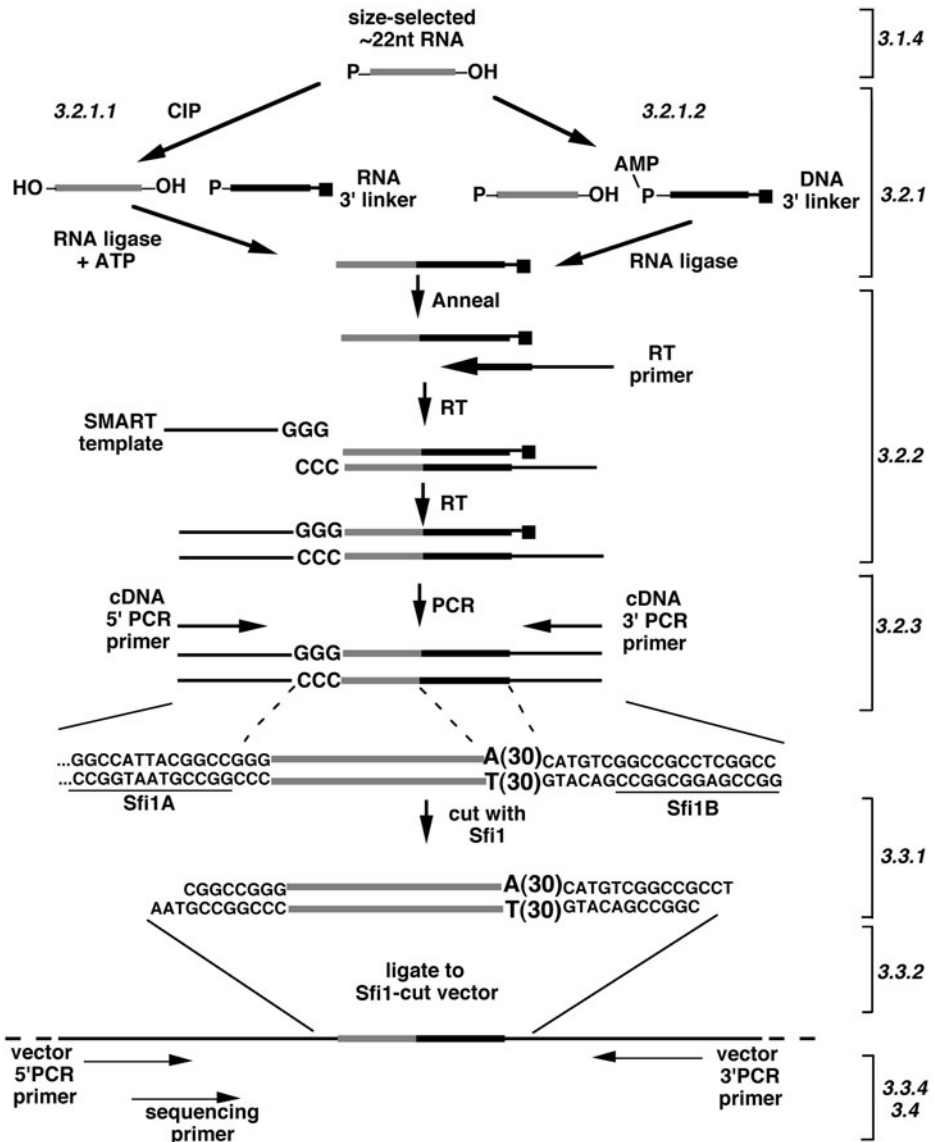


Fig. 1. Flow chart for preparation and analysis of cDNA clones from small RNAs. Brackets on the right correspond to the steps of the protocol described in the indicated sections of the text. Two optional methods for coupling the 3' linker oligonucleotide described in Subheadings 3.2.1.1. or 3.2.1.2 are indicated.

2.8. Gel Purification of cDNA

1. 10X Agarose gel-loading buffer: 200 mM Tris-HCl, pH 7.5; 10 mM EDTA; 40% glycerol.
2. Agarose TBE gel: 2 g of agarose melted into 100 mL of 89 mM Tris base; 89 mM boric acid; and 2 mM EDTA, pH 8.3.
3. EtBr stock (10 mg/mL).
4. Owl Separations model B1 gel box or equivalent.
5. *Hae*III-digested Phi-X DNA (100 ng/ μ L).
6. Bio-Rad Mini Spin Filters (cat no. 732-027).
7. Gel elution buffer: 20 mM Tris-HCl, pH 7.5; 5 mM EDTA; 400 mM Na acetate.

2.9. Restriction Enzyme Cutting of cDNA

1. New England Biolabs restriction buffer 2: 250 mM NaCl, 10 mM Tris-HCl, pH 7.9, 10 mM MgCl₂, 1 mM DTT.
2. *Sfi*I restriction endonuclease (20 U/ μ L) (New England Biolabs).
3. BSA (4 mg/mL).
4. Agarose gel-running solutions (*see* Subheading 2.8.).

2.10. Ligation to Vector

1. *Sfi*I-cut λ cloning vector (500 ng/mL) (Clontech λ TriplEx-2).
2. 10X T4 DNA ligase buffer: 500 mM Tris-HCl, pH 7.5, 100 mM MgCl₂, 100 mM DTT, 250 μ g/mL of BSA.
3. 10 mM ATP.
4. T4 DNA ligase (400 U/ μ L) (New England Biolabs).
5. 16°C Incubator.

2.11. Packaging of Ligation

1. λ packaging extract (e.g., Stratagene Gigapack III).

2.12. Characterization of Insert Efficiency and Size

1. λ Phage dilution buffer: 100 mM NaCl, 10 mM MgSO₄, 35 mM Tris-HCl, pH 7.5.
2. λ phage host bacteria strain (e.g., XL1-blue or BM25.8).
3. 100 μ M Stock of vector 5' PCR primer (**Fig. 2**): 5'-CTCCGAGATCTGGAC GAGC-3'.
4. 100 μ M stock of vector 3' PCR primer (**Fig. 2**): 5'-TAATACGACTCACTAT AGGGC-3'.
5. 750 μ L of 2X PCR mix (made fresh): 275 μ L of water, 150 μ L of 10X *Taq* buffer, 120 μ L of 25 mM MgCl₂, 150 μ L of dNTPs (2 mM each), 15 μ L of 100 μ M vector 5' PCR primer, 15 μ L of 100 μ M vector 3' PCR primer, 5 μ L of *Taq* polymerase (2.5 U/ μ L).

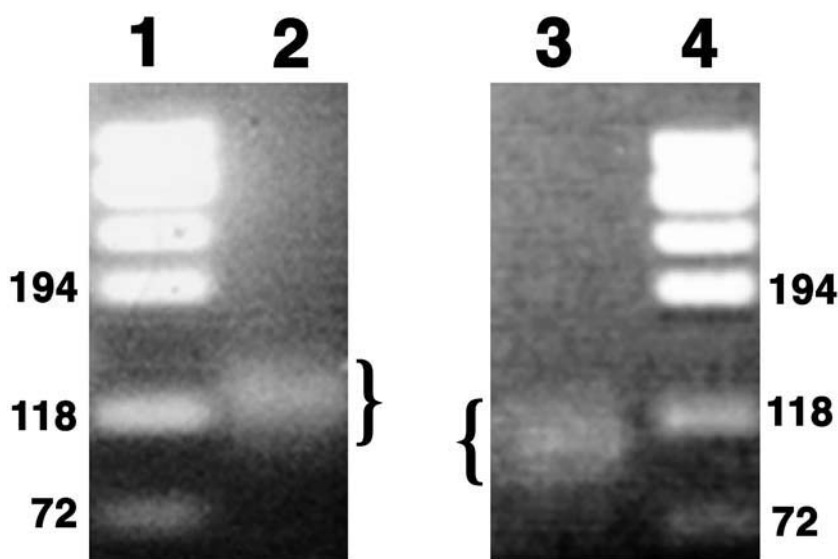


Fig. 2. Agarose gel analysis of cDNAs generated by RT and PCR. cDNAs were prepared from size-fractionated ~22-nt RNA as described in **Subheading 3.2.** and analyzed on a 2% agarose gel, along with 1 μ g of Phi-X *Hae*III makers (lanes 1 and 4). Lane 2, uncut PCR product (brace); lane 3, PCR product after digestion with *Sfi*I restriction endonuclease (brace). Judging from the fluorescence signals, the *Sfi*I-digested cDNA seems similar to the 118-nt Phi-X *Hae*III band, which contains about 60 ng of DNA.

2.13. Amplification and Storage of Library

1. λ phage dilution buffer: 100 mM NaCl, 10 mM MgSO_4 , 35 mM Tris-HCl, pH 7.5.

2.14. cDNA Sequencing

1. Sequencing primer oligonucleotide (**Fig. 2**): 5'-CTCGGGAAGCGGCCAT TGTG-3'

2.15. Distinguishing miRNAs From Other cDNA Sequences

1. Internet connection and browser to access online database search services:
 - a. NCBI Blast: www.ncbi.nlm.nih.gov/BLAST/.
 - b. Michael Zuker's mfold server: www.bioinfo.rpi.edu/~zukerm/rna/.
 - c. Rfam: www.sanger.ac.uk/Software/Rfam/mirna/index.shtml.
 - d. Annotated genome databases.
 - e. Wormbase: <http://wormbase.org/>.

- f. BDGP: www.fruitfly.org/.
- g. Ensemble: www.ensembl.org/.

2.16. Northern Blot Analysis

1. Electroblothing apparatus compatible with gel format (*see Subheading 2.3.*).
2. GeneScreenPlus (Perkin-Elmer) charged nylon membrane.
3. Whatman 3MM blotting paper.
4. Integrated DNA Technologies Starfire Oligo Labeling Kit.
5. Amersham (cat. no. AA0074) or NEN (cat. no. NEG012Z) α -³²P dATP (6000 Ci/mmol) (note that other brands of buffers can inhibit polymerase).
6. Sephadex G25 medium for spin column.
7. 1 M Na₂HPO₄, pH 7.2, stock: Dissolve 268.07 g Na₂PO₄ in 800 mL H₂O. Add 4 mL of 85% phosphoric acid. Adjust the volume to 1 L.
8. Northern blot hybridization buffer: 7% SDS; 0.2 M Na₂HPO₄, pH 7.2.
9. 5X SSPE: Dissolve 175.2 g of NaCl, 27.6 g of NaH₂PO₄•H₂O, and 7.4 g of Na₂EDTA in 3 L of water. Adjust the pH to 7.4 with NaOH and then adjust the volume to 4 L with water; final [Na] of 5X SSPE = 0.8 M.
10. 2X SSPE, 0.1% SDS.
11. 0.5X SSPE, 0.1% SDS.
12. Phosphorimager or X-ray film.

3. Methods

In **Subheading 3.1.**, total RNA is prepared from cells or tissues (or, e.g., whole worms), fractionated by size, and the ~22-nt size fraction is used as a template for cDNA synthesis, described in **Subheading 3.2.** An RNA ligation step is then used to add a linker to the 3' end of the RNA to permit primed cDNA synthesis. The cDNA is amplified by PCR and digested with a restriction enzyme in preparation for cloning, described in **Subheading 3.3.** **Subheading 3.4.** describes the sequencing of cDNA clones and computational analysis of their genomic loci in order to identify the most likely candidates for new miRNA genes. **Subheading 3.5.** describes the detection and quantitation of small RNA transcripts in samples of RNA extracted from cells.

3.1. Preparation of RNA

The following RNA extraction protocol is designed for extracting total RNA from whole nematodes, but it can be used essentially as described for cells grown in culture, or small animals such as insect larvae. For some tissues, or animals with an exoskeleton, frozen samples will need to be ground in a mortar and pestle prior to the Trizol step. It is critical to take steps to avoid contaminating nuclease activity. Thus, clean nitrile gloves should be worn at all times and should be changed often. The cloning of small RNAs requires a fractiona-

tion step (*see Subheading 3.1.2.*) to remove the larger RNAs, which, because of their abundance, would vastly outnumber the relatively rare miRNAs.

3.1.1. Total RNA Extraction

The following protocol lists quantities of reagents appropriate for a 1-mL volume of starting material (the sample pellet); volumes should be scaled up or down according to the actual sample volume.

1. Centrifuge samples of live cells (or, in our example, worms) in a 15-mL polypropylene tube: Remove excess buffer, note the volume of the pellet, and quickly freeze the sample by immersing the tube in liquid nitrogen. Store at -80°C .
2. For a 1-mL pellet, add 4 mL of Gibco Trizol to the frozen pellet, and vortex vigorously for 15 min at room temperature (*see Note 6*).
3. Add 0.85 mL of chloroform and vortex for 1 min.
4. Centrifuge in a Sorvall HS4 swinging-bucket rotor for 10 min at 5000 rpm, 4°C .
5. Transfer the aqueous (top) phase into a clean polypropylene tube, add an equal volume (equal to the total volume of the aqueous phase) of phenol/chloroform (50/50), and vortex for 2 min.
6. Centrifuge in a Sorvall HS4 swinging-bucket rotor for 10 min at 5000 rpm, 4°C .
7. Transfer the aqueous phase to a clean polypropylene tube, add an equal volume of isopropanol, and chill the tube on ice for 10 min. Centrifuge in a Sorvall HS4 swinging-bucket rotor for 10 min at 5000 rpm, 4°C .
8. Carefully pour off the supernatant, and gently rinse the pellet with ice-cold 70% ethanol/water. Allow the pellet to air-dry at room temperature, and then dissolve it in 0.5 mL of TES buffer. It can take considerable time and effort to dissolve the RNA; pipet the suspension up and down with a 1-mL micropipet tip until the cloudiness of the solution dissipates.
9. Add 2 mL of Trizol, and vortex vigorously for several minutes to ensure that the pellet is thoroughly dissolved.
10. Add 425 μL of chloroform and vortex for 2 min. Split the mixture into two 1.5-mL microcentrifuge tubes, and spin for 10 min at 12,000 rpm in a microcentrifuge.
11. Transfer the aqueous phases to a clean tube, add an equal volume of phenol/chloroform, vortex for 30 s, and spin for 10 min at 12,000 rpm in a microcentrifuge.
12. Transfer the aqueous phase to a clean tube, and add 1/10 vol of 3 M sodium acetate, followed by 1 vol of isopropanol. Chill on ice for 1 h, and recover the precipitate by centrifugation.
13. Add 1 mL of ice-cold 70% ethanol, spin in a refrigerated microcentrifuge for 30 s, and decant all the extra ethanol from the tube. Place the tube on its side on the benchtop to dry.
14. When the pellet appears dry, dissolve it in 0.5 mL of TE buffer. Homogenize the pellet with a micropipet tip, as in **step 8**. If the pellet resists dissolution, add one or two additional 100- μL portions of TE and continue mixing.
15. Add an equal volume of phenol/chloroform, vortex for 30 s, and spin for 10 min at 12,000 rpm in a microcentrifuge.

16. Transfer the aqueous phase to a clean tube, and repeat the phenol/chloroform extraction cycle until the aqueous/organic interface is clean after centrifugation.
17. Add to the aqueous phase an equal volume of chloroform, vortex, and separate the phases by centrifugation. Transfer the aqueous phase to a clean tube, and add 1/10 vol of 3 M sodium acetate followed by 3 vol of ethanol. Mix at room temperature for a few minutes, and centrifuge for 5–10 min at 12,000 rpm in a micro-centrifuge.
18. Rinse the pellet in 70% ethanol, allow it to dry at room temperature, and dissolve in 200–500 μ L of ice-cold RNase-free TE or water (use only as much volume as needed to fully dissolve the sample). Remove 1 μ L and dilute it 1/500 in 0.5 mL of water and save on ice (for reading the OD₂₆₀ in **step 20**).
19. Immediately divide the sample into separate tubes of 50- to 100- μ L aliquots and store at -80°C .
20. Read the OD₂₆₀ of the 1/500 dilution; 1 OD₂₆₀ = 35 $\mu\text{g/mL}$.

3.1.2. Preparation of Internal Size Marker for Size Selection

RNAs of about 22 nt in length are separated from other RNAs by filtration and denaturing acrylamide gel electrophoresis. To facilitate the recovery of the ~22-nt material, and to help monitor subsequent purification and processing steps, a ^{32}P -labeled marker oligonucleotide is included in the RNA sample. The RNA oligonucleotide marker should be 22 nt in length and should have approximately equal proportions of the four ribonucleotide bases. Because cDNAs corresponding to the internal marker oligonucleotide can be cloned, just like any other ~22-nt RNA in the sample, its sequence should be chosen so that it is not a close match to any sequence in the genome of the organism under study. For *C. elegans*, the oligonucleotide sequence 5'-GCCAAGCGTG TAGGGATCCAAA-3' was used. The marker is labeled with ^{32}P by T7 transcription in the presence of ^{32}P UTP. The T7 template is prepared by annealing two DNA oligonucleotides to produce a double-stranded T7 promoter linked to a single-stranded DNA complementary to the marker sequence.

1. Anneal the T7 template by mixing in a clean 0.5-mL tube 20 μL of 100 μM T7 top-strand oligonucleotide and 20 μL of 100 μM marker template oligonucleotide. Heat the mixture to 95°C for 2 min. Transfer the tube to room temperature for 10 min.
2. Set up a 20- μL T7 transcription reaction by mixing 5 μL of water, 2 μL of 10X T7 transcription buffer, 1 μL of annealed template, 1 μL of 10 mM ATP, 1 μL of 10 mM CTP, 1 μL of 10 mM GTP, 1 μL of 200 μM UTP, 2 μL α - ^{32}P UTP (10 $\mu\text{Ci}/\mu\text{L}$; 3000 Ci/mM), 2 μL of 100 mM DTT, and 2 μL of T7 polymerase (20 U). Incubate at 37°C 1 h.
3. Add the following to the transcription reaction from **step 2**: 1 μL of RNase-free DNase I (2 U/ μL). Incubate at 37°C for 15 min.
4. Add the following: 2 μL of 40 $\mu\text{g/mL}$ glycogen, 80 μL of water, 15 μL of 3 M Na acetate, and 300 μL of ethanol. Mix, and then spin at room temperature for 10 min at 12,000 rpm.

5. Decant the supernatant (dispose of unincorporated $\alpha^{32}\text{P}$ UTP properly), and add 1 mL of cold 70% ethanol. Spin for 1 min in a refrigerated microcentrifuge. Decant the supernatant again, and place the tube on its side on the bench.
6. Dissolve the pellet in 10 μL of water, and store at -20°C .

3.1.3. Gel Purification of Marker RNA Oligonucleotide

1. Prepare a 12.5% acrylamide 8 M urea gel, 4 cm wide by 1.5 mm thick by 8 cm long, using a comb with a 1-cm-wide lane (the Bio-Rad Protean II minigel system or equivalent is recommended). Accordingly, mix 7.5 mL of Sequagel concentrate (National Diagnostics), 6 mL of Sequagel diluent, and 1.5 mL of Sequagel buffer. Degas by vacuum for 1 min, and then add 150 μL of 10% ammonium persulfate and 3.5 μL of TEMED. Mix gently and pour the gel.
2. To the 10- μL sample of transcription product from **Subheading 3.1.2., step 6**, add 10 μL of denaturing gel sample buffer. Heat the sample to 65°C for 5 min, chill the tube on ice, and load onto the gel. Electrophorese at 18 V/cm (150 V for an 8-cm gel) until the faster dye reaches the bottom of the gel.
3. Remove the front glass plate and wrap the gel and rear plate in Saran Wrap and expose to X-ray film or a phosphorimager plate. Mark or note the coordinates of the corners of the glass plate for future repositioning.
4. After exposing the autoradiogram, mark the location of the 22-nt marker, and reposition the autoradiogram over the gel (still covered with Saran Wrap) using the previously noted locations of the corners of the plate. Mark on the Saran Wrap the position of the full-length 22-nt product (this should run about halfway between the two dyes). Cut out a rectangle of the gel encompassing the full-length product.
5. Place the gel slice in a clean 1.5-mL microcentrifuge tube, and homogenize it with the small end of a 1000- μL plastic pipet tip that has been melted shut to a rounded, roughly spherical shape at the tip by briefly passing the tip through an alcohol burner or small Bunsen burner flame. Be sure to grind the gel slice into small particles.
6. Add 600 μL (at least four times the volume of the gel slice) of gel elution buffer, cap the tube tightly, and agitate it vigorously for 10 min using a vortex mixer.
7. Spin the gel slice homogenate through a Bio-Rad Mini Spin Filter, and monitor the yield of ^{32}P in the eluate compared to the gel remnant (column retentate).
8. If less than about 80% of the radioactivity is eluted, save the eluate at -20°C , transfer the acrylamide gel particles to a fresh microcentrifuge tube, and repeat **steps 6 and 7**.
9. Pool the eluates, and for each 400 μL of sample, add 1 μL of 20 $\mu\text{g/mL}$ glycogen, and 1 mL of ethanol. Centrifuge in a refrigerated microcentrifuge for 10 min; decant the supernatant (monitor the recovery of radioactivity using a Geiger counter); air-dry the pellet at room temperature; and dissolve in 400 μL of 20 mM Tris-HCl (pH 7.5), 300 mM Na acetate.
10. Add 1 mL of ethanol and centrifuge for 10 min in a refrigerated microcentrifuge. Decant the supernatant (monitor the recovery of radioactivity using a Geiger counter), air-dry, and dissolve in 40 μL of water. Store at -80°C .

3.1.4. Preparation of Size-Selected RNA From Total RNA

1. Start with a 100- μ L sample of total RNA at a concentration of about 4 mg/mL. Add 300 μ L of denaturing gel sample buffer. Mix about 200,000 cpm of the 22-nt internal marker with the sample, heat to 65°C for 5 min, and cool on ice. Split the sample into two 200- μ L portions, and centrifuge each portion through separate 0.5-mL Microcon YM-100 columns (Millipore, Bedford, MA) according to the manufacturer's instructions. Monitor the recovery of RNAs of about 22 nt by detecting the ^{32} P marker using a Geiger counter. After the first spin, add 200 μ L of fresh denaturing buffer to the top of each column, mix gently by pipetting up and down, transfer the sample to a fresh Microcon YM-100 column, and centrifuge it again. Pool the eluates. This material is enriched for RNAs shorter than approx 100 nt (see **Note 7**).
2. Add Na acetate to a final concentration of 0.3 M, and add 1 vol of isopropanol. Chill the tubes on ice for 10 min, and recover the precipitate by centrifugation. Rinse the RNA pellets with chilled 70% ethanol, allow the pellets to dry, dissolve them in 50 μ L of denaturing gel sample buffer, and pool the samples into one tube. Store frozen at -20°C or process immediately by gel electrophoresis.
3. Load the sample onto a 12% acrylamide gel in a 2-cm-wide well. Electrophorese the sample until the faster dye is at the bottom of the gel, and elute the radioactive material as described in **Subheading 3.1.3**. (but include about 3 to 4 mm of gel above and below the marker, in order to recover RNAs in the range of about 18–25 nt). The pooled gel eluates containing size-selected RNAs may also contain some small particles of acrylamide, which unless removed, can inhibit the subsequent RNA ligation step (**Subheading 3.2.1**). The Bio Rad Mini Spin Filter step (**Subheading 3.1.3**, **step 7**) is critical for removing these and other particulate contaminants. For each 400 μ L of filtered gel eluate, add 20 μ g of glycogen and 1 mL of ethanol. Place the sample at -20°C for at least 1 h. Recover the RNA/glycogen precipitate by centrifuging in a refrigerated microcentrifuge (monitor the recovery with a Geiger counter), wash the pellets with 70% ethanol, air-dry, dissolve the pooled samples in 20 μ L of water, and immediately store at -80°C.

3.2. Preparation of cDNA

There are several alternative methods for preparing the cDNA from the size-selected RNA sample (**3–6,13**). All of these require the use of T4 RNA ligase to attach a linker of known sequence at the 3' end of the RNA sample (**Subheading 3.2.1**). This 3' linker is then used to prime first-strand reverse transcription (**Subheading 3.2.2**).

3.2.1. Ligation of 3' Linker Oligonucleotide

Two alternative approaches are described for ligation of the 3' linker. Each approach has its advantages and disadvantages. One approach, the ATP-mediated ligation (**Subheading 3.2.1.1**), uses more readily available materials but

requires that the RNA substrate be pretreated with CIP to remove the 5' phosphate (and hence prevent circularization of the RNA during RNA ligation step). In addition, the 3' linker should have three ribonucleotide linkages at its 5' end. The other approach, the activated oligonucleotide ligation (**Subheading 3.2.1.2.**), bypasses the need for phosphatase treatment by using a 3' linker that is preactivated with an adenylyl group, obviating the need to include ATP in the ligation reaction (**4,19**). Conveniently, the activated oligonucleotide can be entirely DNA (*see Note 8*) and is therefore less expensive to synthesize than is RNA. Some researchers find that the activated oligonucleotide ligation method results in a more efficient yield of ligated product. However, the adenylyl group must be added to the linker oligonucleotide, and the activated oligonucleotide must be gel purified, which adds additional steps to the process. These alternative ligation procedures, and their relationship to the rest of the process, are summarized in **Fig. 1**.

3.2.1.1. ATP-MEDIATED LIGATION

1. Set up the following 20- μ L reaction to remove the 5' phosphates: 20 μ L of small RNA (in water; *see Subheading 3.1.3.*), 6 μ L of water, 3 μ L of 10X New England Biolabs buffer 3, 1 μ L of New England Biolabs CIP (10 U). Incubate at 37°C for 1 h. Add 70 μ L of Tris-HCl, pH 7.5.
2. Extract three times with phenol/chloroform, two times with chloroform.
3. Add Na acetate to a final concentration of 0.3 M, and 3 vol of ethanol. Incubate at least 2 h at -20°C. Centrifuge at 12,000 rpm in a refrigerated microcentrifuge; monitor the pellet with a Geiger counter. Essentially all the radioactivity should precipitate. Wash the pellet twice with ice-cold 70% ethanol, air-dry, and dissolve in 5 μ L of water; store at -80°C.
4. Set up the following 20- μ L ligation reaction: 11 μ L of CIP-treated size-fractionated RNAs, 2 μ L of 10X T4 RNA ligase buffer, 2 μ L of 60 μ M ATP, 2 μ L of DMSO, 1 μ L of 100 μ M RNA 3' linker oligonucleotide, 2 μ L of T4 RNA ligase (20 Units; Amersham). Incubate at 37°C for 1 h (*see Note 9*). Stop the ligation by adding 20 μ L of denaturing gel sample buffer. The stopped ligation reaction can be stored at -20°C or processed immediately by gel electrophoresis.
5. Load the ligation (40 μ L) on a 12.5% acrylamide 8 M urea gel and electrophorese as described in **Subheading 3.1.3**. Include a control lane with unligated ³²P-labeled internal marker oligonucleotide.
6. Expose the gel, with the corners of the glass plate marked, as described in **Subheading 3.1.3**. The ligated material is detected as a slower-moving radioactive band compared to the unligated oligonucleotide (**Fig. 3**). Elute the ligated RNA from the acrylamide gel slice as described in **Subheading 3.1.4**. (but include an extra 3–4 mm of gel above and below the labeled band of ligation product). Precipitate the RNA from the eluate with ethanol and 20 μ g of glycogen carrier. Recover the precipitate by centrifugation (use a Geiger counter to check the recovery of radioactivity).

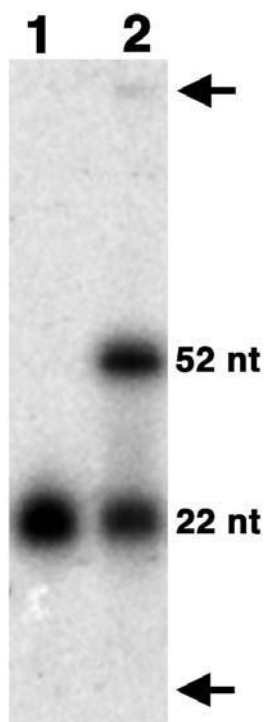


Fig. 3. Gel electrophoretic separation of unligated ~22-nt RNA (lane 1) and same sample ligated to a 30-nt linker RNA oligonucleotide (lane 2). The sample of small RNA, combined with a 22-nt ^{32}P -labeled marker, was treated with RNA ligase and fractionated by electrophoresis on a 12.5% acrylamide 8 *M* urea gel, as described in **Subheading 3.2.1**. Arrows indicate the top and bottom edges of the gel.

7. Thoroughly wash the pellet with 70% ethanol, allow the pellet to dry, and dissolve in 100 μL of water. Add 10 μL of 3 *M* Na acetate and 300 μL of ethanol, spin for 10 min in a refrigerated microcentrifuge, wash the pellet again, dry, and dissolve in 10 μL of water; store at -80°C .

3.2.1.2. ACTIVATED OLIGONUCLEOTIDE LIGATION

1. Set up a 100- μL adenylation reaction: 40 μL of DNA 3' linker oligonucleotide (240 μg), 20 μL of 1 *M* adenosine monophosphate, 20 μL of 1 *M* carbodiimide, 20 μL of 1 *M* methylimidazole (pH 6.0). Incubate at 25°C for 18 h. Add 10 μL of 3 *M* Na acetate and 300 μL of ethanol, and spin immediately for 10 min at room temperature. Decant the supernatant, wash the pellet with 70% ethanol, and air-dry. Dissolve in 40 μL of water; store at -20°C or process immediately by gel electrophoresis.

2. Add 40 μL of denaturing gel sample buffer. Electrophorese on a 12.5% acrylamide gel as described in **Subheading 3.1.3.**, with the untreated oligonucleotide as a marker. Detect the oligonucleotide by shadowing with a short-wave ultraviolet (UV) source over an X-ray film intensifying screen, or by staining the gel briefly with 0.5 $\mu\text{g/mL}$ of EtBr (1/20,000 dilution of 10 mg/mL stock), and visualize the oligonucleotide by fluorescence under long-wave illumination.
3. Cut out the region of the gel containing full-length oligonucleotide (the adenylylated oligonucleotide should run slightly slower than the untreated oligonucleotide). Elute the material from the acrylamide gel slice as described in **Subheading 3.1.3.**
4. Chloroform extract the eluate, ethanol precipitate the material, recover the precipitate by centrifugation, wash the pellet with 70% ethanol, air-dry, and dissolve in 50–100 μL of water. Read the OD_{260} of a sample of the solution to determine the concentration.
5. Set up the following 20- μL ligation reaction: 12 μL of size-fractionated RNAs, 2 μL of 10X T4 RNA ligase buffer, 2 μL of DMSO, 2 μL of 100 μM adenylylated 3' linker DNA oligonucleotide, 2 μL of T4 RNA ligase (20 U) (Amersham). Incubate at 37°C 1 h. Stop the ligation by adding denaturing gel sample buffer. The stopped ligation reaction can be stored at –20°C, or processed immediately by gel electrophoresis.
6. Electrophorese on a 12.5% acrylamide urea gel, and recover the ligation product as described in **Subheading 3.2.1.1., steps 5–7.**

3.2.2. Reverse Transcription

This procedure is adapted from the Clontech SMART manual for construction of cDNA libraries using the SMART strategy. The SMART cDNA cloning kit is recommended, but the protocol is written so that the worker can proceed without the kit using generic materials.

1. Set up the following 5- μL first-strand cDNA synthesis: 3 μL of ligation product (from **Subheading 3.2.1.1.** or **3.2.1.2.**), 1 μL of 10 μM stock of SMART template, 1 μL of 10 μM stock of reverse transcriptase (RT) primer. Mix and spin the tube briefly in a microcentrifuge to collect the contents at the bottom. Incubate at 72°C for 2 min. Cool on ice for 2 min. Spin the tube briefly in a microcentrifuge to collect the contents at the bottom.
2. Add the following (to make a final volume of 10 μL): 2.0 μL of 5X first-strand RT buffer, 1.0 μL of 20 mM DTT, 1.0 μL of dNTP mix (10 mM each dNTP), 1.0 μL of RNase H-deficient MMLV Reverse Transcriptase (Clontech Power-script). Set up an identical reaction but without RT (“minus RT control”). Incubate the tubes at 42°C for 1 h. Place the tube on ice.

3.2.3. PCR Amplification of cDNA

1. Set up the following 100- μL PCR reaction: 2 μL of first-strand cDNA, or minus RT control (from **Subheading 3.2.2.**), 74 μL of water, 10 μL of 10X PCR buffer, 10 μL of dNTP mix (2 mM each), 2 μL of cDNA 5' PCR primer, 2 μL of cDNA

3' PCR primer, 2 μ L of 2.5 U/mL of *Taq* polymerase. Mix the contents by gently flicking the tube. Centrifuge briefly to collect the contents at the bottom of the tube. Add two drops of mineral oil if necessary (or, preferably, use a thermocycler with a hot bonnet and corresponding tubes). Cap the tube, and place it in a preheated (95°C) thermal cycler.

2. Run the following PCR program:
 - a. Step 1: 96°C for 1 min.
 - b. Step 2: 96°C for 10 s.
 - c. Step 3: 50°C for 10 s.
 - d. Step 4: 72°C for 20 s.
 - e. Step 5: 25 cycles to **step 2**.
 - f. Step 6: 72°C for 3 min.
 - g. Step 7: 10°C indefinitely.
3. After the PCR reaction is complete, set up the following proteinase K digestion: 100 μ L of the amplified cDNA (2 to 3 μ g) and 4 μ L of proteinase K (20 mg/mL). Mix the contents and spin the tube briefly. Incubate at 45°C for 20 min.
4. Add 100 μ L of phenol:chloroform and vortex for 1 min. Centrifuge at 12,000g for 5 min. Transfer the top (aqueous) layer to a clean tube.
5. Add 100 μ L of phenol:chloroform to the aqueous layer. Vortex for 1 min. Centrifuge at 12,000g for 5 min.
6. Transfer the top (aqueous) layer to a clean tube.
7. Add 10 μ L of 3 M Na acetate, 2 μ L of glycogen (20 mg/mL), and 300 μ L of ethanol. Immediately centrifuge at 12,000g for 20 min at room temperature.
8. Decant the supernatant, wash the pellet with 70% ethanol, air-dry the pellet, and resuspend in 40 μ L of TE buffer.

3.2.4. Gel Purification of cDNA

1. Prepare a 10-cm-long 2% agarose slab gel, 1X TBE buffer, and 0.5 μ g/mL of EtBr.
2. Add 4 μ L of 10X agarose gel sample buffer to the cDNA sample (from **Subheading 3.2.3., step 8**), and load the sample onto the gel alongside of 10 μ L (1 μ g) of *Hae*III-digested Phi-X DNA size markers and the minus RT control. Electrophoresis at 100 V for 30 min.
3. View the DNA fragments in the gel under illumination with a long-wave UV source in a dark room. Mark the position of the PCR product, and excise using a clean scalpel.
4. Elute the DNA from the gel fragment by pulverizing the fragment, adding 5 vol of gel elution buffer and incubating with gentle agitation at room temperature for at least 4 h.
5. Transfer the gel fragments to a Bio-Rad Mini Spin Filter, and centrifuge at room temperature at 12,000g for 10 min.
6. For each 400 μ L of eluate, add 20 μ g of glycogen and 1 mL of ethanol. Chill at -20°C for at least 1 h, and recover the precipitate by centrifugation. Wash the pellet with 70% ethanol, and dissolve it in 30–50 μ L of TE buffer.

3.3. Cloning cDNAs in a λ Phage Vector

The gel-purified material generated in **Subheading 3.2.4.** is suitable for cloning in a variety of plasmid or phage vectors and subsequent sequencing. The RT-PCR method previously described places distinct, known sequences at the 5' and 3' ends of the ~22-nt cDNA sequence, and these sequences contain asymmetric *Sfi*I sites that permit directional cloning in vectors (λ TriplEx2 or plasmid pTriplEx2) with the appropriate compatible *Sfi*I sites. If the worker elects to use one of these vectors, by choosing the appropriate sequencing primer, the cDNA sequence can be read explicitly in a 5' to 3' direction (*see Note 10*).

3.3.1. Restriction Enzyme Cutting of cDNA

1. Set up a 40- μ L *Sfi*I digestion reaction: 31 μ L of cDNA PCR product (from **Subheading 3.2.4.**), 4 μ L of 10X New England Biolabs buffer 2, 4 μ L of *Sfi*I restriction endonuclease (20 U/ μ L), 1 μ L of 4 mg/mL BSA. Mix well. Incubate at 50°C for 2 h.
2. Add 4 μ L of 10X agarose gel-loading buffer, and mix well.
3. The *Sfi*I digestion products can be cloned directly, but the recovery of inserts will be more efficient if the desired *Sfi*I fragment is gel purified away from the end fragments. To do this, load the sample from **step 2** on a 2% agarose gel, and follow the electrophoresis and gel extraction protocol described in **Subheading 3.2.4.** Dissolve the final sample in 10 μ L of TE buffer.
4. Remove a 1- μ L sample, and electrophorese in 2% agarose with 0.5 μ g/mL of EtBr and 500 ng of Phi-X *Hae*III markers as size and quantitation standards (**Fig. 2**). Estimate the amount of insert fragment based on a visual comparison between the fluorescence level of the insert and the smaller marker bands (*see legend to Fig. 2*).

3.3.2. Ligation to Vector

1. Set up four 5- μ L ligation reactions containing increasing amounts of insert, and constant vector (in this case, 0.5 μ g of λ arms per ligation): 1 μ L of cDNA (0, 2.5, 5.0, or 50 ng), 1 μ L of 500 ng/mL stock of vector, 0.5 μ L of 10X T4 DNA ligase buffer, 0.5 μ L of 10 mM ATP, 0.5 μ L of T4 DNA ligase (200 U), 1.5 μ L of water. Incubate the tubes at 16°C overnight.

3.3.3. Packaging of Ligation

1. Perform a separate packaging reaction for each of the ligations. Commercial packaging extracts are recommended, such as Gigapack III. Follow the manufacturer's instructions.
2. Titer each of the resulting libraries using standard microbiological methods. The unamplified libraries (packaging reactions themselves) can be stored at 4°C for 2 wk.

3. Repeat the ligation using the ratio of cDNA to vector (of the initial three ligations) that gave the best results. Scale up the volumes of all reagents according to the amount of cDNA used. Then package and titer this scaled-up ligation.

3.3.4. Characterization of Insert Efficiency and Size

1. Plate out a portion of the library and pick several dozen plaques into microtiter wells containing 50 μ L of phage dilution buffer. One convenient way to pick a plaque is to use a plastic pipet tip and pipettor to suck up a core of agar containing the plaque and squirt the agar into the microtiter well. Pick plaques from the wild-type phage (no insert) as controls. Let the phage diffuse out overnight at 4°C.
2. Dilute the phage suspension 1 : 10 by transferring 5 μ L to a fresh microtiter well that contains 45 μ L of phage dilution buffer.
3. For each phage dilution set up the following PCR reaction: 10 μ L of 2X PCR mix (*see Subheading 2.12.* for recipe) and 10 μ L of phage dilution. Run the same PCR program as in **Subheading 3.2.3., step 2.**
4. Run 10 μ L of the PCR reactions directly on a 2% agarose gel with molecular weight markers. The PCR products from phage with an insert should be about 75 bp longer than those from phage without an insert (**Fig. 1**). Note the frequency of inserts of the proper size for each of the three libraries. Greater than 90% inserts can be expected. Pool the libraries that appear satisfactory, and use a portion of the pooled library to produce an amplified version of the library (*see Subheading 3.3.5.*).

3.3.5. Amplification and Storage of Library

1. Based on the titer of the unamplified library (from **Subheading 3.3.3., step 2**), plate out enough to produce about 6 to 7×10^4 plaques/150-mm plate. There is no need to plate out the whole library at once; for example, half may be plated, and the rest saved in case something goes wrong.
2. Incubate at 37°C for 6–18 h, or until the plaques become confluent.
3. Add 10 mL of 1X λ dilution buffer to each plate, and let the plates sit at 4°C overnight.
4. Bring the plates out to room temperature, set them on a platform shaker, and gently swirl for about 1 h. Collect the buffer in a sterile tube. This is the phage suspension.
5. Add 1/10 vol of chloroform to the lysate and vortex for 2 min.
6. Centrifuge in a Sorvall GSA rotor at 5000g for 10 min. Collect the supernatant and store at 4°C.
7. Determine the titer of the amplified library. The titer is expected to be about 10^{10} PFU/mL. The amplified library can be stored at 4°C for up to 6 mo. For long-term storage, make 1-mL aliquots and place at –80°C.

3.4. cDNA Sequencing

Sequencing of the clones is carried out using standard methods. Primers from within the vector are chosen so that the insert is at least 25–30 bp from

the primer (**Fig. 1**). If a poly-A linker is used in the cloning step (as in our example), the sequencing primer should be chosen so that the polymerase reads from the side of the cDNA opposite to the linker (**Fig. 1**), because sequencing across long homopolymer stretches can be inaccurate. In our example, a phage vector is used, and thus the most convenient method for producing material for sequencing is to pick the plaques into microtiter wells and amplify the insert by PCR as described in **Subheading 3.3.4., step 1**. The amplified PCR product can then be sequenced by standard methods for sequencing PCR fragments, using the sequencing primer (**Fig. 1**). Do not use the same primers for PCR amplification and sequencing; the sequencing primer(s) should be nested between the PCR primers. Obtain about 200 sequences in order to assess the quality of the library before scaling up and sequencing more clones.

3.5. Distinguishing miRNAs From Other cDNA Sequences

One of the most critical parts of the process of identifying miRNAs by cDNA cloning is analysis of the cDNA sequences to distinguish bona fide miRNA candidates from sequences that come from other kinds of small non-coding RNAs, particularly siRNAs and “tiny noncoding RNAs” (tncRNAs), which are similar in size to miRNAs (6,10). The cDNA clones produced from the procedure in **Subheading 3.4.** will also include fragments of known RNAs such as tRNAs, URNAs, rRNAs, and mRNAs. The first step in the analysis of cDNA sequences is therefore a Blast search of the sequences against an annotated sequence database (**Subheading 3.5.1.**). Sequences that match the genome of the organism under study can also be checked against comprehensive genomic annotations (**Subheading 3.5.2.**) to help distinguish noncoding sequences from those that lie in protein coding sequences. Previously identified miRNAs are filtered out by comparison with known miRNAs using the Rfam miRNA clearinghouse Web site (**Subheading 3.5.3.**). Finally, candidate miRNAs are tested by RNA folding prediction (**Subheading 3.5.4.**) for their potential to fold back into a hairpin precursor typical of miRNAs.

3.5.1. Blast Search of Genomic Databases

1. Using the NCBI Blast Web interface, or equivalent, perform a Blast search of the genomic sequence of the organism from which the cDNAs were isolated. Use the cDNA sequences as queries. Keep only those sequences that match the genome precisely (see **Note 11**).
2. By hand, or using a suitable automatic script, screen the precisely matched hits for sequences annotated as known noncoding RNAs, mRNAs, coding sequence, and so on. Keep only those sequences that do not match previously-known RNA or protein genes. Conserved sequences such as tRNAs and rRNAs can also be flagged by searching the entire NCBI database; sometimes these RNA genes may not be thoroughly annotated in any one specific genome database.

3.5.2. Screening of Genomic Annotations

1. Further check the locations of the cDNA sequences using the appropriate online annotated genome sequence display such as Wormbase (for *C. elegans*), BDGP (for *Drosophila*), or Ensemble (for mammals).

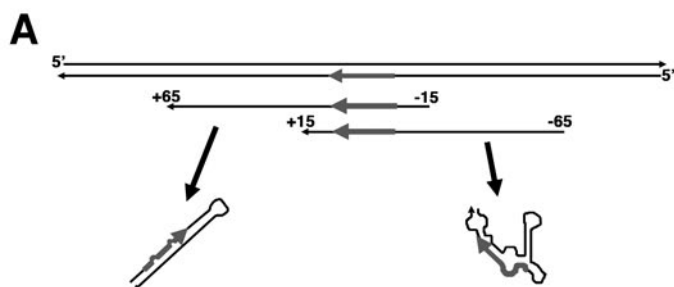
3.5.3. Screening for Previously Identified miRNAs

1. Consult the Rfam miRNA clearinghouse Web site and follow the instructions there for comparing your sequences with known miRNAs.

3.5.4. Prediction of RNA Structure

1. By hand, or using an automatic script, extract two segments of the genomic sequence surrounding the cDNA (be sure to extract from the same strand of genomic DNA as the cDNA) (see **Fig. 4A**). These sequence segments are as follows: one segment starting 15 nt before the 5' end of the cDNA sequence and ending 65 nt beyond the 3' end of the cDNA sequence, and another segment starting 65 nt before the 5' end of the cDNA sequence and ending 15 nt beyond the 3' end of the cDNA sequence.
2. Test both of these sequences for predicted RNA secondary structure using Michael Zuker's mfold program (20) at the mfold Web server.
3. Evaluate the predicted secondary structures visually, or using an automatic script, according to the following rules (14):
 - a. The cDNA sequence should be involved in base pairing with nucleotides outside of the cDNA sequence (no part of the cDNA should fold back on itself).

Fig. 4. (see facing page) Analysis of potential precursor secondary structures for miRNA candidates. (A) Two potential miRNA precursor sequences are extracted from the genomic sequence of the same strand of genomic DNA that contains the cDNA sequence (thick arrows). The two precursor candidate sequences are analyzed using an RNA secondary structure prediction program, such as mfold. A satisfactory miRNA-like fold is detected in this case from the putative precursor transcript extending from 15 nt before the 5' end of the cDNA to 65 nt beyond the 3' end of the cDNA. (B,C) Examples of miRNAs (bold) and their predicted precursor secondary structures. *mir-228* (B) folds back against the 3' side of the hairpin, and in this case, the -15/+65 sequence folded satisfactorily; *mir-230* (C) folds back against the 5' side of the hairpin, and thus the -65/+15 precursor candidate folded satisfactorily. (D,E) examples of cDNAs (bold) for which no satisfactory miRNA-like hairpin was identified; these RNAs are not classified as miRNAs but are given the generic tncRNA designation. The best fold for *tncR27* (D) contained a large bulge in the stem opposite to the cDNA sequence, which is atypical of miRNA precursors. The best fold for *tncR30* (E) involves <16 nt of the cDNA in base pairing.



B

5' A C C A A CUA CAU
GUUCGC **AUGG ACUG AUGA UUC** CGG UG A
CAGGCG UACC UGGC UACU AGG GCC AC A
 A A A - C AG- AGC

C

5' A - U U AACAU
UGCCGUC CUUGGUCG GCGAUU AAUAUUA CAG \
AUGGCAG **GGACCAGC** **UGUUGA** **UUAUGAU** GUU A
 A **G** - U AAAGG

D

5' U GU----- AG AG G
UGGAUUU CA AAUGG CACGGU ACA GUU A
GUUUGAA GU UUACC GUGUUA UGU CAG G
- AAAACU U A- -- A

E

5'-- AG- A G - A CACAACCCU
CUUG UAG UAUC UGUA UG GGUUU \
GAAC GUC AUAG ACGU GC CCAAA U
AG AUG A - U - AUUUCUACC

- b. At least 16 nt of the cDNA sequence should be involved in base pairing.
 - c. There should not be any large asymmetric bulges opposite to the cDNA sequence.
 - d. Preferably, the terminal nucleotides of the cDNA should be involved in base pairing.
 - e. The structure fitting these rules should be the lowest free-energy structure predicted by mfold.
4. Predicted hairpin precursors are classified as either 5', to indicate the location of the ~22-nt miRNA candidate in the 5' end of the hairpin (see example shown in **Fig. 4B**), or 3', to indicate the location of the miRNA in the 3' end of the hairpin (see **Fig. 4C**).

3.6. Northern Blot Analysis

A critical criterion for the identification of miRNAs is the demonstration that a ~22-nt transcript is detectable *in vivo*. Having cloned the sequence from a size-fractionated sample of total RNA is not in itself a sufficient criterion, since fragments of larger RNAs can streak through the ~22-nt fraction on the prep gel used in **Subheading 3.1.4**. Therefore, a detection method based on molecular weight fractionation and hybridization must be employed using the candidate miRNA sequence as a probe. Northern blot hybridization is the method of choice, since the detection of a hybridizing band of the right size signifies an accumulating species *in vivo*. In addition, Northern blots can simultaneously detect the larger precursor species. Since miRNAs are generated by Dicer cleavage of the precursor, in organisms such as *Arabidopsis* or *C. elegans*, where Dicer-deficient mutants are available, (or in cells where Dicer can be knocked down by RNAi), analysis of RNA from Dicer-deficient samples can be used to confirm that the small RNA is generated from the presumed precursor; the level of the precursor may increase in RNA from samples with reduced Dicer activity. However, Dicer dependence is not a definitive characteristic of miRNAs, since other small RNAs, including siRNAs and tncRNAs, are produced by Dicer activity (**10**).

3.6.1. Electrophoretic Separation of RNA

1. Prepare a 1.5-mm-thick 12.5% acrylamide 8 M urea gel as described in **Subheading 3.1.3**, using a Teflon comb with teeth 0.5 cm wide.
2. Dissolve RNA samples (10–40 µg of total RNA) in 10–30 µL of denaturing gel sample buffer, heat to 90°C for 5 min, chill on ice, and electrophorese at 120 V for approx 1.5 h with stirring (place a stir bar in the buffer reservoir and place the apparatus on a magnetic stirrer). Stop the run when the bromophenol blue dye (the faster dye) reaches the bottom of the gel. Twenty-two nucleotide RNAs run about midway between the xylene cyanol and bromophenol blue dyes.
3. Stain the gel with 0.5 µg/mL of EtBr, and take a picture of the gel using an exposure such that the 5S RNA band is not overexposed. The fluorescence signal of

the 5S RNA band can be used to standardize the various lanes for amounts of RNA loaded.

3.6.2. Electrophoretic Transfer of RNA to Filter

It is recommended that an electrophoretic transfer apparatus be employed that is manufactured to be compatible with the gel apparatus used to run the gel. For example, the Bio-Rad Protean II minigel system includes a transfer apparatus. The manufacturer's instructions should be followed; the salient steps are as follows:

1. Wearing clean gloves, cut four rectangles of blotting paper about 1.5 cm larger than the gel. Also cut a rectangular piece of charged nylon hybridization membrane precisely the size of the gel (*see Note 12*).
2. Soak the hybridization membrane and the blotting paper in 0.5X TBE.
3. Prepare a sandwich consisting of the following (bottom to top):
 - a. The spacer sponge that comes with the transfer apparatus.
 - b. Two layers of wet blotting paper.
 - c. The gel (note which side is up, and which way is left or right).
 - d. The hybridization membrane (squeeze out any bubbles between the gel and the membrane by gently rolling a clean glass pipet over the wet membrane).
 - e. Two more layers of wet blotting paper.
 - f. The other sponge that comes with the transfer apparatus.
4. Immediately secure the sandwich between the clamping device part of the apparatus, and place the clamped sandwich into the buffer chamber submerged in 0.5% TBE. Be sure that the gel is facing the negative electrode and the filter is facing the positive electrode.
5. Place a frozen slab of 0.5X TBE in the buffer tank and surround the tank with ice water during transfer. Have a stir bar going in the tank to ensure a uniform thermal environment around the gel sandwich. Transfer at 80 V for 1 h. The current will be high and will increase during the run. *These cooling measures are absolutely essential for a successful transfer.*
6. After transfer, mark the wells with a pencil, disassemble the sandwich, place the filter RNA side up on a dry piece of filter paper, and irradiate it with 120 mJ of UV radiation.
7. Bake the filter at 80°C for 1 h.

3.6.3. Probe Design and Preparation

1. Design an oligonucleotide probe with sequence precisely antisense to the cDNA sequence. The Integrated DNA Technologies Starfire Oligonucleotide Labeling Kit is recommended (www.idtdna.com). Gel purification of the oligonucleotide is not necessary prior to labeling.
2. Label the Starfire oligonucleotide with 6000 Ci/mmol of α -³²P dATP according to the manufacturer's instructions. The probe can be used directly after the spin column with no further preparation; it should be used within 24 h.

3.6.4. Hybridization

1. Wearing clean gloves, wet the membrane with 5 mL of 7% SDS, 0.2 M Na₂PO₄ (pH 7.0) and place it in a hybridization bag or tube.
2. Add the probe (without denaturation), cap the tube or seal the bag, and incubate at 35°C overnight (*see Note 13*).
3. Pour off the hybridization solution, and add 20 mL of a solution of 2X SSPE, 0.1% SDS at room temperature. Let the filter agitate for a few minutes.
4. Pour off the wash solution and add another 20 mL of 2X SSPE, 0.1% SDS and incubate at 35°C (or other appropriate wash temperature; *see Note 13*) for 30 min.
5. Pour off the wash solution and repeat **step 4**.
6. Pour off the wash solution and add 20 mL of 0.5X SSPE, 0.1% SDS and incubate at the appropriate wash temperature for 30 min.
7. Pour off the wash solution and repeat **step 6**.
8. Remove the filter from the bag or tube using forceps and place it on a piece of Saran Wrap. Fold the Saran Wrap over the filter (do not let the filter dry out) and seal the edges of the Saran Wrap to make the interior airtight, so that the filter does not dry out, and the moisture within does not leak out and harm the detection device.
9. Expose the filter to X-ray film or a phosphorimager screen. Quantify signals with the phosphorimager software.
10. The relative efficiency of loading each lane of the gel can be standardized by the relative quantity of 5S RNA fluorescence (**Subheading 3.6.1., step 3**). Alternatively, to standardize loading and transfer efficiency, hybridize the filter to a probe designed to detect a tRNA or U6 RNA. miRNA level in a particular sample is expressed as a ratio of the radioactivity (minus background) in the miRNA band to the level of reference RNA (5S, tRNA, or U6) in the same lane of the gel.
11. Northern filters can be stripped of probe by immersing them in 1 L of boiling 0.1% SDS in a large beaker. Take the beaker off the heat and let it cool to room temperature with swirling. Keep the filters wet, and expose them again to determine whether the stripping is complete. A filter can be stripped and reused three to four times. Store stripped filters wrapped in Saran Wrap at -20°C.

4. Notes

1. All solutions, tubes, and pipet tips that contact RNA should be scrupulously kept RNase free. Solutions should be prepared in new plastic tubes and bottles using deionized water that was collected directly from the deionizer into a new plastic bottle. Bottles and tubes should be handled only with gloved hands. Pipet tips, tubes, and other supplies designated for RNA work should be clearly marked "RNA only," and their packages should be opened only while wearing gloves. Glass pipets and tubes used for RNA work should be baked prior to use.
2. Siliconized tubes are particularly advantageous for handling the very small volumes and small quantities of material without excess loss as a result of sticking to the sides of the tubes.

3. All solutions, buffers, enzymes, oligonucleotides, and so on should be stored at -20°C , unless otherwise specified.
4. Other gel formats can be used. The major consideration is to be sure to use a gel of sufficient thickness (1.5 mm) and casting comb dimensions (at least 5 mm wide) to allow up to 40 μg of RNA to be run per lane without overloading the gel.
5. The RNA 3' linker oligonucleotide is a 30mer of polyadenosine, with six 5' ribonucleotides, followed by 24 deoxyribonucleotides; the 3' end is blocked with a dideoxy modification and the 5' end is phosphorylated. The DNA 3' linker oligonucleotide is a 30mer of polydeoxyadenosine; the 3' end is blocked with a dideoxy modification and the 5' end is phosphorylated. The sequence of the RNA or DNA 3' linker oligonucleotides is not particularly important as long as the cDNA 3' PCR primer is matched to it. In this protocol, a polyadenosine oligonucleotide is used, but other sequences have also been used (4,5). The RNA 3' linker is employed for the ATP-mediated ligation option (*see Subheading 3.2.1.1.*) and the DNA 3' linker is used for the activated oligonucleotide option (*see Subheading 3.2.1.2.*).
6. It is critical to vortex the sample for the full 15 min. This should be carried out in a fume hood. Several samples can be processed in parallel using a vortex mixer with a manifold attachment that permits multiple tubes to be firmly affixed to the mixer during vigorous vortexing.
7. If the RNA is relatively concentrated, the solution can resist flowing through the membrane, requiring spin times of 15–20 min. A small amount of liquid always will remain on the membrane owing to the angle of the column in the rotor.
8. The presence of the adenylyl group bypasses the usual substrate specificity of RNA ligase for the 3' donor to be RNA (19).
9. There is some flexibility regarding times and temperatures of the ligation reaction. The reaction can be carried out for up to 2 h at 37°C , but often extended reaction time does not seem to increase ligation efficiency. Overnight ligation at 15°C has also been used.
10. Even without taking advantage of directional cloning, the asymmetric linker sequences allow the cDNA sequence to be unambiguously read by reference to the linker sequences, regardless of the orientation of the cloned insert relative to the vector (and sequencing primers). Therefore, a very straightforward approach is to clone the PCR product directly in a T/A cloning vector such as Invitrogen's TOPO system, bypassing **Subheading 3.3.** entirely. However, directional cloning is described here, since many workers may prefer the simplicity of interpreting the sequencing data. Asymmetric *Sfi*I sites will be employed, although other directional cloning methods can be used by designing the linker and/or cDNA PCR primer sequences to incorporate other restriction enzymes with asymmetric sites, such as *Ban*I (4). *Ban*I is particularly useful because it produces compatible sticky ends, permitting the ligation of cDNAs into concatamers. This variation on the method allows several different cDNA sequences to be obtained from a single sequence read (4), which saves sequencing costs. However, extra steps are

required prior to cloning to prepare the concatamers, and it is advisable to clone single inserts first, either by T/A cloning or after restriction enzyme digestion. Then, after the worker is satisfied with the rest of the procedure, he or she can introduce the concatamerization steps. Concatamerization is not described here, since for most purposes, single inserts will be entirely satisfactory, albeit more expensive. It is assumed that the worker has access to the materials and equipment for, and experience with, standard phage and plasmid cloning. For workers who wish to clone concatamers of small RNA cDNAs, excellent protocols are described elsewhere (4,5,21).

Plasmid vectors can be more straightforward to sequence, and are more readily adaptable for high-throughput sequencing. Some phage vectors, including the λ TriplEx2 phage used in this example, can be converted to a plasmid by CRE-mediated excision (see Clontech's instructions). However, before using the plasmid conversion for extensive sequencing, a pilot conversion should be carried out and plasmid preps should be performed to check that the plasmid is present in high enough copy number for efficient sequencing.

11. Cloning artifacts can sometimes result in ambiguities at the ends of the cDNA sequence. In particular, the SMART cDNA synthesis strategy can sometimes result in an extra G at the 5' end of the cDNA sequence, and if there happens to be a G in that position in the genomic sequence, a G can be erroneously assigned as an additional 5' nucleotide. Multiple clones corresponding to the same cDNA can clarify this particular ambiguity, but single-read sequences with an apparent 5' G should be provisionally scored as lacking the G. Another artifact can result from the particular sequence of the 3' linker. In the example described here, poly-A linker is used, and thus the resulting cDNA sequences cannot be unambiguously scored as containing 3' terminal As, even in the event that these match the genomic sequence. This issue can eventually be resolved by primer extension analysis of the small RNA *in vivo*.
12. A critical factor in detecting miRNAs is the choice of charged nylon membrane. GenescreenPlus (Perkin-Elmer) has been used with success, as has Nytran Supercharge, Osmonics Magnagraph, and Bio-Rad Zeta-Probe. A significant complicating factor is that some lots of a particular brand of membrane can work well, whereas other lots of the same membrane may not detect miRNAs at all. Some lots will detect the 22-nt band very well but not the ~65-nt precursors. It is recommended that if possible a sample of the lot of membrane be compared with a sample of membrane from a verified lot prior to making a purchase. If a verified lot is not available, get samples of charged nylon from several laboratories and compare them by probing to a published miRNA (e.g., *mir-1*) that is expected to be relatively abundant. Abundant miRNAs can usually be visualized in less than 4 h on a phosphorimager.
13. The necessarily short probes required for detecting ~22-nt RNAs can have very low GC content and hence low melting temperatures. Hybridization should be performed at 35°C, and the conditions of the subsequent washes should be adjusted according to the length and base composition of the probe. We use the

following formula to estimate the melting temperature for a DNA probe of length L (nucleotides) and of GC content M (where M is from 0 to 1): $T = C - (500/L) + 48.7M + 5.9M^2 \cdot C = 73.7$ for 2X SSPE, 63.4 for 0.5X SSPE, and 56.8 for 0.2X SSPE. A rule of thumb is to wash at 10° less than the melting temperature. However, the formula given is just a rough guide, and in some cases, it can be advisable to determine wash conditions empirically. Start with a wash at low stringency, such as 35°C in 2X SSPE; expose the membrane; and then wash with progressively higher-stringency buffers (0.5X SSPE, 0.2X SSPE at 35°C ; then 0.2X SSPE at 36, 38, and 42°C), exposing the membrane in between each step up in stringency. In this manner, the worker can assess empirically the highest-stringency wash that yields a good signal and low background for a probe of given base composition. With progressively stringent washes, distinct expression patterns can be detected for miRNAs that differ by only one base, indicating that the method can be quite discriminating.

Acknowledgments

We thank members of the Ambros laboratory for helpful criticism. The work on small RNAs was supported by a Public Health Service (PHS) grant GM34028.

References

- 1 Eddy, S. R. (2002) Computational genomics of noncoding RNA genes. *Cell* **109**, 137–140.
- 2 Ambros, V. (2001) MicroRNAs: tiny regulators with great potential. *Cell* **107**, 823–826.
- 3 Lee, R. C. and Ambros, V. (2001) An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **294**, 862–864.
- 4 Lau, N. C., Lim, L. P., Weinstein, E. G., and Bartel, D. P. (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294**, 858–862.
- 5 Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001) Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853–858.
- 6 Llave, C., Kasschau, K. D., Rector, M. A., and Carrington, J. C. (2002) Endogenous and silencing-associated small RNAs in plants. *Plant Cell* **14**, 1605–1619.
- 7 Lagos-Quintana, M., Rauhut, R., Meyer, J., Borkhardt, A., and Tuschl, T. (2003) New microRNAs from mouse and human. *RNA* **9**, 175–179.
- 10 Lagos-Quintana, M., Rauhut, R., Yalcin, A., Meyer, J., Lendeckel, W., and Tuschl, T. (2002) Identification of tissue-specific microRNAs from mouse. *Curr. Biol.* **12**, 735–739.
- 9 Lim, L. P., Lau, N. C., Weinstein, E. G., Abdelhakim, A., Yekta, S., Rhoades, M. W., Burge, C. B., and Bartel, D. P. (2003) The microRNAs of *Caenorhabditis elegans*. *Genes Dev.* **17**, 991–1008
- 10 Ambros, V., Lee, R. C., Lavanway, A., Williams, P. T., and Jewell, D. (2003) microRNAs and other tiny endogenous RNAs in *C. elegans*. *Curr. Biol.* **13**, 807–818.

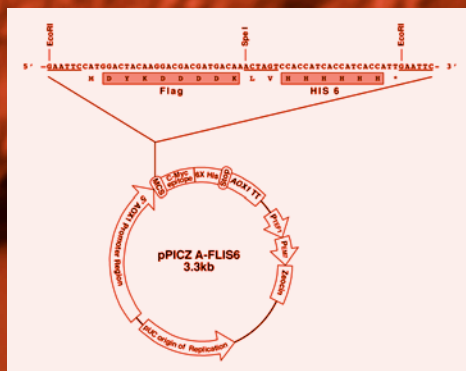
- 11 Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993) The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843–854.
- 12 Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., and Ruvkun, G. (2000) The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **403**, 901–906.
- 13 Elbashir, S. M., Lendeckel, W., and Tuschl, T. (2001) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* **15**, 188–200.
- 14 Ambros, V., Bartel, B., Bartel, D. P., et al. (2003) A uniform system for microRNA annotation. *RNA* **9**, 277–279.
- 15 Llave, C., Xie, Z., Kasschau, K. D., and Carrington, J. C. (2002) Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. *Science* **297**, 2053–2056.
- 16 Olsen, P. H. and Ambros, V. (1999) The *lin-4* regulatory RNA controls developmental timing in *C. elegans* by blocking LIN-14 protein synthesis after the initiation of translation. *Dev. Biol.* **216**, 671–680.
- 17 Seggerson, K., Tang, L., and Moss, E. G. (2002) Two genetic circuits repress the *Caenorhabditis elegans* heterochronic gene *lin-28* after translation initiation. *Dev. Biol.* **243**, 215–225.
- 18 Lim, L. P., Glasner, M. E., Yekta, S., Burge, C. B., and Bartel, D. P. (2003) Vertebrate microRNA genes. *Science* **299**, 1540.
- 19 Chu, B. C. and Orgel, L. E. (1984) Preparation of ligation intermediates and related polynucleotide pyrophosphates. *Biochim. Biophys. Acta* **782**, 103–105.
- 21 Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**, 911–940.
21. Pfeffer, S., Lagos-Quintana, M., and Tuschl, T. (2003) Cloning of small RNA molecules, in *Current Protocols in Molecular Biology* (Ausubel, F. M., Brent, R., Chanda, V. B., Kingston, R. E., Moore, D. D., Seidmann, J. G., Smith, J. A., and Struhl, K., eds.), Wiley and Sons, New York.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



A Historical Perspective on RNA Editing

How the Peculiar and Bizarre Became Mainstream

Donna J. Koslowsky

Summary

The known examples of RNA editing now encompass a variety of alterations of RNA primary sequence that arise from base modifications, nucleotide insertions or deletions, and nucleotide replacements. Hence, the definition of RNA editing has evolved as new systems have been described. This chapter presents a historical perspective on some of the pivotal discoveries that helped direct the current avenues of research in the field of RNA editing.

Key Words: RNA editing; RNA processing; deaminase; trypanosomes; plastids.

1. Introduction

The trypanosomes *Leishmania* and *Trypanosoma* are ancient eukaryotes that have long been considered bizarre because of the repertoire of unique “tricks” they use to survive in their insect and mammalian hosts (1,2). However, the initial report describing the “editing” of a mitochondrial transcript by the insertion of four uridines surprised even seasoned trypanosome investigators (3). This report was soon followed by reports of mRNA U-insertions involving first tens, then hundreds of inserted uridylates (4–6). These editing events not only corrected internal frameshifts and created AUG initiation and UAG/UAA stop codons, but they actually created the open reading frames (ORFs) of many of the highly conserved electron transport proteins. The idea that >50% of the sequence information in an mRNA transcript was not found in its cognate gene was astounding and forced biologists to reappraise the basic genetic concept of gene-protein relationships. While this extreme form of editing remains limited to the

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

mitochondria of trypanosomes, its discovery opened the door to a new level of gene regulation that proved to be widespread among the biological universe.

Within 5 yr of Benne's seminal article (3), RNA-editing events, involving vastly different mechanisms, had been described in a number of organisms. In mammals, two forms of RNA editing resulting from base modification rather than nucleotide insertions were initially described: the deamination of cytidine to uridine ($C \rightarrow U$) in the mRNA encoding apolipoprotein B (apoB), and of adenosine to inosine ($A \rightarrow I$) in transcripts encoding glutamate-sensitive ion channels (7–9). These base modifications were shown to be determining events in the expression of protein variants involved in lipoprotein assembly and higher brain functions, respectively. In the paramyxoviruses, it was shown that the insertion of G residues was responsible for the shift in reading frame that allowed the production of two different proteins (10,11). In plants, both mitochondrial and chloroplast transcripts were found to undergo numerous $C \rightarrow U$ transitions that converted anomalous codons to ones that specified highly conserved amino acid residues (12–14). In the slime mold *Physarum polycephalum*, multiple C-insertions were shown to correct numerous frameshifts in mRNAs (15). The discovery of these very different forms of RNA editing, each with readily perceivable impacts on cell and whole organism physiology, established RNA editing as an important new regulatory process in the overall scheme of RNA maturation. Now, almost two decades later, a diverse number of viral RNAs, mRNAs, and structural and catalytic RNAs are known to be affected by RNA editing across an equally broad spectrum of organisms (Table 1).

The known examples of RNA editing encompass a variety of alterations of RNA primary sequence that arise from base modifications, nucleotide insertions or deletions, and nucleotide replacements. The definition of RNA editing has therefore evolved as new systems have been described. Initially, editing was defined as an event that directly changes base-pairing potential and thus the genetic content of an mRNA. This definition had to be expanded with the discovery of similar nucleotide changes in both tRNAs and rRNAs (16–18). RNA-editing events were found to be necessary to correct stem mismatches and restore conserved structural elements known to be important in the modulation of translational efficiency and fidelity. RNA editing could even change tRNA identity by base conversion in the anticodon loop (18). Editing is now described as any process (other than splicing or capping) that changes the sequence of an RNA transcript from that encoded by the corresponding gene (19).

This broadening of the definition has invariably linked editing with the RNA modification field. The boundaries between RNA editing and RNA modification have always been less than distinct: the $C \rightarrow U$ and $A \rightarrow I$ editing conversions are highly reminiscent of well-established RNA modification reactions, and many "classic" RNA modifications have the same functional consequences

Table 1
Defining a New Level of Genetic Control in RNA Editing

Year	Editing type	Organism	Location	Target RNA	Result	Reference
1986	U-insertions	<i>T. brucei</i>	Mitochondria	coxII mRNA	Frameshift correction via insertion of four uridines	3
1987	C → U	<i>C. fasciculata</i> Mammals	Nucleus	apoB mRNA	Introduction of stop codon to create apoB-48; first report of tissue-specific editing to create two proteins	7,8
1987	U-insertion	<i>T. brucei</i>	Mitochondria	Cyb mRNA	First report of developmental regulation of U-insertions and creation of AUG initiation codon	5
1988	U-insertions U-deletions	<i>T. brucei</i>	Mitochondria	COIII mRNA	Creation of open reading; first report of pan-editing with insertion of >300 nt	4
1988	G-insertion	Paramyxovirus	—	Viral RNA	Frameshift to alternative ORF via insertion of two Gs	10
1989	C → U	Plant	Mitochondria	mRNAs	Multiple codon changes to conserved amino acids	12–14
1990	U → C	Plant	Mitochondria	mRNAs	First report of U → C editing	51,52
1991	A → I	Mammals	Nucleus	GluR mRNAs	First report of specific A → I editing in mRNA	9
1991	C → U	Plant	Chloroplasts	mRNAs	First report of editing in chloroplasts	50
1991	C-insertions	<i>Physarum</i>	Mitochondria	mRNA	First report of C-insertions correcting multiple frameshifts	15
1993	nt-insertion + C → U	<i>Physarum</i>	Mitochondria	mRNA	First report of both nt-insertional editing and C → U editing in one transcript	53
1993	nt-replacements	<i>A. castellanii</i>	Mitochondria	tRNAs	Creation of conserved aminoacyl acceptor stem in tRNAs	16
1994	nt-insertion	<i>Physarum</i>	Mitochondria	rRNA	First report of mixed nt-insertional editing	17
1994	U → C	Mammals	Nucleus	WT1 mRNA	Editing in Wilms tumor susceptibility gene	235
1995	C → U	Marsupial	Mitochondria	tRNAs	C → U in anticodon; change in tRNA identity	18
1996	A → I	<i>Drosophila</i>	Nuclear	4f-rnp mRNA	First report of editing in <i>Drosophila</i>	237
1996	A-insertion	Ebola virus		Glycoprotein	Frameshift generation of C-terminal extension	238
1997	A → I	Mammals	Nucleus	Serotonin 5-HT receptor mRNAs	First report of editing in G-protein-coupled receptor; affects signaling efficiency	61
1997	A → I	Squid (<i>L. peali</i>)	Nuclear	mRNAs	First report of editing in squid neurons	66
1999	A → I	<i>C. elegans</i>	Nuclear	mRNAs	Development of editing system in <i>C. elegans</i>	211
1999	C → U	<i>Dictyostelium</i>	Mitochondrial	rRNA	First report of editing in <i>Dictyostelium</i>	58
1999	U-insertions	<i>Teratocephalus lirellus</i>	Mitochondrial	Cytochrome- <i>b</i> mRNA	First report of insertional RNA editing in metazoan mitochondria	230
2000	G → A, C → U	HIV	—	Viral RNA	Controversial report of hyperediting in HIV	232
2002	U → C	Mouse	Mitochondrial	Chimeric RNA	First report of editing in mammalian mitochondria	231
2002	A → G, C → U, U → C	Dinoflagellates	Mitochondrial	mRNAs	First report of extensive substitutional editing in dinoflagellates	234

as RNA editing (*see* Chapter 19) (20). This linkage between the editing and modification fields was strengthened when two common RNA modifications, 2'-*O*-methylation and the conversion of uridine to pseudouridine (*see* Chapter 20), were found (like trypanosome RNA editing) to be guided by small RNAs (21–23). This phenomenon, first discovered in the modification of RNA(NOi) in the nucleolus of vertebrates, has subsequently been described in a wide number of organisms including protozoans; metazoans; and, more recently, members of the Archaea (24,25). The repertoire of RNAs that undergo guided modifications has also expanded and now includes almost all known classes of RNAs (24,26–29). RNA-guided modifications involve a small nucleolar RNA (snoRNA) and several associated proteins, one of which catalyzes the modification reaction (*see* Chapter 21) (30–32). While the largest groups of snoRNAs direct site-specific modifications, other snoRNAs are known to be involved in pre-rRNA cleavage (32). Base pairing of at least two snoRNAs, U3 and U14, is known to be required for the major pre-rRNA cleavages (33,34). Although it is unclear how these snoRNAs are used during the cleavage events, it has been suggested that U3 may work by presenting the pre-rRNA to the endonuclease in the correct structure for cleavage (32). This is very reminiscent of trypanosome RNA editing in which the guide RNA (gRNA) is also responsible for directing an endonuclease to cleave at a specific editing site. Because of this similarity, it has been proposed that the gRNA-directed editing observed in trypanosomes may have evolved by co-opting components of an early pre-rRNA processing machine (35).

The idea that RNA editing evolved because of its ability to co-opt preexisting RNA modification and processing components is not unprecedented. The recent crystal structure of the yeast mitochondrial transcription factor sc-mtTFB indicates that it is structurally homologous to a large group of DNA/RNA methyltransferases (36). The human orthologs h-mtTFBM1 and h-mtTFBM2 were also identified and shown to share a striking similarity to rRNA dimethyl transferases (37,38). rRNA dimethyl transferases utilize *S*-adenosyl methionine (SAM) as the donor of the methyl group that is transferred to the target RNA, and the crystal structure of sc-mtTFB indicates that many of the amino acid residues that bind SAM are conserved (36). Recently, human mt-TFB1 was shown to have rRNA methyltransferase activity in that it could both bind SAM and methylate a conserved stem loop in bacterial 16S rRNA (39). This conserved site appears to be methylated in the human mitochondrial 13S rRNA; hence, this transcription factor may be a bifunctional protein that plays roles in both mitochondrial transcription initiation and RNA modification. Many examples of different metabolic pathways utilizing domains drafted from RNA modification proteins exist, which is not surprising considering the antiquity of the RNA modification systems (40).

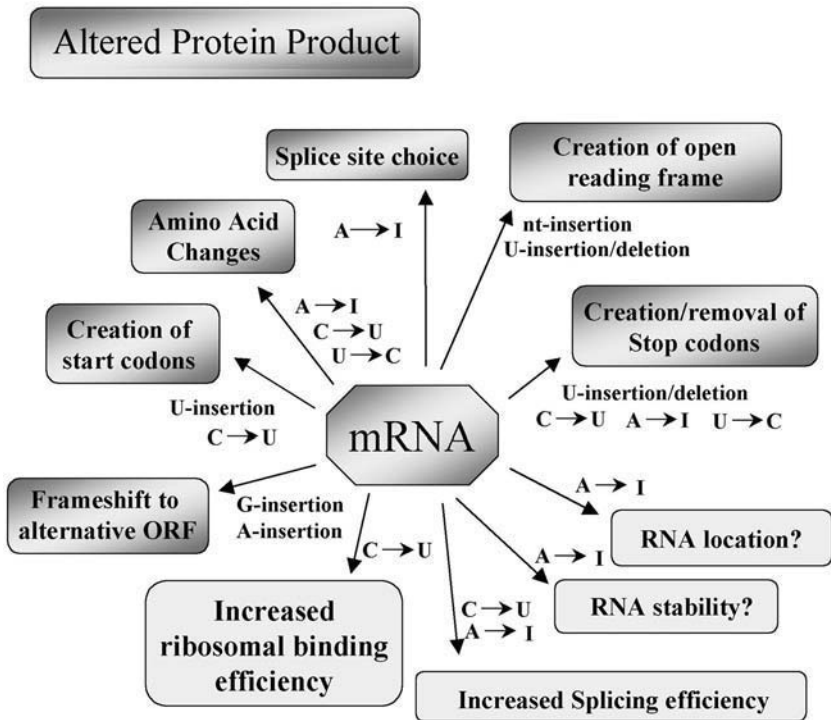
The chapters in this section describe methods used to investigate RNA editing and modification in very different systems that utilize very different mechanisms. Good, up-to-date reviews on each system are available, and, in addition, each chapter gives an introduction to its particulars (41–45). The intention of this chapter is to give the reader, especially new investigators to the field, a historical perspective on some of the pivotal discoveries that helped direct the current avenues of research in the field of RNA editing. For a good historical perspective on RNA modification, see the chapter by Bryon Lanes in **ref. 46**. In following this history, it is clear that the researchers in this field are not linked by any commonality of organism or biochemical mechanism, but by the types of questions that are asked. What RNA transcripts are targeted? How are the editing sites specified? What is the mechanism? Is it regulated? What are the physiological consequences of the editing events? Because of these analogous questions, the chronological developments of the systems tend to parallel one another. RNA substrates have to be identified and the editing characterized. A biological assay that allows manipulation of the system has to be developed. Components of the editing machinery have to be identified and their roles in the editing process determined. Finally, how the editing process interfaces and interacts with other metabolic pathways must be determined.

2. RNA Editing: The Early Years—Defining a New Level of Genetic Control

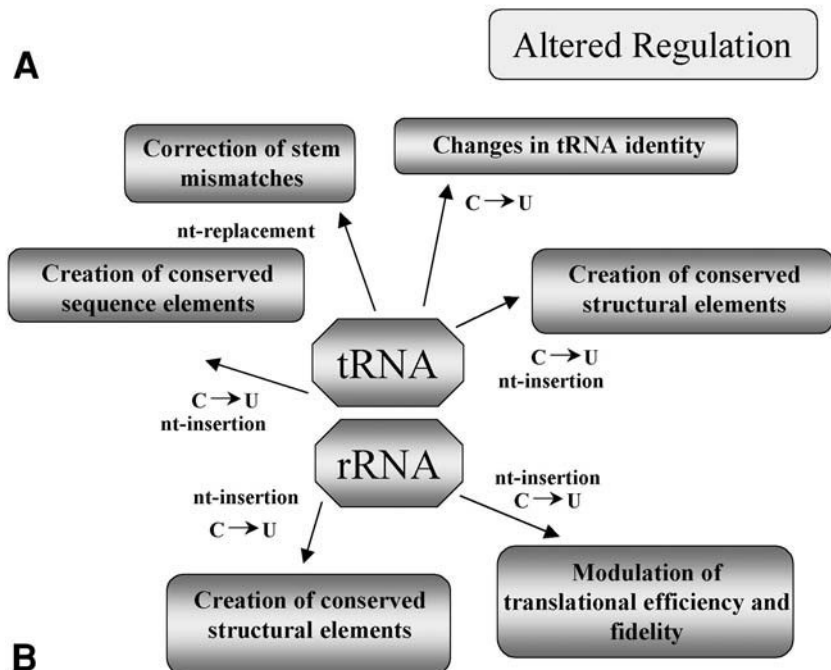
The first years in RNA editing were pivotal in defining the breadth of targets that could be affected, the number of different mechanisms employed, and the functional consequences of the editing events (**Table 1, Fig. 1**).

2.1. Lesson 1: Editing Can Explain Peculiarities Found in Genes and Genomes

With the onset of large-scale DNA sequencing, it became obvious that many genomes, especially organellar genomes, had genetic anomalies that were difficult to explain. The sequencing of plant mitochondrial and chloroplast DNAs indicated either that the predicted protein products of some genes contained variant amino acids at positions completely conserved among other species, or that their codon usage differed from the “universal” genetic code. The initial sequence analyses of the mitochondrial maxicircles of three kinetoplastid species, *Leishmania tarentolae*, *Trypanosoma brucei*, and *Crithidia fasciculata*, revealed several unusual features. These included internal frameshifts in protein-coding regions that were conserved among the three species, the absence of ATG codons for translation initiation for many of the genes, and the absence of some highly conserved electron transport proteins that exist in the mitochondria of all other organisms (47).



A



B

Rob Benne's description of the "editing" of the cytochrome oxidase subunit II (coxII) transcript that corrected a -1 frameshift was the first indication that a mechanism existed that could repair these genetic peculiarities (3). His discovery was soon followed by several reports that stunned biologists. In 1987, Feagin et al. (5) described the addition of 34 uridines at several different sites within the 5' end of apocytochrome-*b* (Cyb). This report was particularly exciting because it reported that the U-insertions not only created a conventional AUG initiation codon, but the insertions appeared to be related to gene regulation. Mitochondrial function is highly regulated in those trypanosomes that cycle through both mammalian and insect hosts (48). Bloodstream forms rely on glycolysis for their energy production and lack cytochromes and Krebs cycle enzymes, while insect forms contain a fully functional mitochondrion. In Cyb, the U-insertions were found only in insect forms, which have a complete mitochondrial respiratory system, but were absent in the bloodstream forms, indicating that the editing process was regulated during the complex life cycle. In a 1988 issue of *Cell*, Shaw et al. (6) described the addition and deletion of multiple uridines into the 5' ends of a number of trypanosome transcripts, regenerating both conserved amino acid sequences and AUG initiation codons. In the same issue, Feagin et al. (4) reported the first example of "pan-editing," involving hundreds of nucleotide changes and the creation of the entire COIII ORF by RNA editing. Defining these editing events in trypanosomes resolved many of the anomalies initially described and began the race to determine whether peculiarities found in other organellar genomes could be resolved by nucleotide changes in the RNA transcripts.

2.2. Lesson 2: RNA Editing Can Have Significant Consequences in Protein Function

The reports of editing in the mitochondrial transcripts in trypanosomes introduced two important concepts: (1) RNA transcripts did not have to be identical copies of the exonic sequence of a gene; and (2) the recoding of an RNA transcript could be regulated, introducing another level where gene expression might be controlled. These concepts were quickly reinforced with the first reports of RNA editing outside of the trypanosome system. The next two systems described the deamination of a C to a U in the mammalian apoB transcript and the insertion of two G residues within the P gene transcript of paramyxoviruses, both of which resulted in the regulated production of two protein

Fig. 1. (see opposite page) (A) Diverse functions of RNA editing in mRNAs; (B) functions of RNA editing in tRNAs and rRNAs.

isoforms (7,8,10). In the mammalian apoB mRNA, researchers found that some cellular transcripts contained a uridine at nucleotide 6666, instead of the encoded cytidine. This C $\rightarrow$ U transition converts a glutamine codon (CAA) to an in-frame stop codon (UAA) and generates apoB-48, a protein isoform of apoB-100 that lacks the binding domain for the low-density lipoprotein receptor. In humans, this editing event is tissue specific occurring only in intestinal cells, and plays a critical role in lipid metabolism (49). In paramyxoviruses, two alternative reading frames can be accessed by the cotranscriptional programmed insertion of G residues (10,11). The production of both proteins is critical for the life cycle of the virus.

2.3. Lesson 3: Editing Is Widespread in Eukaryotes

The initial reports of editing, while affecting a number of evolutionarily diverse organisms, were still considered by many researchers to be an interesting phenomenon of limited relevance. The impact of RNA editing was broadened, however, with the discovery of editing in plants and in additional nuclear transcripts from vertebrates. In 1989, three laboratories independently reported that multiple C $\rightarrow$ U transitions could explain the apparent coding anomalies in plant mitochondria (12–14). Within 2 yr of the initial plant reports, editing was also reported in chloroplasts, and it was clear that editing played a major role in plant organellar biology (50). Although the vast majority of plant editing events involved C $\rightarrow$ U changes in mRNAs, reminiscent of editing in apoB, the reverse reaction, U $\rightarrow$ C conversions, were also observed (51,52).

In 1991, the report of A $\rightarrow$ I editing in mRNAs encoding subunits for glutamate-gated ion channels described yet another unique way that functionally distinct proteins could be produced from a single gene (9). Because inosine pairs preferentially with cytosine, A $\rightarrow$ I editing can result in specific changes in the amino acid coding potential of an mRNA with dramatic effects on protein function. The initial editing event described in the GluR-B subunit caused a single CAG (glutamine) to CIG (arginine) change that tightly controls the Ca²⁺ permeability of the receptor. Hence, editing plays a crucial role in normal function of the central nervous system. This report not only introduced a fourth mechanism by which codons could be recoded, but it firmly established the importance of RNA editing in regulating protein function in vertebrates.

2.4. Lesson 4: Editing Can Affect Structural as Well as Coding RNAs

The next few years saw the impact of RNA editing broaden even further with the description of editing in several different systems that included structural as well as coding RNAs. These include both editing of the mitochondrial tRNA transcripts in *Acanthamoeba castellanii* and the extensive editing of mitochondrial adenosine triphosphate (ATP) synthase mRNAs in *P. poly-*

cephalum (15,16). Editing in the mitochondria of *Acanthamoeba* was the first report of editing limited to a structural gene that had clear biological implications. The mitochondrial DNA of *A. castellanii* had a cluster of five tRNA genes, four of which contained sequences that would result in extensive mismatched base pairs in the acceptor stem of the tRNA transcript. Direct sequencing of the tRNAs indicated that the mismatches were corrected to restore base pairing by nucleotide changes in the 5' half of the stem (16). The RNA editing characterized involved both pyrimidine-to-purine (U $\rightarrow$ A and U $\rightarrow$ G) and purine-to-purine (A $\rightarrow$ G) changes, suggesting that it involved either base or nucleotide replacement, rather than base modification.

Similar to trypanosome mitochondrial transcripts, the *Physarum* ATP synthase mRNAs require the insertion of multiple nucleotides to correct numerous -1 frameshifts and generate a translatable message. However, instead of U-insertions, this RNA is edited at 54 sites by the insertion of cytidine residues. While the initial description of editing in *Physarum* involved an mRNA and only C-insertions, it soon became clear that editing in this organism's mitochondria could involve insertion events involving all four nucleotides and that it affected all three types of transcripts: mRNAs, tRNAs, and rRNAs (17, 53–55). C-insertions into four different tRNAs all increase base pairing in helical regions, and one U-insertion into the mt-tRNA^{Glu} restores highly conserved primary sequence in the T Ψ C loop (55). In the rRNAs, nucleotide insertions along their entire length are required to form important structural elements known to be involved in translational fidelity and efficiency (17). These first reports of editing in tRNAs and rRNAs were soon followed by a report of tRNA editing that did not affect structural elements, but instead changed the identity of the tRNA (18). In marsupials, a C $\rightarrow$ U transition in the anticodon allows the tRNA^{Asp} to be specifically charged with aspartate (18,56). The unedited tRNA^{Asp} is charged with glycine (56). A similar C $\rightarrow$ U transition in the anticodon of an mt-tRNA has also been observed in trypanosomes (57).

2.5. Lesson 5: Editing Can Generate Proteomic Diversity

While examples of editing in noncoding RNAs have slowly accumulated (24,58), the bulk of editing processes affect mRNAs. Regulation of the editing process has the potential to create multiple protein variants, and most of the characterized systems are known to exploit editing to generate multiple protein isoforms. In apoB, the introduction of a stop codon generates a protein variant that lacks an important functional domain (49). Viruses use editing to frameshift to an alternative ORF and produce proteins with different carboxyl ends (59). In kinetoplast editing, mRNA transcripts can have distinct domains that are independently edited and differentially regulated, potentially producing different protein isoforms during the complex life cycle. For example, in *T. brucei*,

the NADH dehydrogenase subunit 7 (ND7) has two distinct editing domains that are separated by a 59-nt unedited region (60). The 5' domain is edited in both bloodstream and insect forms, but the 3' domain is completely edited only in the bloodstream form. Editing of the 5' domain creates an in-frame stop codon that is only removed by complete editing of the 3' domain.

The importance of the ability of RNA editing to increase proteomic diversity was driven home with the characterization of editing in an additional class of neuron-specific RNA transcripts, the 5-HT_{2C}-serotonin receptor (61). This important G-protein-coupled receptor was found to undergo A → I editing at five different nucleotide positions and allows the production of at least 12 different protein isoforms. Different brain regions were shown to express certain edited RNAs, suggesting the tissue or cell-specific regulation of specific editing combinations (61). Editing of the 5-HT_{2C}-serotonin receptor is thought to modulate receptor:G-protein coupling and may affect a variety of physiological processes (62,63). The subsequent identification of A → I editing in important genes found in the nervous systems of *Drosophila* and squid indicate that RNA editing may be essential for neuronal function in a broad array of organisms (64–66). The nervous system requires a high degree of molecular complexity, and with the completion of the human genome sequence, it is clear that mechanisms must exist that will generate a substantially larger number of protein functions than the available genes (67–69). While alternative splicing is used extensively as a way to increase proteomic diversity in the nervous system, RNA editing adds a combinatorial aspect to the generation of protein variants that is mind-boggling.

2.6. Lesson 6: Misediting Can Have Severe Physiological Consequences

Partially edited and misedited transcripts were detected early in the characterization of editing in the kinetoplastid and plant systems. In *T. brucei*, the extensively edited mRNAs are characterized by large populations of heterogeneously sized transcripts (4,60,70). The largest transcripts are the size of fully edited RNAs, and the smaller transcripts correspond to unedited and presumably partially edited RNAs. These transcripts all have the same general features: they are edited in the 3' region and unedited in the 5' region. Most of these mRNAs have fully edited sequence that transitions to unedited sequence through a junction region that has clearly undergone editing events but does not match the mature sequence (71–73). Whether these partially edited molecules were editing intermediates or “mistakes” generated considerable debate, especially when legitimate errors were identified (71,72,74). However, because full editing at the 5' end is required for formation of the AUG initiation codon, there do not appear to be any real physiological consequences of “misediting.”

In plants, the situation is more complex because editing proceeds without a strong directional bias (75–77). Incompletely edited transcripts have been characterized, and since editing is infrequently required to create a start codon, these can be templates for translation (78). While translation of partially edited transcripts would result in nonfunctional protein, no mutant phenotype appears to be associated with the translation of these partially edited mRNAs. However, transgenic plants engineered to express unedited ATP 9 protein in the cytoplasm, with subsequent import into the mitochondria, can induce male sterility (79,80). This suggests that synthesis of proteins from unedited or partially edited mRNAs may not always be without functional consequences.

In contrast to the organellar systems, misediting in nuclear encoded mRNAs is associated with a number of disease phenotypes, and reports have accumulated of both aberrant editing in new, previously unknown substrates, and altered editing in known substrates with detrimental physiological consequences. In 1996, Skuse and colleagues (81,82) found that human neurofibromatosis 1 (NF1) mRNA can undergo a C → U modification that introduces a premature stop codon. Neurofibromatosis type I is a dominantly inherited disease that predisposes affected individuals to various forms of neoplasia (83). The NF1 gene encodes neurofibromin, a tumor suppressor that functions as a GTPase-activating protein (84). Introduction of a premature stop codon truncates the protein and results in loss of tumor suppressor function (81). Analyses of peripheral nerve sheath tumor samples from patients with NF1 indicate that ~20% of the tumors demonstrated 3–12% C → U editing in NF1 RNA (85). These tumors were found to express the catalytic deaminase responsible for editing apoB RNA, suggesting that aberrant editing by this enzyme may account for some of the phenotypic variability associated with this disease.

An aberrant editing event was also shown to affect the function of hematopoietic cell phosphatase (PTPN6), an important modulator of myeloid cell signaling (86). PTPN6 is a tyrosine phosphatase that downregulates a broad spectrum of growth receptors, and mutations in this gene are known to have oncogenic potential in mice (87–89). An investigation of PTPN6 mRNA in CD34⁺/CD117⁺ blasts isolated from acute myeloid leukemia patients found a significant increase in the level of transcripts that had retained an intron in the SH2 domain. Sequence analyses indicated that these transcripts had multiple A → I conversions, with the main editing event involving A⁷⁸⁶⁶, the putative branch site of the retained intron (86). The level of the intron-retaining splice variant was significantly increased over levels in normal bone marrow mononuclear cells, suggesting that aberrant editing in PTPN6 may be involved in leukemogenesis.

Recently, three different groups have reported the altered editing of known A → I substrates in patients who manifest specific disease phenotypes. In patients with amyotrophic lateral sclerosis (ALS), a progressive neurodegener-

ative disease, a significant reduction in A $\rightarrow$ I editing was shown at the critical Q/R site in GluR2 mRNA (90). This reduction in editing was selective to the spinal ventral gray tissue of ALS patients and appears to be tightly linked to the etiology of ALS. A reduction in editing at the Q/R site in GluR2 was also observed in tissues from malignant human brain tumors (91). In these gliomas, alterations in editing were also observed in other known editing substrates and were correlated with a decrease in the activity of the editing enzyme adenosine deaminase that acts on RNA 2 (ADAR2). These findings suggest that loss of editing activity may play a role in tumor progression and provide a model explaining the occurrence of epileptic seizures associated with malignant gliomas. In the third report, Gurevich et al. (92) examined the expression of 5-HT_{2C} mRNA (serotonin 2C receptor) in the prefrontal cortex of depressed individuals who had attempted suicide. They discovered that the pattern of editing found in the most abundant mRNAs was significantly different from the pattern found in healthy control subjects. The observed alterations in editing site preferences were opposite those observed in mice treated with the antidepressant drug fluoxetine (Prozac), suggesting that aberrant editing of the serotonin 2C receptor plays a significant role in depression.

3. RNA Editing: The Middle Years—Defining *cis*-Acting Sequences and *trans*-Acting Factors

Characterization of RNA editing reactions and elucidation of the underlying biochemical mechanisms required the development of in vitro editing systems that allow the manipulation of both the RNA and protein components. In the mammalian apoB C $\rightarrow$ U and A $\rightarrow$ I systems, the availability of genetic tools and the early development of in vitro assays allowed rapid advances in the understanding of the regulatory sequences and cellular factors that mediate these editing events. For the organellar systems, in vitro assays proved more difficult to develop; hence, characterizations of these systems lag somewhat behind. Because of the wide evolutionary diversity of the editing mechanisms, it is difficult to draw analogies among the different systems. However, one reason RNA editing is so remarkable is the level of precision and specificity that is observed across the different editing systems. For all of the described systems, the editing machinery must recognize the correct RNA and then precisely alter specific editing sites. Therefore, I concentrate here on some of the early discoveries that helped elucidate our understanding of the *cis*-elements that are involved in editing site selection.

3.1. Lesson 7: Editing Site Recognition Involves Both Primary Sequence and Structural Elements

For some of the systems, sequence elements surrounding the editing site gave valuable initial clues as to the mechanism of site selection. In apoB, the

sequence just 3' to the edited C was highly conserved in a number of mammalian species, suggesting that this region contained a *cis*-element important for editing site recognition. The early development of an *in vitro* system allowed researchers to quickly show that this conserved sequence, now called the "mooring sequence," is critical for apoB editing (93–95). An assay developed by Driscoll et al. (93) utilized a ³²P-end-labeled 35-nt primer complementary to sequence located just downstream of the target C and primer extension with reverse transcriptase in the presence of high concentrations of dideoxy-GTP. Extension of apoB RNAs that were unedited would stop at the first upstream C (the editing site). Transcripts that had undergone a C → U editing event would terminate at the second C in the synthetic substrate, generating a product 5 nt longer. Using this assay, Driscoll et al. (93) demonstrated that an S100 extract from McArdle 7777 cells, a cell line that produced both apoB-100 and apoB-48, could specifically edit a synthetic apoB transcript. Similar S100 extracts prepared from two other cell lines that did not produce apoB-48 did not contain any editing activity.

This first RNA-editing *in vitro* assay was simple, sensitive, and quantitative, setting the benchmark for *in vitro* assays in the other systems. It allowed a number of studies that refined researchers' understanding of the elements that provide target specificity in apoB. While a basal level of editing requires only that the 11-nt mooring sequence be located 4 nt (the spacer element) downstream of the edited C, additional elements located both 5' and 3' are necessary for efficient editing (94–97). Recently, studies have also indicated that both secondary structure and location within the transcript may also influence the editing process (98–100). Two structural models of sequences surrounding the edited C have been proposed, both involving formation of a conserved stem loop with the targeted cytidine exposed. In one model, the mooring sequence is paired with the upstream 5' regulator region, exposing the target C in a single-stranded loop (Fig. 2A) (99). In the second model, the mooring sequence is paired with the distal 3'-efficiency element with the target C exposed just upstream of the stem (Fig. 2B) (98). Both models suggest that the more distal elements may work by conferring an optimal structure on the editing domain. Editing efficiency is also influenced by the proximity of intron splice sites, which may be the reason that the target C⁶⁶⁶ is located within the middle of an unusually large exon (7.5 kb) (100–102). All of these characteristics combine to greatly increase the specificity of apoB editing and decrease the possibility of misediting, which can have significant physiological consequences.

In kinetoplastid U-insertion/deletion editing, the ability to precisely insert (or delete) thousands of U residues at hundreds of different editing sites was an incredible enigma. After the initial reports of relatively modest editing (four inserted Us to correct a frameshift, minimum U-insertion at the 5' end to gen-

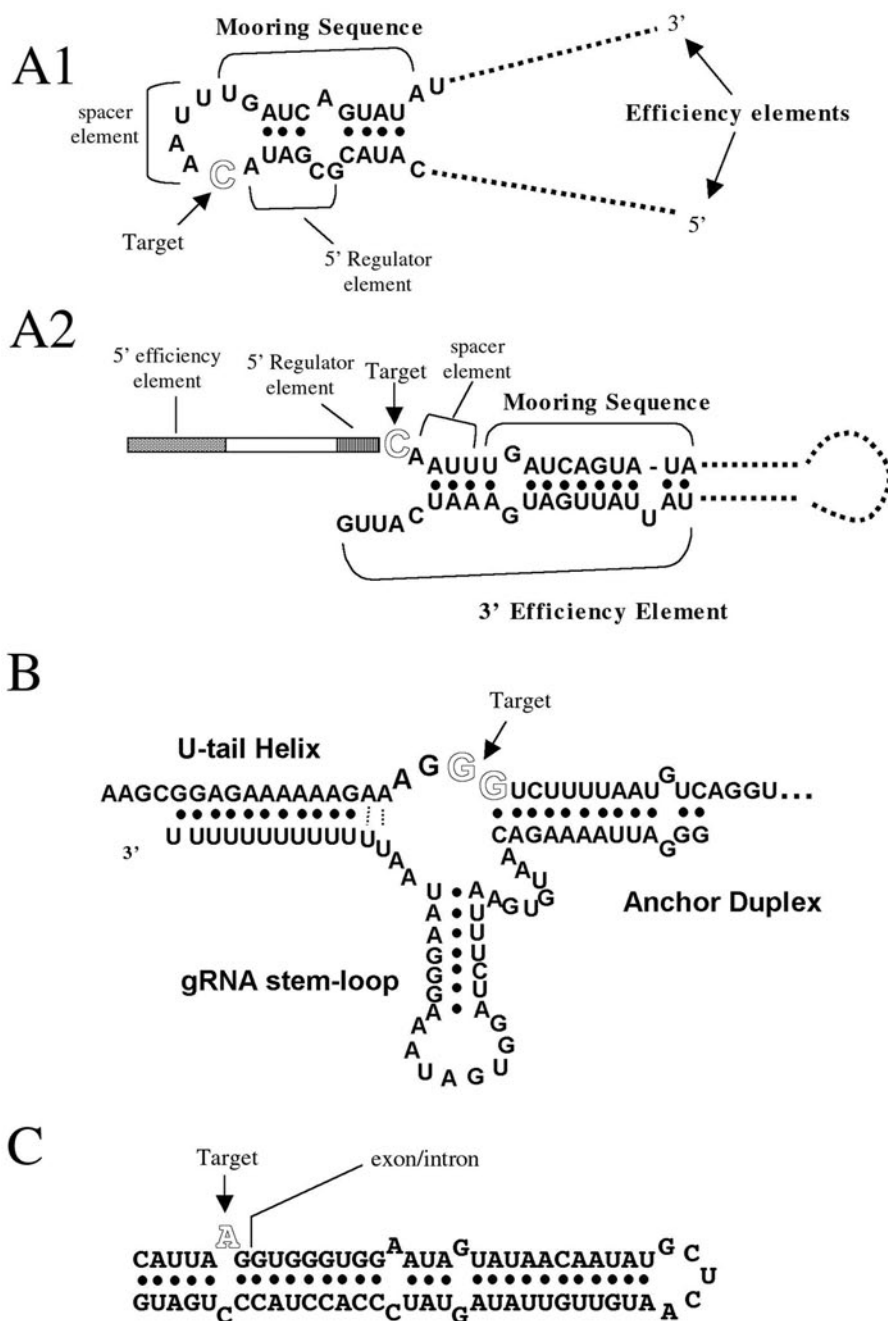


Fig. 2. Sequence and structural elements important for editing site recognition. (A) Model of important recognition elements for C $\rightarrow$ U editing in apolipoprotein B. (B) Alternative structural model for apoB editing. The tripartite sequence motifs (mooring sequence, spacer element, and 5' regulator element) and 5' and 3' efficiency elements are indicated, along with the proposed secondary structures. (C) Secondary structure

erate initiation codons), it was suggested that primary sequence or secondary structure might specify the sites of nucleotide insertion or deletion. However, the description of the insertion of hundreds of nucleotides into COIII made such a mechanism seem highly unlikely. In 1990, researchers breathed a sigh of relief when Blum et al. (103) reported their discovery of a population of small RNAs that contained the information needed to edit the mRNAs. This discovery was pivotal in the understanding of editing in this system. The gRNA characteristics—a 5' end with sequence that can base pair just downstream of a preedited region, a guiding region that is complementary (including G:U base pairs) to the mature sequence, and a nonencoded 3' oligo(U) tail—gave a myriad of clues to the editing process, allowing Blum et al. (103) to propose the first (surprisingly precognitive) model for gRNA-directed RNA editing. In their model, the 5' end of the gRNA forms a duplex with the unedited mRNA precursor (the anchor duplex). The first mismatch upstream from this anchor region could then be targeted for a cascade of enzymatic reactions, including an endonucleolytic cut, a terminal uridylyate addition by a terminal uridylyate transferase, and finally religation of the edited site (103). The 3' U-tail was proposed to be involved in stabilizing the interaction by forming an upstream (3' anchor) (Fig. 2C) (104). Complete editing by one gRNA would create the anchor sequence for the next gRNA and would lead to the overall 3' to 5' polarity of editing observed in the partially edited molecules (105). The presence of a U-tail on the gRNAs, however, immediately suggested a second model for RNA editing, an RNA-driven transesterification (106,107). Using reverse transcriptase and the polymerase chain reaction, Blum et al. (107) detected chimeric molecules of gRNAs covalently linked via their 3' oligo(U) tail to editing sites within their cognate mRNAs. They proposed that these chimeras were editing intermediates and that successive transesterifications result in the transfer of uridine residues from the gRNA 3' oligo(U) tail to the editing site (107). This model led to the development of several in vitro assays based on detection of the chimeric intermediate (108,109).

Four years after the discovery of gRNAs, an in vitro editing assay that allowed a full round of editing was finally developed (see Chapter 13) (110). Seiwert et al. (110) had observed that ATPase6 pre-mRNAs involved in in vitro chimera formation with gRNA (gA6-14) had undergone editing at the first pre-

Fig. 2. (continued) prediction for interaction of gRNA with mRNA in trypanosome U-insertion/deletion editing. The gRNA is shown below the mRNA with the three helices defining the editing site indicated. (D) Secondary structure prediction for portion of GluR-B pre-mRNA. The targeted A is shown in outline font, and the position of the exon/intron boundary is indicated.

dicted editing site. Using this substrate pair and a poison primer-extension assay similar to that developed for apo-B editing, they were able to show that the gRNA was able to direct editing of the mRNA. Fine-tuning of this assay allowed the direct testing of both models, leading to the acceptance of the enzyme cascade pathway as the mechanism of kinetoplastid RNA editing (*111,112*). A number of studies have now investigated the sequence and structural requirements for optimal gRNA directed editing (*113–119*). What is interesting is that in *T. brucei*, the A6 + gA6-14 pair originally used in developing the in vitro assay is the only mRNA/gRNA pair that works. It is unclear why the A6 sequence is optimal for in vitro editing; however, initial experiments indicate that mRNA secondary structure may limit gRNA accessibility in vitro (unpublished results).

In both the trypanosome and apoB systems, editing is dependent on primary sequences that flank the editing site. In contrast, the subunit-specific A $\rightarrow$ I editing observed in glutamate receptor transcripts indicated that site selection could not be dependent on immediate flanking sequence (*9*). The AMPA receptor units GluR-A to GluR-D are closely related and have nearly identical nucleotide sequences surrounding the identified editing site in GluR-B. While these subunits are often coexpressed in the same neuronal populations, only GluR-B undergoes editing, and this subunit specificity cannot be explained by either the presence of a *cis*-element or by a gRNA. Sommer et al. (*9*) suggested that editing site selection therefore either was dependent on secondary structure or involved unique intronic sequences. Two years later, Higuchi et al. (*120*) showed that the proximal part of the intron immediately downstream of the editing site was required for Q/R editing. This section of the intron was part of an imperfect inverted repeat with complementarity to the exon centered on the unedited codon (**Fig. 2D**). Because of the requirement for an extended RNA duplex, these investigators suggested that editing of GluR-B involved an ubiquitous double-stranded RNA (dsRNA) adenosine deaminase that had been originally identified and characterized as a dsRNA unwinding activity (called dsRAD or DRADA, now known as ADAR1) (*121–125*). Several studies that had been conducted to characterize the substrate preferences of this enzyme suggested that it might be responsible for editing of the glutamate receptor subunits, but no in vivo substrate had been proven (*125,126*). In 1994, three different laboratories described the purification of the dsRNA adenosine deaminase (*127–129*), which allowed the development of in vitro assays utilizing recombinant protein (see Chapter 11) (*130–132*). Using this type of assay, Maas et al. (*132*) demonstrated that purified recombinant ADAR1 was capable of specifically editing adenosines in GluR pre-mRNAs. Surprisingly, it showed considerable substrate selectivity for certain editing sites in that it could edit the R/G site of GluR-B and the Q/R site of GluR6, but not the Q/R site of

GluR-B. By comparing the different structures for the different editing sites, it was clear that the structural environment must be one determinant for editing site selection and that the mismatches, bulges, and loops that interrupt the helical structure are important for selectivity. How the enzyme utilizes these structural disruptions to orient its selection is unclear.

The editing found in plant organelles is similar to the trypanosome system in that editing involves directed changes at hundreds of different sites (133,134). Examination of the sequences immediately surrounding the editing targets did not reveal any obvious shared sequence or secondary structure, and by analogy to trypanosome editing, it seemed reasonable that small gRNAs could also be involved in editing site selection. However, despite considerable effort, gRNAs have not been identified. The development of an in vitro system for plant mitochondrial editing has proven to be notoriously difficult. With no mitochondrial gene transformation system in place, researchers had to rely on naturally occurring recombination events that resulted in the insertion of gene fragments in aberrant locations. A comparison of the editing patterns observed in the transcribed pseudogenes with editing in the intact gene copy suggested that important *cis*-elements reside upstream (5') of the editing site with the amount of upstream sequence required dependent on the editing site analyzed (135–140).

In plant chloroplasts, the availability of chloroplast transformation techniques allowed a more extensive and systematic analysis of the *cis*-elements involved in target recognition (see Chapter 18) (134,141). The results of several studies suggest that the major *cis*-acting elements are similar to those that act in the mitochondria in that they reside 5' to the edited site (142–145). Both the identity of the nucleotides directly adjacent to the editing site and the distance from the upstream *cis*-element also appear to be critical to the editing process (142,144–146). The observation that high-level expression of a transgenic RNA could specifically inhibit editing of the endogenous site, but not affect editing at other sites, suggests that site-specific *trans*-acting factors exist for each individual editing site (143). The lack of any evidence for RNA factors and the incredibility of the idea that a different protein was needed for every editing site led Chateigner-Boutin and Hanson (147) to determine whether overexpression of a transgene containing one editing site had any effect on any of the other known sites (31 known tobacco chloroplast editing sites). Their results indicate that the important *cis*-elements can be grouped into clusters with conserved nucleotides located upstream (within 30 nt) of the target. In addition, they were able to detect sequences 5' of several mitochondrial editing sites that had similar sequence elements, suggesting that editing in the two organellar compartments may share some *trans*-acting proteins.

Recently, two methodologies have been developed that should allow considerable progress in our understanding of plant organellar editing: a novel

mitochondrial electroporation technique and the development of an in vitro chloroplast editing assay (*see* Chapter 17) (**148,149**). The ability to transfect isolated mitochondria will allow a more systematic analysis of the target requirements for mitochondrial editing. Using this technique, one editing site (COII) has been investigated in detail indicating that only sequences immediately flanking the editing site, 16 nt upstream and 6 nt located immediately downstream, were essential in editing of the target C (**150**). This is surprisingly similar to the *cis*-element requirements for editing of the chloroplast *psbL* where a 16-nt element upstream and a 9-nt element downstream were essential for target recognition (**142,149**). Using the new chloroplast in vitro assay, the *cis*-elements for two different chloroplast editing sites were examined in detail (**151**). In these elegant experiments, a series of competitor RNAs was constructed by scanning mutagenesis in which each successive 5-nt block in the upstream -40 to -1 region was substituted with its complementary nucleotide sequence. These experiments indicate that a 10-nt sequence from -15 to -6 is essential for editing of *psbE*. For *petB*, the essential element was longer and located between -20 and -6. An examination of editing in the mutated transcripts directly indicated that while nucleotides in the -1 to -5 region are not critical for binding of the transacting factor, they are critical for the editing event.

In other systems, it has been proven even more difficult to understand the signals involved in directing the editing process. In both the paramyxoviruses and *Physarum*, the insertion of nucleotides at specific editing sites occurs cotranscriptionally, suggesting that aspects of the DNA template as well as *cis*-elements in the RNA may play a role in site selection. Editing in the paramyxoviruses is fairly simple, with only G nucleotides inserted at a single site within a short run of guanylates (**10,11,152**). This is suggestive of a polymerase stuttering mechanism (**153,154**) that involves the reiterative copying of the same template base, and good evidence exists that the insertions occur by such a pseudotemplated transcription process (for reviews *see* **refs. 59** and **155**). In contrast, RNA editing in *Physarum* is quite complex. Single nucleotides or dinucleotides are added to approx 1000 different sites in mRNAs, the rRNAs, and some tRNAs. The nucleotide insertions are accurate, efficient, and tightly coupled to RNA synthesis (**156–158**). However, there is no evidence for a pseudotemplated (polymerase stuttering) mechanism, and it is unclear where the information that determines editing site selection resides (**159–162**).

4. RNA Editing Today—Realizing the Complexities

Today, the different editing systems are all at different stages in their characterization. Most progress has been made in the nuclear C → U and A → I editing systems, for which the proteins necessary and sufficient for editing have

been identified and full editing has been reconstituted in vitro (*132,163,164*). The availability of strong genetic tools has allowed researchers to investigate gene function using both gain- and loss-of-function strategies and to begin to decipher some of the regulatory mechanisms that control the editing process (for reviews, see refs. *41,46*, and *165–168*). In addition, the availability of genome databases has allowed the identification of homologs of the editing catalytic enzymes with surprising results.

4.1. Lesson 8: Editing Can Play Numerous Roles in the Cellular Function of Eukaryotic Organisms

Four years after development of the in vitro assay, the enzyme responsible for editing apoB mRNA was finally identified (*169,170*). Apobec1 (apoB mRNA editing enzyme catalytic polypeptide 1) is a zinc-dependent cytidine deaminase that requires at least 1 auxiliary protein (*170,171*). While the catalytic component was first cloned using rat cDNA, homologs were quickly identified and characterized in rabbits, humans, and mice (*172–175*) (for reviews, see refs. *176–178*). This allowed the targeted deletion of the apobec-1 gene, and it was definitively shown that this enzyme was required for the C – U editing of apoB and the production of apoB-48 (*179–181*). Reconstitution of editing in vitro, however, required identification and cloning of the auxiliary factor(s) necessary for activity. Editing of apoB mRNA was known to be catalyzed by a multiprotein complex; the number of proteins, however, as well as their identity and function, was controversial. The low abundance and instability of the complexes made isolation of the necessary auxiliary factor technically difficult, and only recently has one protein, apobec-1 complementation factor (ACF), been identified and shown to be both necessary and sufficient for apoB-editing activity (see Chapter 12) (*163,164*). While ACF is sufficient for editing, a number of proteins have now been identified that interact with ACF and apobec-1 to modulate activity (*182–185*).

Loss-of-function studies confirmed the role of apobec-1 in the generation of apoB-48; since the mice showed no other phenotype, they did not generate any clues to possible auxiliary roles. Over-expression of apobec-1, however, resulted in hepatic dysplasia and hepatocellular carcinoma, thought to be caused by promiscuous editing of cytidines (*101,186,187*). This hypothesis has recently come under question with the identification of an apobec-1 homolog involved in generating antibody diversity (*188*). The identification of activation-induced cytidine deaminase (AID) generated considerable excitement when it was found to be the master gene involved in controlling class switch recombination, somatic hypermutation, and gene conversion, all modifications of vertebrate immunoglobulin (Ig) genes involved in generating Ig diversity (*189–195*). How AID is involved in these processes is unclear: while it has the ability to deam-

inate cytidines in vitro, it is not known whether AID has editing activity in vivo. However, the expression of AID in *E. coli* generates a biased mutator phenotype, suggesting that AID may work by causing dC → dU deaminations in the DNA (196). Recently, it was demonstrated that apobec-1, as well as two other identified apobec homologs, can also act as DNA mutators in *E. coli* (196,197). The three apobec family members, apobec-1, apobec-3C, and apobec-3G, show surprising local target sequence specificity in their dC deaminase activity (197,198). This opens up the possibility that unregulated expression of apobec-1 could result in a DNA-targeted activity that plays a role in cancer. Defining the physiological roles of the newly identified apobec homologs should prove to be very interesting (184,199). The sites of expression of the new apobec family members indicate specific, if unknown, functions in different tissues (198). Several lines of evidence link tumor formation with apobec expression, suggesting that the enzymes may play a role in growth or cell-cycle control (85,198). In addition, a protein with significant homology to apobec-1, CEM15, appears to be a significant component of an innate antiviral phenotype present in human T-lymphocytes (200). Taken together, these data suggest that cytidine deaminases can play several different roles in cell physiology.

Characterization of the ADARs has paralleled the apoB system. ADAR1, the second editing enzyme cloned, is now known to belong to a family of adenosine deaminases that have been characterized as to their substrate and editing site selectivity, tissue specificity, and cellular location (for reviews, *see refs. 165 and 201*). ADARs appear to be ubiquitous in the metazoa, and enzyme orthologues have been cloned in a number of genetically malleable organisms including *Drosophila* and *C. elegans* (202,203).

While ADAR knockouts are viable in *Drosophila* and *C. elegans*, the transgenic animals are not normal and it is clear that the ADARs play an important role in the nervous system (203,204). By contrast, ADARs appear to be absolutely essential in mammals and appear to play a wider role in development. In mice, ADAR1 is required for embryonic erythropoiesis, and knocking out even a single allele leads to embryonic death by d 14 (205,206). ADAR2 heterozygote mice with one functional null allele are viable; homozygotes, however, suffer repeated episodes of epileptic seizures and die shortly after birth (205). These studies are complemented by investigations into the functional significance of each editing event, using mice that are engineered to contain specific editing-incompetent alleles (207–209).

Similar to the C → U editing story, the ADARs appear to play a number of different roles in cellular function. The embryonic lethality of the ADAR1 null mutant heterozygotes suggests that important physiological targets of ADAR1 still need to be identified. Indeed, the amount of inosine residues detected in different mammalian tissues suggests that A → I editing may play a more

prominent role in regulating gene expression than previously thought (210). Recently, Bass and colleagues (211–213) have developed a method for detecting inosine-containing RNAs (described in Chapter 10) and have identified a number of new potential ADAR targets in both *C. elegans* and human brain tissue. Interestingly, in all of the new targets identified (10 in *C. elegans* and 19 in human brain tissue), the editing sites are located in noncoding regions including 3′ untranslated regions, introns, and a noncoding RNA (213). This suggests that A → I editing may influence RNA stability, transport, or translation of an mRNA. In addition, the regulated expression of ADAR1 behind an interferon-inducible promoter suggests that A → I editing also plays a role in the interferon-mediated antiviral response (165,214). A → I modifications have been found in numerous viral RNAs, with both specific editing (hepatitis delta virus) and extensive editing, referred to as hypermutation, reported (for review, see refs. 214,215, and 236). Identification of a ribonuclease specific for inosine-containing RNAs suggests that ADAR1 may be able to target viral RNAs for destruction (216).

While characterization of the organellar editing systems lags behind, these systems are poised to make significant progress in the next few years. In the trypanosome system, the recent development of molecular genetic techniques and the genome-sequencing projects have combined to allow very rapid progress both in identifying the protein subunits involved in the editing process and in determining their roles by functional gene knockouts (for reviews, see refs. 43–45 and 217). Currently, more than 20 proteins associated with the editing complex (editosome) have been identified, and significant progress is being made in assigning functions to the identified complex proteins (218–222) (see also Chapter 14). In addition, a number of editing-associated proteins have been identified with activities important for resolving and promoting RNA/RNA interactions (223–226). In plant chloroplasts, the very recent development of a sensitive in vitro assay has allowed the first identifications of proteins associated with editing at specific sites (151) (see also Chapter 17). The availability of genome sequence and strong genetic tools should help facilitate the subsequent cloning and characterization of the editing proteins in this system. Many of the other editing systems, including tRNA editing in *Acanthamoeba* (see Chapter 16), mitochondrial editing in *Physarum* (see Chapter 15), and viral RNA editing in the paramyxoviruses, have all made significant strides that increase our understanding of the mechanisms involved in the editing process (59,161,227).

While it is clear that much progress has been made in our understanding of the mechanisms and roles of RNA editing, it is also evident that there is still much to be discovered. Identifying the components involved in the editing process only marks the beginning of our understanding of this important level

of genetic control. Crucial questions concerning the developmental regulation of editing and how the editing process interacts with and interfaces with other RNA processes still need to be answered. In addition, new RNA editing (or putative RNA editing) events, some involving new mechanisms as well as new substrates, are still being described (58,228–234). Each report adds to the breadth of examples of the diverse roles RNA editing plays in cellular function, and each report moves the process of RNA editing firmly into the realm of mainstream RNA-processing events.

References

- 1 Vickerman, K. (1994) The evolutionary expansion of the trypanosomatid flagellates. *Int. J. Parasitol.* **24**, 1317–1331.
- 2 Donelson, J. E., Gardner, M. J., and El-Sayed, N. M. (1999) More surprises from Kinetoplastida. *Proc. Natl. Acad. Sci. USA* **96**, 2579–2581.
- 3 Benne, R., Van den Burg, J., Brakenhoff, J. P., Sloof, P., Van Boom, J. H., and Tromp, M. C. (1986) Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell* **46**, 819–826.
- 4 Feagin, J. E., Abraham, J. M., and Stuart, K. (1988) Extensive editing of the cytochrome c oxidase III transcript in *Trypanosoma brucei*. *Cell* **53**, 413–422.
- 5 Feagin, J. E., Jasmer, D., and Stuart, K. (1987) Developmentally regulated addition of nucleotides within apocytochrome b transcripts in *Trypanosoma brucei*. *Cell* **49**, 337–345.
- 6 Shaw, J. M., Feagin, J. E., Stuart, K., and Simpson, L. (1988) Editing of kinetoplastid mitochondrial mRNAs by uridine addition and deletion generates conserved amino acid sequences and AUG initiation codons. *Cell* **53**, 401–411.
- 7 Chen, S. H., Habib, G., Yang, C. Y., Gu, Z. W., Lee, B. R., Weng, S. A., Silberman, S. R., Cai, S. J., Deslypere, J. P., and Rosseneu, M. (1987) Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in-frame stop codon. *Science* **238**, 363–366.
- 10 Powell, L. M., Wallis, S. C., Pease, R. J., Edwards, Y. H., Knott, T. J., and Scott, J. (1987) A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. *Cell* **50**, 831–840.
- 9 Sommer, B., Kohler, M., Sprengel, R., and Seeburg, P. H. (1991) RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. *Cell* **67**, 111–119.
- 10 Thomas, S. M., Lamb, R. A., and Paterson, R. G. (1988) Two mRNAs that differ by two non-templated nucleotides encode the amino coterminal proteins P and V of the paramyxovirus SV5. *Cell* **54**, 891–902.
- 11 Cattaneo, R., Kaelin, K., Baczko, K., and Billeter, M. A. (1989) Measles virus editing provides an additional cysteine-rich protein. *Cell* **56**, 759–764.

- 12 Gualberto, J. M., Lamattina, L., Bonnard, G., Weil, J.-H., and Grienemberger, J.-M. (1989) RNA editing in wheat mitochondria results in the conservation of protein sequences. *Nature* **341**, 660–662.
- 13 Covello, P. S. and Gray, M. W. (1989) RNA editing in plant mitochondria. *Nature* **341**, 662–666.
- 14 Hiesel, R., Wissinger, B., Schuster, W., and Brennicke, A. (1989) RNA editing in plant mitochondria. *Science* **246**, 1632–1634.
- 15 Mahendran, R., Spottswood, M. R., and Miller, D. L. (1991) RNA editing by cytosine insertion in mitochondria of *Physarum polycephalum*. *Nature* **349**, 434–438.
- 16 Lonergan, K. M. and Gray, M. W. (1993) Editing of transfer RNAs in *Acanthamoeba castellanii* mitochondria. *Science* **259**, 812–816.
- 17 Mahendran, R., Spottswood, M. S., Ghate, A., Ling, M.-L., Jeng, K., and Miller, D. L. (1994) Editing of the mitochondrial small subunit rRNA in *Physarum polycephalum*. *EMBO J.* **13**, 232–240.
- 18 Morl, M., Dorner, M., and Paabo, S. (1995) C to U editing and modifications during the maturation of the mitochondrial tRNA^{Asp} in marsupials. *Nucleic Acids Res.* **23**, 3380–3384.
- 19 Gray, M. W. and Covello, P. S. (1993) RNA editing in plant mitochondria and chloroplasts. *FASEB J.* **7**, 64–71.
20. Grosjean, H. (1996) *EMBO Workshop on RNA Editing*, Maastricht, Netherlands.
- 21 Kiss-Laszlo, Z., Henry, Y., Bachellerie, J.-P., Caizergues-Ferrer, M., and Kiss, T. (1996) Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell* **85**, 1077–1088.
- 22 Tycowski, K. T., Smith, C. M., Shu, M. D., and Steitz, J. A. (1996) A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in *Xenopus*. *Proc. Natl. Acad. Sci. USA* **93**, 14,480–14,485.
- 22 Ni, J., Tien, A. L., and Fournier, M. J. (1997) Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell* **89**, 565–573.
- 27 Ben-Shlomo, H., Levitan, A., Shay, N. E., Goncharov, I., and Michaeli, S. (1999) RNA editing associated with the generation of two distinct conformations of the trypanosomatid *Leptomonas collosoma* 7SL RNA. *J. Biol. Chem.* **274**, 25,642–25,650.
- 25 Omer, A. D., Lowe, T. M., Russell, A. G., Ebhardt, H., Eddy, S. R., and Dennis, P. P. (2000) Homologs of small nucleolar RNAs in Archaea. *Science* **288**, 517–522.
- 26 Tycowski, K. T., You, Z.-H., Graham, P. J., and Steitz, J. A. (1998) Modification of U6 spliceosomal RNA is guided by other small RNAs. *Mol. Cell* **2**, 629–638.
- 27 Cavaille, J., Buiting, K., Kieffmann, M., Lalande, M., Brannan, C. I., Horsthemke, B., Bachellerie, J.-P., Brosius, J., and Huttenhofer, A. (2000) Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. *Proc. Natl. Acad. Sci. USA* **97**, 14,311–14,316.
- 28 Filipowicz, W. (2000) Imprinted expression of small nucleolar RNAs in brain: time for RNomics. *Proc. Natl. Acad. Sci. USA* **97**, 14,035–14,037.

- 29 Darzacq, X., Jady, B. E., Verheggen, C., Kiss, A. M., Bertrand, E., and Kiss, T. (2002) Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. *EMBO J.* **21**, 2746–2756.
- 30 Kiss, T. (2002) Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell* **109**, 145–148.
- 31 Decatur, W. A. and Fournier, M. J. (2003) RNA-guided nucleotide modification of ribosomal and other RNAs. *J. Biol. Chem.* **278**, 695–698.
- 32 Tollervey, D. and Kiss, T. (1997) Function and synthesis of small nucleolar RNAs. *Curr. Opin. Cell Biol.* **9**, 337–342.
- 33 Beltrame, M. and Tollervey, D. (1995) Base-pairing between U3 and the pre-ribosomal RNA is required for 18S rRNA synthesis. *EMBO J.* **14**, 4350–4356.
- 34 Liang, W.-Q. and Fournier, M. J. (1995) U14 base-pairs with 18S rRNA: a novel snoRNA interaction required for rRNA processing. *Genes Dev.* **9**, 2433–2443.
- 35 Stoltzfus, A. (1999) On the possibility of constructive neutral evolution. *J. Mol. Evol.* **49**, 169–181.
- 36 Schubot, F. D., Chen, C.-J., Rose, J. P., Dailey, T. A., Dailey, H. A., and Wang, B.-C. (2001) Crystal structure of the transcription factor sc-mtTFB offers insights into mitochondrial transcription. *Protein Sci.* **10**, 1980–1988.
- 37 Falkenberg, M., Gaspari, M., Rantanen, A., Trifunovic, A., Larsson, N.-G., and Gustafsson, C. M. (2002) Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA. *Nat. Genet.* **31**, 289–293.
- 38 McCulloch, V., Seidel-Rogol, B. L., and Shadel, G. S. (2002) A human mitochondrial transcription factor is related to RNA adenine methyltransferases and binds S-adenosylmethionine. *Mol. Cell. Biol.* **22**, 1116–1125.
- 39 Seidel-Rogol, B. L., McCulloch, V., and Shadel, G. S. (2003) Human mitochondrial transcription factor B1 methylates ribosomal RNA at a conserved stem-loop. *Nat. Genet.* **33**, 23–24.
- 40 Anantharaman, V., Koonin, E. V., and Aravind, L. (2002) Comparative genomics and evolution of proteins involved in RNA metabolism. *Nucleic Acids Res.* **30**, 1427–1464.
41. Bass, B. L., ed. (2001) RNA editing, in *Frontiers in Molecular Biology*, vol. 34 (Hames, B. D. and Glover, D. M., eds.), Oxford University Press, New York.
- 42 Samuel, C. E. (2003) RNA editing minireview series. *J. Biol. Chem.* **278**, 1389, 1390.
- 43 Stuart, K. and Panigrahi, A. K. (2002) RNA editing: complexity and complications. *Mol. Microbiol.* **45**, 591–596.
- 44 Madison-Antenucci, S., Grams, J., and Hajduk, S. L. (2002) Editing machines: the complexities of trypanosome RNA editing. *Cell* **108**, 435–438.
- 45 Simpson, L., Sbicego, S., and Aphasizhev, R. (2003) Uridine insertion/deletion RNA editing in trypanosome mitochondria: a complex business. *RNA* **9**, 265–276.
46. Grosjean, H. and Benne, R., eds. (1998) *Modification and Editing of RNA*. ASM, Washington D.C.
47. Benne, R., Burg, J. V. D., Brakenhoff, J., Vries, B. F. D., Nederlof, P., Sloof, P., and Voogd, A. (1985) Mitochondrial genes in trypanosomes: abnormal initiator triplets,

- a conserved frameshift in the gene for cytochrome oxidase subunit II and evidence for a novel mechanism of gene expression, in *Achievements and Perspectives of Mitochondrial Research*, vol. II (Quagliariello, E., Slater, E. C., Palmieri, F., Saccone, C., and Kroon, A. M., eds.), Elsevier, Amsterdam, pp. 325–336.
- 48 Vickerman, K. (1985) Developmental cycles and biology of pathogenic trypanosomes. *Br. Med. Bull.* **41**, 105–114.
- 49 Chan, L. (1992) Apolipoprotein b, the major protein component of triglyceride-rich and low density lipoproteins. *J. Biol. Chem.* **267**, 25,621–25,624.
- 50 Hoch, B., Maier, R. M., Appel, K., Igloi, G. L., and Kossel, H. (1991) Editing of a chloroplast mRNA by creation of an initiation codon. *Nature* **353**, 178–180.
- 51 Schuster, W., Hiesel, R., Wissinger, B., and Brennicke, A. (1990) RNA editing in the cytochrome b locus of the higher plant *Oenothera berteriana* includes a U-to-C transition. *Mol. Cell. Biol.* **10**, 2428–2431.
- 52 Gualberto, J. M., Weil, J.-H., and Grienemberger, J.-M. (1990) Editing of the wheat coxIII transcript: evidence for twelve C to U and one U to C conversions and for sequence similarities around editing sites. *Nucleic Acids Res.* **18**, 3771–3776.
- 53 Gott, J. M., Visomirski, L. M., and Hunter, J. L. (1993) Substitutional and insertional RNA editing of the cytochrome c oxidase subunit 1 mRNA of *Physarum polycephalum*. *J. Biol. Chem.* **268**, 25,483–25,486.
- 54 Miller, D., Mahendran, R., Spottswood, M., Costandy, H., Wang, S., Ling, M.-L., and Yang, N. (1993) Insertional editing in mitochondria of *Physarum*. *Semin. Cell Biol.* **4**, 261–266.
- 55 Antes, T., Costandy, H., Mahendran, R., Spottswood, M., and Miller, D. (1998) Insertional editing of mitochondrial tRNAs of *Physarum polycephalum* and *Didymium nigripes*. *Mol. Cell. Biol.* **18**, 7521–7527.
- 56 Borner, G. V., Morl, M., Janke, A., and Paabo, S. (1996) RNA editing changes the identity of a mitochondrial tRNA in marsupials. *EMBO J.* **15**, 5949–5957.
- 57 Alfonzo, J. D., Blanc, V., Estevez, A. M., Rubio, M. A., and Simpson, L. (1999) C to U editing of the anticodon of imported mitochondrial tRNA (Trp) allows decoding of the UGA stop codon in *Leishmania tarentolae*. *EMBO J.* **18**, 7056–7062.
- 58 Barth, C., Greferath, U., Kotsifas, M., and Fisher, P. R. (1999) Polycistronic transcription and editing of the mitochondrial small subunit (SSU) ribosomal RNA in *Dictyostelium discoideum*. *Curr. Genet.* **36**, 55–61.
59. Hausmann, S., Garcin, D., and Kolakofsky, K. (2001) Paramyxovirus RNA polymerase stuttering, in *RNA Editing*, vol. 34 (Bass, B., ed.), Oxford University Press, New York, pp. 139–156.
- 60 Koslowsky, D. J., Bhat, G. J., Perrollaz, A. L., Feagin, J. E., and Stuart, K. (1990) The MURF3 gene of *T. brucei* contains multiple domains of extensive editing and is homologous to a subunit of NADH dehydrogenase. *Cell* **62**, 901–911.
- 61 Burns, C. M., Chu, H., Rueter, S. M., Hutchinson, L. K., and Canton, H. (1997) Regulation of serotonin-2C receptor G-protein coupling by RNA editing. *Nature* **387**, 303–308.

- 62 Niswender, C. M., Copeland, S. C., Herrick-Davis, K., Emeson, R. B., and Sanders-Bush, E. (1999) RNA editing of the human serotonin 5-hydroxytryptamine 2C receptor silences constitutive activity. *J. Biol. Chem.* **274**, 9472–9478.
- 63 Wang, Q., O'Brien, P. J., Chen, C.-X., Cho, D.-S. C., Murray, J. M., and Nishikura, K. (2000) Altered G protein-coupling functions of RNA editing isoform and splicing variant serotonin_{2C} receptors. *J. Neurochem.* **74**, 1290–1300.
- 64 Hanrahan, C. J., Palladino, M. J., Bonneau, L. J., and Reenan, R. A. (1999) RNA editing of a *Drosophila* sodium channel gene. *Ann. NY Acad. Sci.* **868**, 51–66.
- 65 Grauso, M., Reenan, R. A., Culetto, E., and Sattelle, D. B. (2002) Novel putative nicotinic acetylcholine receptor subunit genes, D α 5, D α 6 and D α 7, in *Drosophila melanogaster* identify a new and highly conserved target of adenosine deaminase acting on RNA-mediated A-to-I pre-mRNA editing. *Genetics* **160**, 1519–1533.
- 66 Patton, D. E., Silva, T., and Bezanilla, F. (1997) RNA editing generates a diverse array of transcripts encoding squid Kv2K⁺ channels with altered functional properties. *Neuron* **19**, 711–722.
- 67 Baltimore, D. (2001) Our genome unveiled. *Nature* **409**, 814–816.
- 68 Black, D. L. (2000) Protein diversity from alternative splicing: a challenge for bioinformatics and post-genome biology. *Cell* **103**, 367–370.
- 69 Grabowski, P. J. and Black, D. L. (2001) Alternative RNA splicing in the nervous system. *Prog. Neurobiol.* **65**, 289–308.
- 70 Bhat, G. J., Koslowsky, D. J., Feagin, J. E., Smiley, B. L., and Stuart, K. (1990) An extensively edited mitochondrial transcript in kinetoplasts encodes a protein homologous to ATPase subunit 6. *Cell* **61**, 885–894.
- 71 Decker, C. J. and Sollner Webb, B. (1990) RNA editing involves indiscriminate U changes throughout precisely defined editing domains. *Cell* **61**, 1001–1011.
- 72 Koslowsky, D. J., Bhat, G. J., Read, L. K., and Stuart, K. (1991) Cycles of progressive realignment of gRNA with mRNA in RNA editing. *Cell* **67**, 537–546.
- 73 Sturm, N. R. and Simpson, L. (1990) Partially edited mRNAs for cytochrome b and subunit III of cytochrome oxidase from leishmania tarentolae mitochondria: RNA editing intermediates. *Cell* **61**, 871–878.
- 74 Sturm, N. R., Maslov, D. A., Blum, B., and Simpson, L. (1992) Generation of unexpected editing patterns in leishmania tarentolae mitochondrial mRNAs: misediting produced by misguiding. *Cell* **70**, 469–476.
- 75 Schuster, W., Wissinger, B., Unseld, M., and Brennicke, A. (1990) Transcripts of the NADH-dehydrogenase subunit 3 gene are differentially edited in Oenothera mitochondria. *EMBO J.* **9**, 263–269.
- 76 Yang, A. J. and Mulligan, R. M. (1991) RNA editing intermediates of cox2 transcripts in maize mitochondria. *Mol. Cell. Biol.* **11**, 4278–4281.
- 77 Marchfelder, A., Binder, S., Brennicke, A., and Knoop, V. (1998) RNA editing by base conversion in plant organellar RNAs, in *Modification and Editing of RNA* (Grosjean, H. and Benne, R., eds.), ASM, Washington, DC, pp. 307–323.
- 78 Phreaner, C. G., Williams, M. A., and Mulligan, R. M. (1996) Incomplete editing of rps12 transcripts results in the synthesis of polymorphic polypeptides in plant mitochondria. *Plant Cell* **8**, 107–117.

- 79 Hernould, M., Suharsono, S., Litvak, S., Araya, A., and Mouras, A. (1993) Male-sterility induction in transgenic tobacco plants with an unedited *atp9* mitochondrial gene from wheat. *Proc. Natl. Acad. Sci. USA* **90**, 2370–2374.
- 80 Zabaleta, E., Mouras, A., Hernould, M., Suharsano, M., and Araya, A. (1996) Transgenic male-sterile plant induced by an unedited *atp9* gene is respoed to fertility by inhibiting its expression with antisense RNA. *Proc. Natl. Acad. Sci. USA* **93**, 11,259–11,263.
- 81 Cappione, A. J., French, B. L., and Skuse, G. R. (1997) A potential role for NF1 mRNA editing in the pathogenesis of NF1 tumors. *Am. J. Human Genet.* **60**, 305–312.
- 82 Skuse, G. R., Cappione, A. J., Sowden, M., Metheny, L. J., and Smith, H. C. (1996) The neurofibromatosis type I messenger RNA undergoes base-modification RNA editing. *Nucleic Acids Res.* **24**, 478–486.
- 83 Huson, S. M., Compston, D. A. S., Clark, P., and Harper, P. S. (1989) A genetic study of von Recklinghausen neurofibromatosis in south east Wales. I. Prevalence, fitness, mutation rate, and effect of parental transmission on severity. *J. Med. Genet.* **26**, 704–711.
- 84 Xu, G., O'Connell, P., Viskochil, D., Cawthon, R., Robertson, M., Culver, M., Dunn, D., Stevens, J., Gesteland, R., White, R., and Weiss, R. (1990) The neurofibromatosis type 1 gene encodes a protein related to GAP. *Cell* **62**, 599–608.
- 85 Mukhopadhyay, D., Anant, S., Lee, R. M., Kennedy, S., Viskichil, D., and Davidson, N. O. (2002) C to U editing of neurofibromatosis 1 mRNA occurs in tumors that express both the type II transcript and apobec-1, the catalytic subunit of the apolipoprotein B mRNA-editing enzyme. *Am. J. Human Genet.* **70**, 38–50.
- 86 Beghini, A., Ripamonti, C. B., Peterlongo, P., Roversi, G., Cairoli, R., Morra, E., and Larizza, L. (2000) RNA hyperediting and alternative splicing of hematopoietic cell phosphatase (PTPN6) gene in acute myeloid leukemia. *Hum. Mol. Genet.* **9**, 2297–2304.
- 87 Yi, T., Cleveland, J. L., and Ihle, J. N. (1992) Proteins tyrosine phosphatase containing SH2 domains: characterization, preferential expression in hematopoietic cells, and localization to human chromosome 12p12-p13. *Mol. Cell. Biol.* **12**, 836–846.
- 88 Uesugi, Y., Fuse, I., Toba, K., Kishi, K., Koike, T., and Aizawa, Y. (1999) Involvement of SHP-1, a phosphotyrosine phosphatase, during myeloid cell differentiation in acute promyelocytic leukemia cell lines. *Eur. J. Haematol.* **62**, 239–245.
- 89 Kozlowski, M., Kozlowski, M., Mlinaric-Rascan, I., Feng, G. S., Shen, R., Pawson, T., and Siminovitch, K. A. (1993) Expression and catalytic activity of the tyrosine phosphatase PTP1C is severely impaired in motheaten and viable motheaten mice. *J. Exp. Med.* **178**, 2157–2163.
- 90 Takuma, H., Kwak, S., Yoshizawa, T., and Kanazawa, I. (1999) Reduction of GluR2 RNA editing, a molecular change that increases calcium influx through AMPA receptors, selective in the spinal ventral gray of patients with amyotrophic lateral sclerosis. *Ann. Neurol.* **46**, 806–815.

- 91 Maas, S., Patt, S., Schrey, M., and Rich, A. (2001) Underediting of glutamate receptor GluR-B mRNA in malignant gliomas. *Proc. Natl. Acad. Sci. USA* **98**, 14,687–14,692.
- 92 Gurevich, I., Tamir, H., Arango, V., Dwork, A. J., Mann, J. J., and Schmauss, C. (2002) Altered editing of serotonin 2C receptor pre-mRNA in the prefrontal cortex of depressed suicide victims. *Neuron* **34**, 349–356.
- 93 Driscoll, D. M., Wynne, J. K., Wallis, S. C., and Scott, J. (1989) An in vitro system for the editing of apolipoprotein B mRNA. *Cell* **58**, 519–525.
- 94 Shah, R. R., Knott, T. J., Legros, J. E., Navaratnam, N., Greeve, J. C., and Scott, J. (1991) Sequence requirements for the editing of apolipoprotein B mRNA. *J. Biol. Chem.* **266**, 16,301–16,304.
- 95 Backus, J. W. and Smith, H. C. (1991) Apolipoprotein B mRNA sequences 3' of the editing site are necessary and sufficient for editing and editosome assembly. *Nucleic Acids Res.* **19**, 6781–6786.
- 96 Driscoll, D. M., Lakhe-Reddy, S., Oleksa, L. M., and Martinez, D. (1993) Induction of RNA editing at heterologous sites by sequences in apolipoprotein B mRNA. *Mol. Cell. Biol.* **13**, 7288–7294.
- 97 Hersberger, M. and Innerarity, T. L. (1998) Two efficiency elements flanking the editing site of cytidine 6666 in the apolipoprotein B mRNA support mooring-dependent editing. *J. Biol. Chem.* **273**, 9435–1998.
- 98 Hersberger, M., Patarroyo-White, S., Arnold, K. S., and Innerarity, T. L. (1999) Phylogenetic analysis of the apolipoprotein B mRNA-editing region. *J. Biol. Chem.* **274**, 34,590–34,597.
- 99 Richardson, N., Navaratnam, N., and Scott, J. (1998) Secondary structure for the apolipoprotein B mRNA editing site. *J. Biol. Chem.* **273**, 31,707–31,717.
- 100 Sowden, M. P. and Smith, H. C. (2001) Commitment of apolipoprotein b RNA to the splicing pathway regulates cytidine-to-uridine editing-site utilization. *Biochem. J.* **359**, 697–705.
- 101 Sowden, M., Hamm, J. K., and Smith, H. C. (1996) Overexpression of APOBEC-1 results in mooring sequence-dependent promiscuous RNA editing. *J. Biol. Chem.* **271**, 3011–3017.
- 102 Sowden, M., Hamm, J. K., Spinelli, S., and Smith, H. C. (1996) Determinants involved in regulating the proportion of edited apolipoprotein B RNAs. *RNA* **2**, 274–288.
- 103 Blum, B., Bakalara, N., and Simpson, L. (1990) A model for RNA editing in kinetoplastid mitochondria: “Guide” RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell* **60**, 189–198.
- 104 Blum, B. and Simpson, L. (1990) Guide RNAs in kinetoplastid mitochondria have a nonencoded 3' oligo-(U) tail involved in recognition of the pre-edited region. *Cell* **62**, 391–397.
- 105 Maslov, D. A. and Simpson, L. (1992) The polarity of editing within a multiple gRNA-edited domain is due to formation of anchors for upstream gRNAs by downstream editing. *Cell* **70**, 459–467.
- 106 Cech, T. R. (1991) RNA editing: world's smallest introns? *Cell* **64**, 667–669.

- 107 Blum, B., Sturm, N. R., Simpson, A. M., and Simpson, L. (1991) Chimeric gRNA-mRNA molecules with oligo(U) tails covalently linked at sites of RNA editing suggest that addition occurs by transesterification. *Cell* **65**, 543–550.
108. Koslowsky, D. J., Goring, H. U., Morales, T. H., and Stuart, K. (1992) In vitro guide RNA/mRNA chimera formation in *Trypanosoma brucei* RNA editing. *Nature* **356**, 807–809 (see comments).
- 109 Harris, M. E. and Hajduk, S. L. (1992) Kinetoplastid RNA editing: in vitro formation of cytochrome b gRNA-mRNA chimeras from synthetic substrate RNAs. *Cell* **68**, 1–20.
- 110 Seiwert, S. D. and Stuart, K. (1994) RNA editing: transfer of genetic information from gRNA to precursor mRNA in vitro. *Science* **266**, 114–117.
- 111 Seiwert, S. D., Heidmann, S., and Stuart, K. (1996) Direct visualization of uridylylate deletion in vitro suggests a mechanism for kinetoplastid RNA editing. *Cell* **84**, 831–841.
112. Kable, M. L., Seiwert, S. D., Heidmann, S., and Stuart, K. (1996) RNA editing: a mechanism for gRNA-specified uridylylate insertion into precursor mRNA [published erratum appears in *Science* 1996 Oct. 4;274(5284):21]. *Science* **273**, 1189–1195 (see comments).
- 113 Cruz-Reyes, J., Zhelonkina, A., Rusche, L., and Sollner-Webb, B. (2001) Trypanosome RNA editing: simple guide features enhance U deletion 100-fold. *Mol. Cell. Biol.* **21**, 884–892.
- 114 Igo, R. P., Lawson, S. D., and Stuart, K. (2002) RNA Sequence and base pairing effects on insertion editing in *Trypanosoma brucei*. *Mol. Cell. Biol.* **22**, 1567–1576.
- 115 Igo, R. P., Plazzo, S. S., Burgess, M. L. K., Panigrahi, A. K., and Stuart, K. (2000) Uridylate addition and RNA ligation contribute to the specificity of kinetoplastid insertion RNA editing. *Mol. Cell. Biol.* **20**, 8447–8457.
- 116 Leung, S. S. and Koslowsky, D. J. (1999) Mapping contacts between gRNA and mRNA in trypanosome RNA editing. *Nucleic Acids Res.* **27**, 778–787.
- 117 Leung, S. S. and Koslowsky, D. J. (2001) RNA editing in *Trypanosoma brucei*: characterization of gRNA U-tail interactions with partially edited mRNA substrates. *Nucleic Acids Res.* **29**, 703–709.
- 118 Leung, S. S. and Koslowsky, D. J. (2001) Interactions of mRNAs and gRNAs involved in trypanosome mitochondrial RNA editing: structure probing of an mRNA bound to its cognate gRNA. *RNA* **7**, 1803–1816.
- 119 Pai, R. D., Oppgaard, L. M., and Connell, G. J. (2003) Sequence and structural requirements for optimal guide RNA-directed insertional editing within *Leishmania tarentolae*. *RNA* **9**, 469–483.
- 120 Higuchi, M., Single, F. N., Kohler, M., Sommer, B., Sprengel, R., and Seeburg, P. H. (1993) RNA editing of AMPA receptor subunit GluR-B: a base-paired intron-exon structure determines position and efficiency. *Cell* **75**, 1361–1370.
- 121 Bass, B. L. and Weintraub, H. (1988) An unwinding activity that covalently modifies its double-stranded RNA substrate. *Cell* **55**, 1089–1098.
- 122 Polson, A. G., Crain, P. F., Pomerantz, S. C., McCloskey, J. A., and Bass, B. L. (1991) The mechanism of adenosine to inosine conversion by the double-stranded

- RNA unwinding/modifying activity: a high-performance liquid chromatography-mass spectrometry analysis. *Biochemistry* **30**, 11,507–11,514.
- 123 Wagner, R. W., Smith, J. E., Cooperman, B. S., and Nishikura, K. (1989) A double-stranded RNA unwinding activity introduces structural alterations by means of adenosine to inosine conversions in mammalian cells and *Xenopus* eggs. *Proc. Natl. Acad. Sci. USA* **86**, 2647–2651.
- 124 Wagner, R. W., Yoo, C., Wrabetz, L., Kamholz, J., Buchhalter, J., Hassan, N. F., Khalili, K., Kim, S. U., Perussia, B., McMorris, F. A., and Nishikura, K. (1990) Double-stranded RNA unwinding and modifying activity is detected ubiquitously in primary tissues and cell lines. *Mol. Cell. Biol.* **10**, 5586–5590.
- 125 Nishikura, K., Yoo, C., Kim, U., Murray, J. M., Estes, P. A., Cash, F. E., and Liebhaber, S. A. (1991) Substrate specificity of the dsRNA unwinding/modifying activity. *EMBO J.* **10**, 3523–3532.
- 126 Polson, A. G. and Bass, B. L. (1994) Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase. *EMBO J.* **13**, 5701–5711.
- 127 Hough, R. F. and Bass, B. L. (1994) Purification of the *Xenopus laevis* double-stranded RNA adenosine deaminase. *J. Biol. Chem.* **269**, 9933–9939.
- 128 Kim, U., Garner, T. L., Sanford, T., Speicher, D., Murray, J. M., and Nishikura, K. (1994) Purification and characterization of double-stranded RNA adenosine deaminase from bovine nuclear extracts. *J. Biol. Chem.* **269**, 13,480–13,489.
- 129 O'Connell, M. A. and Keller, W. (1994) Purification and properties of double-stranded RNA-specific adenosine deaminase from calf thymus. *Proc. Natl. Acad. Sci. USA* **91**, 10,596–10,600.
- 130 Hurst, S. R., Hough, R. F., Aruscavage, P. J., and Bass, B. L. (1995) Deamination of mammalian glutamate receptor RNA by *Xenopus* dsRNA adenosine deaminase: similarities to *in vivo* RNA editing. *RNA* **1**, 1051–1060.
- 131 Dabiri, D. A., Lai, F., Drakas, R., and Nishikura, K. (1996) Editing of the GLuR-B ion channel RNA *in vitro* by recombinant double-stranded RNA adenosine deaminase. *EMBO J.* **15**, 34–45.
- 132 Maas, S., Melcher, T., Herb, A., Seeburg, P. H., Keller, W., Krause, S., Higuchi, M., and O'Connell, M. A. (1996) Structural requirements for RNA editing in glutamate receptor pre-mRNAs by recombinant double-stranded RNA adenosine deaminase. *J. Biol. Chem.* **271**, 12,221–12,226.
- 133 Giege, P. and Brennicke, A. (1999) RNA editing in Arabidopsis mitochondria effects 441 C to U changes on ORFs. *Proc. Natl. Acad. Sci. USA* **96**, 15,324–15,329.
- 134 Bock, R. (2001) RNA editing in plant mitochondria and chloroplasts, in *RNA Editing*, vol. 34 (Bass, B. L., ed.), Oxford University Press, New York, pp. 38–59.
- 135 Moneger, F., Smart, C., and Leaver, C. (1994) Nuclear restoration of cytoplasmic male sterility in sunflower is associated with the tissue-specific regulation of a novel mitochondrial gene. *EMBO J.* **13**, 8–17.
- 136 Kumar, R. and Levings, C. S. 3rd (1993) RNA editing of a chimeric maize mitochondrial gene transcript is sequence specific. *Curr. Genet.* **23**, 154–159.

- 137 Kubo, N. and Kadowaki, K. (1997) Involvement of 5' flanking sequence for specifying RNA editing sites in plant mitochondria. *FEBS Lett.* **413**, 40–44.
- 138 Kubo, T. and Mikami, T. (1996) A duplicated sequence in sugarbeet mitochondrial transcripts is differentially edited: analysis of orfB and its derivative orf324 mRNAs. *Biochim. Biophys. Acta* **1307**, 259–262.
- 139 Williams, M., Kutcher, B. M., and Mulligan, R. M. (1998) Editing site recognition in plant mitochondria: the importance of 5' flanking sequences. *Plant Mol. Biol.* **36**, 229–237.
- 140. Mulligan, R. M., Williams, M. A., and Shanahan, M. T. (1999) RNA editing site recognition in higher plant mitochondria. *J. Hered.* **90**, 338–344.
- 141 Svab, Z. and Maliga, P. (1993) High-frequency plastid transformation in tobacco by selection for a chimeric *aadA* gene. *Proc. Natl. Acad. Sci. USA* **90**, 913–917.
- 142 Chaudhuri, S. and Maliga, P. (1996) Sequences directing C to U editing of the plastid psbL mRNA are located within a 22 nucleotide segment spanning the editing site. *EMBO J.* **15**, 5958–5964.
- 143 Chaudhuri, S., Carrer, H., and Maliga, P. (1995) Site-specific factors mediate RNA editing of the psbL mRNA in tobacco plastids. *EMBO J.* **12**, 2951–2957.
- 144 Bock, R., Hermann, M., and Fuchs, M. (1997) Identification of critical nucleotide positions for plastid RNA editing site recognition. *RNA* **3**, 1194–1200.
- 145 Bock, R., Hermann, M., and Kossel, H. (1996) In vivo dissection of cis-acting determinants for plastid RNA editing. *EMBO J.* **15**, 5052–5059.
- 146 Hermann, M., and Bock, R. (1999) Transfer of plastid RNA-editing activity to novel sites suggests a critical role for spacing in editing-site recognition. *Proc. Natl. Acad. Sci. USA* **96**, 4856–4861.
- 147 Chateigner-Bourin, A.-L. and Hanson, M. R. (2002) Cross-competition in transgenic chloroplasts expressing single editing sites reveals shared cis-elements. *Mol. Cell. Biol.* **22**, 8448–8456.
- 148 Farre, J.-C. and Araya, A. (2001) Gene expression in isolated plant mitochondria: high fidelity of transcription, splicing and editing of a transgene product in electroporated organelles. *Nucleic Acids Res.* **29**, 2484–2491.
- 149 Hirose, T. and Sugiura, M. (2001) Involvement of a site-specific trans-acting factor and a common RNA-binding protein in the editing of chloroplast mRNAs: development of a chloroplast in vitro RNA editing system. *EMBO J.* **20**, 1144–1152.
- 150 Farre, J.-C., Leon, G., Jordana, X., and Araya, A. (2001) cis recognition elements in plant mitochondrion RNA editing. *Mol. Cell. Biol.* **21**, 6731–6737.
- 151 Miyamoto, T., Obodata, J., and Sugiura, M. (2002) Recognition of RNA editing sites is directed by unique proteins in chloroplasts: biochemical identification of cis-acting elements and trans-acting factors involved in RNA editing in tobacco and pea chloroplasts. *Mol. Cell. Biol.* **22**, 6726–6734.
- 152 Vidal, S., Curran, J., and Kolakofsky, D. (1990) Editing of the sendai virus P/C mRNA by G insertion occurs during mRNA synthesis via a virus-encoded activity. *J. Virol.* **64**, 239–246.
- 153 Vidal, S., Curran, J., and Kolakofsky, D. (1990) A stuttering model for paramyxovirus P mRNA editing. *EMBO J.* **9**, 2017–2022.

- 154 Jacques, J.-P. and Kolakofsky, D. (1991) Pseudo-templated transcription in prokaryotic and eukaryotic organisms. *Genes Dev.* **5**, 707–713.
155. Kolakofsky, D. and Hausmann, S. (1998) Cotranscriptional paramyxovirus mRNA editing: a contradiction in terms? in *Modification and Editing of RNA* (Grosjean, H. and Benne, R., eds.), ASM, Washington, DC, pp. 413–420.
- 156 Visomirski-Robic, L. M. and Gott, J. M. (1995) Accurate and efficient insertional RNA editing in isolated *Physarum* mitochondria. *RNA* **1**, 681–691.
- 157 Visomirski-Robic, L. M. and Gott, J. M. (1997) Insertional editing of nascent mitochondrial RNAs in *Physarum*. *Proc. Natl. Acad. Sci. USA* **94**, 4324–4329.
- 158 Cheng, Y.-W., Visomirski-Robic, L. M., and Gott, J. M. (2001) Non-templated addition of nucleotides to the 3' end of nascent RNA during RNA editing in *Physarum*. *EMBO J.* **20**, 1405–1414.
- 159 Cheng, Y.-W. and Gott, J. M. (2000) Transcription and RNA editing in a soluble *in vitro* system from *Physarum* mitochondria. *Nucleic Acids Res.* **28**, 3695–3701.
160. Gott, J. M. (2001) RNA editing in *Physarum polycephalum*, in *RNA Editing*, vol. 34 (Bass, B., ed.), Oxford University Press, New York, pp. 20–37.
- 161 Byrne, E. M. and Gott, J. M. (2002) Cotranscriptional editing of *Physarum* mitochondrial RNA requires local features of the native template. *RNA* **8**, 1174–1185.
- 162 Byrne, E. M., Stout, A., and Gott, J. M. (2002) Editing site recognition and nucleotide insertion are separable processes in *Physarum* mitochondria. *EMBO J.* **21**, 6154–6161.
- 163 Mehta, A., Kinter, M. T., Sherman, N. E., and Driscoll, D. M. (2000) Molecular cloning of apobec-1 complementation factor, a novel RNA-binding protein involved in the editing of apolipoprotein B mRNA. *Mol. Cell. Biol.* **20**, 1846–1854.
- 164 Lellek, H., Kirsten, R., Diehl, I., Apostel, F., Buck, F., and Greeve, J. (2000) Purification and molecular cloning of a novel essential component of the apolipoprotein B mRNA editing enzyme-complex. *J. Biol. Chem.* **275**, 19,848–19,856.
- 165 Bass, B. L. (2002) RNA editing by adenosine deaminases that act on RNA. *Ann. Rev. Biochem.* **71**, 817–846.
- 166 Davidson, N. O. (2002) The challenge of target sequence specificity in C → U RNA editing. *J. Clin. Invest.* **109**, 291–294.
- 167 Blanc, V. and Davidson, N. O. (2003) C-to-U RNA editing: mechanisms leading to genetic diversity. *J. Biol. Chem.* **278**, 1395–1398.
- 168 Maas, S., Rich, A., and Nishikura, K. (2003) A-to-I RNA editing: recent news and residual mysteries. *J. Biol. Chem.* **278**, 1391–1394.
- 169 Teng, B., Burant, C. F., and Davidson, N. O. (1993) Molecular cloning of an apolipoprotein B mRNA editing protein. *Science* **260**, 1816–1819.
- 170 Navaratnam, N., Morrison, J. R., Bhattacharya, S., et al. (1993) The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. *J. Biol. Chem.* **268**, 20,909–20,912.
- 171 Davidson, N. O., Innerarity, T. L., Scott, J., Smith, H. C., Driscoll, D. M., Teng, B.-B., and Chan, L. (1995) Proposed nomenclature for the catalytic subunit of the mammalian apolipoprotein B mRNA editing enzyme: apobec-1. *RNA* **1**, 3–5.

- 172 Yamanaka, S., Poksay, K. S., Balestra, M. E., Zeng, G.-Q., and Innerarity, T. L. (1994) Cloning and mutagenesis of the rabbit ApoB mRNA editing protein. A zinc motif is essential for catalytic activity, and noncatalytic auxiliary factor(s) of the editing complex are widely distributed. *J. Biol. Chem.* **269**, 21,725–21,734.
- 173 Lau, P. P., Zhu, H.-J., Baldini, A., Charnsangavej, C., and Chan, L. (1994) Dimeric structure of a human apolipoprotein B mRNA editing protein and cloning and chromosomal localization of its gene. *Proc. Natl. Acad. Sci. USA* **91**, 8522–8526.
- 174 Hadjiagapiou, C., Giannoni, F., Funahashi, T., Skarosi, S. F., and Davidson, N. O. (1994) Molecular cloning of a human small intestinal apolipoprotein B mRNA editing protein. *Nucleic Acids Res.* **22**, 1874–1879.
- 175 Nakamuta, M., Oka, K., Krushkal, J., et al. (1995) Alternative mRNA splicing and differential promoter utilization determine tissue-specific expression of the apolipoprotein B mRNA-editing protein (apobec1) gene in mice. *J. Biol. Chem.* **270**, 13,042–13,056.
176. Chang, B. H.-J., Lau, P. P., and Chan, L. (1998) Apolipoprotein B mRNA editing, in *Modification and Editing of RNA* (Grosjean, H. and Benne, R., eds.), ASM, Washington, DC, pp. 325–342.
177. Driscoll, D. M. and Innerarity, T. L. (2001) RNA editing by cytidine deamination in mammals, in *RNA Editing*, vol. 34 (Bass, B. L., ed.), Oxford University Press, New York, pp. 61–76.
- 178 Chan, L., Chang, B. H.-J., Nakamuta, M., Li, W.-H., and Smith, L. C. (1997) Apobec-1 and apolipoprotein B mRNA editing. *Biochim. Biophys. Acta* **1345**, 11–26.
- 179 Nakamuta, M., Chang, B. H.-J., Zsigmond, E., et al. (1996) Complete phenotypic characterization of apobec-1 knockout mice with a wild-type genetic background and a human apolipoprotein B transgenic background, and restoration of apolipoprotein B mRNA editing by somatic gene transfer of apobec-1. *J. Biol. Chem.* **271**, 25,981–25,988.
- 180 Hirano, K.-I., Young, S. G., Farese, R. V., Ng, J., Sande, E., Warburton, C., Powell-Braxton, L. M., and Davidson, N. O. (1996) Targeted disruption of the mouse apobec-1 gene abolishes apolipoprotein B mRNA editing and eliminates apolipoprotein B48. *J. Biol. Chem.* **271**, 9887–9890.
- 181 Morrison, J. R., Paszty, C., Stevens, M. E., Hughes, S. D., Forte, T., Scott, J., and Rubin, E. M. (1996) Apolipoprotein B RNA editing enzyme-deficient mice are viable despite alterations in lipoprotein metabolism. *Proc. Natl. Acad. Sci. USA* **93**, 7154–7159.
- 182 Lau, P. P., Villanueva, H., Kobayashi, K., Nakamuta, M., Chang, B. H.-J., and Chan, L. (2001) A DNAJ protein, apobec-1-binding protein-2, modulates apolipoprotein B mRNA editing. *J. Biol. Chem.* **276**, 46,445–46,452.
- 183 Anant, S., Henderson, J. O., Mukhopadhyay, D., Navaratnam, N., Kennedy, S., Min, J., and Davidson, N. O. (2001) Novel role for RNA-binding protein CUGBP2 in mammalian RNA editing. *J. Biol. Chem.* **275**, 47,338–47,351.

- 184 Anant, S., Mukhopadhyay, D., Sankaranand, V., Kennedy, S., Henderson, J. O., and Davidson, N. O. (2001) ARCD-1, an apobec-1-related cytidine deaminase, exerts a dominant negative effect on C to U editing. *Am. J. Physiol. Cell Physiol.* **281**, C1904–C1916.
- 185 Blanc, V., Navaratnam, N., Henderson, J. O., Anant, S., Kennedy, S., Jarmuz, A., Scott, J., and Davidson, N. O. (2001) Identification of GRY-RBP as an apolipoprotein B RNA-binding protein that interacts with both apobec-1 and apobec-1 complementation factor to modulate C to U editing. *J. Biol. Chem.* **276**, 10,272–10,283.
- 186 Yamanaka, S., Balestra, M. E., Ferrell, L. D., Fan, J., Arnold, K. S., Taylor, S., Taylor, J. M., and Innerarity, T. L. (1995) Apolipoprotein B mRNA-editing protein induces hepatocellular carcinoma and dysplasia in transgenic animals. *Proc. Natl. Acad. Sci. USA* **92**, 8483–8487.
- 187 Yamanaka, S., Poksay, K. S., Arnold, K. S., and Innerarity, T. L. (1997) A novel translational repressor mRNA is edited extensively in livers containing tumors caused by the transgene expression of the apoB mRNA-editing enzyme. *Genes Dev.* **11**, 321–333.
- 188 Muramatsu, M., Sankaranand, V. S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N. O., and Honjo, T. (1999) Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J. Biol. Chem.* **274**, 18,470–18,476.
- 189 Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., and Honjo, T. (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**, 553–563.
- 190 Revy, P., Muto, T., Levy, Y., et al. (2000) Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the hyper-IgM syndrome (HIGM2). *Cell* **102**, 565–575.
- 191 Tian, M. and Alt, F. W. (2000) RNA editing meets DNA shuffling. *Nature* **407**, 31–33.
- 192 Longacre, A. and Storb, U. (2000) A novel cytidine deaminase affects antibody diversity. *Cell* **102**, 541–544.
- 193 Neuberger, M. S. and Scott, J. (2000) RNA editing AIDs antibody diversification? *Science* **289**, 1705–1707.
- 194 Arakawa, H., Hauschild, J., and Buerstedde, J. M. (2002) Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion. *Science* **295**, 1301–1306.
- 195 Fugmann, S. D. and Schatz, D. G. (2002) One AID to unite them all. *Science* **295**, 1244–1245.
- 196 Petersen-Mahrt, S. K., Harris, R. S., and Neuberger, M. S. (2002) AID mutates *E. coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature* **418**, 99–104.
- 197 Harris, R. S., Petersen-Mahrt, S. K., and Neuberger, M. S. (2002) RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol. Cell* **10**, 1247–1253.

- 198 Jarmuz, A., Chester, A., Bayliss, J., Gisbourne, J., Dunham, I., Scott, J., and Navaratnam, N. (2002) An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics* **79**, 285–296.
- 199 Liao, W., Hong, S.-H., Chan, B. H.-J., Rudolph, F. B., Clark, S. C., and Chan, L. (1999) APOBEC-2, a cardiac- and skeletal muscle-specific member of the cytidine deaminase supergene family. *Biochem. Biophys. Res. Comm.* **260**, 398–404.
- 200 Sheehy, A. M. (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418**, 646–650.
201. Hough, R. F. and Bass, B. L. (2001) Adenosine deaminases that act on RNA, in *RNA Editing*, vol. 34 (Bass, B. L., ed.), Oxford University Press, New York, pp. 77–108.
- 202 Palladino, M. J., Keegan, L. P., O'Connell, M. A., and Reenan, R. A. (2000) dADAR, a *Drosophila* double-stranded RNA-specific adenosine deaminase is highly developmentally regulated and is itself a target for RNA editing. *RNA* **6**, 1004–1018.
- 203 Tonkin, L. A., Saccomanno, L., Morse, D. P., Brodigan, T., Drause, M., and Bass, B. L. (2002) RNA editing by ADARs is important for normal behavior in *Caenorhabditis elegans*. *EMBO J.* **21**, 6025–6035.
- 204 Palladino, M. J., Keegan, L. P., O'Connell, M. A., and Reenan, R. A. (2000) A-to-I pre-mRNA editing in *Drosophila* is primarily involved in adult nervous system function and integrity. *Cell* **102**, 437–449.
- 205 Higuchi, M., Maas, S., Single, F. N., Hartner, J., Rozov, A., Burnashev, N., Feldmeyer, D., Sprengle, R., and Seeburg, P. H. (2000) Point mutation in an AMPS receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. *Nature* **406**, 78–81.
- 206 Wang, Q., Khillan, J., Gadue, P., and Nishikura, K. (2000) Requirement of the RNA editing deaminase ADAR1 gene for embryonic erythropoiesis. *Science* **290**, 1765–1768.
- 207 Brusa, R., Zimmermann, F., Koh, D.-S., Feldmeyer, D., Gass, P., Seeburg, P. H., and Sprengel, R. (1995) Early-onset epilepsy and postnatal lethality associated with an editing-deficient GluR-B allele in mice. *Science* **270**, 1677–1680.
- 208 Feldmeyer, D., Kask, K., Brusa, R., Kornau, H.-C., Kolhekar, R., Rozov, A., Burnashev, N., Jensen, V., Hvalby, O., Sprengel, R., and Seeburg, P. H. (1999) Neurological dysfunctions in mice expressing different levels of the Q/R site-unedited AMPAR subunit GluR-B. *Nat. Neurosci.* **2**, 57–64.
- 209 Vissel, B., Royle, G. A., Christie, B. R., et al. (2001) The role of RNA editing of kainate receptors in synaptic plasticity and seizures. *Neuron* **29**, 217–227.
- 210 Paul, M. S. and Bass, B. L. (1998) Inosine exists in mRNA at tissue-specific levels and is most abundant in brain mRNA. *EMBO J.* **17**, 1120–1127.
- 211 Morse, D. P. and Bass, B. L. (1999) Long RNA hairpins that contain inosine are present in *Caenorhabditis elegans* poly(A)⁺ RNA. *Proc. Natl. Acad. Sci. USA* **96**, 6048–6053.
- 212 Morse, D. P. and Bass, B. L. (1997) Detection of inosine in messenger RNA by inosine-specific cleavage. *Biochemistry* **36**, 8429–8434.

- 213 Morse, D. P., Aruscavage, P. J., and Bass, B. L. (2002) RNA hairpins in noncoding regions of human brain and *Caenorhabditis elegans* mRNA are edited by adenosine deaminases that act on RNA. *Proc. Natl. Acad. Sci. USA* **99**, 7906–7911.
- 214 Samuel, C. E. (2001) Antiviral actions of interferons. *Clin. Microbiol. Rev.* **14**, 778–809.
- 215 Cattaneo, R. (1994) Biased A-to-I hypermutation of animal RNA virus genomes. *Curr. Opin. Genet. Devel.* **4**, 895–900.
- 216 Scadden, A. D. J. and Smith, C. W. J. (1997) A ribonuclease specific for inosine-containing RNA: a potential role in antiviral defense. *EMBO J.* **16**, 2140–2149.
- 217 Beverley, S. M. (2003) Protozoomics: trypanosomatid parasite genetics comes of age. *Nat. Rev.* **4**, 11–19.
- 218 Panigrahi, A. K., Schnaufer, A., Ernst, N. L., Wang, B., Carmean, N., Salavati, R., and Stuart, K. (2003) Identification of novel components of *Trypanosoma brucei* editosomes. *RNA* **9**, 484–492.
- 219 Schnaufer, A., Panigrahi, A. K., Panicucci, B., Igo, R. P. Jr, Salavati, R., and Stuart, K. (2001) An RNA ligase essential for RNA editing and survival of the bloodstream form of *Trypanosoma brucei*. *Science* **291**, 2159–2162.
- 220 Aphasizhev, R., Sbicego, S., Peris, M., Jang, S.-H., Aphasizheva, I., Simpson, A. M., Rivlin, A., and Simpson, L. (2002) Trypanosome mitochondrial 3' terminal uridylyl transferase (TUTase): the key enzyme in U-insertion/deletion RNA editing. *Cell* **108**, 637–648.
- 221 Huang, C. E., Cruz-Reyes, J., Zhelonkina, A. G., O'Hearn, S., Wirtz, E., and Sollner-Webb, B. (2001) Roles for ligases in the RNA editing complex of *Trypanosoma brucei*; band IV is needed for U-deletion and RNA repair. *EMBO J.* **20**, 4694–4704.
- 222 Huang, C. E., O'Hearn, S. F., and Sollner-Webb, B. (2002) Assembly and function of the RNA editing complex in *Trypanosoma brucei* requires band III protein. *Mol. Cell. Biol.* **22**, 3194–3203.
- 223 Missel, A., Souza, A. E., Norskau, G., and Goring, H. U. (1997) Disruption of a gene encoding a novel mitochondrial DEAD-box protein in *Trypanosoma brucei* affects edited mRNAs. *Mol. Cell. Biol.* **17**, 4895–4903.
- 224 Muller, U. F. and Goring, H. U. (2002) Mechanism of the gBP21-mediated RNA/RNA annealing reaction: matchmaking and charge reduction. *Nucleic Acids Res.* **30**, 447–455.
- 225 Pelletier, M. and Read, L. K. (2003) RBP16 is a multifunctional gene regulatory protein involved in editing and stabilization of specific mitochondrial mRNAs in *Trypanosoma brucei*. *RNA* **9**, 457–468.
- 226 Aphasizhev, R., Aphasizheva, I., Nelson, R. E., and Simpson, L. (2003) A 100-kD complex of two RNA-binding proteins from mitochondria of *Leishmania tarentolae* catalyzes RNA annealing and interacts with several RNA editing components. *RNA* **9**, 62–76.
- 227 Price, D. H. and Gray, M. W. (1999) A novel nucleotide incorporation activity implicated in the editing of mitochondrial transfer RNAs in *Acanthamoeba castellanii*. *RNA* **5**, 302–317.

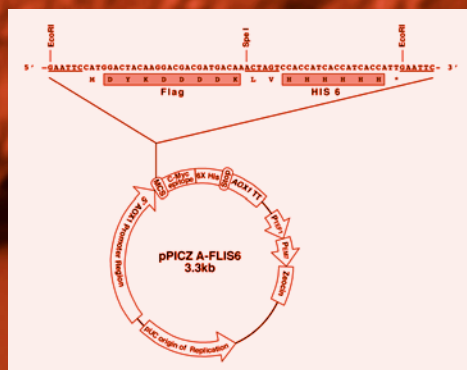
- 228 Lavrov, D. V., Brown, W. M., and Boore, J. L. (2000) A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus*. *Proc. Natl. Acad. Sci. USA* **97**, 13,738–13,742.
- 229 Harlid, A., Janke, A., and Arnason, U. (1998) The complete mitochondrial genome of *Rhea americana* and early avian divergences. *J. Mol. Evol.* **46**, 669–679.
- 230 Vanfleteren, J. R. and Vierstraete, A. R. (1999) Insertional RNA editing in meta-zoan mitochondria: the cytochrome b gene in the nematode *Teratocephalus lirellus*. *RNA* **5**, 622–624.
- 231 Villegas, J., Muller, I., Arredondo, J., Pinto, R., and Burzio, L. O. (2002) A putative RNA editing from U to C in a mouse mitochondrial transcript. *Nucleic Acids Res.* **30**, 1895–1901.
- 232 Bourara, K., Litvak, S., and Araya, A. (2000) Generation of G-to-A and C-to-U changes in HIV-1 transcripts by RNA editing. *Science* **289**, 1564–1566.
- 233 Martinez, I. and Melero, J. A. (2002) A model for the generation of multiple A to G transitions in the human respiratory syncytial virus genome: predicted RNA secondary structures as substrates for adenosine deaminases that act on RNA. *J. Gen. Virol.* **83**, 1445–1455.
- 234 Lin, S., Zhang, H., Spencer, D. F., Norman, J. E., and Gray, M. W. (2002) Widespread and extensive editing of mitochondrial mRNAs in dinoflagellates. *J. Mol. Biol.* **320**, 727–739.
- 235 Sharma, P. M., Bowman, M., and Madden, S. L. (1994) RNA editing in the Wilm's tumor susceptibility gene, WT1. *Genes Dev.* **8**, 720–731.
- 236 Casey, J. L. and Gerin, J. L. (1995) Hepatitis D virus RNA editing: specific modification of adenosine in the antigenomic RNA. *J. Virol.* **69**, 7593–7600.
- 237 Petschek, J. P., Mermer, M. J., Scheckelhoff, M. R., Simone, A., and Vaughn, J. C. (1996) RNA editing in *Drosophila* 4f-rnp gene nuclear transcripts by multiple A-to-G conversions. *J. Mol. Biol.* **259**, 885–890.
- 238 Sanchez, A., Trappier, S. G., Mahy, B. W. J., Peters, C. J., and Nichol, S. T. (1996) The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proc. Natl. Acad. Sci. USA* **93**, 3602–3607.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



Identification of Substrates for Adenosine Deaminases That Act on RNA

Daniel P. Morse

Summary

Adenosine deaminases that acts on RNA (ADARs) are RNA-editing enzymes that convert adenosine to inosine in double-stranded RNA. This chapter provides a detailed protocol for identifying inosine-containing RNAs. Candidate ADAR substrates are identified by cleaving poly (A)⁺ RNA specifically after inosine and using differential display to detect cleaved molecules. To confirm the presence of inosine, each individual candidate substrate is amplified by reverse transcriptase polymerase chain reaction (RT-PCR) and the PCR product is directly sequenced. Sites that contain inosine at the RNA level appear as a mixture of adenosine and guanosine in the cDNA. The relative peak areas provide an estimate of the extent of editing at each site.

Key Words: Adenosine deaminases that act on RNA; RNA editing; inosine; substrate; differential display; RNase T1; reverse transcriptase polymerase chain reaction.

1. Introduction

Adenosine deaminases that act on RNA (ADARs) convert adenosines to inosines within double-stranded regions of RNA (*1–3*). Since inosine is read as guanosine by the translational machinery, ADARs can produce codon changes in mRNAs resulting in the synthesis of multiple protein isoforms with distinct functional properties. The best studied examples of this type of editing are found in several glutamate receptor (GluR) pre-mRNAs (*4*), 5HT_{2C} serotonin receptor pre-mRNA (*5*), and the antigenome of hepatitis delta virus (*6*). In addition, the Bass lab's detection of abundant editing in noncoding regions of mRNAs suggests that ADARs may regulate gene expression (*7,8*). Analysis

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

of knockout animals indicated that ADARs are important for the proper functioning of the nervous system in mice (**9,10**), flies (**11**), and worms (**12**) and are required for red blood cell maturation in mice (**13**). Recently, ADARs have been implicated in the control of the RNA interference pathway in *Caenorhabditis elegans* (**14**). ADAR activity has been observed in every metazoan examined, and surprisingly high levels of inosine have been detected in mRNA derived from multiple rat tissues (**15**). Thus, there are probably many substrates and biologic roles for ADARs remaining to be discovered.

Many of the known ADAR substrates were discovered by noticing discrepancies between genomic and cDNA sequences. In each case, adenosines within genomic DNA appeared as guanosines in the corresponding cDNA. Such A-to-G changes are diagnostic of A-to-I changes at the RNA level because inosine pairs with cytosine during cDNA synthesis. Here, I describe a method to systematically identify poly (A)⁺ RNA that contains inosine.

2. Materials

Materials are subdivided by method and any materials used in more than one method are listed only once. Vendors are included only when the source affected the outcome of the procedure.

2.1. Preparation of RNA

1. Transcription templates for control RNA.
2. IpG dinucleotide.
3. T7 RNA polymerase, NTP mix, transcription buffer (supplied with enzyme).
4. T4 polynucleotide kinase, [γ -³²P]ATP, kinase buffer (supplied with enzyme).
5. Bridging oligonucleotide, T4 DNA ligase, 5X ligase buffer: 250 mM Tris-HCl, pH 7.6, 50 mM MgCl₂, 5 mM ATP, 5 mM dithiothreitol (DTT), 25% PEG-8000.
6. Phenol:chloroform:isoamyl alcohol (25:24:1, v/v).
7. Source for cellular RNA (organism, tissue, or cells), or RNA from vendor.
8. Denaturing solution for total RNA prep: 4 M guanidine thiocyanate, 25 mM sodium citrate, 0.5% sarkosyl, 0.1 M β -mercaptoethanol (add fresh).
9. Oligo(dT) cellulose.
10. Dimethylsulfoxide (DMSO), buffered LiCl: 1 M LiCl; 50 mM EDTA; 2% sodium dodecyl sulfate, 10 mM Tris-HCl, pH 6.5.
11. 0.5 M Sodium acetate, pH 5.5.
12. 50 mM Sodium periodate (freshly dissolved; keep in dark).
13. 2% (v/v) Ethylene glycol.

2.2. Inosine-Specific Cleavage of RNA and Postcleavage Processing

1. Siliconized 1.6-mL microcentrifuge tubes (Phenix).
2. Carrier RNA (e.g., total yeast RNA).
3. 250 mM Sodium phosphate, pH 7.0.

4. AG 501-X8 ion-exchange resin (Bio-Rad, Hercules, CA).
5. 40% Deionized glyoxal.
6. 1 M Sodium borate, pH 7.5.
7. 1.4 M Sodium borate, pH 7.5.
8. 100 mM Tris-HCl, pH 7.8.
9. Tris-borate buffer: 10 mM Tris-HCl, pH 7.8, 1 M sodium borate, pH 7.5 (made by dilution of 1.4 M sodium borate and 100 mM Tris-HCl).
10. RNase T1 (Gibco-BRL, Gaithersburg, MD).
11. Proteinase K (15 $\mu\text{g}/\mu\text{L}$) (Boehringer Mannheim).
12. T4 polynucleotide kinase (30 U/ μL) (USB), 10X kinase buffer: 200 mM Tris-HCl, pH 8.0, 100 mM MgCl_2 .
13. Poly(A) polymerase (500 U/ μL) (USB), 5X poly(A) polymerase buffer: 100 mM Tris-HCl, pH 7.9, 250 mM KCl, 3.5 mM MnCl_2 , 1 mM EDTA, 500 $\mu\text{g}/\text{mL}$ of bovine serum albumin, 50% glycerol.
14. 10 mM ATP.
15. 1 mM Cordycepin triphosphate.

2.3. Arbitrarily Primed Reverse Transcriptase Polymerase Chain Reaction

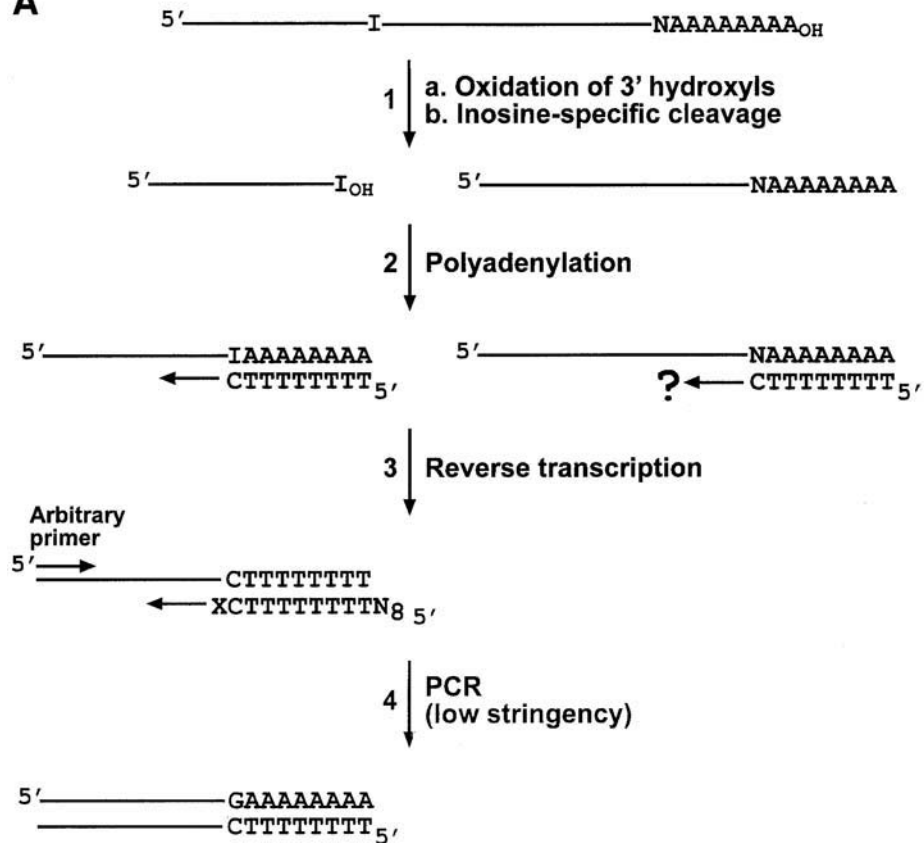
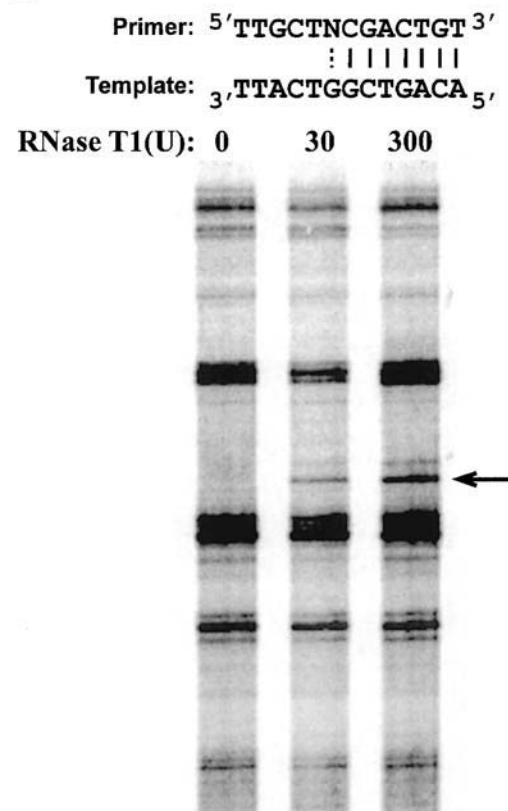
1. MLV reverse transcriptase (RT).
2. 5X MLV RT buffer: 250 mM Tris-HCl, pH 8.3, 375 mM KCl, 15 mM MgCl_2 .
3. T_{12}C RT primer.
4. 1 M NaOH.
5. 1 M HCl.
6. Amplitaq Gold (5 U/ μg) (Perkin-Elmer).
7. 10X polymerase chain reaction (PCR) buffer: 100 mM Tris-HCl, pH 8.3, 500 mM KCl, 0.1% gelatin.
8. [α - ^{33}P]dATP (2000–4000 Ci/mmol).
9. Downstream primers, arbitrary primers.
10. Formamide gel-loading buffer: 95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol.
11. Solutions and equipment for pouring and running sequencing gels and agarose gels.
12. Fluorescent stickers (e.g., Glogos from Stratagene).
13. X-ray film.
14. DNA mass ladder.

2.4. Confirmation of Candidates

1. mRNA and genomic DNA from single individual (or genetically identical individuals).
2. PCR primers specific for candidate ADAR substrates.

3. Methods

The method outlined in **Fig. 1A** combines a technique for cleaving RNA specifically after inosine (*see Subheading 3.2.*) with a differential display pro-

A**B**

protocol for amplification and detection of cleaved transcripts (*see Subheading 3.3.*). Two samples of highly purified poly (A)⁺ RNA are subjected to identical protocols except that the RNA-cleaving enzyme (RNase T1) is not added to one of the samples. If desired, a synthetic inosine-containing RNA can be added to the cellular RNA as a positive control. If previously discovered ADAR substrates are present in the cellular RNA, these can serve as internal positive controls. The RNA is first treated with sodium periodate to oxidize the 3' hydroxyls (step 1a, **Fig. 1A**). Following inosine-specific cleavage (step 1b, **Fig. 1A**) and postcleavage processing, a poly (A) tail is added to the cleavage sites (step 2, **Fig. 1A**). The previous oxidation step prevents elongation of the original poly (A) tails. The RNA is then amplified by arbitrarily primed RT-PCR as in the differential display method (*16*) (steps 3 and 4, **Fig. 1A**). The resulting labeled PCR products are "displayed" on a sequencing gel. Candidate ADAR substrates are detected as RNase T1-dependent bands. True ADAR substrates are distinguished from false positives by comparing the cDNA and genomic sequences for each candidate (*see Subheading 3.4.*). **Figure 1B** shows the results of applying this method to detect cleavage of a synthetic inosine-containing RNA (*see Subheading 3.1.1.* for a description of this control RNA). **Figure 2** shows a typical differential display gel produced with *C. elegans* poly (A)⁺ RNA.

Unless stated otherwise, phenol:chloroform extractions are performed with 1 vol of phenol:chloroform:isoamyl alcohol (25:24:1), and nucleic acids are precipitated with 1/10 vol of 3 M sodium acetate and 2.5 vol of ethanol. To

Fig. 1. (*see opposite page*) Detection of inosine in RNA. **(A)** Outline of strategy. Step 1a: Poly (A)⁺ RNA is treated with sodium periodate to oxidize 3' hydroxyls. Step 1b: RNA is treated with glyoxal and RNase T1 to cleave RNA specifically after inosine. Step 2: A poly (A) tail is added 3' to the inosine at the cleavage site to create a primer binding site for first-strand cDNA synthesis. Step 3: Reverse transcription with a T₁₂C primer. The question mark indicates that the primer will extend on uncleaved RNA only if N = G. Step 4: Low-stringency PCR. X at the 3' end of the downstream primer is either G, A, C, TG, TA, TC, or TC. N₈ is the 8-bp extension on the 5' end of the downstream primer. Note that the cDNA template is drawn 3' to 5'. **(B)** Detection of single inosine in synthetic control RNA. The control RNA (0.5 fmol) was spiked into 5 µg of yeast RNA and subjected to inosine-specific cleavage using the indicated amounts of RNase T1. The cleaved RNA was then subjected to the protocol in (A). The complementarity between the upstream arbitrary PCR primer and its priming site is shown above the gel; N represents a randomized position. For the downstream PCR primer, X = G because the nucleotide on the 5' side of the inosine in the control RNA was a C. The arrow points to an RNase T1-dependent band whose sequence confirmed that it derived from the control RNA cleaved precisely 3' to its single inosine.

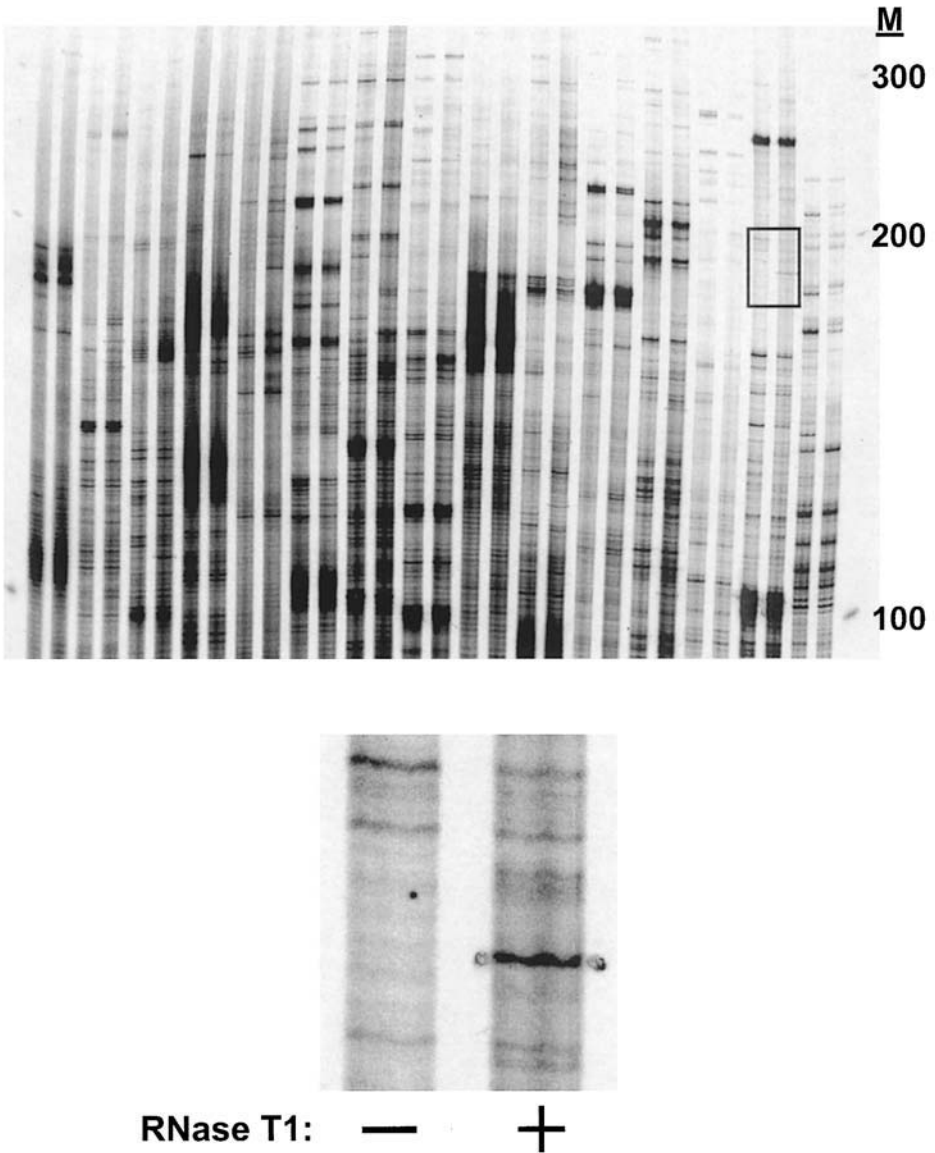


Fig. 2. A differential display gel from the analysis of *C. elegans* poly (A)⁺ RNA is shown, with a region containing an RNase T1-dependent band (boxed) enlarged below. Each pair of lanes corresponds to a different primer pair, minus (left lane) or plus (right lane) RNase T1. The dots on each side of the band are pinholes used to mark its position for excision and elution of the DNA. M, 100-bp ladder.

ensure reproducibility when performing multiple reactions, master mixes of common reagents should be used.

3.1. Preparation of RNA

3.1.1. Synthesis of Control RNA (see **Note 1**)

The use of a synthetic radiolabeled RNA containing a single inosine was critical to the development of most of the procedures in this chapter (**17**). As described in the relevant sections below, this RNA (subsequently referred to as the control RNA) is important for both optimizing the methods and for monitoring the RNA recovery after each step. The control RNA is synthesized by joining two half molecules using the oligo-bridged ligation method of Moore and Sharp (**18**). In this method, two RNA halves are annealed to a complementary DNA oligonucleotide that spans the ligation junction. The nick in the resulting RNA/DNA duplex is sealed with T4 DNA ligase. The two half molecules are produced by *in vitro* transcription from either linearized plasmids or PCR products that contain phage promoters upstream of the desired sequences. Inosine is placed at the 5' end of the 3' half molecule by initiating transcription with an IpG dinucleotide using a 5:1 molar ratio of dinucleotide to GTP in the transcription reaction. Since IpG is not commercially available, my colleagues and I synthesize it on an Applied Biosystems 394 DNA/RNA synthesizer. The 3' half molecule is 5' end labeled using T4 polynucleotide kinase and [γ - ^{32} P] ATP (**19**) and purified on a denaturing (contains urea) polyacrylamide gel (**19**).

1. Combine 75 pmol of labeled 3' half, 100 pmol of 5' half, and 80 pmol of bridging oligodeoxynucleotide in a total of 14 μ L of water.
2. Heat to 100°C for 1 min and place at 42°C.
3. Add 4 μ L of 5X ligase buffer and keep at 42°C for 20 min to anneal the bridge to the RNA molecules.
4. Cool to room temperature on a benchtop.
5. Add 2 μ L of 5 U/ μ L DNA ligase and incubate at room temperature for 4 h.
6. Dilute to 100 μ L with water, extract once with phenol:chloroform, and precipitate nucleic acids.
7. Purify the ligation product on a denaturing polyacrylamide gel and quantify by liquid scintillation counting.

3.1.2. Purification of Cellular RNA

Total RNA and poly (A)⁺ RNA from a variety of organisms and tissues are available commercially but can be expensive. Whether purchased or purified in the laboratory, it is important that the starting material be highly enriched for

poly (A)⁺ RNA and nearly free of rRNA. In my laboratory, no commercially available mRNA purification kit yields RNA of the required quality. We purify poly (A)⁺ RNA from total RNA using the method of Bantle et al. (20). In this method, total RNA is subjected to two (optionally three) rounds of poly (A) selection on an oligo (dT) cellulose column. Prior to the second round, the RNA is heated in DMSO and buffered LiCl, which dissociates mRNA-rRNA complexes. This treatment results in poly (A)⁺ RNA with little, if any, detectable rRNA contamination.

We prepare total RNA using the guanidinium thiocyanate-phenol method (21). Ten milligrams of total RNA is loaded onto a 1-mL oligo (dT) cellulose column. Following the standard protocol for binding, washing, and elution (19), the first round of selection typically yields 100–200 µg of poly (A)⁺ RNA in a 3- to 4-mL volume. To prepare RNA for the second round of selection follow these steps:

1. Divide RNA into eight microfuge tubes and precipitate (precipitating in microfuge tubes facilitates redissolving in a small volume).
2. Dissolve each sample in 50 µL of water, combine into one microfuge tube, and precipitate again.
3. Dissolve RNA in 20 µL of 10 mM Tris-HCl (pH 7.5).
4. Add 180 µL of DMSO and 20 µL of buffered LiCl.
5. Heat at 55°C for 5 min, add 2 mL of loading buffer, and load onto the same column used for the first round of selection.

The second round typically yields 50–100 µg of RNA. The RNA is precipitated twice as before and dissolved to a final concentration of 2 µg/µL. A 5-µg sample of the RNA is run on a formaldehyde gel (19). If significant amounts of rRNA are still visible by ethidium bromide staining, then the procedure followed for the second round can be repeated in a third round of selection. A small amount of rRNA contamination is acceptable but will increase the background of false positives.

3.1.3. Oxidation of 3' Hydroxyls (see **Note 2**)

1. Combine 10 µL (20 µg) of poly (A)⁺ RNA, 3.3 µL of 0.5 M sodium acetate (pH 5.5), and 3.3 µL of 50 mM sodium periodate (freshly dissolved).
2. Incubate for 1 h in the dark at room temperature.
3. Add 16.6 µL of 2% ethylene glycol and continue incubation at room temperature for another 10 min.
4. Dilute to 400 µL with water and precipitate the RNA.
5. Dissolve in 174 µL of water. This is sufficient for four glyoxal reactions containing 5 µg of RNA each (see **Subheading 3.2.3.**).

3.2. Inosine-Specific Cleavage of RNA and Postcleavage Processing

Ribonuclease T1 (RNase T1) cleaves RNA 3' of both guanosine and inosine. The strategy for cleaving RNA specifically after inosine (**Fig. 3**) exploits the ability of the reagent glyoxal to discriminate between these structurally similar nucleotides. In the presence of borate ions, glyoxal forms a stable adduct with guanosine but not with inosine (**22**), and glyoxalated guanosines are resistant to RNase T1 (**23**). Therefore, an inosine-containing RNA that is stably modified by glyoxal is cleaved by RNase T1 only after inosine (**17**).

3.2.1. Optimization

The control RNA is used to determine the optimal conditions for the RNase T1 reaction (**Fig. 4**). For each condition to be tested, 1 fmol of the labeled control RNA (*see Subheading 3.1.1.*) is added to 5 µg of carrier RNA (e.g., total yeast RNA), reacted with glyoxal, and digested with RNase T1. Following the RNase T1 reaction, the products are separated on a 6% denaturing (contains urea) polyacrylamide gel (**19**). Since the control RNA is labeled immediately 5' of the single inosine, only two discrete bands should be visible (*see Fig. 4*). The two bands correspond to uncleaved RNA and the 5' half of the inosine-specific cleavage product. The optimal conditions are those that give the maximum amount of inosine-specific cleavage and the minimum amount of nonspecific cleavage. Nonspecific cleavage appears as an increase in the intensity of the background smear (compared to no enzyme control) coupled with a decrease in the intensities of the two discrete bands. We find that about 80% of the input control RNA can be cleaved specifically after inosine before nonspecific cleavage becomes significant. Optimization should be repeated with each new lot of RNase T1 or whenever a new reagent solution is made.

3.2.2. General Considerations

Each reaction is performed in duplicate to assess reproducibility. Therefore, four 5-µg samples of RNA are processed in parallel: two samples are treated with RNase T1 and two samples are not. Reactions that contain glyoxalated RNA should be performed in siliconized tubes to facilitate dissolving of the RNA following precipitation. Although tubes can be siliconized in the laboratory, I see better results using commercially available presiliconized tubes such as those available from Phenix.

3.2.3. Glyoxal Reaction (*see Note 3*)

1. Combine 43.5 µL (5 µg) of poly (A)⁺ RNA diluted in water, 1 µL (1 fmol) of ³²P-labeled control RNA diluted in water, 4 µL of 250 mM sodium phosphate (pH 7.0), 50 µL of DMSO, and 1.5 µL of 40% deionized glyoxal.

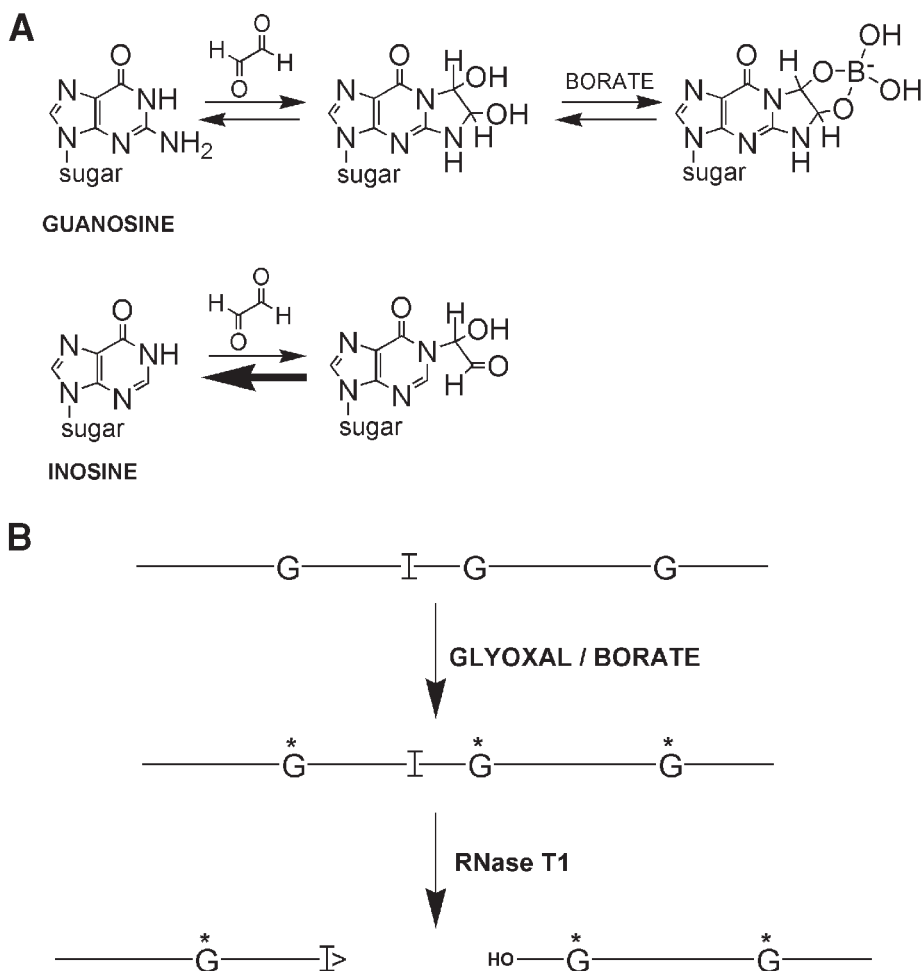


Fig. 3. Strategy for inosine-specific cleavage of RNA. (A) The diagram shows the reaction of glyoxal ($C_2H_2O_2$) with guanosine and inosine, and stabilization of the guanosine adduct with borate. As shown, inosine does not react stably with glyoxal. (B) Scheme for inosine-specific RNase T1 cleavage of glyoxalated RNA. Asterisks mark sites of glyoxalated guanosines that are resistant to RNase T1. The RNase T1 fragment that contains the inosine is shown with a 2',3' cyclic phosphate (>) because the RNase T1 reaction with inosine frequently does not go to completion. (Reprinted with permission from Morse and Bass [17]. Copyright 1997 American Chemical Society.)

2. Incubate for 45 min at 37°C.
3. Add 100 μ L of 1 M sodium borate (pH 7.5) and precipitate with 500 μ L of ethanol (no sodium acetate is added).
4. Dissolve in 15 μ L of Tris-borate buffer.

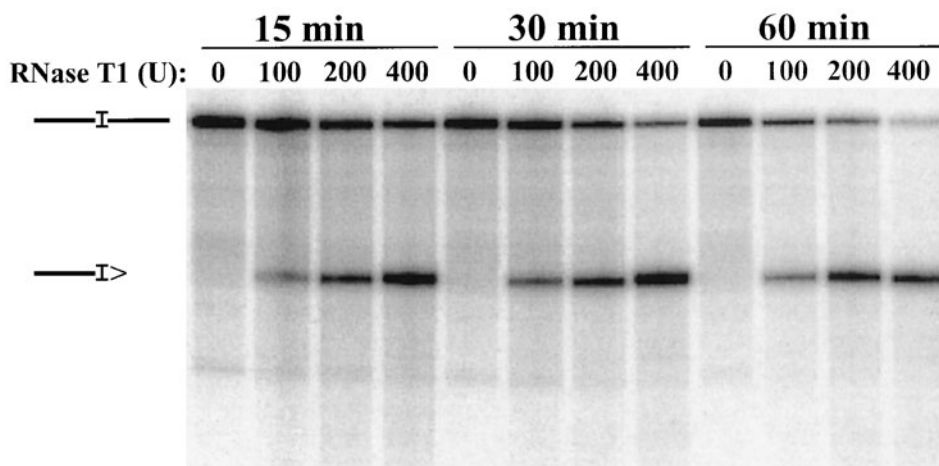


Fig. 4. Optimization of inosine-specific cleavage using synthetic control RNA. For each lane, 1 fmol of RNA containing a single inosine was spiked into 5 μ g of total yeast RNA and treated with glyoxal and borate as described in the text. The glyoxalated RNA was incubated with increasing amounts (U) of RNase T1 and time, as indicated. The PhosphorImager image shows electrophoretically separated starting material and a single product of the size expected for cleavage after the inosine, which accumulated with increasing time and RNase T1. Positions of the full-length and cleaved RNAs (determined in control experiments) are shown on the left. In this experiment, treatment with 400 U of RNase T1 for 30 min was optimal. Treatment with 400 U for 60 min resulted in an unacceptably high level of nonspecific cleavage, as indicated by the decrease in the intensity of the specific cleavage product. (Reprinted with permission from Morse and Bass [17]. Copyright 1997 American Chemical Society.)

3.2.4. Digestion With RNase T1 (see **Note 4**)

1. Add 1 μ L of water (for no RNase T1 controls) or 1 μ L of RNase T1 containing the previously determined optimal number of units (typically 100–400 U of BRL enzyme) to the redissolved glyoxalated RNA.
2. Incubate at 37°C for the previously determined optimal time (typically 30 min).
3. To inactivate RNase T1, add 0.6 μ L of 15 μ g/ μ L proteinase K and incubate for 20 min at 37°C.
4. Add 100 μ L of phenol:chloroform and vortex for 30 s.
5. Add 85 μ L of water and vortex for 30 s.
6. Transfer the aqueous phase to a new tube and extract again with 100 μ L of phenol:chloroform.
7. Dilute to 400 μ L with water, precipitate, and dissolve in 43 μ L of water (or dissolve in 10 μ L of water if the reaction products are to be run on a gel).

3.2.5. Removal of 3' and 2',3'-Cyclic Phosphates (see **Note 5**)

RNase T1 normally produces 3' phosphates via a 2',3'-cyclic phosphate intermediate (**23**). Under conditions used in this protocol, the reaction with inosine does not go to completion, resulting in a mixture of 3' and cyclic phosphates. Removal of phosphates, which allows tailing of the cleavage sites, is accomplished with T4 polynucleotide kinase. In addition to its more well-known activity, this enzyme has both 3' phosphatase (**24**) and 2',3'-cyclic phosphodiesterase (**25**) activities.

1. Combine 43 μL of RNA (from the RNase T1 digestion), 5 μL of 10X RNase T1 buffer, and 2 μL of 30 U/ μL T4 polynucleotide kinase (USB).
2. Incubate for 1 h at 37°C.
3. Add 0.6 μL of 15 $\mu\text{g}/\mu\text{L}$ proteinase K, and incubate for 20 min at 37°C.
4. Dilute to 100 μL with water and extract twice with phenol:chloroform.
5. Dilute to 400 μL with water, precipitate, and dissolve in 46 μL of water. The RNA may not completely dissolve in water at this stage, but it will dissolve once DMSO is added.

3.2.6. Removal of Glyoxal Adducts (see **Note 6**)

Glyoxal must be removed from RNA prior to first-strand cDNA synthesis because RT terminates at glyoxalated guanosines. Glyoxal adducts are unstable at alkaline pH, but the use of high pH results in unacceptable levels of RNA degradation. I remove glyoxal by incubating at neutral pH and high temperature. These conditions result in only a negligible amount of RNA hydrolysis.

1. Combine 46 μL of RNA (from the kinase reaction), 4 μL of 250 mM sodium phosphate (pH 7.0), and 50 μL of DMSO.
2. Incubate for 3 h at 60°C.
3. Dilute to 400 μL with water, precipitate, and dissolve in 13.5 μL of water.

3.2.7. Polyadenylation (see **Note 7**)

1. Combine 13.5 μL of RNA (from **Subheading 3.2.6.**), 4 μL of 5X poly (A) polymerase buffer, 1 μL of 10 mM ATP, 1 μL of 1 mM cordycepin triphosphate, and 0.5 μL of 500 U/ μL poly (A) polymerase (USB).
2. Incubate for 1 h at 30°C.
3. Add 0.5 μL of 15 $\mu\text{g}/\mu\text{L}$ proteinase K and incubate for 20 min at 37°C.
4. Dilute to 100 μL with water, extract twice with phenol:chloroform, precipitate, and dissolve in 11.4 μL of water.

3.3. Arbitrarily Primed RT-PCR

After first-strand cDNA synthesis, small aliquots of the cDNA are amplified in numerous low-stringency PCRs. Each reaction is performed with one of

seven different downstream primers and one of a large collection of arbitrary upstream primers. The downstream primers are of the form GAGA CCACT₁₂CX where X is one of G, A, C, TG, TA, TC, or TT. The upstream primers are 13mers whose sequences are chosen “arbitrarily.” Each PCR amplifies a subset of the cDNA population and the products are “displayed” on a sequencing gel. Candidate ADAR substrates are identified as RNase T1-dependent bands.

3.3.1. First-Strand cDNA Synthesis (see **Note 8**)

1. Combine 11.4 μL of RNA (see **Subheading 3.2.7., step 4**), 4 μL of 5X MLV RT buffer, 1 μL of 10 mM DTT, 1.6 μL of 250 μM dNTPs, and 1 μL of 10 pmol/ μL T₁₂C primer.
2. Incubate for 5 min at 65°C (to denature the RNA) and cool to 37°C.
3. Add 1 μL of 200 U/ μL MLV RT and incubate for 1 h at 37°C.
4. Add 2.5 μL of 1 M NaOH and incubate at 50°C for 30 min to hydrolyze the RNA.
5. Neutralize with 2.5 μL of 1 M HCl and dilute to 1 mL with water.

3.3.2. Arbitrarily Primed PCR (see **Notes 9–13**)

1. Combine 2 μL of cDNA (from first-strand synthesis), 1 μL of 10X PCR buffer, 1 μL of 25 mM MgCl₂, 0.8 μL of 250 μM dNTPs, 1 μL of 10 pmol/ μL downstream primer, 1 μL of 10 pmol/ μL arbitrary primer, 0.2 μL of 10 $\mu\text{Ci}/\mu\text{L}$ [α -³³P] dATP (2000–4000 Ci/mmol), 2.8 μL of water, and 0.25 μL of 5 U/ μL Amplitaq Gold (Perkin-Elmer).
2. Use the following cycling conditions: 94°C for 9 min (to activate the enzyme); 50 cycles of 94°C for 1 min, 40°C for 2 min, and 72°C for 1 min; and 1 cycle of 72°C for 5 min.

3.3.3. Identification of Candidate ADAR Substrates

Six microliters of each PCR is added to 4 μL of formamide gel-loading buffer and loaded onto a 6% sequencing gel (**19**). The samples that differ only in whether or not they were treated with RNase T1 are loaded in adjacent lanes (see **Fig. 2**). The gel is run at 35 mA until the xylene cyanol is about two-thirds of the way down the gel. I use a 100-bp ladder 5' end-labeled with polynucleotide kinase and [γ -³²P]ATP (**19**) as a molecular weight marker. Fluorescent stickers (e.g., Stratagene's Glogos) are placed on the dried gel for later alignment with the X-ray film. An X-ray film is placed on the gel and exposed for about 36 h. The developed film is examined for the presence of bands that are more intense in samples that were treated with RNase T1 (see **Fig. 2**). Each PCR that produces such RNase-T1 dependent bands is repeated using the duplicate pair of cDNA samples. Bands that are reproducibly dependent on RNase T1 represent candidate ADAR substrates.

3.3.4. Elution, Reamplification, and Sequencing of PCR Products (see **Note 14**)

Bands representing candidate ADAR substrates are eluted from the gel, reamplified, and sequenced. Since the poly (A) tails are directly attached to the RNase T1 cleavage sites, the sequences reveal both the identities of the RNAs and the locations of the cleavage sites.

1. Place developed X-ray film on top of the dried gel and align with the fluorescent markers.
2. Use a needle to punch holes through the film into the gel on either side of the band of interest.
3. Cut out the band with a razor blade using the holes as a guide.
4. Use a second X-ray film or a phosphorimager screen to confirm that the correct band was excised.
5. Soak the gel slice in 100 μL of water for 10 min.
6. Heat at 100°C for 15 min, and cool to room temperature.
7. Spin at top speed for 2 min in a microcentrifuge, and remove the water containing the eluted DNA to a fresh tube.
8. Reamplify 3 μL of the eluted DNA in a 50- μL PCR. The reaction components and cycling conditions are identical to those of the original PCR (see **Subheading 3.3.2.**) except that the final dNTP concentration is 200 μM and no isotope is added.
9. Run 10 μL of the PCR on an agarose gel with a DNA mass ladder (BRL) to estimate the yield.
10. Extract the remaining 40 μL of the PCR with phenol:chloroform and precipitate.
11. Gel purify the PCR product and sequence using the PCR primers.

3.4. Confirmation of Candidates

There are two major reasons why not every RNase T1-dependent band yields an ADAR substrate. First, some bands do not represent a single major species but are a mixture of many different sequences. Second, guanosines in glyoxalated RNA are not completely resistant to RNase T1. Therefore, some of the RNase T1-dependent bands are false positives owing to RNAs that have been cleaved after guanosine (see **Subheading 3.5.**). This is not a serious problem because true ADAR substrates (I cleavages) are easily distinguished from false positives (G cleavages) by examining genomic sequences. Since ADARs convert adenosines to inosines, T1 cleavage sites within true ADAR substrates appear as adenosines in the corresponding genomic sequences. Several other characteristics are typical of ADAR substrates (see **Subheading 3.4.3.**). When present in a candidate RNA, these characteristics provide additional evidence for its status as a true ADAR substrate.

3.4.1. Screening Out False Positives

The sequences obtained from the excised bands are used in a BLAST search to identify the corresponding genomic DNA sequences. Genes that contain an adenosine (rather than a guanosine) at the RNase T1 cleavage site are likely to encode ADAR substrates. The T1 cleavage sites are easily identified in the cDNA sequences because they are immediately 5' of the added poly (A) tail.

3.4.2. Confirmation of A-to-G Changes

It is important to confirm independently the sequences of the gene and cDNA for each of the candidates that remains after the Blast search. This is to screen out false positives that could arise from three possible (although unlikely) sources: sequencing errors in the database; RNase T1 cleavage after adenosine in the RNA; or, when using RNA from one or a few individual animals, allelic variation. To avoid false positives owing to allelic variation, the confirmation test described next should be performed with RNA and genomic DNA isolated from the same individual.

To confirm the A-to-G changes, a region surrounding each candidate editing site is amplified from both genomic DNA and cDNA, and their sequences are compared. The sequences found in the previous BLAST searches provide the information needed to design PCR primers. Unclassified PCR products can be sequenced directly or the sequences of multiple individual clones can be determined. These two approaches provide complementary information. The sequences of individual clones reveal the distributions of A-to-G changes within single molecules. By sequencing the PCR products directly, one can estimate the fraction of molecules that are deaminated at each site. A very sensitive and accurate method to measure the efficiency of editing at one particular site is limited primer extension performed on the PCR products (*see* **ref. 26** for an example).

3.4.3. Other Characteristics of ADAR Substrates (*see* **Note 15**)

ADARs usually deaminate multiple adenosines within double-stranded regions of RNA. Thus, most candidate substrates should have the potential to fold into one or more stem-loop structures, and multiple A-to-G changes should be found within these potential structures. RNAs that are predicted to be almost completely double stranded should be deaminated at more sites than those whose structures are frequently interrupted by mismatches, bulges, and loops. ADARs have a 5' nearest neighbor preference: A or U is preferred over C, which is preferred over G (27). These preferences should be reflected in the deamination patterns seen in candidate ADAR substrates. That is, adenosines in good context should be deaminated in a greater fraction of the population than those in poor context.

3.5. Results From *C. elegans* and Human Brain

My colleagues and I have applied the differential display strategy to search for new ADAR substrates in *C. elegans* and in human brain (7,8). The similarities and differences in the results from these two RNA sources may be instructive.

We found that most of the inosine in the poly (A)⁺ RNA from both sources was located in long stem-loop structures found in noncoding regions. Many of the edited regions were repetitive elements (IR elements in *C. elegans* and mostly Alu sequences in human brain) embedded within mRNAs. One interpretation of these results is that the protocol described here is somehow biased for the selection of substrates edited in noncoding regions. The fact that we readily detected editing in coding regions of previously known ADAR substrates expressed in human brain (glutamate receptor and serotonin receptor mRNAs) argues against this possibility. Instead, it is likely that editing in coding regions is the exception rather than the rule. Thus, efficient detection of edited coding regions may require a new strategy.

The efficiencies of the two searches (number of ADAR substrates found per PCR) were surprisingly different. We found, on average, one human brain ADAR substrate for every 5 PCRs, compared with one *C. elegans* substrate for every 70 PCRs. The higher efficiency with human brain RNA was owing to an increased frequency of RNase T1-dependent bands (about one in three PCRs compared with one in seven for *C. elegans*) and to a dramatic decrease in the number of false positives (~4% in human vs ~50% in *C. elegans*). The reason for these differences is not clear, but one possible explanation is that there are a lot more ADAR substrates in human brain than in *C. elegans*.

4. Notes

1. It is not necessary to gel purify the 5' half because any prematurely terminated transcripts will not be ligated to the 3' half. The bridge must have a C in the position that will pair with inosine. The specific activity of the ligation product will be equal to the specific activity of the [γ -³²P]ATP because only phosphorylated 3' half molecules participate in the ligation reaction. In my laboratory, this protocol typically gives 10–40% ligation efficiency. The efficiency drops as the sizes of the RNA halves increase. The control RNA should be as large as possible while maintaining a reasonable ligation efficiency. We have had good success with ligating RNAs on the order of 200 nt in length. For RNAs in this size range, we use a 40-nt bridging oligonucleotide with approx 20 nt of complementarity to both RNAs. The sequence of the control RNA is not important.
2. I have not rigorously tested whether this step is required. Its purpose is to prevent elongation of the original poly (A) tails, which could interfere with tailing of the cleavage sites and reverse transcription.

3. The 40% glyoxal stock solution is deionized with AG 501-X8 until the pH is >5.0. My colleagues and I prepare solutions containing 1 M borate from a stock solution of 1.4 M boric acid that has been adjusted to pH 7.5 with NaOH. The boric acid will not completely dissolve until the pH approaches neutrality. At 37°C the pH of the Tris-HCl in the Tris-borate solution changes to 7.5, which is optimal for the RNase T1 reaction. On addition of ethanol the solution becomes very cloudy, and the resulting pellet is huge. This is normal, but the precipitate should not be cooled excessively or spun for a long time or there will be an even larger pellet. We usually cool at -70°C for about 2 min (cooling is probably unnecessary) and spin at room temperature for 10 min. It takes about 20 min at room temperature to dissolve the pellet. There is no need to vortex until after the 20-min incubation. The labeled control RNA is useful at this stage (and every subsequent precipitation) for confirming that the RNA is completely dissolved.
4. RNase T1 is diluted in water immediately prior to use. Treatment with both proteinase K and phenol:chloroform ensures that RNase T1 is completely inactivated. The first phenol:chloroform extraction is done before diluting the reaction so that glyoxal does not begin to dissociate prior to inactivation of the enzyme. The RNA is diluted to 400 µL prior to precipitation to prevent the coprecipitation of excess borate. Excess borate will inhibit the subsequent removal of glyoxal.
5. The reaction is optimal at pH 8.0 and requires a high concentration of enzyme (~1 U/µL). ATP should not be added. I have found that without proteinase K treatment, much of the RNA is pulled into the organic phase during the phenol:chloroform extraction. This is likely owing to the large amount of enzyme in the reaction. Since borate stabilizes the glyoxal adducts, it is important to remove the residual borate before attempting to remove glyoxal in the next step. This is the purpose of diluting the RNA to 400 µL prior to precipitation at this (and the previous) step.
6. The buffer is the same as that used for the glyoxal reaction (minus glyoxal, of course). The RNA is diluted prior to precipitation to prevent coprecipitation of glyoxal. The RNA pellet should dissolve readily now that glyoxal is removed. The control RNA can be used to confirm that glyoxal has been removed because glyoxalated RNA migrates more slowly in a gel than does unmodified RNA.
7. I include cordycepin (3'-deoxyadenosine) triphosphate in the reaction to limit the lengths of the poly (A) tails. The control RNA can be used to visualize the lengths of the added tails and to monitor the efficiency of the oxidation reaction. If oxidation was successful, only cleaved RNA will be tailed.
8. To enrich for cleaved molecules, the reverse transcription primer has a 3' terminal C that will pair with inosine (or guanosine) at the RNase T1 cleavage site (*see Fig. 1A*). The primer will be extended efficiently from the poly (A) tails on uncleaved RNA only when the nucleotide immediately 5' of the tail (N in *Fig. 1A*) is a G. I find that hydrolysis of the RNA improves the sensitivity of the procedure.
9. Only one pair of cDNA samples (plus and minus RNase T1) is used at this stage. The duplicate cDNA samples are used to verify the reproducibility of any RNase T1-dependent PCR products detected. An arbitrary primer is not the same as a random primer. It has a single fixed sequence that is chosen "arbitrarily." Since

PCR is performed with a low annealing temperature (40°C), each arbitrary primer can be extended from multiple poorly matched priming sites. Each combination of primers produces a unique pattern of bands that represent a subset of the molecules in the original population (*see Fig. 2*). By using a sufficient number of primer pairs, one can in theory sample the entire population.

10. Typically, an arbitrary primer can produce a PCR product if only six to eight of its 3'-most nucleotides are complementary to the template. My colleagues and I have found that by randomizing the eighth position from the 3' end of each arbitrary primer, we could significantly improve sensitivity (*see Fig. 1B*). The one or two extra nucleotides on the 3' end of the downstream primers results in each primer amplifying only a subset of the cleaved RNA molecules. This keeps the number of bands produced in each reaction within a manageable range and improves sensitivity. We found that a downstream primer with a single extra T was extended inefficiently. This problem was overcome by adding a second extra 3' nucleotide ($X = \text{TG, TA, TC, or TT}$; *see Fig. 1A*). The eight extra nucleotides at the 5' end of the downstream primers (GAGACCAG) improve the efficiency of reamplifying RNase T1-dependent bands. The 5' extension can be any sequence, but it is important to avoid palindromes (such as restriction sites) because they result in the production of dimeric molecules in which two PCR products are joined tail to tail.
11. I have found that Amplitaq Gold from Perkin-Elmer, a heat-activated enzyme, greatly improves the sensitivity of this procedure (compared with the original Amplitaq).
12. Since only 2 μL of cDNA is used in each PCR, the 1 mL of cDNA produced from 5 μg of RNA is sufficient for 500 reactions. This corresponds to using about 70 arbitrary primers coupled with the seven different downstream primers ($70 \times 7 = 490$).
13. The synthetic control RNA or internal positive controls (previously discovered ADAR substrates) can be used to monitor the success of the entire procedure. When amplifying a positive control, my colleagues and I mimic typical differential display conditions by using an upstream primer that is complementary to the control RNA at only its last seven nucleotides. The eighth position from the 3' end of the primer is randomized as in all our arbitrary primers (*see Fig. 1B*). We typically design the upstream primers so that they anneal 100–200 nt upstream of an inosine.
14. The gel in **Fig. 2** shows that the PCR products range in size from very small to about 300 bp. My colleagues and I ignore bands that are smaller than 100 nt because these may derive from contaminating tRNA. Bands often appear as doublets (sometimes triplets). We cut these out as a single gel slice because they usually represent the same sequence. The multiple bands may be owing to nontemplated nucleotides added to the 3' ends of the PCR products or to the two strands of the DNA migrating differently in the gel.
15. A candidate ADAR substrate should not be considered a false positive owing to the lack of a detectable secondary structure. It is possible that the required dsRNA is formed by intermolecular base pairing with an antisense transcript or the struc-

ture may be difficult to detect owing to multiple mismatches or a large distance between inverted repeats (see **ref. 28** for an example of the latter). Future experiments could reveal that the 5' nearest neighbor preferences for ADARs from some organisms differ from those observed to date.

Acknowledgments

This work was conducted in the laboratory of Dr. Brenda L. Bass in the Department of Biochemistry and the Howard Hughes Medical Institute at the University of Utah. D. Morse was supported by an National Institutes of Health training grant (CA 09602) and a postdoctoral fellowship from the American Cancer Society (PF 3891).

References

- 1 Bass, B. L. (2002) RNA editing by adenosine deaminases that act on RNA. *Annu. Rev. Biochem.* **71**, 817–846.
- 2 Schaub, M. and Keller, W. (2002) RNA editing by adenosine deaminases generates RNA and protein diversity. *Biochimie* **84**, 791–803.
- 3 Gerber, A. P. and Keller, W. (2001) RNA editing by base deamination: more enzymes, more targets, new mysteries. *Trends Biochem. Sci.* **26**, 376–384.
- 4 Seeburg, P. H. (1996) The role of RNA editing in controlling glutamate receptor channel properties. *J. Neurochem.* **66**, 1–5.
- 5 Burns, C. M., Chu, H., Rueter, S. M., Hutchinson, L. K., Canton, H., Sanders-Bush, E., and Emeson, R. B. (1997) Regulation of serotonin-2C receptor G-protein coupling by RNA editing. *Nature* **387**, 303–308.
- 6 Polson, A. G., Bass, B. L., and Casey, J. L. (1996) RNA editing of hepatitis delta virus antigenome by dsRNA-adenosine deaminase. *Nature* **380**, 454–456.
- 7 Morse, D. P. and Bass, B. L. (1999) Long RNA hairpins that contain inosine are present in *Caenorhabditis elegans* poly(A)+ RNA. *Proc. Natl. Acad. Sci. USA* **96**, 6048–6053.
- 10 Morse, D. P., Aruscavage, P. J., and Bass, B. L. (2002) RNA hairpins in non-coding regions of human brain and *Caenorhabditis elegans* mRNA are edited by adenosine deaminases that act on RNA. *Proc. Natl. Acad. Sci. USA* **99**, 7906–7911.
- 9 Brusa, R., Zimmermann, F., Koh, D. S., Feldmeyer, D., Gass, P., Seeburg, P. H., and Sprengel, R. (1995) Early-onset epilepsy and postnatal lethality associated with an editing-deficient GluR-B allele in mice. *Science* **270**, 1677–1680.
- 10 Higuchi, M., Maas, S., Single, F. N., Hartner, J., Rozov, A., Burnashev, N., Feldmeyer, D., Sprengel, R., and Seeburg, P. H. (2000) Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. *Nature* **406**, 78–81.
- 11 Palladino, M. J., Keegan, L. P., O'Connell, M. A., and Reenan, R. A. (2000) A-to-I pre-mRNA editing in *Drosophila* is primarily involved in adult nervous system function and integrity. *Cell* **102**, 437–449.

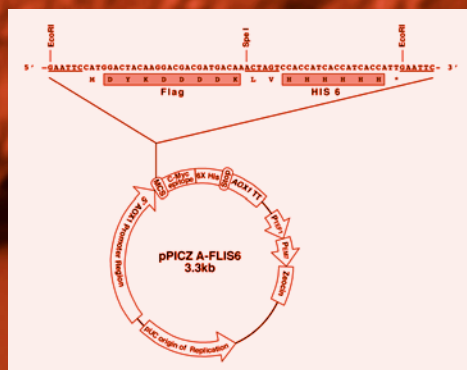
- 12 Tonkin, L. A., Saccomanno, L., Morse, D. P., Brodigan, T., Krause, M., and Bass, B. L. (2002) RNA editing by ADARs is important for normal behavior in *Caenorhabditis elegans*. *EMBO J.* **21**, 6025–6035.
- 13 Wang, Q., Khillan, J., Gadue, P., and Nishikura, K. (2000) Requirement of the RNA editing deaminase ADAR1 gene for embryonic erythropoiesis. *Science* **290**, 1765–1768.
- 14 Knight, S. W. and Bass, B. L. (2002) The role of RNA editing by ADARs in RNAi. *Mol. Cell* **10**, 809–817.
- 15 Paul, M. S. and Bass, B. L. (1998) Inosine exists in mRNA at tissue-specific levels and is most abundant in brain mRNA. *EMBO J.* **17**, 1120–1127.
- 16 Liang, P. and Pardee, A. B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* **257**, 967–971.
- 17 Morse, D. P. and Bass, B. L. (1997) Detection of inosine in messenger RNA by inosine-specific cleavage. *Biochemistry* **36**, 8429–8434.
- 18 Moore, M. J. and Sharp, P. A. (1992) Site-specific modification of pre-mRNA: the 2'-hydroxyl groups at the splice sites. *Science* **256**, 992–997.
- 19 Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- 20 Bantle, J. A., Maxwell, I. H., and Hahn, W. E. (1976) Specificity of oligo (dT)-cellulose chromatography in the isolation of polyadenylated RNA. *Anal. Biochem.* **72**, 413–427.
- 21 Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., and Smith, J. A. (1987) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York.
- 22 Broude, N. E. and Budowsky, E. I. (1971) The reaction of glyoxal with nucleic acid components. 3. Kinetics of the reaction with monomers. *Biochim. Biophys. Acta* **254**, 380–388.
- 22 Whitfeld, P. R. and Witzel, H. (1963) On the mechanism of action of Takadiastase ribonuclease T1. *Biochim. Biophys. Acta* **72**, 338–341.
- 27 Cameron, V. and Uhlenbeck, O. C. (1977) 3'-Phosphatase activity in T4 polynucleotide kinase. *Biochemistry* **16**, 5120–5126.
25. Greer, C. L. and Uhlenbeck, O. C., personal communication.
- 26 Melcher, T., Maas, S., Higuchi, M., Keller, W., and Seeburg, P. H. (1995) Editing of alpha-amino-3-hydroxy-5-methylisoxazole-4-propionic acid receptor GluR-B pre-mRNA in vitro reveals site-selective adenosine to inosine conversion. *J. Biol. Chem.* **270**, 8566–8570.
- 27 Polson, A. G. and Bass, B. L. (1994) Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase. *EMBO J.* **13**, 5701–5711.
28. Herb, A., Higuchi, M., Sprengel, R., and Seeburg, P. H. (1996) Q/R site editing in kainate receptor GluR5 and GluR6 pre-mRNAs requires distant intronic sequences. *Proc. Natl. Acad. Sci. USA* **93**, 1875–1880.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



Purification and Assay of Recombinant ADAR Proteins Expressed in the Yeast *Pichia pastoris* or in *Escherichia coli*

Gillian M. Ring, Mary A. O'Connell, and Liam P. Keegan

Summary

ADARs are found in Metazoans but are not present in yeasts. We have found that the methanol-utilizing yeast *Pichia pastoris* can be used to efficiently express enzymatically active epitope-tagged ADARs. We describe plasmid construction and protein expression procedures for producing *Drosophila* ADAR in this system.

ADAR expression in *Pichia pastoris* uses the methanol-inducible alcohol oxidase *AOX1* promoter for induction. A Zeocin resistance gene on the plasmid is used to select high copy number tandem integrations of the plasmid constructs. Preparation of extracts by grinding cultures in liquid nitrogen and purification protocols using $6 \times$ HIS and FLAG epitope tags are described. Procedures for preparing radiolabeled dsRNA and for assaying the non-specific RNA editing activity of ADARs are described.

ADARs produced in *Escherichia coli* are not enzymatically active. We describe expression of the ADAR dsRNA binding domains in *E. coli* using current versions of the T7 promoter based Studier vectors as well as the purification of the domains.

Key Words: ADAR; RNA-editing; RNA interference; deaminase; dsRNA; *Pichia pastoris*; *Drosophila melanogaster*; *Escherichia coli*; protein overproduction; protein purification; ion channel; nervous system.

1. Introduction

The adenosine deaminases that act on RNA (ADARs) are composed of two or three dsRNA-binding domains and a deaminase domain and have been found so far only in Metazoans (*see review in ref. 1*). Therefore, bacteria and yeasts are free of endogenous ADAR activities. Recombinant ADAR proteins were

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

initially expressed in *Escherichia coli*. This method is very efficient, resulting in large quantities of protein in a matter of hours, particularly when individual domains of ADARs are expressed. However, for unknown reasons, possibly related to posttranslational modifications that do not occur in bacteria, this system yields inactive ADAR proteins. Active ADAR proteins for structure and function studies have been obtained by expression in the methanol-utilizing yeast *Pichia pastoris* using the methanol-inducible promoter of the alcohol oxidase gene *AOX1* (2). ADARs are active after expression in *Saccharomyces cerevisiae* (3). ADARs have also been overexpressed and purified using the baculovirus system in insect cells (4,5). In this chapter, our work with the *Drosophila melanogaster* ADAR protein is used to illustrate the methods utilized for expression of full-length ADAR and individual domains of ADAR in *P. pastoris* (6). Expression of individual ADAR domains in *E. coli* and the subsequent extraction, purification, and assay of the proteins is also discussed.

2. Materials

1. Easysselect™ *Pichia* Expression Kit (Invitrogen, Carlsbad, CA).
2. pET-28a vector (Novagen).
3. hADAR1, hADAR2, and dADAR cDNAs.
4. pGEM-T Easy vector system (Promega, Madison, WI).
5. *P. pastoris* strain KM71H (genotype; *arg4*, *AOX1::arg4*).
6. *E. coli* strains DH5 α and BL21(DE3).
7. Oligonucleotide primers.
8. Expand™ Long Template PCR System (Roche, Indianapolis, IN).
9. Restriction enzymes, T4 DNA ligase (New England Biolabs).
10. Agarose and gel-running equipment.
11. Low-salt Luria Bertani (LB) medium.
12. LB medium.
13. Zeocin™ (Invitrogen).
14. Kanamycin.
15. Low-salt LB plates containing 25 μ g/mL of Zeocin.
16. LB plates containing 50 μ g/mL of kanamycin.
17. YPD: 1% yeast extract, 2% peptone, 2% glucose.
18. Electroporator.
19. 1 M Sorbitol.
20. YPDS plates containing 100, 500, or 1000 μ g/mL of Zeocin.
21. Buffered glycerol-complex medium (BMGY).
22. Buffered methanol-complex medium (BMMY).
23. 10X YNB: 13.4% yeast nitrogen base with ammonium sulfate without amino acids.
24. 500X B: 0.02% biotin.
25. 1 M Potassium phosphate buffer, pH 6.0.

26. 10X Gy: 10% glycerol.
27. 10X M: 5% methanol.
28. Baffled culture flasks.
29. Sterile cheesecloth.
30. Yeast extract and peptone.
31. 5X M9 salts: 64 g of Na_2HPO_4 , 15 g of KH_2PO_4 , 2.5 g of NaCl, 5.0 g of NH_4Cl . Make up to 1 L with dH_2O and autoclave.
32. 1 M CaCl_2 .
33. 40% Glucose.
34. 1 M MgSO_4 .
35. Ultraviolet (UV) spectrophotometer.
36. Isopropyl- β -D-thiogalactopyranoside (IPTG).
37. French press or mortar and pestle.
38. Protease inhibitors: 0.5 mM phenylmethylsulfonyl fluoride (PMSF) (resuspended in ethanol), 0.4 mg/mL of leupeptin, 0.7 mg/mL of pepstatin (resuspended in ethanol), and dithiothreitol (DTT) (omitted in some cases).
39. Sonication buffer for *E. coli*: 50 mM KH_2PO_4 , 500 mM NaCl, 20 mM imidazole plus protease inhibitors, 0.5 mM PMSF, 0.7 mg/mL of pepstatin, and 0.4 mg/mL of leupeptin (all added fresh).
40. Sonicator.
41. Buffer Q/X: X = mM KCl; 50 mM Tris-HCl, pH 7.9; 20% glycerol plus protease inhibitors (e.g., Buffer Q/100: 100 mM KCl, 50 mM Tris-HCl, pH 7.9; 20% glycerol; 0.5 mM PMSF; 0.7 mg/mL of pepstatin, and 0.4 mg/mL of leupeptin).
42. Chromatography equipment.
43. 250 mM Imidazole buffered to pH 7.9 with HCl.
44. Ni^{2+} -NTA resin (Qiagen, Valencia, CA).
45. Ni^{2+} -NTA column load buffer: buffer Q/200 + 10 mM imidazole, 0.5% Triton X-100.
46. Ni^{2+} -NTA column wash buffer: buffer Q/200 + 20 mM imidazole, 0.5% Triton X-100.
47. Ni^{2+} -NTA column elution buffer: buffer Q/200 + 250 mM imidazole, 0.5% Triton X-100.
48. Anti-FLAG M2 affinity purification resin (Sigma).
49. 0.1 M Glycine HCl, pH 3.5.
50. FLAG peptide (Sigma).
51. Centricon concentrators.
52. 4X Laemmli buffer.
53. High-molecular-weight protein standards for sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Bio-rad, Hercules, CA).
54. SDS-PAGE equipment.
55. Anti-tetra HIS antibody (Qiagen).
56. Anti-FLAG antibody (Sigma).
57. pBluescript KS-cat plasmid template for dsRNA synthesis.

58. T3 and T7 RNA polymerases (Stratagene).
59. Nucleotide solutions for unlabeled in vitro transcription: 20X stocks are 10 mM CTP, 10 mM UTP, 2 mM GTP, and 10 mM ATP.
60. 0.1 M CaCl₂.
61. DNase I (Roche).
62. 0.5 M EDTA.
63. 0.25 M EGTA.
64. tRNA (2.5 mg/mL).
65. RNA guard (Amersham Pharmacia Biotech, Piscataway, NJ).
66. 7.5 M Ammonium acetate.
67. Vacuum dryer.
68. 10X TBE.
69. FA-hybridization buffer: 80% deionized formamide; 40 mM Pipes-KOH, pH 6.4; 400 mM NaCl; 1 mM EDTA.
70. RNA elution buffer: 0.75 M ammonium acetate, 10 mM magnesium acetate, 0.1% SDS, 0.1 mM EDTA.
71. P1 nuclease (Roche).
72. Polygram Cel 300C thin-layer chromatography (TLC) plates, 20 × 20 cm (Macherey-Nagel).
73. Inosine 5' monophosphate (Sigma).
74. Chromatography tank for 20 × 20 cm plates (Fisher).
75. X-ray film (Kodak).
76. PhosphorImager and screens.
77. Scintillation counter.

3. Methods

The methods described outline construction of the plasmids for expression of full-length ADAR or individual domains in *P. pastoris* and in *E. coli*, expression of the proteins, protein extraction from *P. pastoris* and *E. coli*, purification of full-length ADAR or dsRNA-binding domains, and preparation of dsRNA substrate and its use in the nonspecific editing assay to determine ADAR activity.

3.1. Expression Plasmids

3.1.1. *P. pastoris* Expression Vector pPICZ A (3329 bp)

The Easyselect *Pichia* expression system (Invitrogen) was originally developed by scientists at Salk Institute Biotechnology/Industry Associates for high-level expression of recombinant proteins (7–9). Sequences inserted into the multiple-cloning site of pPICZ A (**Fig. 1**) are transcribed under the control of the methanol-inducible *AOX1* promoter. Downstream of the multiple-cloning site is the native transcription termination and polyadenylation signal from the *AOX1* gene. This allows efficient 3' mRNA processing and increased mRNA stability. The vector contains a Zeocin resistance gene (a bacterial bleomycin

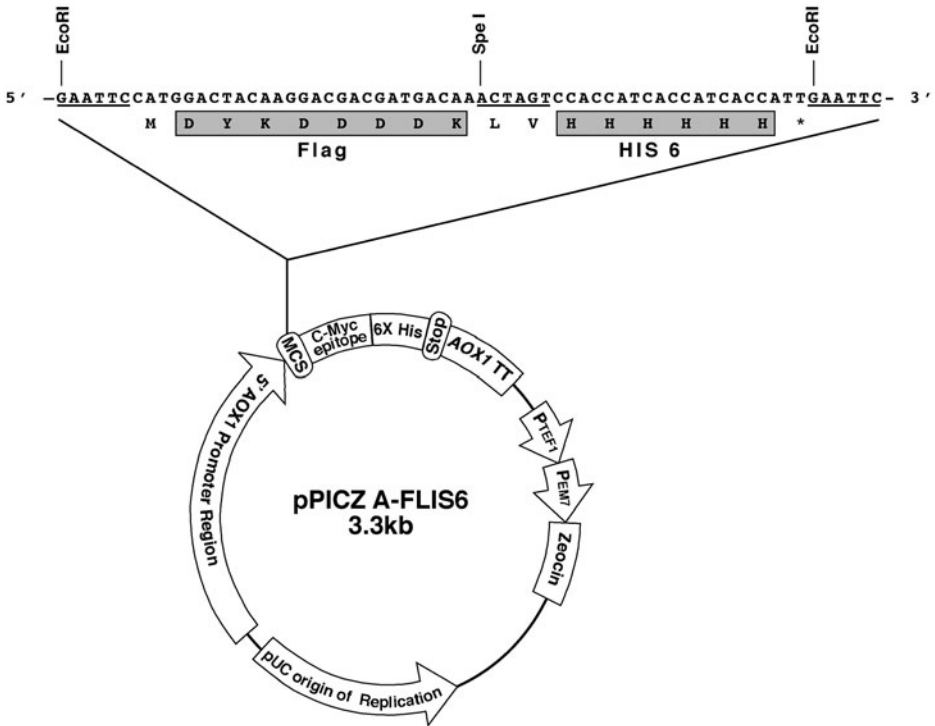


Fig. 1. Schematic of pPICZ A-FLIS6 constructed from the Invitrogen vector. (Copyright Invitrogen Corporation 2003. All rights reserved. Used with permission.)

resistance gene, *ble*) driven by an EM7 constitutive promoter for expression in *E. coli* and by a *TEF1* promoter from *S. cerevisiae* (GenBank accession nos. D12478 and D01130) for expression in *Pichia*. The presence of these different promoters allows for selection of Zeocin-resistant transformants in either *E. coli* or in *P. pastoris*.

Earlier *Pichia* expression vectors from Invitrogen were designed to integrate a single copy of the plasmid construct using gene replacement at the HIS4 locus. The Easyselect pPICZ vectors are designed to obtain multicopy transformants. For transformation into *P. pastoris* the vector is linearized by digestion at one of three unique restriction sites in the 5'AOX1 promoter region. Transformation occurs by integration of one or many copies of the linearized plasmid at the homologous sequence in the chromosomal AOX1 locus in the *Pichia* genome. Transformant lines can then be screened for those with the highest Zeocin resistance and the largest number of tandem integrations at the AOX1 locus (see **Note 1**).

3.1.2. *E. coli* Expression Vector pET28a

Studier and colleagues (**10,11**) constructed the original pET vectors. Recombinant protein is expressed from a T7 promoter in the pET vectors. Expression plasmids are constructed in a nonexpression host and subsequently introduced into an expression host that produces T7 RNA polymerase when induced with IPTG. The BL21(DE3) expression strain is lysogenic for the λ DE3 phage construct that expresses some T7 RNA polymerase from a partially constitutive *lacUV5* promoter even in the absence of IPTG. In addition to the *lac* repressor gene *lacI*, in the bacterial host chromosome the pET plasmids therefore carry additional copies of wild-type *lacI*. In later versions of the pET plasmids, such as pET-28a (**Fig. 2**), a *lac* operator sequence is also placed downstream of the T7 promoter to make a T7lac promoter. The *lac* repressor both represses transcription of T7 RNA polymerase under the control of the *lacUV5* promoter and acts at the plasmid T7lac promoter to block transcription by T7 RNA polymerase. Even with high *lac* repressor and the T7lac promoter, the basal level of expression in BL21(DE3) may prevent transformation of expression plasmids into this expression host if a recombinant protein is sufficiently toxic to *E. coli*. If this occurs, other T7 expression strains that also express lysozyme must be used to reduce T7 polymerase activity. To aid in purification of recombinant protein, the pET-28a vector allows for protein fusions to an N-terminal or to a C-terminal polyhistidine epitope tag. There is also a thrombin cleavage site between the multiple-cloning site and N-terminal polyhistidine epitope tag to allow removal of the epitope tag after purification. A kanamycin resistance gene allows plasmid selection in *E. coli*.

3.1.3. cDNAs

Full-length ADAR coding sequences were amplified using primers that added *SpeI* sites (see **Note 2**). The dsRNA-binding domain coding sequence comprising the first dsRNA-binding domain, linker, and second dsRNA-binding domain was also amplified from full-length *Drosophila* ADAR (*dADAR*) cDNA using oligos that added *NheI* and *SacI* restriction enzyme sites (see **Note 3**). Both polymerase chain reaction (PCR) fragments were subcloned into the pGEM-T Easy vector.

3.1.4. Cloning into pPICZ A

pPICZ A (Invitrogen) was used to construct a derivative called pPICZ A-FLIS6 (see **Fig. 1**) that contains a short open reading frame (ORF) encoding a FLAG epitope tag at the N-terminus and a 6 $\times$ HIS epitope tag at the carboxy terminus inserted at the *EcoRI* site in the pPICZ A multiple-cloning site. An *SpeI* site between the two epitope tags was used to clone full-length *Drosophila* ADAR coding sequence in frame to the epitope tags. The myc and 6 $\times$ HIS

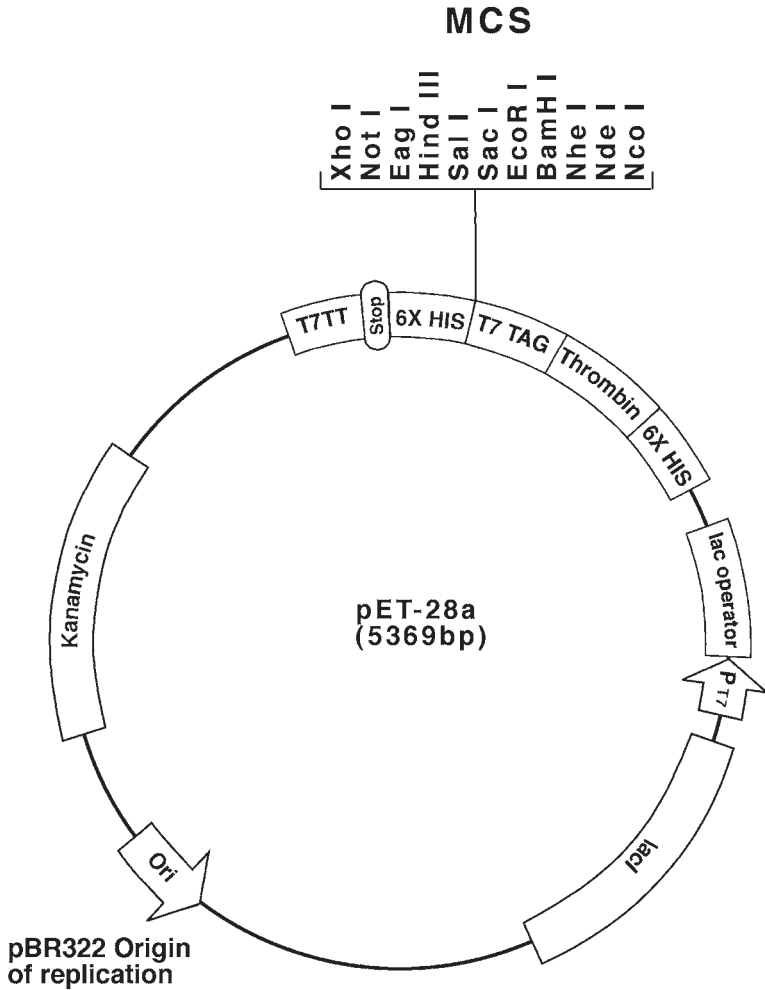


Fig. 2. Schematic of pET-28a vector adapted from EMD Biosciences, Novagen. (Permission to adapt this image was obtained from EMD Biosciences.)

epitope tag sequences present in the original pPICZ A vector are not expressed in this construct.

The DNA was transformed into *E. coli* DH5 α cells, plated on low-salt LB agar containing 25 μ g/mL of Zeocin, and incubated overnight at 37°C. Single colonies were selected and grown overnight in low-salt LB medium with Zeocin. Plasmid DNA was isolated and checked for inserts by restriction enzyme digestion. Positive clones were sequenced to confirm the presence of the correct ORF.

3.1.5. Cloning into pET-28a

PCR products encoding the dsRNA-binding domain were digested with *NheI* and *SacI* and ligated into the corresponding restriction enzyme sites of the digested pET-28a vector (**Fig. 2**). The ligated DNA was transformed in *E. coli* DH5 α cells and spread on an LB agar plate containing 50 μ g/mL of kanamycin. Plates were incubated overnight at 37°C. Single colonies were selected and grown overnight in 2 mL of LB medium containing 50 μ g/mL of kanamycin. Plasmid DNA was purified and an aliquot digested with *NheI* and *SacI* to check for inserts. Positive clones were sequenced to confirm that the construct was correct and that the correct ORF fusing ADAR to the carboxyl-terminal 6 $\times$ HIS epitope tag was present. Positive clones were transformed using standard methods into *E. coli* BL21(DE3) prior to protein expression.

3.1.6. Transformation of *P. pastoris* With pPICZ Plasmid Constructs

1. Digest 5–10 μ g of pPICZ A-FLIS6 *Adar* plasmid DNA with *SacI* or *PmeI* that cleaves within the *AOXI* promoter region 5' to the multiple-cloning site. Check that linearization is complete by electrophoresing a small aliquot on an agarose gel.
2. Extract the DNA with phenol/chloroform, and precipitate with 100% ice-cold ethanol and 3 *M* sodium acetate. Wash with 70% ethanol and dry the pellet. Resuspend in 10 μ L of dH₂O.
3. Prepare electrocompetent *Pichia*. Inoculate 5 mL of YPD with the *Pichia* strain KM71H. Grow overnight in a 50-mL conical flask at 30°C.
4. Inoculate 500 mL of fresh YPD with 1 mL of overnight culture, and grow in a 2-L flask to an OD₆₀₀ of 1.3–1.5.
5. Harvest the cells by centrifuging at 1500g for 5 min at +4°C. Wash twice with 500 mL of ice-cold sterile water and once with 20 mL of ice-cold 1 *M* sorbitol. Finally, resuspend the pellet in 1 mL of 1 *M* sorbitol to give a final volume of approx 1.5 mL (see **Note 4**).
6. Mix 5–10 μ g linearized DNA with 80 μ L of competent cells in an ice-cold 0.2-cm electroporation cuvet. Electroporate at 1.5k V, 25 μ F, 200 Ω (Bio-Rad Gene Pulser). Immediately add 1 mL of ice-cold 1 *M* sorbitol to the cuvet, transfer to a 15-mL sterile tube, and incubate without shaking at 30°C for 1 to 2 h.
7. Plate 25 and 100 μ L on separate YPDS plates containing 100 μ g/mL of Zeocin (see **Note 5**). Incubate for 2–5 d at 30°C until colonies form.

3.1.7. Screening for High Copy Number Transformed *P. pastoris* Lines

1. Pick 10–20 colonies from transformation plates and streak for single colonies on fresh YPDS agar plates with 100, 200, 500, and 1000 μ g/mL of Zeocin (see **Note 6**).
2. Incubate for 2 to 3 d at 30°C. Transformant lines showing the highest resistance to Zeocin should have the highest copy number of the expression construct. Transformant lines with a range of different copy numbers must then be tested by induction of small-scale cultures to determine which line gives the best expression of recombinant protein.

3.2. Protein Expression

The following protocols describe the growth, induction, and harvesting conditions required for both the *P. pastoris* pPICZ A and the *E. coli* pET-28a expression systems.

3.2.1. Expression of Recombinant ADARs in *P. pastoris*

1. Inoculate 5 mL of BMGY (see **Notes 7** and **8**) with a *Pichia* colony and grow overnight at 30°C, 300 rpm in a loosely capped 50-mL Falcon tube.
2. Inoculate 500 mL of fresh BMGY with the 5-mL starter culture. Grow overnight at 30°C on a rotary shaker at 300 rpm in a 2-L flask covered with sterile muslin to allow good aeration (see **Note 9**).
3. Once the OD₆₀₀ of the culture is 2–6 (approx 16–18 h), pellet the cells by centrifuging at 3500 rpm for 5 min at +4°C in a GSA rotor (Sorvall). Remove the BMGY supernatant and resuspend the pellet in BMMY (see **Notes 7** and **8**) to induce expression. Return the culture to the original flask and grow overnight at 30°C, 300 rpm covered with sterile muslin.
4. Add more methanol the following day (final concentration: 1%) to counter any losses owing to evaporation, and grow overnight again as in **step 3**.

3.2.2. Harvest of *P. pastoris* Containing Overexpressed Recombinant ADAR Protein

At this point a small aliquot of the culture can be taken to prepare an SDS-PAGE sample to test for expression of recombinant ADAR (see **Note 10**).

1. Spin each culture at 3500 rpm in a GSA 1500 rotor (Sorval) for 5 min at +4°C and remove the BMMY supernatant.
2. Wash the pellet with 250 mL of autoclaved ice-cold water (see **Note 11**) making sure the pellet is fully resuspended to wash out the BMMY. Spin as in **step 1** and remove the supernatant.
3. Add 40 mL of Buffer Q/200 to each pellet. Divide into two 50-mL Falcon tubes and pellet in a table top centrifuge at 3000 rpm (2000g) for 5 min at +4°C. Weigh the Falcon tubes and write the weight of the pellet on them. At this point either the cells can be broken and protein extracted or the pellet can be frozen in liquid nitrogen and stored at –70°C for an indefinite period of time.

3.2.3. Induction and Harvest of *E. coli* Cells Expressing dADAR dsRNA-Binding Domains

1. Transform pET-28a plasmid containing insert into electrocompetent BL21(DE3) *E. coli*. Plate on LB agar containing 50 µg/mL of kanamycin and incubate overnight at 37°C (see **Note 12**).
2. Pick a single colony and inoculate 10 mL of M9 medium containing 50 µg/mL of kanamycin (see **Note 13**). Grow overnight at 37°C.

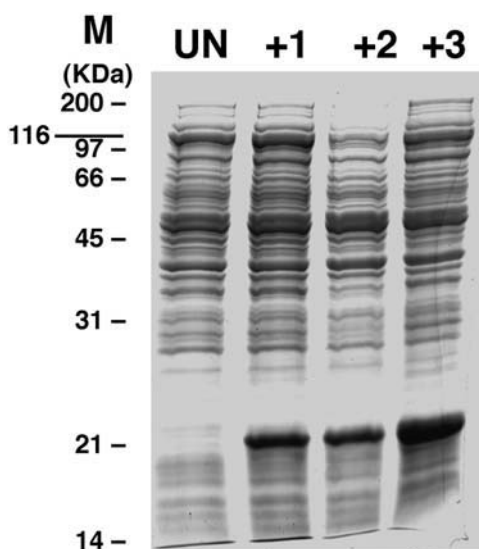


Fig. 3. Twelve percent SDS gel showing expression of 21-kDa *Drosophila* ADAR double-stranded RNA-binding domain in *E. coli*. Samples were taken at 1-, 2-, and 3-h intervals after induction with IPTG. Uninduced (UN) had no IPTG.

3. Inoculate 500 mL of fresh M9 medium containing 50 $\mu\text{g/mL}$ of kanamycin with the 10 mL of starter culture and continue to grow at 37°C.
4. Once the OD_{600} of the culture reaches 0.6, induce expression with 500 μL of 1 M IPTG. Continue growing for another 4 h (**Fig. 3**).
5. Harvest the cells by centrifuging the culture. At this point the cells can be broken immediately or stored at -70°C .

3.3. Protein Extraction

There are a variety of methods to extract protein from cells. It is important to choose a method that is efficient and will result in active, folded protein. To extract active ADAR protein from *P. pastoris*, we have found that breaking cells by grinding in liquid nitrogen provides the best protection against proteolysis.

3.3.1. Grinding *Pichia* in Liquid Nitrogen

1. Resuspend cells in 1 mL of buffer Q/200/g of cell pellet. Protease inhibitors should be freshly added since PMSF hydrolyzes in water.
2. Drop the cells into liquid nitrogen in a mortar or into a grinding machine (Fritsch) precooled with liquid nitrogen, and grind until the nitrogen has evaporated and the mix is a fine, dry powder (see **Note 14**).

3. Before the powder begins to thaw, add more liquid nitrogen and grind again. Grind in this way three times. At this point the frozen powder can be collected in 50-mL Falcon tubes and stored at -70°C .

3.3.2. Lysing *E. coli* by Sonication

1. Resuspend cells in 1 mL/g of sonication buffer with fresh protease inhibitors.
2. Sonicate the cells for 10 s and then incubate on ice for 20 s. Repeat this cycle three times.
3. Ultracentrifuge the samples and transfer the supernatant to a fresh tube. At this point it is recommended that purification be continued rather than freezing the lysate, to avoid proteolysis.

3.4. Protein Purification

The level of purity required depends on the intended use of the protein. ADARs used in activity and binding assays are often sufficiently pure after Ni^{2+} -NTA affinity chromatography. However, these proteins also have a FLAG epitope tag at the amino terminus, so the proteins can be purified further by elution from a FLAG column if required.

3.4.1. Ni^{2+} -NTA Affinity Chromatography of ADAR

1. Centrifuge ground *Pichia* cells for 15 min at 13,000 rpm and $+4^{\circ}\text{C}$ in an SS34 rotor (Sorvall).
2. Remove the supernatant and combine it with Ni^{2+} -NTA resin (Qiagen) previously equilibrated with buffer Q/200, and rotate gently for 1 to 2 h at $+4^{\circ}\text{C}$ (*see Note 15*). No EDTA can be present in this buffer because it would strip the nickel from the resin.
3. Load the resin mix onto a column (Bio-Rad) and collect the flowthrough. Pass over two times.
4. Wash the column twice with 10 mL of wash buffer each time (*see Note 15*) and collect fractions.
5. Elute the protein 10 times with 500 μL of elution buffer each time and collect the 10 fractions.
6. At each stage all fractions are collected and stored at -70°C . Take 50- μL aliquots of crude extract, flowthrough, wash, and each eluate and freeze so that the main aliquot need not be thawed to make samples for SDS-PAGE (**Figs. 4 and 5**) and activity assays (**Fig. 6**).

3.4.2. Anti-FLAG Affinity Column Purification of Recombinant ADARs

To purify further the eluate from the Ni^{2+} -NTA column, an anti-FLAG M2 affinity gel is used in which a purified murine monoclonal antibody is coupled to agarose. Protein is eluted with a FLAG peptide.

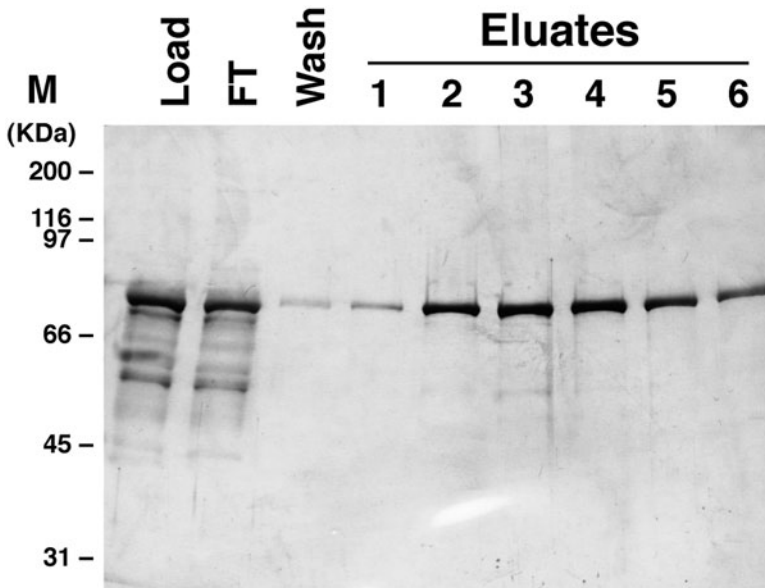


Fig. 4. Anti-FLAG affinity column purification of *Drosophila* ADAR (74 kD) extracted from *P. pastoris*. Fractions resolved on an 8% SDS gel were visualized by silver staining. FT, flowthrough.

1. Pour a minicolumn (0.7 cm diameter) containing a 150- μ L bed volume of anti-FLAG M2 matrix (Sigma) and equilibrate with buffer Q/200.
2. Wash the column by loading three sequential 5-mL aliquots of 0.1 M glycine HCl (pH 3.5) followed by three sequential 5-mL aliquots of buffer Q/200. Do not leave the column in glycine HCl for longer than 20 min.
3. Pool the fractions containing ADAR from the Ni^{2+} -NTA column and load on the anti-FLAG M2 affinity column. Dialysis is not required since the same buffer is used with both columns and the imidazole from the Ni^{2+} -NTA column does not interfere with binding to the FLAG matrix. Collect the FLAG eluate and pass it over the affinity column a further two times to maximize binding to the column.
4. Wash the column three times with 5-mL aliquots of buffer Q/200.
5. Elute the column five times with 500 μ L of elution buffer containing FLAG peptide at 100 μ g/mL in buffer Q/200 each time and collect five fractions.
6. Recycle the column by washing with three sequential 5-mL aliquots of 0.1 M glycine HCl (pH 3.5) followed by three sequential 5-mL aliquots of buffer Q/100 containing 0.2% sodium azide, and store the column in a cold room without draining the last of the buffer completely. The initial part of the glycine HCl wash may be retained in case there is ADAR present.
7. At each stage all fractions are collected and stored at -70°C . Take small aliquots of crude extract, flowthrough, wash, and each eluate to check on SDS-PAGE.

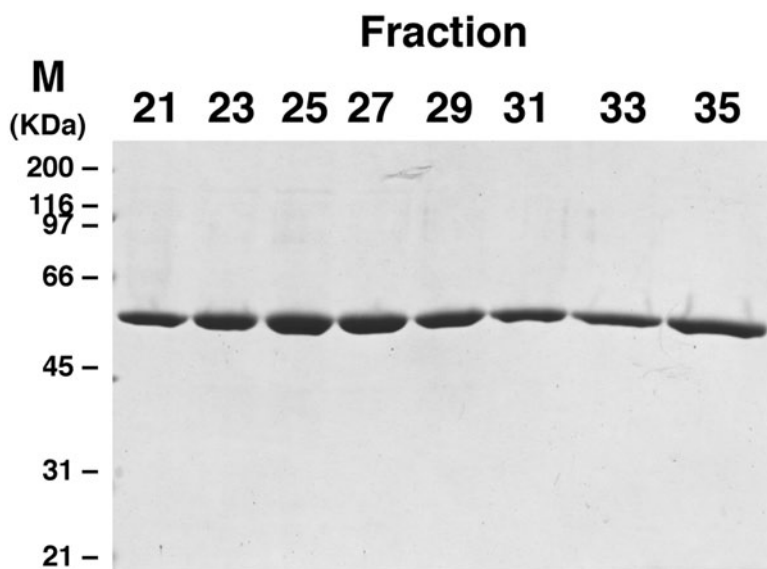


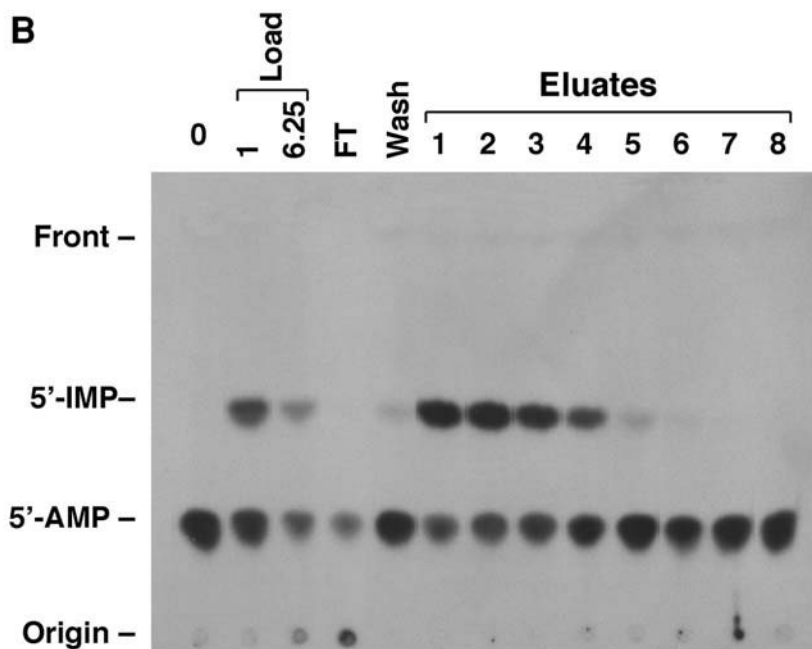
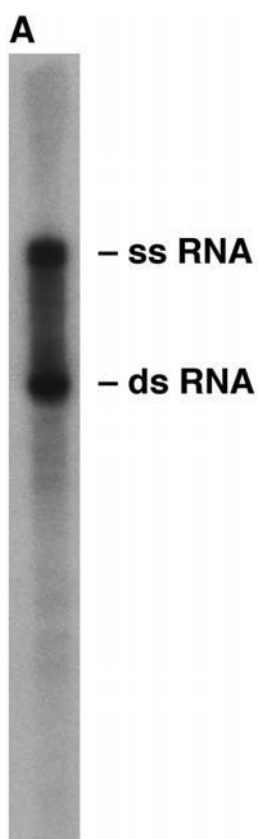
Fig. 5. Ni²⁺-NTA affinity chromatography of *Drosophila* ADAR adenosine deaminase domain extracted from *P. pastoris*. Samples were resolved on an 8% SDS gel and visualized by Coomassie blue staining.

3.4.3. Purification of dADAR dsRNA-Binding Domains From *E. coli*

1. Combine the supernatant with equilibrated Ni²⁺-NTA and rotate gently for 1 to 2 h at +4°C.
2. Load the resin mix onto a column (Bio-Rad) and collect the flowthrough. Pass over three times.
3. Wash the column twice with 50 mL of wash buffer each time and collect fractions.
4. Elute 40 times with 1.5 mL of elution buffer each time. Take aliquots of every second fraction and check on SDS-PAGE.
5. Concentrate the eluates in a centriprep to approx 2 mL and dialyze overnight at +4°C into 150 mM KCl; 50 mM Tris, pH 7.0; and 20% glycerol. A cation-exchange column may also be used to remove any nucleic acid contamination (see Note 16).

3.5. Nonspecific Editing Assay to Determine Activity of Purified ADAR Protein

The ADAR enzymes convert adenosine to inosine in dsRNA. Up to 50% of all the adenosines in the substrate can be deaminated. The RNA is then digested to completion with P1 nuclease, and the mononucleotides are resolved by TLC (12). The preparation of labeled dsRNA substrate and the nonspecific dsRNA adenosine deaminase assay are described here (13).



3.5.1. Preparation of [α - ^{32} P]-Labeled dsRNA Substrate

1. Linearize the pBluescript KS-cat plasmid with *Hind*III. Electrophorese the digest on an agarose gel and purify the linear product (see **Note 17**). Phenol/chloroform extract and ethanol precipitate. Transcribe the sense strand with T7 RNA polymerase using standard protocols to give a 605-nt transcript. Use 0.8–1 μ g of plasmid per transcription.
2. Linearize the pBluescript KS-cat plasmid with *Bam*HI and prepare the linearized template in the same way. Transcribe the antisense strand with T3 RNA polymerase to give a 594-nt transcript.
3. One strand is internally labeled using 4 μ L of [α - ^{32} P]ATP (3000 Ci/mmol). For this transcription, use 1 μ L each of 10 mM CTP, 10 mM UTP, 2 mM GTP, and 2 mM ATP in a 20- μ L reaction. The ATP concentration is 10 mM in the unlabeled transcriptions. Retain 1 μ L of the labeling reaction and count the input radioactivity in a scintillation counter (see **Note 17**).
4. Treat the 20- μ L transcription reaction with DNaseI. Add the following to give a final volume of 50 μ L: 2 μ L of RNA guard, 2.5 μ L of 0.1 M CaCl₂, 6 μ L of DNaseI, 19 μ L of water. Incubate for 15 min at 37°C.
5. Combine labeled and unlabeled strands (50 μ L each). Add 8 μ L of 0.5 M EDTA, 8 μ L of 0.25 M EGTA, 80 μ L of water, and 4 μ L of 5 mg/mL tRNA to give a volume of 200 μ L. Phenol/chloroform extract. To the supernatant add 100 μ L of 7.5 M ammonium acetate and 750 μ L of ethanol. Centrifuge for 30 min at 13,000 rpm (18,000g) in a cooled Eppendorf centrifuge. Wash the pellet twice with 70% ethanol stored at -20°C. Vacuum dry the pellet for 15 min.
6. Resuspend the pellets in 200 μ L of FA hybridization buffer. Heat to 85°C for 5 min, allow to cool slowly on a heating block to 45°C, and anneal at this temperature for 15 h or overnight.
7. Add 600 μ L of ethanol to each 200- μ L annealing reaction. Centrifuge for 30 min at 13,000 rpm (18,000g) in a cooled Eppendorf centrifuge. Wash the pellet twice with 70% ethanol stored at -20°C. Air-dry the pellet for 15 min.

Fig. 6. (see opposite page) Preparation of radiolabeled dsRNA substrate and non-specific dsRNA adenosine deaminase assay. (A) Autoradiograph (5-min exposure) of a 4% nondenaturing polyacrylamide gel resolving annealed dsRNA from labeled and unlabeled ssRNAs. (B) dsRNA adenosine deaminase assays on Ni²⁺-NTA affinity column fractions of *Drosophila* ADAR extracted from *Pichia*. TLC-separated labeled mononucleotides are indicated. Phosphate release by activities in crude fractions appears as a decrease of total label in the lane, and α - ^{32} P may be seen migrating with the solvent front. Because of ribonucleases or other activities in the load, flowthrough (FT), and wash fractions, 1 μ L of load fraction is preferable to 6.25 μ L in this assay. The eluate fractions with 250 mM imidazole were assayed using 6.25 μ L of fraction per assay. The P1 nuclease digestion in this assay was very complete since there was sometimes more label remaining at the origin.

8. Resuspend dsRNA in 10 μ L of TE buffer (10 mM Tris, pH 7.9; 1 mM EDTA) + 1% bromophenol blue, 1% xylene cyanol, and 10% glycerol. Separate the annealed dsRNA from residual ssRNA on a nondenaturing 4% acrylamide ([80:1] acrylamide:bisacrylamide) gel (200 V, 3 h with 0.5X TBE) (see **Note 18**).
9. Expose the gel to X-ray film for 0.5–5 min and excise the labeled dsRNA band (**Fig. 6A**). Count the slice in a scintillation counter to estimate recovery from the band later. Break up the acrylamide and incubate overnight at 37°C in 0.4 mL of RNA elution buffer.
10. Collect as much elution buffer as possible, add 4 μ L of 2.5 mg/mL tRNA, phenol/chloroform extract, scintillation count an aliquot and ethanol precipitate. Recover all of the dsRNA by repeating the elution overnight with another 400 μ L of elution buffer. The dsRNA is safely stored as an ethanol precipitate at –20°C. Measure the proportion of the input [α -³²P]ATP that has been incorporated into dsRNA by scintillation counting an aliquot of each dsRNA eluate (see **Note 19**).

3.5.2. Nonspecific dsRNA Editing Assay for ADAR Activity

1. Calculate the amount of radiolabeled dsRNA to set up the required number of reactions, including one reaction with no ADAR protein, using 200 fmol of adenosine residues per reaction. Take this quantity of ethanol-precipitated dsRNA, centrifuge, wash the pellet with cold 70% ethanol, and dry the pellet. The final reaction volume is 25 μ L and contains 5 mM EDTA, 0.15 mg/mL of tRNA, 0.2 mg/mL of bovine serum albumin, 0.5 μ L of RNA guard, and 200 fmol of dsRNA. Set up 12.5- μ L protein samples containing 1–12.5 μ L of ADAR-containing protein fraction on ice; buffer Q/200 is used to make up the remaining volume. Add 12.5 μ L of resuspended dsRNA substrate and reaction mix to each reaction. Incubate the reactions for 1 h at 37°C.
2. Add 8.3 μ L of 7.5 M ammonium acetate and 300 μ L of ethanol, centrifuge for 30 min at 4°C, wash the pellets with 70% ethanol, and vacuum dry the pellets. Do not overdry the pellets or the nuclease digestion may be incomplete at the next step.
3. Resuspend the pellets in 10 μ L of P1 buffer (30 mM potassium acetate, pH 5.3; 10 mM zinc sulfate) containing 1.5 μ L of 1 mg/mL nuclease P1 for 1 h at 50°C.
4. Prepare a 20 $\times$ 20 cm TLC plate by marking an origin line lightly with a pencil 1.5 cm from the bottom of the plate. Then create vertical lanes (15 lanes 1.3 cm wide or 13 lanes 1.5 cm wide, as in **Fig. 6B**) by using the edge of a spatula to remove the cellulose layer between the lanes. Spot 1 μ L of 20 mg/mL unlabeled inosine at the origin of each lane and allow to dry.
5. Load 5 μ L of P1 nuclease digest at the origin of each lane on the TLC plate, allow to dry and load the second 5 μ L. Resolve the nucleotide mixture using 100 mL of chromatography solvent (saturated $[\text{NH}_4]_2\text{SO}_4$; 100 mM sodium acetate, pH 6.0; and isopropanol [79:19:2]). The spotted sample must be above the solvent. It is not necessary for the solvent front to reach the top of the plates, and usually 3 h is sufficient for a good separation of the nucleotides.

6. Air-dry the plate and expose overnight to X-ray film or to a PhosphorImager screen. 5' α - ^{32}P inosine migrates faster in this system than 5' α - ^{32}P adenosine (**Fig. 6B**). View the unlabeled inosine marker under UV light to verify the position of inosine. Quantitate the radioactivity in the inosine and adenosine spots in each lane either by using the PhosphorImager or by cutting out the area bearing the nucleotide and counting with scintillation fluid in a scintillation counter (*see Note 20*).

4. Notes

1. Since there is no yeast origin of replication in the pPICZ A vector, Zeocin-resistant transformants can only be isolated if recombination occurs between the plasmid and the *Pichia* genome. For every integrated copy of the expression construct, there is a single copy of the Zeocin gene integrated; therefore, the greater the resistance to Zeocin the more copies are integrated.
2. To maximize the translational efficiency of recombinant ADAR proteins in *Pichia*, the initiation ATG should conform with the yeast Kozak consensus sequence: A/YAA/TAATGTCT.
3. In the pET-28a vector, the *Bam*HI site should be avoided because *Bam*HI displays high levels of star activity. For cloning ADARs into the *Spe*I site in pPICZ A-FLIS6, it is helpful, even if the particular ADAR coding sequence does not necessitate it, to add *Nhe*I or *Xba*I sites instead of *Spe*I sites to the PCR products because these produce 5' overhangs compatible with *Spe*I and the ligation can be digested with *Spe*I to select for inserts.
4. Competent *Pichia* cells should be stored on ice and ideally used the same day. However, if stored at +4°C, they can be used 2 or 3 d after they have been made without any noticeable loss in transforming efficiency.
5. All Zeocin plates should be stored in a cool, dark place because the antibiotic breaks down quickly in light. It is best to pour plates as they are required to ensure that the Zeocin is as efficient as possible. Approximately 100 mL of melted YPDS agar is sufficient for three plates.
6. When streaking out colonies to select for multicopy recombinants, always streak out a 100 $\mu\text{g}/\text{mL}$ Zeocin plate in case higher concentrations prove to be toxic.
7. BMGY and BMMY are made up as follows: 1% yeast extract; 2% peptone; 100 mM potassium phosphate, pH 6.0; 1.34% YNB; $4 \times 10^{-5}\%$ biotin; 1% glycerol (for BMGY) or 0.5% methanol (for BMMY). For practical reasons, it is useful to make up these buffers in 5-L batches. Fifty grams of yeast extract and 100 g of peptone are dissolved in 3 L of dH_2O , split into ten 500-mL bottles (300 mL per bottle), and autoclaved. These can be stored at room temperature for long periods of time. For BMGY add to each bottle (*see Note 8*) 50 mL of 1 M potassium phosphate buffer, pH 6.0; 50 mL of 10X YNB; 1 mL of 500X B; 50 mL of 10X GY; 500 μL of 50 mg/mL kanamycin; and ice-cold dH_2O up to 500 mL. For BMMY, add 50 mL of 10X M instead of 10X GY.
8. Stock solutions for BMGY and BMMY are as follows: 10X YNB (dissolve 67 g of yeast nitrogen base with ammonium sulfate in 500 mL of dH_2O and filter ster-

ilize); 500X B (dissolve 20 mg of biotin in 100 mL of dH₂O and filter sterilize); 10X GY (add 100 mL of glycerol to 900 mL of dH₂O and autoclave), 10X M (add 5 mL of methanol to 95 mL of dH₂O and filter sterilize); 1 M potassium phosphate buffer, pH 6.0 (combine 132 mL of 1 M K₂HPO₄ and 868 mL of 1 M KH₂PO₄, adjust to pH 6.0 using phosphoric acid or KOH, and autoclave). All these solutions should be stored at +4°C.

9. Ensure that the flasks are at least four times the volume of the culture since good aeration is critical to obtaining strong induction of the *AOX1* promoter. The use of muslin to cover the top of the culture flask is particularly recommended as cultures grow to a higher OD₆₀₀. Baffled flasks that have a nonflat base may also be used to improve aeration of cultures.
10. Before breaking the entire pellet from large cultures and purifying protein, it is advisable to check that the induction was successful. When harvesting cultures remove 2 mL of culture and spin down the cells. Resuspend in 200 µL of breaking buffer (50 mM sodium phosphate, pH 7.4, 1 mM EDTA, 5% glycerol, 1 mM PMSF). Add an equal volume of acid-washed glass beads (size: 0.5 mm). Vortex for 30 s and then incubate on ice for 30 s. Repeat for a total of eight cycles. Centrifuge at maximum speed for 10 min at 4°C and transfer the supernatant into a clean tube. Take 30 µL of supernatant and mix with 10 µL of 4X Laemmli loading buffer. Boil for 5 min at 95°C and load 10–20 µL into each well. Protein expression is analyzed by Western blot analysis with anti-FLAG antibody at a 1:3000 dilution.
11. It is recommended that when working with *Pichia* a supply of cold autoclaved dH₂O for making buffers be kept on hand.
12. pET-28a expression constructs must always be transformed freshly into BL21(DE3). Problems can be encountered with some batches of IPTG that give ineffective inductions, so it must be clear that transformants are fresh and should induce properly. If the BL21(DE3) transformants are stored on plates, there is a risk of selecting promoter or other mutations that reduce protein expression. For this reason also plasmid DNA should never be recovered from the expression strain. Using nonexpressing strains for plasmid construction and maintenance is one of the advantages the T7 expression system has over systems that use *E. coli* promoters because it prevents any selection against foreign protein expression during plasmid construction. In the past, TATA box mutations in *tac* promoters were readily obtained by careless handling of expression transformants.
13. M9 medium is made as follows: 200 mL of 5X M9 salts, 2 mL of 1 M MgSO₄, 100 mL of 1 M CaCl₂, 10 mL of 40% glucose made up to 1 L with water.
14. It is extremely important that the cells do not thaw during grinding. It is advisable to store the mortar and pestle at –70°C and place on ice when grinding. Pour nitrogen in the mortar repeatedly until the ice is solid and the nitrogen does not boil off quickly. The frozen ice also facilitates grinding by holding the mortar in position. The grinding machine is very helpful for the first round of grinding, but we then grind further in a mortar and pestle. The necessary amount of grinding is determined by experience; cells that remain unbroken will be clear in the bottom

of the pellet when the mixture is centrifuged. An alternative method of breaking *Pichia* cells is to pass them over a French press three times. This is a more efficient method than grinding, with a much greater yield of protein, but it does carry an added risk of the protein being degraded since the cells do not remain frozen.

15. Load, wash, and elution buffers for Ni^{2+} -NTA columns will vary depending on whether full-length ADAR or a domain is being purified. Full-length ADARs require 1 mM DTT for activity and the buffers contain 200 mM KCl, 50 mM Tris (pH 7.9), and 20% glycerol. However, when purifying domains higher salt concentrations up to 500 mM KCl are used. DTT is omitted because it can interfere with binding to Ni^{2+} -NTA, and 0.5% Triton and small amounts of imidazole are used in load and wash buffers to prevent nonspecific binding. Protease inhibitors are added to all buffers fresh each time.
16. The dsRNA-binding motifs of dADAR have a high affinity for dsRNA, so it is advisable to run a cation-exchange column to remove any dsRNA contamination.
17. Any dsRNA over 30 bp can be used in a nonspecific editing assay. However, for efficient editing activity, dsRNA over 100 bp is recommended. The template for the dsRNA that we use in the nonspecific assay is transcribed from the pBS-cat plasmid that contains a portion of the chloramphenicol acetyltransferase coding sequence cloned into the polylinker of Bluescript KS (**13**). It is important to remove any undigested plasmid because this is a favored transcription template and gives a heterogeneous mixture of runaround transcription products. Failed transcriptions are most often the result of deterioration of the linearized template.
18. It is worthwhile to make both labeled and unlabeled transcripts of each strand and combine each labeled strand with the unlabeled transcript of the other strand.
19. An unannealed sample of the labeled strand will aid in identifying the correct dsRNA band that migrates faster than the single strand. This particular dsRNA migrates close to the xylene cyanol marker on these gels.
20. The labeled strand is transcribed in the presence of unlabeled ATP (2000 pmol of ATP and 25 pmol of $[\alpha\text{-}^{32}\text{P}]\text{ATP}$ when 4 μL of label is used). Using 200 fmol of incorporated adenosine per reaction, the yield of dsRNA should be sufficient for at least 500 assays at 6000–8000 cpm/assay. At least 2000 cpm/assay is needed to see the result from an overnight exposure of the TLC plates. It would be possible to obtain much more highly labeled RNA by reducing the concentration of ATP in the transcription reactions. However, characterization of purified ADAR1 showed that a minimum of 200 fmol of incorporated adenosine is required so that the substrate concentration in the assay (8 nM [see **Note 21**]) is above the k_m .
21. Only 50% of the adenosines are converted to inosine at saturation, so if there is 50% conversion in any lane, less ADAR must be used. If the sample has not saturated the assay, calculate the number of activity units where 1 ADAR unit is sufficient to convert 1 fmol of adenosine to inosine/min.

Acknowledgments

We wish to thank Walter Keller, Biozentrum of Basel University, in whose laboratory *Pichia* overexpression of these proteins was initiated; and Jose Gal-

lego, MRC Laboratory of Molecular Biology, Cambridge, for assistance with expression and purification of individual ADAR domains in *E. coli*.

References

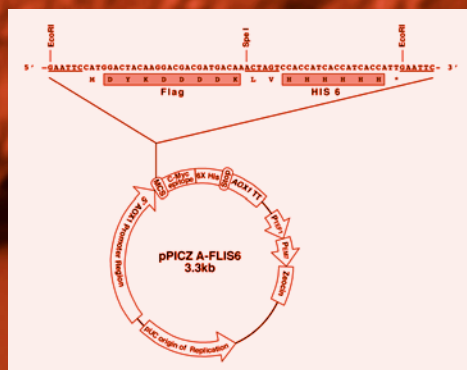
- 1 Keegan, L. P., Gallo, A., and O'Connell, M. A. (2001) The many roles of an RNA editor. *Nat. Rev. Genet.* **2**(11), 869–878.
- 2 O'Connell, M. A., Gerber, A., and Keegan, L. P. (1998) Purification of native and recombinant double-stranded RNA-specific adenosine deaminases. *Methods* **15**, 51–62.
- 3 Ohman, M., Kallman, A. M., and Bass, B. L. (2000) In vitro analysis of the binding of ADAR2 to the pre-mRNA encoding the GluR-B R/G site. *RNA* **6**(5), 687–697.
- 4 Lai, F., Chen, C. X., Lee, V. M., and Nishikura, K. (1997) Editing of glutamate receptor B subunit ion channel RNAs by four alternatively spliced DRADA2 double-stranded RNA adenosine deaminases. *Mol. Cell. Biol.* **17**, 2413–2424.
- 5 Herbert, A., Wagner, S., and Nickerson, J. A. (2002) Induction of protein translation by ADAR1 within living cell nuclei is not dependent on RNA editing. *Mol. Cell* **10**(5), 1235–1246.
- 6 Palladino, M. J., Keegan, L. P., O'Connell, M. A., and Reenan, R. A. (2000) dADAR, a *Drosophila* double-stranded RNA-specific adenosine deaminase is highly developmentally regulated and is itself a target for RNA editing. *RNA* **6**, 1004–1018.
- 7 Cregg, J. M., Barringer, K. J., Hessler, A. Y., and MAdden, K. R. (1985) *Pichia pastoris* as a host system for transformations. *Mol. Cell. Biol.* **5**(12), 3376–3385.
- 8 Cregg, J. M., Vedvick, T. S., and Raschke, W. C. (1993) Recent advances in the expression of foreign genes in *Pichia pastoris*. *Biotechnology* (NY) **11**(8), 905–910.
- 9 Wegner, G. H. (1990) Emerging applications of the methylotrophic yeasts. *FEMS Microbiol. Rev.* **7**(3–4), 279–283.
- 10 Studier, F. W. and Moffatt, B. A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**(1), 113–130.
- 11 Studier, F. W., et al. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89.
- 12 Bass, B. L. and Weintraub, H., (1988) An unwinding activity that covalently modifies its double-strand RNA substrate. *Cell* **55**, 1089–1098.
- 13 O'Connell, M. A. and Keller, W., (1994) Purification and properties of double-stranded RNA-specific adenosine deaminase from calf thymus. *Proc. Natl. Acad. Sci. USA* **91**, 10,596–10,600.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



Isolation of an mRNA-Binding Protein Involved in C-to-U Editing

Carri A. Gerber, Anne Relich, and Donna M. Driscoll

Summary

This chapter describes the technique of RNA affinity chromatography, which is a powerful approach for isolating RNA-binding proteins. This method takes advantage of the fact that sequence-specific RNA-binding proteins often bind their targets with high affinity. Here we outline a protocol for purifying Apobec-1 complementation factor (ACF), the RNA-binding subunit of the apolipoprotein-B (apo-B) mRNA-editing enzyme. ACF was purified using synthetic wild-type and mutant apo-B RNAs, which were coupled to cyanogen bromide (CNBr)-activated Sepharose. The methods are plasmid construction for in vitro transcription, affinity chromatography column preparation, protein purification by RNA affinity chromatography, and analysis of the purified protein.

Key Words: RNA editing; RNA-binding protein; affinity chromatography; Apobec-1 complementation factor.

1. Introduction

RNA affinity chromatography is an extremely useful method for isolating RNA-binding proteins, including proteins involved in RNA editing (*1–5*). There are multiple advantages to this approach. The protein of interest is likely to have a high affinity for its RNA ligand so that substantial purification may be achieved in a single step. This is in contrast to conventional biochemical purification on ion-exchange resins, heparin agarose, or phosphocellulose, which may not separate the protein from other nucleic acid-binding proteins that share similar properties. Furthermore, a mutant RNA that does not bind to the factor of interest can be used to control for specificity. One may be able to identify a

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

specific polypeptide that elutes from the wild-type RNA affinity column but not from the mutant RNA column (4,5). In this case, the protein does not need to be purified to homogeneity before submitting the protein band for microsequence analysis by mass spectrometry (6,7).

Our laboratory used the method of RNA affinity chromatography to purify Apobec-1 Complementation Factor (ACF), a protein that functions as the RNA-binding subunit of the enzyme that edits apolipoprotein-B (apo-B) mRNA. The isolation of ACF is used here to illustrate the protocols for generating wild-type and mutant RNA affinity columns, purifying proteins by RNA affinity chromatography, and analyzing the purified fractions. In our initial studies, we used RNAs labeled with biotinylated nucleotide analogs coupled to streptavidin-coated paramagnetic beads (4). This method can be expensive owing to the cost of the modified nucleotides and paramagnetic beads. More recently, we have utilized RNA coupled to cyanogen bromide (CNBr)-activated Sepharose to purify ACF as well as other RNA-binding proteins ([6], unpublished data). The CNBr-activated Sepharose-RNA method, which is simpler and more economical than the biotin-streptavidin method, is described in this chapter.

2. Materials

1. pGEM3Zf(+) vector (Promega, Madison, WI).
2. Restriction enzymes.
3. Phenol:chloroform:isoamyl alcohol (25:24:1).
4. Sterile water (*see Note 1*).
5. RiboMAX™ Large Scale RNA Production System-T7 (Promega).
6. 25 mM rNTP mix: 25 mM rATP, 25 mM rCTP, 25 mM rGTP, 25 mM rUTP.
7. RNase-free DNase I (Roche, Indianapolis, IN).
8. Micro Bio-Spin columns with Bio-Gel 30 in Tris buffer (Bio-Rad, Hercules, CA).
9. Agarose (Invitrogen, Carlsbad, CA).
10. 1X TAE buffer: 40 mM Tris-acetate; 1 mM EDTA, pH 8.0.
11. Ethidium bromide (EtBr) (10 mg/mL) (Sigma, St. Louis, MO).
12. 6X RNA loading buffer: 50% (w/v) glycerol, 1 mM EDTA, pH 8.0, 0.04% (w/v) bromophenol blue, 0.04% (w/v) xylene cyanol.
13. RNA or DNA molecular weight markers (Fermentas, Hanover, MD).
14. CNBr-activated Sepharose 4B (Amersham Pharmacia Biotech, Piscataway, NJ).
15. 2X Coupling buffer: 0.4 M 2-(N-morpholino)ethanesulfonic acid (MES), pH 6.0, at 4°C.
16. 1X Coupling buffer: 0.2 M MES, pH 6.0, at 4°C.
17. Binding buffer: 20 mM HEPES, pH 7.9, 25 mM KCl, 20% glycerol, 0.2 mM EDTA, 0.5 mM dithiothreitol at 4°C.
18. Partially purified baboon kidney extract (*see Note 2*) (4).
19. Elution buffer 1: binding buffer with 0.2 M NaCl at 4°C.
20. Elution buffer 2: binding buffer with 0.5 M NaCl at 4°C.

21. Elution buffer 3: binding buffer with 1 M NaCl at 4°C.
22. Poly-Prep Chromatography Columns (Bio-Rad).
23. Laboratory equipment: agarose gel electrophoresis equipment, vacuum filtration device with 0.2- μ m-pore-size filter (Millipore, Bedford, MA), and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) equipment.

3. Methods

The methods described next outline construction of the apo-B 100 wild-type and mutant plasmids for in vitro transcription, preparation of the affinity chromatography column for protein purification, purification of ACF by RNA affinity chromatography, and analysis of the purified protein.

3.1. Preparation of Expression Plasmid and Large-Scale Synthesis of RNA

Described in **Subheadings 3.1.1.–3.1.6.** are the steps to construct the expression plasmids and perform large-scale RNA synthesis. This protocol will generate a sufficient amount of synthetic RNA for affinity column preparation.

3.1.1. pGEM3Zf(+) Plasmid

The pGEM3Zf(+) vector (Promega) is a standard cloning vector that contains both SP6 and T7 RNA polymerase promoters. Although this vector is very effective for producing RNA from both promoters, other expression vectors may be used (*see Note 3*). For our studies, a 280-bp fragment of baboon apo-B 100 cDNA (nucleotides 6504–6784) was cloned into the *EcoRI/HindIII* site of pGEM3Zf(+) under the control of the T7 promoter using standard recombinant DNA methods (*see Note 4*). This fragment was chosen because it encompasses the apo-B editing site and contains all of the sequence required to support editing in vitro. To control for specificity, we also constructed a mutant apo-B 100 plasmid by standard site-directed mutagenesis (**4**). The mutant RNA, which contains three point mutations, binds ACF with very low affinity (*see Note 5*).

3.1.2. Plasmid Linearization

Using restriction enzymes, the wild-type and mutant plasmid DNAs are linearized for in vitro transcription.

1. Linearize 50 μ g of the wild-type and mutant apo-B 100 constructs from **Subheading 3.1.1.** with *HindIII*, which cuts in the 3' part of the polylinker but not within either cDNA (*see Note 6*).
2. Purify the DNAs by standard phenol:chloroform extraction and ethanol precipitation. Resuspend each DNA in 200 μ L of sterile water at 250 ng/ μ L.

3. Analyze 1 μL of each linearized sample by agarose gel electrophoresis. Confirm plasmid linearization by comparison to uncut DNA. On a 1% agarose gel, the linearized DNA generally runs slower than the uncut supercoiled plasmid. If uncut molecules of DNA remain, run-on transcripts may be created during RNA synthesis. Recut the DNA if the digestion has not gone to completion.

3.1.3. Large-Scale Synthesis of RNA

The method of RNA synthesis described here is adapted from the protocol for the RiboMAX Large Scale RNA Production System–T7 (Promega).

1. Assemble the reactions in 1.5-mL microcentrifuge tubes at room temperature.
2. For each DNA template, combine 100 μL of 5X T7 transcription buffer, 150 μL of 25 mM rNTP mix, 200 μL of linearized template DNA (50 μg total), and 50 μL of T7 enzyme mix for a final reaction volume of 500 μL .
3. Gently mix the reactions by pipetting and incubate at 37°C for 2–4 h.

3.1.4. Removal of DNA Template After Transcription

The DNA template is removed after transcription by digestion with DNase I.

1. Add 50 U of RNase-free DNase I (Roche) to each reaction.
2. Incubate for 20–30 min at 37°C.
3. Extract RNA by phenol:chloroform extraction.
4. To remove unincorporated nucleotides, pass the aqueous phase from the phenol:chloroform extraction over a Micro Bio-Spin column with Bio-Gel 30 in Tris buffer (Bio-Rad).
5. Store the RNA at -80°C . Since repeated handling can lead to degradation, we recommend aliquoting the RNA into smaller amounts before storage.

3.1.5. Determination of RNA Concentration Using Ultraviolet Spectrophotometry

The RNA concentrations can be quantitated by ultraviolet (UV) light absorbance.

1. For both RNAs, prepare a 1:200 dilution in water.
2. Read the absorbance at a wavelength of 260 nm. One A_{260} unit equals $\sim 40 \mu\text{g/mL}$ of RNA.
3. To determine purity of sample, take a wave scan of 260–280 nm. The ratio of A_{260}/A_{280} should be ≥ 1.8 , which indicates a lack of protein contamination in the sample (*see Note 7*).

A 500- μL in vitro transcription reaction containing 50 μg of linearized DNA generally yields 2 to 3 mg of RNA (*see Note 8*).

3.1.6. Visualization of RNA Using Electrophoresis

Visualize the RNAs by nondenaturing agarose gel electrophoresis to verify the presence of full-length transcripts.

1. Prepare a 2% agarose gel in 1X TAE buffer containing 0.5 $\mu\text{g/mL}$ of EtBr.
2. To 1.5-mL tubes, mix 200–300 ng of RNA and 2–5 μL of 6X RNA loading buffer. Heat the samples at 65°C for 5 min, followed by a quick chill on ice for 5 min.
3. In an RNase-free electrophoresis tank, load the samples and markers onto the agarose gel. RNA size markers are the best way to correctly assess the size of the transcript. If DNA markers are used, keep in mind that an RNA will run faster than a double-stranded DNA fragment of similar size. The full-length wild-type and mutant apo-B transcripts are 280 nucleotides in length. The RNA should appear as a single band on the gel (*see Note 9*).

3.2. Preparation of CNBr-Activated Sepharose RNA Affinity Chromatography Column

Described in **Subheadings 3.2.1.–3.2.4.** are the steps to prepare a single 2-mL CNBr-activated Sepharose (Amersham Pharmacia Biotech) column for RNA affinity chromatography. The method used here is adapted from Amersham Pharmacia Biotech's product insert and the method of Kaminski et al. (8). With the exception of the RNA used, the wild-type and mutant affinity columns are prepared identically.

3.2.1. Preparing Matrix

The CNBr-activated Sepharose is supplied as a freeze-dried powder containing dextran and lactose. The beads must be swollen and washed extensively to remove the sugars before use (*see Note 10*).

1. For a 2-mL column, weigh out ~800 mg of the CNBr-activated Sepharose.
2. In a sterile 15-mL polypropylene tube, swell beads in 8 mL of cold 1 mM HCl on ice.
3. Pour the slurry into a vacuum filtration device fitted with a 0.2- μm -pore-size membrane (Millipore).
4. Wash the beads with 240 mL of cold 1 mM HCl.
5. Wash the beads with 20 mL of cold sterile water.
6. Wash the beads with 20 mL of cold 1X coupling buffer.
7. Scrape the beads into a sterile 15-mL polypropylene tube and resuspend in 4 mL of cold 1X coupling buffer. Allow the beads to settle on ice and remove the supernatant.

3.2.2. Coupling RNA to Matrix

The imidocarbonate groups of the CNBr-activated Sepharose react with the amino groups of the RNA, resulting in stable covalent linkages between the CNBr-activated Sepharose matrix and RNA.

1. For 2 mL of beads, use 2.0 mg of RNA (synthesized in **Subheading 3.1.**) diluted with 1 vol of cold 2X coupling buffer.
2. Further dilute the RNA to 4 mL with cold 1X binding buffer.
3. Add the RNA to the beads from **Subheading 3.2.1.** and rotate end over end for 16 h at 4°C.

3.2.3. Blocking the Remaining Active Groups

The noncovalently bound RNA must be washed away from the matrix to avoid complications during RNA affinity chromatography. Furthermore, any remaining imidocarbonate groups must be blocked to avoid the linkage of proteins to the matrix.

1. Remove unbound RNA by washing beads with 10 mL of cold sterile water.
2. Block the remaining active groups with the addition of 8 mL of a cold small primary amine such as 0.1 M Tris-HCl, pH 7.8. Rotate the tube end over end for 2 h at 4°C.
3. Remove excess blocking solution by washing the beads with 20 mL of cold binding buffer.
4. Resuspend the beads in 4 mL of cold binding buffer.

3.2.4. Pouring RNA Affinity Chromatography Column

1. To prepare the RNA affinity column, place an empty Poly-Prep Chromatography Column (Bio-Rad) in an appropriate holder allowing sufficient room to work below the column for fraction collection (*see Note 11*).
2. Pour the bead slurry from **Subheading 3.2.3.** into the column and allow the column to pack with the force of gravity (*see Note 12*).
3. For a 2-mL column, wash the column with 4 mL of cold binding buffer and then with 2 mL of cold elution 3 buffer.
4. Equilibrate the column with 10 mL of cold binding buffer. The column is now ready for use.

3.3. RNA Affinity Chromatography

Described in **Subheading 3.3.1.** are the steps for purifying ACF from partially purified tissue extracts. This simple method, which uses gravity flow and step elution, does not require any special laboratory equipment. The protocol is for a single column.

3.3.1. Binding Extract, Washing Column, and Eluting Protein

In the following protocol, protein extract is applied to the RNA affinity column. Unbound and weakly bound proteins are washed from the column. Following rigorous washing, the bound proteins are step eluted from the column with buffers containing increasing concentrations of salt (*see Note 13*).

1. Dilute 50 mg of partially purified baboon kidney S100 extract (*see Note 14*) to 2 mL with cold binding buffer, and apply to the RNA affinity column prepared in **Subheading 3.2.4**. Collect the drop-by-drop flowthrough (*see Note 15*).
2. Wash the column with 20 mL of cold binding buffer. Collect the drop-by-drop flowthrough in 2-mL fractions.
3. Elute with 2 mL of cold elution buffer 1. Collect the drop-by-drop flowthrough in ~0.5-mL fractions.
4. Elute with 2 mL of cold elution buffer 2. Collect the drop-by-drop flowthrough in ~0.5-mL fractions.
5. Elute with 2 mL of cold elution buffer 3. Collect the drop-by-drop flowthrough in ~0.5-mL fractions. After fractions are collected from the wild-type and mutant RNA affinity columns, you can begin analyzing the eluted proteins (*see Notes 16 and 17*).

3.4. Identification of ACF

Twenty-three fractions from each column (*see Note 18*) were collected as described in **Subheading 3.3**. To identify the fraction(s) that contain ACF, it is useful to have a functional and sensitive RNA-binding assay. We previously developed a UV crosslinking assay for ACF detection (**4**). This assay essentially transfers a radiolabel from a ^{32}P -UTP body-labeled wild-type apo-B RNA to ACF. When the crosslinked samples are analyzed by SDS-PAGE and autoradiography, ACF is clearly detected as a radiolabeled protein of 65 kDa (*see Note 19*). We routinely detect ACF crosslinking activity in the second and third fractions eluted from the wild-type apo-B RNA column with elution buffer 2 (0.5 M NaCl), but not in fractions eluted with elution buffer 1 (0.2 M NaCl) or 3 (1 M NaCl). No ACF crosslinking activity is detected in any of the fractions eluted from the mutant RNA column. Instead, ACF is found in the flowthrough fraction that contains the unbound proteins from the mutant RNA column (*see Note 20*).

In the absence of an RNA-binding assay, the eluted proteins can be visualized directly by SDS-PAGE and Coomassie staining (*see Note 21*). Comparison of proteins eluted from the wild-type and mutant RNA affinity columns can distinguish between specific and nonspecific RNA-binding proteins. For example, there are several proteins in common that eluted with 0.5 M NaCl from both the wild-type and mutant apo-B RNA affinity columns. However, one unique protein band was specifically found in the 0.5 M NaCl eluate from the wild-type RNA column and was not detected in the comparable fraction from the mutant RNA column (**4**). The size of this polypeptide, 65 kDa, correlated with the size of ACF detected in the UV crosslinking assay. Once this protein was identified, the band was excised from the gel for protein microsequence analysis by mass spectrometry (*see Note 22*).

4. Notes

1. To reduce RNase contamination, many laboratories use diethylpyrocarbonate-treated water. We do not feel that to be necessary. To minimize contamination, we change gloves frequently and utilize autoclaved MilliQ (Millipore) purified dH₂O, tips, and tubes in our protocols.
2. Although one can use crude protein extracts for this protocol, we obtain better results if we use a partially purified extract. The fold-purification does not have to be substantial. To partially purify ACF, a baboon kidney S100 extract was precipitated with 15–30% ammonium sulfate. The ammonium sulfate fraction was further purified by gel filtration on Sephacryl S300 (4). This scheme resulted in only ~10-fold purification of ACF, but it significantly improved the results of the RNA affinity chromatography.
3. An alternative approach to subcloning by restriction digestion is through oligonucleotide design and polymerase chain reaction (PCR) amplification of the desired DNA sequence. *Taq* polymerase adds a single deoxyadenosine to the 3' ends of PCR products. Using commercially available vector systems such as the TOPO TA Cloning® kit (Invitrogen), the PCR product can easily be TA cloned into vectors such as pCR®II-TOPO® (Invitrogen) that have 3' T overhangs.
4. The optimal size of RNA for generating RNA affinity columns on CNBr-activated Sepharose is between 100 and 400 nucleotides. Unlike some methods that use biotinylated nucleotide analogs, there is no need for a linker sequence between the RNA ligand and the affinity beads.
5. By employing both wild-type and mutant RNAs, one can distinguish specific binding over background. For this approach to be successful, the protein should have at least a 50-fold higher affinity for the wild-type RNA than the mutant RNA. If a point mutation does not significantly reduce the binding affinity, one can generate a more deleterious mutation by mutating or deleting several nucleotides from the putative binding site. If a mutant RNA is not available, the antisense transcript or an irrelevant RNA can be used to distinguish between specific and nonspecific RNA-binding proteins.
6. Avoid using restriction enzymes that produce 3' overhangs. Using such enzymes may result in the appearance of extraneous transcripts. If these enzymes must be used, 3' overhangs should be made blunt using DNA Polymerase I Large (Klenow) Fragment (Promega).
7. If the ratio is lower than 1.8, repeat the phenol:chloroform extraction to remove contaminating proteins.
8. A significantly lower yield than the expected amount suggests a failed transcription reaction. The enzyme should be tested with a control plasmid. If the yield is significantly higher, this suggests that the RNA may still contain free nucleotides. The solution should be passed over another spin column.
9. If multiple bands are seen on the gel, there may be incomplete transcripts present in the sample. Such transcripts may be the result of premature termination of RNA

synthesis. It has been shown that incubating the transcription reaction at 30°C can increase the amount of full-length transcripts produced. The presence of a smear on the gel instead of a single band may be owing to degradation of the RNA in the sample. This is most often seen after repeated handling of a sample. The RNA should be resynthesized and aliquoted into smaller volumes before storage.

10. During the entire RNA affinity procedure (from bead preparation to column elution), it is important to keep the beads and fractions on ice as much as possible. This will reduce the amount of RNA degradation on the column and protein degradation in the fractions. If space permits, perform all steps involving the beads in a 4°C room.
11. An alternative to using a gravity flow column is using a chromatography system, such as the BioCad Sprint (PE Biosystems, Framingham, MA). Chromatography systems allow for larger columns, faster flow rates, serial columns, and protein UV absorbance profile.
12. To preserve column integrity, ensure that the top of the bead bed does not dry out during the entire procedure. To avoid this, have all materials prepared before starting and never leave a running column unattended.
13. Stepwise column elution may not give sufficient resolution if many proteins bind to the RNA affinity column. As an alternative, the bound proteins can be eluted using a linear gradient of 0 to 1 *M* NaCl. One can also try eluting with buffers that contain different salts (e.g., KCl or MgCl₂). If the protein of interest does not elute from the column, the salt concentration of the elution buffer can be increased up to 4 *M*. Alternatively, one can destroy the column with ribonucleases to free the bound protein.
14. In purifying different RNA-binding proteins by RNA affinity chromatography, we have used between 10 and 100 mg of crude extract or partially purified protein. The amount of protein to be added to the column depends on the concentration of the protein of interest. If too much extract is added, the specific binding of a low-abundance protein may be competed by other abundant proteins that bind the RNA ligand with low affinity. Furthermore, if the column becomes saturated, the bulk of the protein of interest will be present in the flowthrough fraction. If a functional RNA-binding assay is available, we suggest performing small-scale pilot experiments using different amounts of extracts and a constant amount of RNA affinity beads. The start and unbound fractions can be analyzed to optimize the bead:protein ratio.
15. During this procedure you will accumulate 23 fractions: one 2-mL flowthrough fraction that contains the unbound proteins, ten 2-mL wash fractions, four 0.5-mL fractions from elution buffer 1, four 0.5-mL fractions from elution buffer 2, and four 0.5-mL fractions from elution buffer 3. As described earlier, it is very important to keep these fractions on ice as much as possible.
16. Because of the high salt concentrations in the elution buffers and increased volume of the eluted fractions, it may be necessary to desalt and concentrate the eluted fractions before analysis. To prepare these fractions, one should use a

desalting and concentrating column such as a Microcon Centrifugal Filter Column (Millipore).

17. RNA affinity columns can be reused. To regenerate the column; first, wash the column with 20 mL of cold elution buffer 3 and then wash the column with 20 mL of cold sterile water. The column is now ready for immediate use. Columns can be stored for up to 6 mo at 4°C if in 20% methanol (to inhibit bacterial growth) and sealed properly.
18. For each elution buffer, four 0.5-mL fractions are obtained. The first 0.5 mL displaces the previous buffer from the column. The eluted proteins are generally found in the second and third 0.5-mL fractions. The fourth fraction usually does not contain detectable protein. Therefore, it may only be necessary to assay the second and third fractions.
19. As an alternative to UV crosslinking, the electrophoretic mobility shift assay can be utilized to analyze RNA-binding activity.
20. Although ACF is detected in the unbound fraction from the mutant RNA column, other RNA-binding proteins may bind to the mutant RNA column with low affinity. In this case, the RNA-binding protein may elute from the mutant RNA column at a lower salt concentration than from the wild-type RNA column (5).
21. The method used to visualize proteins after SDS-PAGE will depend on the amount of protein obtained. For those fortunate enough to obtain a highly concentrated sample, individual desalted/concentrated fractions can be run on SDS-PAGE and stained with GelCode Blue (Pierce, Rockford, IL). However, if the protein concentration is low, it may be necessary to pool and concentrate several fractions as described in **Note 16**. These concentrated samples can then be run on SDS-PAGE and stained with a more sensitive stain such as Silver Stain Plus (Bio-Rad). The type of stain chosen must be compatible with microsequence analysis by mass spectrometry if that is the goal of the experiment.
22. The initial round of purification may not yield a preparation that is clean enough to excise a band for peptide sequencing. In this case, one can utilize a “preclear” RNA affinity column to remove nonspecific binding proteins. This preclear column can be made with a mutant RNA, the antisense transcript or an unrelated RNA or the mutant RNA. Once the sample is passed over the “pre-clear” column, the flowthrough containing the unbound proteins can be purified on the wild-type RNA column.

Acknowledgments

We wish to thank Drs. Laurent Chavatte and Lisa Middleton for critical review of the manuscript. This work was supported by Public Health Services (PHS) grant HL45478.

References

- 1 Sela-Brown, A., Silver, J., Brewer, G., and Naveh-Many, T. (2000) Identification of AUF1 as a parathyroid hormone mRNA 3'-untranslated region-binding protein

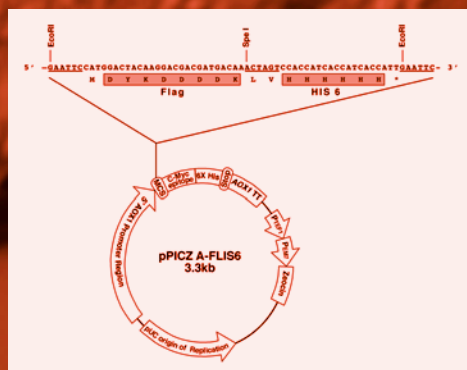
- that determines parathyroid hormone mRNA stability. *J. Biol. Chem.* **275**(10), 7424–7429.
- 2 Neupert, B., Thompson, N. A., Meyer, C., and Kuhn, L. C. (1990) A high yield affinity purification method for specific RNA-binding proteins: isolation of the iron regulatory factor from human placenta. *Nucleic Acids Res.* **18**(1), 51–55.
 - 3 Rouault, T. A., Hentze, M. W., Haile, D. J., Harford, J. B., and Klausner, R. D. (1989) The iron-responsive element binding protein: a method for the affinity purification of a regulatory RNA-binding protein. *Proc. Natl. Acad. Sci. USA* **86**(15), 5768–5772.
 - 4 Mehta, A. and Driscoll, D. (1998) A sequence-specific RNA-binding protein complements apobec-1 to edit apolipoprotein B mRNA. *Mol. Cell. Biol.* **18**(8), 4426–4432.
 - 5 Copeland, P. R. and Driscoll, D. M. (1999) Purification, redox sensitivity, and RNA binding properties of SECIS-binding protein 2, a protein involved in selenoprotein biosynthesis. *J. Biol. Chem.* **274**(36), 25,447–25,454.
 - 6 Copeland, P. R., Fletcher, J. E., Carlson, B. A., Hatfield, D. L., and Driscoll, D. M. (2000) A novel RNA binding protein, SBP2, is required for the translation of mammalian selenoprotein mRNAs. *EMBO J.* **19**(2), 306–314.
 - 7 Mehta, A., Kinter, M. T., Sherman, N. E., and Driscoll, D. M. (2000) Molecular cloning of apobec-1 complementation factor, a novel RNA-binding protein involved in the editing of apolipoprotein B mRNA. *Mol. Cell. Biol.* **20**(5), 1846–1854.
 - 8 Kaminski, A., Hunt, S. L., Patton, J. G., and Jackson, R. J. (1995) Direct evidence that polypyrimidine tract binding protein (PTB) is essential for internal initiation of translation of encephalomyocarditis virus RNA. *RNA* **1**(9), 924–938.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



In Vitro Assays for Kinetoplastid U Insertion–Deletion Editing and Associated Activities

Kenneth Stuart, Reza Salavati, Robert P. Igo, Jr.,
Nancy Lewis Ernst, Setareh S. Palazzo, and Bingbing Wang

Summary

This chapter describes biochemical assays that have been used in analyzing RNA editing in kinetoplastid mitochondria and to characterize the general mechanism of editing by the editosome. Studies using these assays have shown that the characteristics of each activity contribute to editing site selection, U addition and removal, and RNA ligation resulting in accurately edited mRNAs.

Key Words: Kinetoplastid RNA editing; U insertion and deletion editing; editing activities; guide RNA.

1. Introduction

Kinetoplastid RNA (kRNA) editing entails maturation of mitochondrial mRNAs by the insertion and deletion of uridylates (Us) as specified by small guide RNAs (gRNAs). RNA editing is essential to the production of functional mRNAs since it creates start and stop codons, and determines the mature protein coding sequence. It can be extensive, even accounting for more than half of the sequence.

In vitro studies using mitochondrial extract from the kinetoplastid *Trypanosoma brucei* indicate that RNA editing occurs by endonucleolytic cleavage of the precursor mRNA at the editing site; U addition by 3'-terminal uridylyl transferase (TUTase) activity or removal by 3' exouridylylase activity, both at the 3' end of the 5' cleavage fragment; and subsequent ligation of the RNA fragments by RNA ligase activity. Each gRNA has a 3' oligo (U) tail that is

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

added posttranscriptionally. The function of the U tail is unknown but it may stabilize the gRNA/pre-mRNA interaction. The complex that catalyzes editing, the editosome, contains the aforementioned catalytic activities and associated RNA helicase activity, which may affect gRNA/mRNA interactions and/or displace gRNA after its use.

We describe here biochemical assays that have been used in analyzing RNA editing in kinetoplastid mitochondria and to characterize the general mechanism of editing by the editosome. Studies using these assays have shown that the characteristics of each activity contribute to editing site selection, U addition and removal, and RNA ligation resulting in accurately edited mRNAs.

2. Materials

All solutions must be prepared under RNase-free conditions, using diethylpyrocarbonate (DEPC)-treated H₂O.

2.1. Transcription

1. 10X T7 transcription buffer (**I**): 400 mM Tris-HCl, pH 7.6; 240 mM MgCl₂; 20 mM spermidine; 0.1% Triton X 100. Store at -20°C.
2. rNTP mix: 25 mM each rATP, rCTP, rGTP, rUTP (Pharmacia; >98% triphosphate) in DEPC-H₂O. Store at -20°C.
3. 1 M Dithiothreitol (DTT) in DEPC-H₂O. Store at -70°C in 50-μL aliquots.
4. RQ1 DNase (Promega, Madison, WI).
5. RNasin (40 U/μL) (Promega).
6. T7 RNA polymerase (80 U/μL) (Promega).
7. 7 M Urea-loading buffer: 7 M urea, 1X TBE, 0.05% bromophenol blue, 0.05% xylene cyanol.
8. Equipment for denaturing (7 M urea) polyacrylamide gel electrophoresis (PAGE).
9. Razor blades (sterile or new).
10. Polyacrylamide gel elution buffer: 0.3 M NaOAc, pH 5.2; 0.1% sodium dodecyl sulfate (SDS); 1 mM EDTA.

2.2. Full-Round In Vitro RNA Editing

1. [³²P] substrate RNA, 0.25 μM, labeled by ligation of [³²P]pCp to transcribed RNA. For the insertion editing assay, the RNA substrate is A6-eES1 (5'-GGAAAG GUUAGGGGGAGGAGAGAAGAAAGGGAAAGUUGUGAUUGGAGUUAUA GAUACUUACCUGGCAUC-3'). The T7 transcription template for A6-eES1 is prepared by polymerase chain reaction (PCR) of plasmid A6eES-1 with primers EcoR1 T7 (5'-CGGCGGAATTCTGTAATACGACTCAC-3') and 3'TAG A6 (5'-GATGCCAGGTAAGTATTC-3') (**2**). For deletion editing, the RNA substrate is A6short/TAG.1 (5'-GGAAAGGUUAGGGGGAGGAGAGAAGAAAGGGA AAGUUGUGAUUUUUGGAGUUAUAGAAUACUUACCUGGCAUC-3'). The T7 transcription template for A6short/TAG.1 is prepared by PCR of plasmid

- A6/TAG.1 with primers EcoRI T7 (5'-CGGCGGAATTCTGTAAATACGACTCAC-3') and 3'TAG A6 (5'-GATGCCAGGTAAGTATTC-3') (3).
- gRNA, 0.5 μ M. For the insertion editing assay, gA6[14]+A21 (5'-GGAUAUACU AUAACUCCGAUAAACGAAUCAGAUUUUGACAGUGAUUAUGAUAAUUUAU UUUUUUUUUUUUUUUUU-3') is prepared from the T7 transcription of the gA6[14]+A21 template digested by *Dra*I (2). For the deletion editing assay, gA6[14] Δ 16G (5'-GGAUAUACU AUAACUCCAUAAACGAAUCAGAUUUUGA CAGUGAUUAUGAUAAUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU-3') is prepared by PCR of gA6[14]NX (4) with primers T7gA6 Δ 16G (5'-GTAATACGACTCAC TATAGGATATACTATATAACTCCATAACGAATC-3') and T3 (5'-AATTAACC CTCCTAAAG-3').
 - 2X HHE buffer: 50 mM HEPES, pH 7.9; 20 mM Mg(OAc)₂, 100 mM KCl, 1 mM DTT, 2 mM EDTA. Filter sterilize and store at -20°C in 500- μ L aliquots.
 - 2X HHE without KCl (if using a high-salt mitochondrial extract).
 - 0.1 M CaCl₂.
 - 30 mM and 200 μ M ATP (from 100 mM; Promega); store at -70°C.
 - 2 mM UTP (from 100 mM; Promega); store at -70°C.
 - Yeast RNA (Torula type VI; Sigma, St. Louis, MO), 500 ng/ μ L in H₂O; store at -70°C in 1-mL aliquots.
 - Glycogen (5 μ g/ μ L) (from 20 mg/mL; Roche, Indianapolis, IN); store at -20°C.
 - Stop buffer: 130 mM EDTA, 2.5% SDS.
 - Phenol:CHCl₃:isoamyl alcohol, 25:24:1 (v:v:v) (PCI).
 - 3 M NaOAc, pH 5.2.
 - Cold 100 and 70% EtOH.
 - RNase T₁ buffer: 7 M urea; 33 mM sodium citrate, pH 5.5, 1.7 mM EDTA, 1 mg/mL of yeast tRNA (Invitrogen, Carlsbad, CA), 0.05% bromophenol blue; 0.05% xylene cyanol. Store at -20°C in 100- μ L aliquots.
 - 10X Alkaline hydrolysis buffer: 0.5 M NaHCO₃, pH 9.3, 10 mM EDTA, 5 mg/mL of yeast tRNA.
 - 7 M Urea-loading buffer (see Subheading 2.1., item 7).

2.3. Hammerhead Ribozyme Assay

The materials are the same as for the full-round reaction, plus the following:

- For deletion editing, the edited RNA substrate (positive control) (5'-GGGGG AAAGGUUAGGGGGAGGAGAGAAAGGGAAAGUUGUGACUGAUGAGUC CGUGAGGACGAAACAAUAGAUAUCAAUGU-3') is generated from oligonucleotide dA6RZ 5'-ACATTTGATCTATTGTTTCGTCCTCACGGACTCATCA GTCACAACCTTTCCCTTTCTCTCCTCCCCCTAACCTTTCCCTTATAGTGAG TCGTATTA-3', the preedited RNA substrate (5'-GGGGGAAAGGUUAGG GGGAGGAGAGAAAGGGAAAGUUGUGACUUUUUGAUGAGUCCGUGAG GACGAAACAAUAGAUAUCAAUGU-3') is generated from oligonucleotide dpre-A6RZ 5'-ACATTTGATCTATTGTTTCGTCCTCACGGACTCATCAAAAAGT

CACAACTTTCCCTTTCTCTCCTCCCCCTAACCTTTCCCCCTATAGTGAGT CGTATTA-3', and the guide RNA (5'-GGGGGAUAUACACGGACUCAUCAC CCUCACAACUUUCCCUUGACAGUGAUUAUGAUAAUUUUUUUUUUUUUUUUUUUUU-3') is generated from oligonucleotide dgA6RZ 5'-AAAAAAAAA AAAAAATAATTATCATATCACTGTCAAGGGAAAGTTGTGAGGGTGATG AGTCCGTGTATATCCCCCTATAGTGAGTCGTATTA-3' in combination with a T7 oligonucleotide. The sequence complementary to the T7 promoter sequence is underlined.

2. [³²P]Substrate RNA, 40 pmol, labeled by ligation of [³²P]pCp to (5'-CUAUUGU CUCACAACUUC-3') (oligos and so on).
3. 10X Ribozyme buffer: 100 mM MgCl₂, 500 mM Tris-HCl, pH 7.9, and 100 mM DTT.
4. G-25 column (Pharmacia).

2.4. Precleaved RNA Editing

Reagents are as for the full-round reaction except the RNAs, plus the following:

1. 5' [³²P]-Labeled 5' fragment RNA, 0.25 μM (0.25 pmol/μL). The canonical 5' fragment for precleaved insertion editing is 5'CL18 (5'-GGAAGUAUGA GACGUAGG-3'). The transcription template is prepared by PCR of the oligonucleotide pair 5'CL18-Tmpl (5'-GGCGGAATTCTGTAATACGACTCACTATAG GAAGTATGAGACGTAGG-3' and complementary sequence) using primers EcoRI T7 and 5'CL18-3' (5'-CCTACGTCTCATACTTCCTATAG-3'). The typical 5' fragment for precleaved deletion editing is U5-5'CL (5'-GGAAAGGG AAAGUUGUGAUUUU-3'). The PCR template for U5-5'CL is prepared by annealing the U5-5' antisense oligonucleotide (5'-AAAATCACAACTTTCC CTTTCCTATAGTGAGTCGTATTAC-3') with the EcoRI T7 sense oligonucleotide (5'-CGGCGGAATTCTGTAATACGACTCAC-3'). The overhangs are filled by the Klenow enzyme, and PCR is performed using EcoRI T7 and antisense primer U5-5'short (5'-AAAATCACAACTTTCCCT-3').
2. 3' Fragment RNA, 1 μM (1 pmol/μL). The usual 3' fragment for precleaved insertion editing is 3'CL13pp (5'-pAUUGGAGUUAUAGp-3'), which contains a 5' and 3' monophosphate group (Oligos etc.). For precleaved deletion editing, U5-3'CL (5'-pGCGAGUUAUAGAAUAp-3'), which also contains a 5' and 3' monophosphate, is used.
3. gRNA, 0.5 μM (0.5 pmol/μL). The typical gRNA for precleaved insertion of two Us is gPCA6-2A (5'-GGAUAUACUAUAACUCCGAUAACCUACGUCUCAU ACUUCC-3'). The transcription template for gPCA6 is prepared by PCR of gPCA6-2A-Tmpl (5'-CGGCGGAATTCTGTAATACGACTCACTATAGGATAT ACTATAACTCCGATAACCTACGTCTCATACTTCC-3' and complementary sequence) with primers EcoRI T7 and gPCA6-2A-3' (5'-GGAAGTATGAGACG TAGGTTATCGGAGT-3'). The gRNA for deletion of four Us is gA6[14]PC-del (5'-GGUUCUAUAACUCGCUCACAACUUUCCCUUCC-3'). The gA6[14]PC-del gRNA transcription template is prepared by annealing the A6comp1 antisense

oligonucleotide (5'-GGAAAGGGAAAGTTGTGAGCGAGTTATAGAACCTATAGAACCTATAGTGAGTCGTATTAC-3') with EcoRI T7 and extended with Klenow enzyme.

4. Competitor RNAs, 5 μ M (5 pmol/ μ L). The competitor RNA for precleaved insertion editing is 5'-GGAAGUAUGAGACGUAGGAUUGGAGUUAUAG-3' and for precleaved deletion is 5'-GGAAAGGGAAAGUUGUGAUUUU-3'.

2.5. Endonuclease Activity

1. 5' [³²P]-Labeled CYb preedited mRNA. Preedited mRNA (5'-GUUAAGAAUAAUGGUUAUAAAUUUUAUUAUAAAAGCGGAGAAAAAGAAAGGGUCUUUUAAUGUCAGGUUGUUUAUUAUAGAAUUAUAGG-3') is synthesized by T7 RNA polymerase from a *Bam*HI-linearized DNA template (5,6). The CYb gRNA (5'-ACUGACAUUAAAAGACAAUUAUAAAUUU-3') that directs insertion of two, one, and three U residues at ES1, ES2, and ES3, respectively, is synthesized by heating two oligonucleotides (5' AGCTTAATACGACTCAC TATAGGG 3' and 5' AAATTTATATTGTCTTTTAATGTCAGTCCCTATAGT GAGTCGTATTA 3') to 65° and slowly cooling down; the T7 promoter is underlined (7).
2. 5X Tris buffer: 50 mM Tris-HCl, pH 7.5, 250 mM KCl, 50 mM MgCl₂, 2.5 mM DTT. Filter sterilize and store at -20°C in 500- μ L aliquots.
3. 5X Tris without KCl (if using a high-salt mitochondrial extract).

2.6.3. 3' Exouridylylase Activity

The materials are the same as for the precleaved deletion editing reaction (see Subheading 2.4.).

2.7. TUTase Activities

2.7.1. UTP Incorporation into Yeast tRNA

1. 10X Buffer: 250 mM HEPES, pH 7.9, 100 mM Mg(OAc)₂, 500 mM KCl, 10 mM EDTA, 5 mM DTT.
2. 10X Buffer without KCl if using a high-salt mitochondrial extract.
3. Yeast tRNA (Sigma): 1 mg/mL in water.
4. α -[³²P]-UTP (800 Ci/mmol).
5. Whatman GF/C glass fiber disks.
6. 10% Trichloroacetic acid (TCA) with 50 mM disodium pyrophosphate at 4°C: Dissolve 1 kg of TCA in 454 mL of distilled H₂O for 100% TCA. For 1 L, dilute 100 mL of 100% TCA with 900 mL of distilled H₂O and add 11.1 g of disodium pyrophosphate. Store at 4°C.
7. 95% Ethanol at 4°C.

2.7.2. U Addition to RNA Primer

The materials are the same as for precleaved insertion editing (see Subheading 2.4.).

2.8. Ligase

2.8.1. Adenylation Assay

1. 5X buffer: 125 mM HEPES, pH 7.9, 50 mM magnesium acetate, 2.5 mM DTT.
2. Dimethylsulfoxide (DMSO) (Sigma).
3. [α - ^{32} P] ATP (3000 Ci/mmol) (Amersham, Piscataway, NJ).
4. Beta-block or other shielding equipment.
5. Apparatus for running 10% SDS-PAGE gel.
6. 10% SDS-PAGE gel.
7. 1X SDS-PAGE (Laemmli) loading dye: 100 mM Tris-HCl, pH 6.8, 4% (w/v) SDS (electrophoresis grade), 0.2% (w/v) bromophenol blue, 20% (v/v) glycerol, 200 mM DTT.
8. 1X SDS-PAGE running buffer: 25 M Tris base, 250 mM glycine, pH 8.3, 0.1% SDS.
9. Equipment for drying gel.
10. Phosphorimager cassettes or Kodak X-OMAT film.

2.8.2. Precleaved RNA Ligation Assay

Materials are as for the precleaved editing reaction minus UTP and yeast RNA. gRNAs are transcribed from PCR products derived from gpCA6-3A-Tmpl: (5'-GGCGGAATTCTGTAATACGACTCACTATAGGATATACTATAAC TCCGA TAAACCTACGTCTCATACTTCC- 3' and complementary sequence) with primers EcoRI T7 (5'-CGGCGGAATTCTGTAATACGACTCACTATAG-3') and gPCA6-0A, 1A, 2A, 3A (5'-GGAAGTATGAGACGTAGG[T]_nATCG GAGT- 3', where *n* is equivalent to the number of A nucleotides in the gRNA).

3. Methods

3.1. Large-Scale T7 Transcription of Short RNAs for Editing Experiments (see Notes 1–6)

Template DNA should contain T7 promoter followed by the sequence to be transcribed (note that this sequence will necessarily begin with a G) for runoff transcription. For optimal yield, use short PCR product rather than linearized plasmid. The T7 promoter should be preceded by several 5' nucleotides to facilitate binding of the RNA polymerase. Very short transcripts (as used for the precleaved editing assay) require a very high promoter sequence concentration.

1. Assemble the transcription reaction at room temperature and incubate overnight at 37°C: 55 μ L of PCR product template; 30 μ L of rNTPs, 25 mM each; 10 μ L of 10X T7 transcription buffer; 1 μ L of 1 M DTT; 1 μ L of RNasin; and 3 μ L of T7 RNA polymerase (80 U/ μ L) for a total volume of 100 μ L.
2. Add 5 μ L of RQ1 DNase and incubate for 15 min at 37°C.
3. Add 300 μ L of DEPC-H₂O, 40 μ L of 3 M NaOAc (pH 5.2), and 1 mL of cold 100% EtOH. Mix and *immediately* spin for 15 min at 4°C in a microfuge.

Remove the supernatant and wash with 1 mL of 70% (or 85%) EtOH. Air-dry the pellet and resuspend in 40 μ L of 7 M urea-loading buffer. Heat the sample to 65°C for 1 min to ensure resuspension of the pellet.

4. Pour a 9% denaturing (7 M urea, 1X TBE) polyacrylamide (19.5:1 acrylamide:bis-acrylamide) gel; use 0.75-mm spacers and a comb. Load the sample on one lane of the gel, and run the gel until the expected transcription product is fairly close to the bottom (bromophenol blue runs at about 13–15 nt, and xylene cyanol at about 60–70 nt). If RNAs are very short (<25 nt), use a 15% gel.
5. Sandwich the gel between two layers of plastic wrap, place on top of a reflection screen, and locate a band of RNA by ultraviolet (UV) shadowing with a short-wave UV lamp. Trace the band by drawing on the plastic wrap with a felt-tip pen. Use a razor blade to cut the band from the gel. Place the gel slice in a 1.5-mL microfuge tube, and add 400 μ L of elution buffer. Elute overnight at room temperature on a rotator or with gentle agitation.
6. Transfer the eluate to a new centrifuge tube. (About 50 μ L will have been absorbed by the gel slice.) Rinse the gel slice with 50 μ L of additional elution buffer, and add this buffer to the eluate. Add 1 mL of cold 100% EtOH; mix well and precipitate at –20°C for at least 1 h. Resuspend the pellet in 10–40 μ L of RNase-free H₂O, depending on the expected yield.

3.2. Full-Round In Vitro RNA Editing

The in vitro system described next faithfully edits synthetic pre-mRNA by the insertion or deletion of Us as specified by synthetic gRNAs (**Fig. 1**) (2,3,8). Editing requires mitochondrial extract, Mg⁺², and hydrolysis of the α/β bond of ATP. Added KCl supports editing up to 100 mM (above which it is inhibitory), and addition of 5 mM CaCl₂ enhances the production of edited RNA. The temperature and pH optima were 27°C and 7.9, respectively. Editing of the synthetic pre-mRNA is dependent on, and directed by, exogenously supplied cognate gRNA. Noncognate gRNA does not support in vitro editing.

The efficiency of in vitro editing varies among different preparations of *T. brucei* mitochondrial extract. Glycerol gradient-enriched extract contains much material that is irrelevant or potentially inhibitory to editing. For instance, it may contain inhibitors of ligase activity, since the adenylation activity of this extract increases on dilution of the extract (9). Most extracts edit 3–5% of the input pre-mRNA, although some edit as much as 25%. Preliminary experiments suggest that multiple rounds of editing occur at low frequency in vitro (unpublished data). For example, a cDNA cloned from an in vitro insertion assay was edited at both ES1 and ES2 of ND7 pre-mRNA. In addition, experiments with A6 pre-mRNA revealed evidence of RNA edited at ES1 and cleaved at ES2.

1. Combine 1 μ L each of gRNA and [³²P]substrate RNA per reaction to be performed, plus enough for one reaction extra. (Final amounts will be 0.5 pmol/

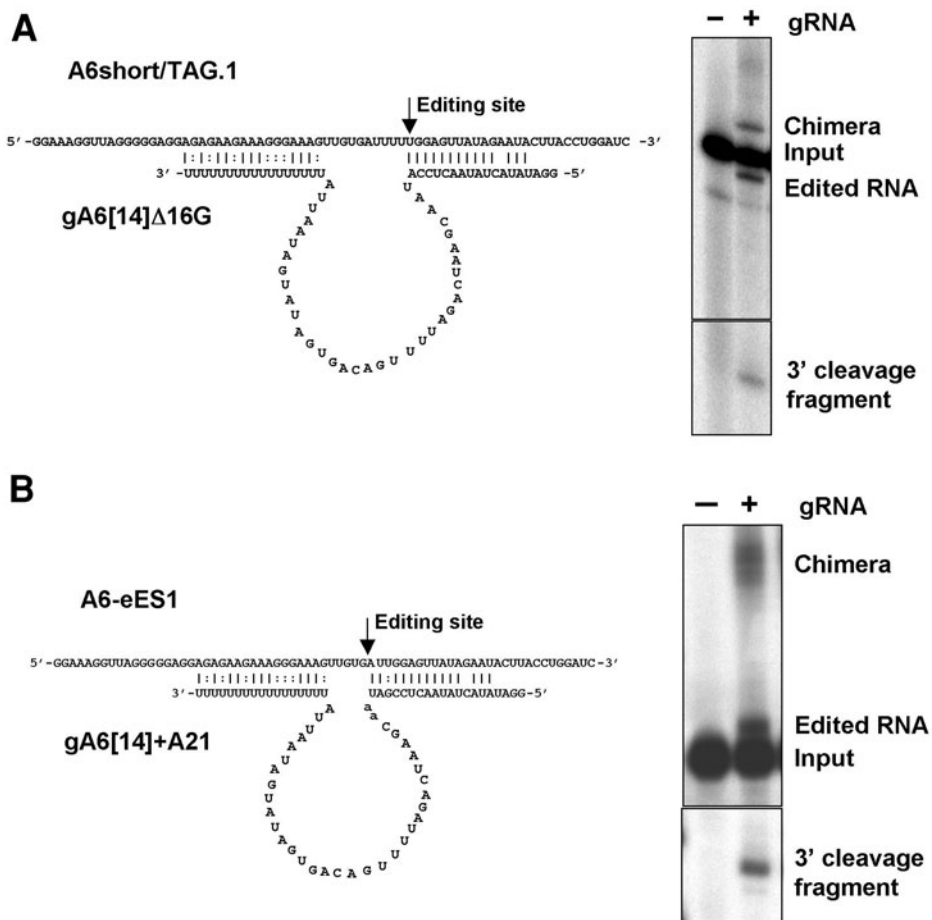


Fig. 1. Full-round deletion and insertion RNA editing. **(A)** In vitro deletion editing RNAs (**left**) and editing reactions containing 3'-labeled A6short/TAG.1 substrate RNA in absence or presence of gA6[14] Δ 16G gRNA and column-purified mitochondrial extract (**right**). **(B)** In vitro insertion editing RNAs (**left**) and editing reactions containing 3'-labeled A6-eES1 substrate RNA in absence or presence of gA6[14]+21A gRNA and glycerol gradient-purified mitochondrial extract (**right**). Input RNA, edited product, 3' cleavage product, and chimeras are indicated.

reaction of gRNA and 0.25 pmol/reaction of substrate.) Incubate at 65°C for 2 min and let sit at room temperature for at least 10 min.

2. Assemble the master mix (*see* **Notes 7–9**). Make enough for all desired reactions plus two extra. For insertion assays, use 2X HHE with 60 mM KCl; for deletion, use 2X HHE with 100 mM KCl.

<u>Component</u>	<u>Volume per deletion reaction (μL)</u>	<u>Volume per insertion reaction (μL)</u>
DEPC- H_2O	6.7	4.0
2X HHE	11.5	11.5
0.1 M $CaCl_2$	1.5	1.5
30 mM ATP	0.3	—
200 μ M ATP	—	1.5
2 mM UTP	—	1.5
Torula RNA	1.0	1.0

3. Assemble the reaction at room temperature: 21 μ L of master mix, 2 μ L of gRNA + substrate RNA, 7 μ L of mitochondrial extract, for a total volume of 30 μ L. Incubate for 3 h at 28°C in a heat block.
4. Add 2 μ L of stop buffer to each reaction and then add 100 μ L of H_2O . Extract once with 135 μ L of PCI. Add 2 μ L (10 μ g) of glycogen and 13.5 μ L (0.1 vol) of 3 M NaOAc, pH 5.2. Mix and add 350 μ L (2.5 vol) of cold 100% EtOH. Mix by inversion and incubate for 30 min at -70°C (or at least 1 h at -20°C).
5. Spin for 30 min in a microfuge at 4°C, wash with 150 μ L of cold 85% EtOH, spin for 5 min at 4°C, remove the supernatant, and dry in a SpeedVac. Resuspend the RNA in 12 μ L of 7 M urea-running buffer.
6. For RNase T₁ digest, dilute 1 μ L of RNase T₁ (4°C; Roche) to 200 μ L with H_2O . Combine 3 μ L of RNase T₁ buffer, 1 μ L of H_2O , 1 μ L of [^{32}P]-substrate RNA, and 1 μ L of diluted RNase T₁ in a 0.5-mL tube. Incubate at 55°C for 15 min, cool on ice, and add 10 μ L of 7 M urea-running buffer.
7. For partial alkaline hydrolysis ladder, combine 1 μ L of alkaline hydrolysis buffer, 1 μ L of labeled substrate RNA, and 8 μ L of H_2O in a 0.5-mL microcentrifuge tube. Incubate for 2 to 3 min at 94°C. It may be necessary to adjust the time of incubation, or the amount of buffer added, to optimize the partial hydrolysis. Cool on ice; add 6 μ L of 10 M urea-loading buffer.
8. Incubate the editing reactions for 2 min at 100°C immediately before loading on a 9% polyacrylamide gel. Load 4 μ L of each reaction, plus 8 μ L each of the RNase T₁ digest and the partial alkaline hydrolysis ladder. The length of gel will depend on the desired resolution. For large (about 70-nt) transcripts, we have used custom-made long (60-cm) gel plates. For RNAs <50 nt, an S1 (Bio-Rad, Hercules, CA) sequencing apparatus, about 40 cm long, is adequate to separate RNAs at a resolution of 1 nt.
9. Remove the gel from one plate and transfer to a piece of Whatman paper. Cover with plastic wrap. Dry a thin gel (0.35 mm) for at least 30 min at 80°C, or a thick gel (0.75 mm) for at least 2 h. Expose on a PhosphorImager screen overnight or autoradiography film for several days at -70°C with an intensifying screen.

3.3. Hammerhead Ribozyme Substrate for In Vitro RNA Editing

The development of an in vitro assay system for kRNA editing provided a tool for monitoring editing activity in the purification and analysis of the

T. brucei editosome. This assay has a low detection limit, and thus the development of a more sensitive specific assay was desirable. The highly efficient cleavage of a substrate RNA by a hammerhead ribozyme, which may be created from a preedited inactive ribozyme by *in vitro* RNA editing, makes it an attractive candidate for such a sensitive *in vitro* RNA editing assay. The hammerhead ribozyme is the smallest catalytic RNA motif that can cleave substrate RNA efficiently either *in cis* or *in trans* at a specific phosphodiester bond. The 5' cleavage product has a 2', 3' cyclic phosphate end while the 3' cleavage product has a 5' hydroxyl terminus. Any substrate RNA with an NHH sequence appropriately positioned relative to the base-paired stem can be cleaved by the ribozyme at the phosphodiester bond 3' to the second H, where N is any nucleotide and H can be either A, C, or U.

The ribozyme assay described here entails the conversion of an inactive ribozyme to an active ribozyme by accurate *in vitro* editing in which three uridylates are removed as directed by the appropriate gRNA (**Fig. 2**). The edited functional ribozyme is then used to cleave its targeted substrate. For convenience the conversion of the inactive ribozyme to its active form by editing and the cleavage of the substrate by the now-active ribozyme are performed in a single tube.

1. For annealing, combine 1 pmol of pre-A6RZ (1 μ L) and 2.5 pmol of gA6RZ (1 μ L) for each reaction. Incubate at 70°C for 3 min and cool to room temperature. Add 0.5 pmol of 3' [³²P]pCp-labeled SubA6RZ (1 μ L).
2. Assemble the master mix (*see Note 9*): 7 μ L of DEPC-H₂O, 0.2 μ L of RNasin, 1 μ L of yeast *Torula* RNA (500 ng/ μ L), 11.5 μ L of 2X HHE, 0.3 μ L of ATP (100 mM), for a total volume of 20 μ L.
3. Assemble the reaction: 20 μ L of master mix, 3 μ L of pre-A6RZ + gA6RZ + SubA6RZ, 7 μ L of mitochondrial extract, for a total volume of 30 μ L. Incubate for 1–3 h at 28°C in a heat block.
4. Assemble a positive control: 8 μ L of A6RZ (0.5 pmol) + labeled SubA6RZ (0.5 pmol) + DEPC-H₂O, 1 μ L of 10X ribozyme buffer, 1 μ L of RNasin, for a total volume of 10 μ L. Incubate for 1 h at 28°C in a heat block.
5. Add 2 μ L of stop buffer to each reaction. Phenol chloroform extract and ethanol precipitate RNAs as described for full-round editing (*see Subheading 3.2., step 4*).
6. Resuspend RNAs in 10 μ L of 10 M urea-running buffer and load 5 μ L onto a 20% polyacrylamide gel. The gel glass plate is 20 cm long to separate cleavage product (10 nt) from the input RNA (18 nt).
7. Remove the gel from both plates and sandwich in plastic wrap. Carefully remove all liquid and bits of gel from outside of the plastic wrap. It is very difficult to dry a denaturing 20% gel. However, the RNAs diffuse very little in a 20% gel kept at room temperature overnight, so it is possible to expose a PhosphorImager screen overnight without losing resolution.

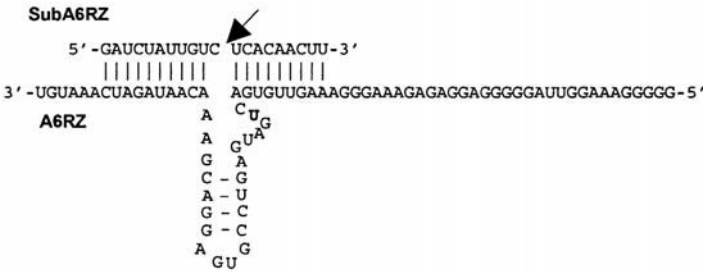
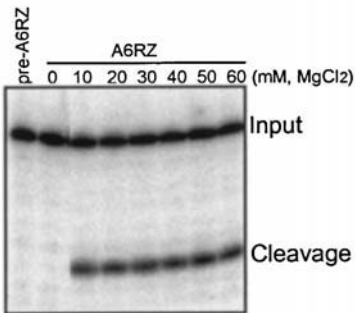
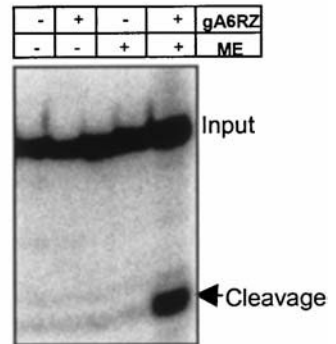
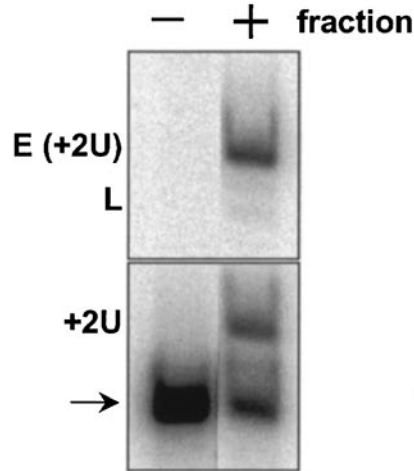
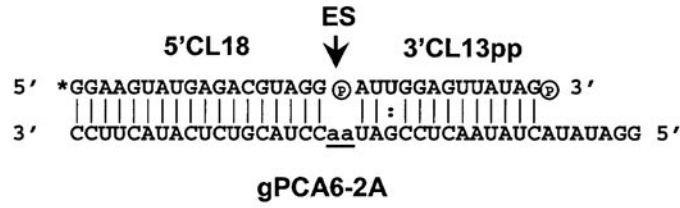
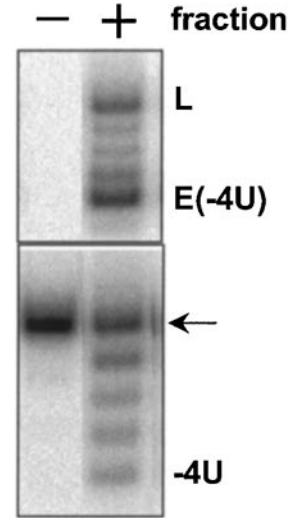
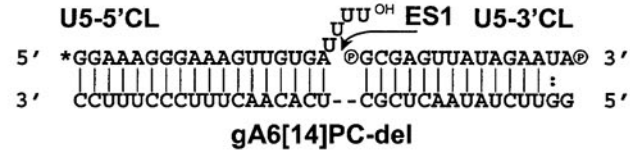
A**B****C**

Fig. 2. (A) Ribozyme A6RZ is shown in association with its substrate SubA6RZ with the cleavage site of SubA6RZ indicated by an arrow. The conserved 5'-CUGA-3' of A6RZ in the catalytic core is essential for ribozyme activity and the U in bold indicates ES1 where four Us are removed from pre-A6RZ by editing. (B) SubA6RZ cleavage by synthetic A6RZ. Labeled SubA6RZ was incubated with pre-A6RZ (10 mM MgCl₂) or with A6RZ at a serial concentration of MgCl₂. Cleavage was generated only when both MgCl₂ and A6RZ were present in the reaction. (C) The cleavage product was generated on incubation of labeled SubA6RZ with unlabeled pre-A6RZ, gA6RZ, and mitochondrial extract (ME).

3.4. Precleaved RNA Editing

The ability to mimic editing of pre-mRNA following the endonucleolytic cleavage step was assessed in the *in vitro* precleaved editing assay in which synthetic pre-mRNA is provided as 5' and 3' cleavage fragments of the pre-mRNA (Fig. 3) (10). This assay has been particularly useful in examining the U addition, U deletion, and RNA ligation catalytic steps of editing. These precleaved mRNAs were incubated with synthetic cognate gRNA and mito-

A**B**

chondrial extract containing the editing complexes. The production of edited RNA was assayed as previously described. The substrate RNAs are based on A6 pre-mRNA that is cleaved either at editing site 1, where gA6 specifies the removal of two Us, or at editing site 2, where the same gRNA specifies addition of two Us, to examine U deletion and insertion, respectively. The 3' pre-mRNA fragments have a 5' monophosphate to mimic the RNA resulting from endoribonucleolytic cleavage in vitro. The 3' end of the 3' pre-mRNA fragment was blocked by addition of a 3' monophosphate.

1. Each editing reaction requires 1 μL each of 0.25 μM [^{32}P]-labeled 5' RNA fragment, 1 μM 3' RNA fragment, and 0.5 μM gRNA (specific RNAs used for precleaved editing assays are listed in **Subheading 2.**) (see **Notes 1–13**). Prepare a sufficient amount of these RNAs for the desired number of reactions plus an additional two reactions. Incubate at 65°C for 2 min, and leave at room temperature for at least 10 min.
2. Assemble the master mix. Make enough for all desired reactions and two additional reactions. *Optional:* For reactions with no added ATP, substitute 0.3 μL of H_2O per reaction. The protocol is for reactions in which 7 μL of extract will be used in each reaction. For reactions in which a smaller volume of extract will be used, the volume should be adjusted with H_2O and 2X HHE (no KCl). *The total volume of the master mix will be 27 μL minus the volume of mitochondrial extract added.*

<u>Component</u>	<u>Volume per reaction</u>	<u>Final concentration in reaction</u>
DEPC- H_2O	3.0	—
2X HHE, no KCl	11.5	1X HHE
0.1 M CaCl_2	1.5	5 mM CaCl_2
200 μM ATP	1.5	10 μM ATP
2 mM UTP (insertion assay only)	1.5	0.1 mM UTP
<i>Torula</i> RNA	1.0	16.7 ng/ μL <i>Torula</i> RNA
Total volume	20.0 μL	

Fig. 3. (see opposite page) Precleaved insertion and deletion RNA editing. (A) Precleaved insertion editing RNAs (**top**); precleaved insertion editing reactions performed with 5'-labeled 5'CL18, 3'CL13pp, and gPC-A6-2A gRNA that specifies the insertion of 2 Us with column-purified mitochondrial extract (**bottom**). Input RNA (arrow), 5' fragment with two added Us (+2U), edited product with two inserted Us (E[+2U]), and ligation of the 5' and 3' fragments without added Us (L) are indicated. (B) Precleaved deletion editing RNAs (**top**); precleaved deletion editing reactions performed with 5'-labeled U5-5'CL, U5-3'CL, and gA6[14]PC-del gRNA that specifies the deletion of four Us with column-purified mitochondrial extract (**bottom**). Input RNA (arrow), 5' fragment with four Us removed (-4U), edited product (E[-4U]), and ligation of the 5' and 3' fragments without Us removed are indicated.

3. Add annealed RNA mix to the master mix. Aliquot this into 1.5-mL tubes at room temperature. The total volume of the reaction, including the extract, should be 30 μ L. Mix by pipetting after adding each ingredient: 23 μ L of master mix + RNA, 7 μ L of mitochondrial extract, for a total volume of 30 μ L. Incubate for 3 h at 28°C in a heat block.
4. Add 2 μ L of stop buffer to each reaction and then add 100 μ L of H₂O. Extract once with 135 μ L of PCI. Add 2 μ L of 5 μ M competitor RNA (*see* **Notes 10–13**). Add 2 μ L of 5 μ g/ μ L glycogen and 13.5 μ L (0.1 vol) of 3 M NaOAc, pH 5.2. Mix and add 350 μ L (2.5 vol) of cold 100% EtOH. Mix by inversion and incubate for at least 1 h at –20°C.
5. Spin for 30 min in a microfuge at 4°C, wash with 150 μ L of cold 85% EtOH, spin for 5 min at 4°C, remove the supernatant, and dry in a SpeedVac. Resuspend the RNA in 20 μ L of 7 M urea-loading buffer.
6. Run 5–10 μ L of each reaction on a 14.5% polyacrylamide gel. Run until the bromophenol blue dye runs off the end of the gel, mark the location of the xylene cyanol, and allow the xylene cyanol to run an additional 1 to 2 in. The run will take about 4 h.
7. Remove the gel from both plates and sandwich in plastic wrap. Carefully remove all liquid and bits of gel from outside of the plastic wrap. It is very difficult to dry a denaturing 14.5% gel. However, the RNA will diffuse very little in a 14.5% gel kept at room temperature overnight, so it is possible to expose a PhosphorImager screen overnight without losing resolution.

3.5. Assays for Editing-Associated Activities

3.5.1. Site-Specific Endonuclease Activity

The proteins that catalyze the endoribonucleolytic cleavage during RNA editing have not yet been identified. The most likely candidates at the time of this writing are one or more of five related proteins, which have RNase III or RNase III-like motifs (**11**). Mitochondrial extract containing editosomes catalyzes endoribonucleolytic cleavage of pre-mRNA. The cleavage occurs 5' of the predicted anchor duplex between the pre-mRNA and gRNA but does not invariably occur at the first unpaired nucleotide 5' to the anchor duplex (**12**). This suggests that substrate recognition and selection of cleavage probably entails recognition of an RNA duplex and adjacent unpaired region. Such substrate characteristics often occur immediately 5' to the initial, or extended, anchor duplex but may also occur elsewhere as the pre-mRNA and gRNA undergo a variety of interactions during editing. A realignment model of editing has been proposed to account for the characteristics of partially edited pre-mRNA (**13**) and is consistent with endonucleolytic cleavage at sites other than immediately 5' to the initial or extended anchor duplex. The cleavage reaction is not well characterized, but it does not entail production of a chimeric reaction intermediate formed by a transesterification reaction in which the

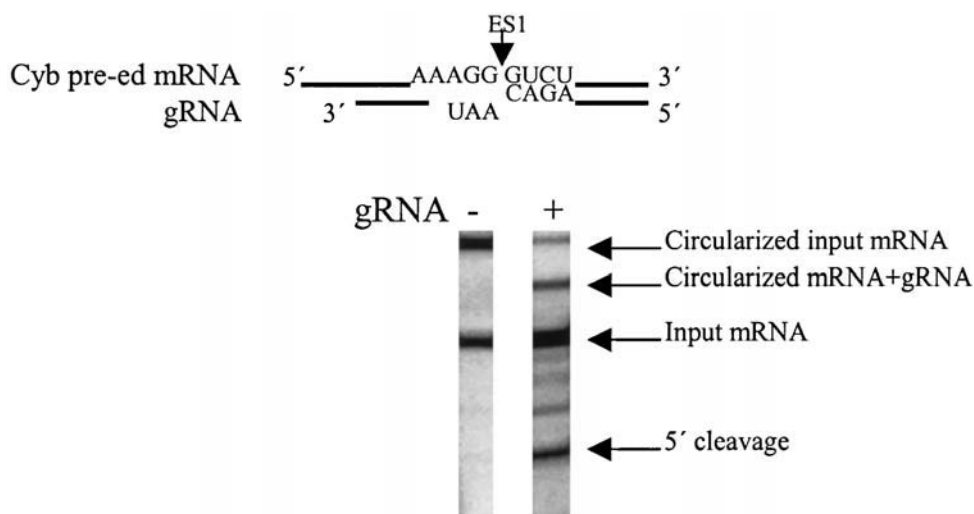


Fig. 4. CYb endonuclease assay. (**Top**) The 104-nt CYb preedited mRNA base pairing with 30-nt CYb gRNA with the first ES1 cleavage site indicated by arrow; (**bottom**) autoradiogram of gRNA-dependent cleavage of 5'-labeled CYb preedited mRNA in presence of peak of Superose 6 fraction enriched for RNA editing and presence (+) or absence of gRNA (-).

3' OH of the gRNA attacks the phosphodiester bond at the editing site leading to a covalent linkage with the 3' fragment, as has been previously proposed. This is evident because 3' cleavage products are produced during in vitro editing with kinetics, indicating that they are intermediates that are generated prior to the edited RNA (3).

Characterization of RNA-editing endonuclease and related activities in *T. brucei* editing complexes have been described in several mitochondrial extracts (14–16). There are two general classes of endoribonuclease activities. One cofractionates with the in vitro editing activity and requires gRNA, whereas the other fractionates away from the in vitro editing activity and is not dependent on added gRNA. This indicates that the former activity has a role in editing, and editing site selection, while the latter is likely to have other roles. The assay described here uses CYb pre-mRNA at the first editing site (Fig. 4).

1. Combine 1 μ L each of gRNA with the [^{32} P] CYb preedited mRNA per reaction to be performed, and enough for one additional reaction. The final amounts are to be 0.5 pmol/reaction of gRNA and 0.25 pmol/reaction of substrate.
2. Assemble the reaction at room temperature: 2 μ L of 5X Tris buffer, 2 μ L of gRNA + substrate RNA, 5 μ L of mitochondrial extract, 1 μ L of DEPC- H_2O , for a total volume of 10 μ L. Incubate for 40 min at 28°C in a heat block, and proceed with steps 4–9 of the full-round in vitro RNA editing in Subheading 3.2.

3.5.2. 3' Exouridylylase Activity (see **Note 14**)

The protocol is the same as for the precleaved deletion assay, but leave out the ATP (see **Subheading 3.4.**).

3.5.3. TUTase Activities (see **Note 15**)

Terminal uridylyl transferases transfer UMP from UTP to the 3' hydroxyl of an RNA substrate. The 3' TUTase activity in trypanosomes was first observed in whole-cell extracts in 1987 (**17**) and subsequently in isolated mitochondria (**18**). Since that time, Us have been shown to be added in insertion RNA editing and posttranscriptionally to the 3' ends of gRNAs and mitochondrial rRNAs. Recently, two trypanosomatid TUTases that are involved in RNA editing were identified that differ in their substrate specificities and catalytic properties. The larger 3' TUTase, RNA editing TUTase 1, or RET1, is essential for editing, as determined by gene inactivation experiments (**19**), but it is not a stable component of the complex that contains the endonuclease ExoUase and RNA ligase activities (**20**). RET1 catalyzes the processive addition of many Us to single-stranded RNA (ssRNA) without a strong preference for the 3' terminal nucleotide (**20**). By contrast, RET2 is an integral component of the editosome and primarily adds a single U to an ssRNA primer and the specified number of Us to a double-stranded precleaved insertion editing substrate (**21**). Hence, RET2 likely adds the Us to mRNA during editing, and while RET1 may add Us to mRNA during editing, it may also add the oligo (U) tails to gRNAs.

TUTase activity is readily detected by incorporation of radiolabeled UTP to a nonspecific RNA primer such as yeast tRNA that results in acid-precipitable RNAs that can be retained on glass filters and quantified by scintillation counting. This assay is simple and sensitive and can be readily quantified. However, it is obviously biased toward detecting activities that add many Us, such as RET1, since the amount of activity is based on the number of incorporated Us. Alternative substrates include the 5' fragment from the precleaved assay and this RNA along with the 3' fragment and cognate gRNA from the precleaved insertion editing assay, thus mimicking a U addition editing substrate following cleavage. Use of these substrates followed by resolution of the products on denaturing polyacrylamide gels (**Fig. 5**) makes it possible to qualitatively distinguish between the two TUTases.

3.5.3.1. UTP INCORPORATION INTO YEAST RNA

1. Prepare the master mix (plus two additional reactions) at room temperature and aliquot to tubes: 2 μ L of 10X buffer, 1 μ L of yeast tRNA, 0.5 μ L of α -[^{32} P]-UTP, 11.5 μ L of DEPC H₂O, for a final volume of 15 μ L.

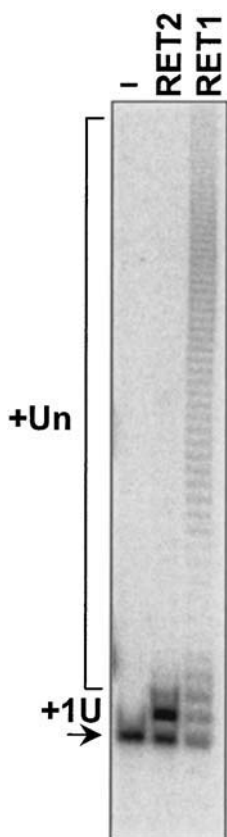


Fig. 5. TUTase activity with 5'-labeled 5'CL18. U addition was assayed with 5'-labeled 5'CL18 using column-purified mitochondrial extracts containing primarily RET2 or RET1. Products with one or many added Us are designated by +1U and +Un, respectively. The arrow indicates input RNA and the dash indicates RNA incubated without added fraction.

2. Add mitochondrial extract and mix by pipeting: 15 μ L of master mix, 5 μ L of mitochondrial extract, for a total volume of 20 μ L.
3. Incubate at 27°C for 30–60 min.
4. Spot 5 μ L onto glass fiber disks (Whatman GF/C) on filter paper.
5. Wash the disks three times with cold 10% TCA with 50 mM disodium pyrophosphate for 10 min each (approx 80–100 mL/10 disks).
6. Rinse the disks once with cold 95% ethanol.
7. Dry and count in a scintillation counter.

3.5.3.2. U ADDITION TO SINGLE AND DOUBLE-STRANDED RNA SUBSTRATES

Assay U addition to 5'CL18 alone or precleaved insertion editing substrate as described for precleaved insertion editing assays (see **Subheading 2.4.**).

3.5.4. RNA Ligase Activity

Shortly after the discovery of RNA editing, whole-cell *T. brucei* lysates were found to ligate RNA (17). When mitochondrial extracts were incubated with α -[^{32} P] ATP, two bands of 48 and 52 kDa were visible on SDS-PAGE gels and identified as potential RNA ligases based on protein adenylation, which is a reaction step common to ligases (22). Incubation of the adenylylated proteins with ligatable RNA substrates resulted in deadenylylation, giving further evidence of their identity as ligases (22). The two adenylylatable bands cosedimented with RNA-editing activity as well as precleaved editing (9,23), indicating an association with RNA editing. Following purification of the editosome, the genes for these ligases were cloned (24–26) and their function as ligases was demonstrated in vitro (26,27). Conditional inactivation of the gene for the larger ligase, RNA editing ligase 1 (REL1), resulted in cessation of editing in bloodstream forms and cell death (26,27). However, inactivation of expression of the second ligase, REL2, by RNA interference did not affect cell growth in either procyclic or bloodstream forms (28). Why there are two RNA ligases, one essential and the other not, may be related to the existence of insertion and deletion editing subcomplexes (29,30), and the possibility that REL1 can compensate for REL2, but not the reverse.

3.5.4.1. ADENYLATION ASSAY

1. Prepare the master mix: 2 μ L of 5X buffer, 2 μ L of DMSO, 0.5 μ L of α - ^{32} P ATP (3000 Ci/mmol), 10.5 μ L of DEPC-H₂O, for a total volume of 20 μ L.
2. Add 5 μ L of mitochondrial extract and incubate at 28°C in a heat block for 15 min.
3. Stop the reactions by adding 7 μ L of 3X SDS-PAGE dye.
4. Boil the reactions for 5 min, ensuring that the caps will not burst by using lid locks. Briefly centrifuge at room temperature to collect evaporated material on the caps and to prevent contamination.
5. Load 10 μ L onto a 10% SDS-PAGE gel. Run until the bromophenol blue dye is at the bottom of the gel or has just run off.
6. Cover the gel in plastic wrap and dry for about 1 h with heat and vacuum.
7. Expose the gel to film or a phosphorimager cassette.

3.5.4.2. PRECLEAVED RNA LIGATION ASSAY

This assay is essentially identical to the precleaved insertion editing assay (see **Subheading 3.4.**) except for the exclusion of UTP and Torula RNA in the master mix and the use of ArtgA6 0A as the gRNA.

3.5.5. RNA Helicase

A role for RNA helicase mHEL61p has been suggested since null mutants of procyclic forms have reduced editing and grow slowly, suggesting that this protein has a role in editing but is not essential, perhaps owing to the presence of another helicase. Encouragingly, the protein has been identified as a component of the editing complex (20). The helicase assay uses gRNA gA6[14] hybridized to its cognate domain of the edited ATPase 6 mRNA (31).

1. Combine 1 μ L each of gRNA and [^{32} P]substrate RNA per reaction to be performed, plus enough for one extra reaction. (Final amounts will be 0.5 pmol/reaction of gRNA and 0.25 pmol/reaction of substrate.) Incubate at 65°C for 2 min and let sit at room temperature for at least 10 min.
2. Combine the following: 10 μ L of 2X HHE, 2 μ L of gRNA + substrate RNA, 2 μ L of 10 mM ATP, 5 μ L of mitochondrial extract, 1 μ L of DEPC-H₂O, for a total volume of 20 μ L. Incubate for 30–60 min at 28°C in a heat block and proceed with **steps 4–9** of the full-round in vitro RNA editing in **Subheading 3.2**.

4. Notes

1. For best results, the PCR product should contain several micrograms of DNA, especially for short transcripts (<40 nt). Low yield is most often caused by too little template DNA.
2. Mutagenic PCR is possible.
3. The PCR product does not need to be cleaned up after the PCR reaction.
4. T7 polymerase adds an extra random nucleotide to the 3' end of about half of transcripts. If the identity of the 3' terminal nucleotide is important (i.e., for 5' cleavage fragments), run the gel long enough to achieve separation of the desired sequence from the +1-nt sequence. The RNA and dyes may migrate more slowly than usual owing to the large quantity of solutes present in the transcription product.
5. A ladder of partial-length transcripts may be visible in the gel under UV light, which may help to identify the desired-length product. The desired product should appear as two closely spaced bands (that differ in size by 1 nt).
6. Before using T7 transcripts for precleaved editing reactions, the transcript should be tested for proper length and, if possible, 3' end sequence. Length may be verified by ligating a radiolabeled 5' fragment to a 3' fragment with a 5' phosphate, and subsequent RNase T₁ digest (*see Subheading 3.2.*). A 3'-terminal G may be identified directly by T₁ digest, while a 3'-terminal G or A may be identified by nuclease P1 digest (15).
7. Note the difference in ATP concentration between insertion and deletion reactions (32).
8. Several different RNA substrates can be used in this reaction.
9. This formula assumes use of 7 μ L of mitochondrial extract per reaction. If other volumes of extract are used, adjust the H₂O and 2X HHE as necessary. If extracts

with high salt (200 mM KCl or greater) are used, add 2X HHE *without* KCl to the master mix, rather than 2X HHE with KCl.

10. The 5' cleavage fragment may be labeled either by guanylyltransferase (capping) or by removal of the 5' phosphate followed by addition of a radiolabeled phosphate using polynucleotide kinase. The latter method generally generates RNA with a much higher specific activity, but the presence of a 5' monophosphate allows circularization by RNA ligase in *T. brucei* mitochondrial extract. Thus, the gRNA must bind the 5' cleavage fragment tightly, and an excess to the 5' fragment.
11. If the 5' RNA fragment is labeled by the phosphatase-kinase method, it is possible to use a two- to fivefold dilution of this RNA since it has such a high specific activity. In this case, the concentrations of the other RNAs are not changed.
12. The A6-like substrate previously described binds tightly to its gRNA, so it is necessary, even in a denaturing gel, to have competitor RNA in the sample when it is loaded onto the gel.
13. Adjust the final KCl concentration to as close to 50 mM as possible for optimal activity. A KCl concentration between 12 and 60 mM is acceptable. If addition of KCl is required, such as owing to extract that is low in KCl, add it from a solution of 0.5 M KCl.
14. Mitochondrial extract may ligate the cleavage fragments even in the absence of added ATP.
15. This assay is designed for mitochondrial extract containing 50 mM KCl. When using extracts with varying KCl concentrations, use 10X buffer without KCl and adjust the final KCl concentration to between 15 and 60 mM with 100 mM KCl.

Acknowledgment

This work was supported by National Institute of Health grant AI/14102.

References

- 1 Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA template. *Nucleic Acids Res.* **15**, 8783–8798.
- 2 Kable, M. L., Seiwert, S. D., Heidmann, S., and Stuart, K. (1996) RNA editing: a mechanism for gRNA-specified uridylyte insertion into precursor mRNA. *Science* **273**, 1189–1195.
- 3 Seiwert, S. D., Heidmann, S., Stuart, K. D., and Stuart, K. (1996) Direct visualization of uridylyte deletion in vitro suggests a mechanism for kinetoplastid RNA editing. *Cell* **84**, 831–841.
- 4 Read, L. K., Göringer, H. U., and Stuart, K. (1994) Assembly of mitochondrial ribonucleoprotein complexes involves specific guide RNA (gRNA)-binding proteins and gRNA domains but does not require preedited mRNA. *Mol. Cell. Biol.* **14**, 2629–2639.
- 5 Decker, C. J. and Sollner-Webb, B. (1990) RNA editing involves indiscriminate U changes throughout precisely defined editing domains. *Cell* **61**, 1001–1011.

- 6 Harris, M., Decker, C., Sollner-Webb, B., and Hajduk, S. (1992) Specific cleavage of pre-edited mRNAs in trypanosome mitochondrial extracts. *Mol. Cell. Biol.* **12**, 2591–2598.
- 7 Piller, K. J., Decker, C. J., Rusché, L. N., and Sollner-Webb, B. (1995) *Trypanosoma brucei* mitochondrial guide RNA-mRNA chimera-forming activity cofractionates with an editing-domain-specific endonuclease and RNA ligase and is mimicked by heterologous nuclease and RNA ligase. *Mol. Cell. Biol.* **15**, 2925–2932.
- 8 Burgess, M. L. K., Heidmann, S., and Stuart, K. D. (1999) Kinetoplastid RNA editing does not require the terminal 3' hydroxyl of guide RNA but modifications to the guide RNA terminus can inhibit *in vitro* U insertion. *RNA* **5**, 883–892.
- 9 Corell, R. A., Read, L. K., Riley, G. R., et al. (1996) Complexes from *Trypanosoma brucei* that exhibit deletion editing and other editing-associated properties. *Mol. Cell. Biol.* **16**, 1410–1418.
- 10 Igo, R. P. Jr., Palazzo, S. S., Burgess, M. L. K., Panigrahi, A. K., and Stuart, K. (2000) Uridylate addition and RNA ligation contribute to the specificity of kinetoplastid insertion RNA editing. *Mol. Cell. Biol.* **20**, 8447–8457.
- 11 Stuart, K. and Panigrahi, A. K. (2002) RNA editing: complexity and complications. *Mol. Microbiol.* **45**, 591–596.
- 12 Lawson, S., Igo, R. P. Jr., Salavati, R., and Stuart, K. D. (2000) The specificity of nucleotide removal during RNA editing in *Trypanosoma brucei*. *RNA* **7**, 1–10.
- 13 Koslowsky, D. J., Bhat, G. J., Read, L. K., and Stuart, K. (1991) Cycles of progressive realignment of gRNA with mRNA in RNA editing. *Cell* **67**, 537–546.
- 14 Salavati, R., Panigrahi, A. K., Morach, B., Palazzo, S. S., Igo, R. P. Jr., and Stuart, K. D. (2002) Endoribonuclease activities of *Trypanosoma brucei* mitochondria. *Mol. Biochem. Parasitol.* **120**, 23–31.
- 15 Piller, K. J., Rusché, L. N., Cruz-Reyes, J., and Sollner-Webb, B. (1997) Resolution of the RNA editing gRNA-directed endonuclease from two other endonucleases of *Trypanosoma brucei* mitochondria. *RNA* **3**, 279–290.
- 16 Adler, B. K. and Hajduk, S. L. (1997) Guide RNA requirement for editing-site-specific endonucleolytic cleavage of preedited mRNA by mitochondrial ribonucleoprotein particles in *Trypanosoma brucei*. *Mol. Cell. Biol.* **17**, 5377–5385.
- 17 White, T. C. and Borst, P. (1987) RNA end-labeling and RNA ligase activities can produce a circular rRNA in whole cell extracts from trypanosomes. *Nucleic Acids Res.* **15**, 3275–3290.
- 18 Bakalara, N., Simpson, A. M., and Simpson, L. (1989) The *Leishmania* kinetoplast-mitochondrion contains terminal uridylyltransferase and RNA ligase activities. *J. Biol. Chem.* **264**, 18,679–18,686.
- 19 Aphasizhev, R., Sbicego, S., Peris, M., Jang, S. H., Aphasizheva, I., Simpson, A. M., Rivlin, A., and Simpson, L. (2002) Trypanosome mitochondrial 3' terminal uridylyl transferase (TUTase): the key enzyme in U-insertion/deletion RNA editing. *Cell* **108**, 637–648.
- 20 Panigrahi, A. K., Schnaufer, A., Ernst, N., Wang, B., Carmean, N., Salavati, R., and Stuart, K. (2003) Identification of novel components of *Trypanosoma brucei* editosomes. *RNA* **9**, 484–492.

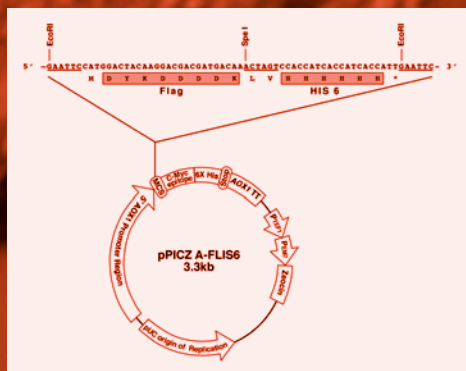
- 21 Ernst, N. L., Panicucci, B., Igo, R. P. Jr., Panigrahi, A. K., Salavati, R., Stuart, K. (2003) TbMP57 is a 3' terminal uridylyl transferase (TUTase) of the *Trypanosoma brucei* editosome. *Mol. Cell.* **11**(6), 1525–1536.
- 22 Sabatini, R. and Hajduk, S. L. (1995) RNA ligase and its involvement in guide RNA/mRNA chimera formation. Evidence for a cleavage-ligation mechanism of *Trypanosoma brucei* mRNA editing. *J. Biol. Chem.* **270**, 7233–7240.
23. Rusché, L. N., Cruz-Reyes, J., Piller, K. J., and Sollner-Webb, B. (1997) Purification of a functional enzymatic editing complex from *Trypanosoma brucei* mitochondria. *EMBO J.* **16**, 4069–4081.
- 27 Panigrahi, A. K., Gygi, S., Ernst, N., et al. (2001) Association of two novel proteins, TbMP52 and TbMP48, with the *Trypanosoma brucei* RNA editing complex. *Mol. Cell. Biol.* **21**, 380–389.
- 25 McManus, M. T., Shimamura, M., Grams, J., and Hajduk, S. L. (2001) Identification of candidate mitochondrial RNA editing ligases from *Trypanosoma brucei*. *RNA* **7**, 167–175.
26. Rusché, L. N., Huang, C. E., Piller, K. J., Hemann, M., Wirtz, E., and Sollner-Webb, B. (2001) The two RNA ligases of the *Trypanosoma brucei* RNA editing complex: cloning the essential Band IV gene and identifying the Band V gene. *Mol. Cell. Biol.* **21**, 979–989.
- 27 Schnauffer, A., Panigrahi, A. K., Panicucci, B., Igo, R. P. Jr., Salavati, R., and Stuart, K. D. (2001) An RNA ligase essential for RNA editing and survival of the bloodstream form of *Trypanosoma brucei*. *Science* **291**, 2159–2162.
- 28 Drozd, M., Palazzo, S. S., Salavati, R., O'Rear, J., Clayton, C., and Stuart, K. (2002) TbMP81 is required for RNA editing in *Trypanosoma brucei*. *EMBO J.* **21**, 1791–1799.
- 29 Cruz-Reyes, J., Zhelonkina, A. G., Huang, C. E., and Sollner-Webb, B. (2002) Distinct functions of two RNA ligases in active *Trypanosoma brucei* RNA editing complexes. *Mol. Cell. Biol.* **22**, 4652–4660.
- 30 Stuart, K. and Panigrahi, A. K. (2002) RNA editing: complexity and complications. *Mol. Microbiol.* **45**, 591–596.
- 31 Missel, A., Souza, A. E., Norskau, G., and Göringer, H. U. (1997) Disruption of a gene encoding a novel mitochondrial DEAD-box protein in *Trypanosoma brucei* affects edited mRNAs. *Mol. Cell. Biol.* **17**, 4895–4903.
32. Cruz-Reyes, J., Rusche, L. N., Piller, K. J., and Sollner-Webb, B. (1998) *T. brucei* RNA editing: adenosine nucleotides inversely affect U-deletion and U-insertion reactions at mRNA cleavage. *Mol. Cell* **1**, 401–409.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



Identification and Characterization of Trypanosome RNA-Editing Complex Components

Kenneth Stuart, Aswini K. Panigrahi, and Achim Schnauffer

Summary

This chapter describes the methods used to purify the RNA-editing complex, to identify the proteins by mass spectrometry, and to demonstrate the functions for some of the proteins.

Key Words: Chromatography; editosome; mass spectrometry; monoclonal antibodies; protein; RNA editing; tandem affinity purification tag.

1. Introduction

Most mitochondrial mRNAs in trypanosomes are post transcriptionally edited, and the mature functional mRNAs are generated by the insertion and deletion of numerous uridylates (Us). This type of editing process is catalyzed by a series of coordinated enzymatic steps. The steps are endonucleolytic cleavage of the pre-mRNA followed by addition or removal of one or more Us at the 3' end of the resultant 5' cleavage fragment for insertion or deletion editing, respectively. Then the edited 5' RNA fragment is ligated to the 3' RNA fragment by RNA ligase. The editing sites and the number of Us to be inserted and deleted are specified by small guide RNAs (gRNAs). The editosome is a multiprotein complex that catalyzes these steps of RNA editing. Recent progress in the purification of this complex led to the identification of a number of its components, and the functions of some of the component proteins have been demonstrated (*see reviews in refs. 1–3*). This chapter describes the methods used to purify the RNA-editing complex, to identify the proteins by mass spectrometry (MS), and to demonstrate the functions for some of the proteins.

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

2. Materials

2.1. Equipment and Supplies

1. Autoradiography system/PhosphorImager.
2. Centricon-YM50 membrane (Amicon).
3. Centrifuge bottles and tubes.
4. Centrifuge machines.
5. Dounce and pestle.
6. Dynabeads M-450 goat antimouse IgG (Dynal).
7. Electroporation cuvet.
8. Electroporator.
9. Filters.
10. Fast protein liquid chromatography (FPLC) system.
11. Gradient maker.
12. Incubator/shaker incubator.
13. LCQ mass spectrometer.
14. Magnetic stand.
15. Polyvinyl difluoride (PVDF) membrane.
16. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) system.
17. Speedvac.
18. Syringe and needle.
19. TC flasks.
20. X-ray film/phosphor screen.

2.2. Chromatography Columns

1. C18 microcapillary column (packed in-house, can be purchased from New Objective).
2. Econo-Columns (5 mL) (Bio-Rad, Hercules, CA).
3. Econo-Pac columns (20 mL) (Bio-Rad).
4. Hi-Trap SP Sepharose HR 5-mL column (Pharmacia).
5. Hi-Trap Q Sepharose HR 1-mL column (Pharmacia).
6. Superose 6 HR (10/30) column (Pharmacia).

2.3. Reagents

1. Acetic acid.
2. Acetonitrile.
3. Anti-6xHis antibody (Clontech).
4. Antimouse IgG HRP conjugate (Bio-Rad).
5. [α - ^{32} P]-ATP.
6. Bovine serum albumin (BSA).
7. "Complete mini" protease inhibitor tablets (EDTA-free; Roche, Indianapolis, IN).
8. Coomassie blue/silver nitrate/Sypro-Ruby staining supplies.
9. ECL or ECL Plus kit (Amersham).
10. Fetal bovine serum (FBS).

11. Formic acid.
12. G418.
13. Hemin.
14. Hygromycin.
15. IgG Sepharose beads ("fast flow"; Amersham).
16. Nonfat milk powder.
17. Percoll.
18. Phleomycin.
19. Protease inhibitors (leupeptin, pepstatin, and pefabloc).
20. Tetracycline hydrochloride.
21. TNT cell-free-coupled transcription/translation system (Promega, Madison, WI).
22. DNase.
23. TEV protease (Invitrogen, Carlsbad, CA).

2.4. Media and Buffers

1. 2X SDS-PAGE sample buffer: 100 mM Tris-HCl, pH 6.8; 200 mM dithiothreitol (DTT); 20% glycerol; 4% SDS; 0.01% bromophenol blue.
2. Calmodulin binding buffer (CBB): 10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1% NP40, 10 mM β -mercaptoethanol, 1 mM Mg acetate, 1 mM imidazole, 2 mM CaCl_2 . Add β -mercaptoethanol just before use.
3. Calmodulin elution buffer (CEB): 10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1% NP40, 10 mM β -mercaptoethanol, 1 mM Mg acetate, 1 mM imidazole, 2 mM EGTA. Add β -mercaptoethanol just before use.
4. CytoMix: 25 mM HEPES, pH 7.6; 120 mM KCl, 0.15 mM CaCl_2 , 10 mM $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$, pH 7.6, 2 mM EDTA, 5 mM MgCl_2 , 6 mM glucose.
5. DTE buffer: 1 mM Tris-HCl, pH 8.0, 1 mM EDTA.
6. HHE buffer: 25 mM HEPES, pH 7.9, 10 mM $\text{Mg}(\text{OAc})_2$, 50 mM KCl, 1 mM EDTA.
7. IP buffer: 10 mM Tris, pH 7.2, 10 mM MgCl_2 , 200 mM KCl, 0.1% Triton X-100.
8. IPP150: 10 mM Tris HCl, pH 8.0, 150 mM NaCl, 0.1% NP40.
9. Phosphate-buffered saline (PBS): 10 mM phosphate buffer, pH 7.2, 150 mM NaCl.
10. PBS-G: 10 mM phosphate buffer, pH 7.2, 150 mM NaCl, 6 mM glucose.
11. PBST: 10 mM phosphate buffer, pH 7.2, 150 mM NaCl, 0.05% Tween-20.
12. Q-A buffer: 10 mM Tris-HCl, pH 8.3, 10 mM MgCl_2 , 50 mM KCl, 1 mM DTT.
13. Q-B buffer: 10 mM Tris-HCl, pH 8.3, 10 mM MgCl_2 , 1 M KCl, 1 mM DTT.
14. S6 buffer: 10 mM Tris-HCl, pH 7.0, 10 mM MgCl_2 , 200 mM KCl, 1 mM DTT.
15. SBG buffer: 0.2 M phosphate buffer, pH 7.9, 150 mM NaCl, 20 mM glucose.
16. SDM-79 medium (JRH Biosciences).
17. SP-A buffer: 10 mM Tris-HCl, pH 7.0, 10 mM MgCl_2 , 50 mM KCl, 1 mM DTT.
18. SP-B buffer: 10 mM Tris-HCl, pH 7.0, 10 mM MgCl_2 , 1 M KCl, 1 mM DTT.
19. STM buffer: 20 mM Tris-HCl, pH 8.0, 250 mM sucrose, 2 mM MgCl_2 .
20. TEV cleavage buffer (TEVCB): 10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1% NP40, 0.5 mM EDTA, 1 mM DTT. Add DTT just before use.

21. Trypsinization solution: sequencing-grade modified trypsin (Promega) (12.5 ng/μL), 50 mM NH₄-HCO₃.

3. Methods

The editing complexes were initially isolated from *Trypanosoma brucei* cells by conventional column chromatography techniques, and a panel of monoclonal antibodies (MAbs) was generated against the proteins (4,5). MAbs against component proteins were used to affinity purify the complexes. The proteins were identified by MS. Genetic approaches were used to incorporate tags into some editosome proteins in *T. brucei* cells, and these tagged proteins were used to efficiently purify functional complexes. Sixteen proteins that form the stable core of the editosome, which contains the multiple catalytic activities, were consistently detected in these editosomes purified by different means (6).

3.1. Fractionation of Mitochondria

3.1.1. Cell Culture

1. Grow *T. brucei* procyclic forms (strain IsTaR 1.7a) to log phase in vitro in SDM-79 medium containing 10% FBS and hemin (7.5 mg/mL) at 27°C with shaking at 80–100 rpm (7). The culture can be maintained in a 10-mL volume in 25-cm² TC flasks (at upright position).
2. Expand the culture 10-fold when the cells reach a density of 3 to 4 × 10⁷ per mL.
3. For large-scale preparations, grow several 1-L cultures in 2-L glass flasks (the doubling time is about 17 h).

3.1.2. Isolation of Mitochondria

The following method is used to isolate the mitochondrial vesicles from procyclic cells (8).

1. Harvest the cells by centrifuging at 6000g for 10 min at 4°C, and wash the cells once in SBG buffer (for 8- to 12-L cultures use 300 mL of buffer).
2. Resuspend the pellet in DTE buffer (the total number of cells divided by 10⁹ equals the number of milliliters of DTE to be added), and quickly disrupt the cells with a sterile 100-mL Dounce homogenizer using a tight-fitting pestle B for five strokes (examine the cells microscopically to ensure lysis).
3. Immediately add sucrose to a final concentration of 0.25 M (167 μL of 60% sucrose/mL of DTE used).
4. Centrifuge the slurry at 15,000g for 10 min at 4°C and pour off the supernatant.
5. Resuspend the pellet in STM buffer (0.167 mL/mL of lysate used) and add MgCl₂ to 3 mM, CaCl₂ to 0.3 mM, and DNase to 9 μg/mL of final concentrations.
6. Incubate on ice for 60 min and add an equal volume of STE. Centrifuge as in **step 4** and pour off the supernatant.
7. Resuspend the pellet in 70% Percoll (4 mL/2 L of culture), and homogenize using a small Dounce homogenizer with a tight-fitting pestle B for five strokes.

8. Pour 32-mL linear Percoll gradients of 20–35% and layer 5 mL of the lysate under it using a long needle.
9. Centrifuge at 103,900g for 60 min at 4°C, and collect the mitochondrial fraction using a syringe and an 18-gage needle (the fraction is the broad smear of density range of 1.052–1.069 g/mL between the two prominent bands).
10. Wash the mitochondria three to four times with 1X STE buffer (following each wash the mitochondrial vesicles can be pelleted at 32,530g for 15 min). The mitochondrial vesicles can be stored at –70°C.

3.1.3. Lysis

The mitochondrial vesicles are lysed with nonionic detergent Triton X-100 for subsequent fractionation (4).

1. Resuspend the vesicles isolated from 10–12 L of cultured cells ($\sim 2 \times 10^{11}$ cells) in 30 mL of SP-A buffer. Add a cocktail of protease inhibitors (10 μ g/mL of lupeptin, 5 μ g/mL of pepstatin, and 1 μ M Pefabloc) to the buffer and add nonionic detergent Triton X-100 to a final concentration of 0.5% (from 10% stock solution).
2. Carry out lysis for 15 min at 4°C with bidirectional mixing, and clear the lysate by centrifuging at 17,500g for 30 min at 4°C. (The editing complexes enriched from this preparation are pure and concentrated enough for generating materials for in vitro assays. However, larger preparations are required for generating MAbs.)

3.1.4. Column Chromatography

The cleared mitochondrial lysate from **Subheading 3.1.3.** is fractionated by sequential ion-exchange and gel filtration column chromatography (4).

1. Filter the lysate through a 0.2- μ m membrane and load onto a 5-mL SP Sepharose HR column (Pharmacia) at a 1 mL/min flow rate. All chromatographic steps are carried out at 4°C. (We use an automated LKB-Pharmacia FPLC system.)
2. Wash away the unbound proteins with five column volumes of buffer SP-A.
3. Elute the bound proteins with a 40-mL linear salt gradient of 50–330 mM KCl, followed by a 40-mL linear gradient to 1 M KCl (using buffer SP-B) at a 1 mL/min flow rate.
4. Collect 2-mL fractions and assay for in vitro deletion RNA-editing activity (as described in Chapter 13).
5. Pool all positive fractions (9–19) for further fractionation on a Q Sepharose column.
6. Join two Hi-Trap Q 1-mL columns (Pharmacia) in series and equilibrate with buffer Q-A.
7. Dilute the pooled fractions from the SP Sepharose column and adjust the pH to the same conditions as for buffer Q-A. Load onto the Q-Sepharose column at a 1 mL/min flow rate and wash away the unbound proteins with five column volumes of buffer Q-A.

8. Carry out elution with a 16-mL linear salt gradient to 330 mM KCl, followed by a 14-mL linear gradient to 1 M KCl (using buffer Q-B) at a 0.5 mL/min flow rate.
9. Collect 1-mL fractions and assay for deletion RNA-editing activity.
10. Pool the positive fractions (11–20) (they can be stored with 10% glycerol at -70°C , without further purification for routine in vitro assays).
11. To further purify the editing complex for protein identification and MAb production, treat the sample with 0.1% Triton X-100 and concentrate to 1/10 to 1/20 volume using a Centricon-YM50 membrane (Amicon) at 3000g.
12. Size fractionate the complexes using a Superose 6 HR (10/30) column (Pharmacia). Equilibrate the column with S6 buffer, and in each run load a 250- μL sample onto the column.
13. Fractionate the samples at a 0.2 mL/min flow rate, and collect 500- μL fractions and assay for deletion RNA editing. Most editing activities elute in fractions 19–22.

3.1.5. Density Gradient Centrifugation

The peak positive fractions from Superose 6 (19–22) are pooled and concentrated as in **Subheading 3.1.4.** and sedimented in a 10–30% linear glycerol gradient (4).

1. Prepare an 11-mL linear gradient in 10 mM Tris (pH 7.0), 10 mM MgCl_2 , and 100 mM KCl and layer 500 μL of sample on top of it.
2. Centrifuge the tubes at 38,000 rpm for 5 h at 4°C (SW40 rotor; Beckman).
3. Collect 500- μL fractions from the top and assay for deletion RNA-editing activity. The editing complex that is isolated by column chromatography sediments at $\sim 20\text{S}$ in the gradient (4).

3.2. Immunoaffinity Purification

3.2.1. Monoclonal Antibodies

MAbs were produced at Biologics Production Facility, Fred Hutchinson Cancer Research Center, Seattle, WA.

1. Fractionate the RNA-editing complex in multiple batches as previously described, from a total of 1.64×10^{12} cells.
2. Pool and concentrate the peak editing fractions from glycerol gradients using a Centricon-YM50 membrane, and use as the immunogen.

3.2.2. Immunoprecipitation

Batches of MAbs that are specific for four different editosome proteins are available from our laboratory. Three of these MAbs immunoprecipitate

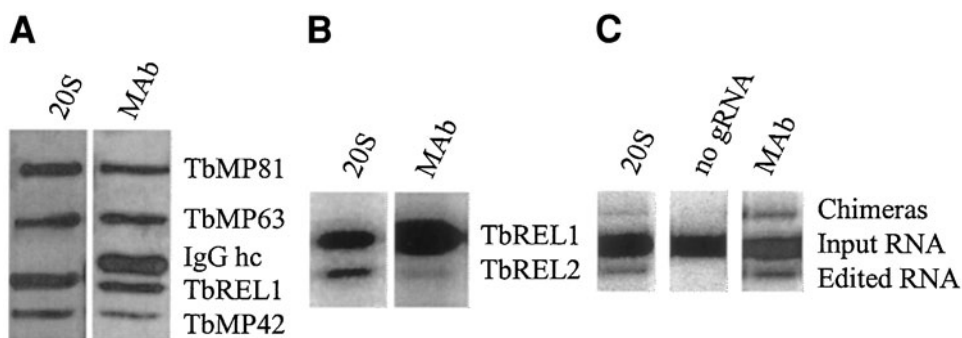


Fig. 1. MAb immunoaffinity purification of editing complexes. (A) Western analysis of 20S and MAb affinity-purified complexes showing presence of editosome proteins TbMP81, TbMP63, TbREL1, and TbMP42. The IgG heavy chain of the MAb is indicated. (B) The presence of TbREL1 and TbREL2 in the complexes is shown by autoadenylation in the presence of radioactive ATP. (C) In vitro deletion RNA-editing assay showing that isolated complexes are functional. Input pre-mRNA, edited RNA, and chimeras resulting from ligation of gRNA with 3' pre-mRNA are indicated. The no-gRNA lane indicates negative control.

functional editing complexes from mitochondrial lysate (5). The efficacy and purity of these complexes are very high when the purifications are carried out from mitochondrial 20S fractions.

1. Lyse the mitochondria with 1% Triton X-100, clear by centrifugation, and fractionate on a 10–30% glycerol gradient as previously described.
2. Couple the immunomagnetic beads (Dynabeads M-450) coated with goat anti-mouse IgG to MAbs from tissue culture supernatants. Incubate 4×10^7 beads with 1 mL of tissue culture supernatant (or 10 μ L of purified MAbs) and 1% BSA at 4°C with bidirectional mixing for 1 h.
3. Wash the beads three times with IP buffer, and incubate the antibody-bound beads with pooled mitochondrial 20S fraction (~50 μ g of proteins) at 4°C with bidirectional mixing for 1 h in 1X IP buffer and 1% BSA. Then wash the beads four times (each wash of 5-min duration with bidirectional mixing) with IP buffer. The complexes bound to the beads can be directly assayed for editing activities, the presence of editosome proteins by adenylation, and Western analyses. **Figure 1** shows the results obtained from affinity-purified editing complexes using anti-TbMP63 MAb.
4. For identification of proteins by MS, purify in a five-fold concentration and separate the proteins on a 10% SDS-PAGE gel.
5. Stain the proteins (with coomassie blue, silver nitrate, or Sypro Ruby), and analyze individual protein bands by mass spectrometry.

3.3. Tandem Affinity Purification Strategy

The tandem affinity purification (TAP) procedure was originally developed for yeast (9) and later adapted for other organisms, including trypanosomes (10). Briefly, the protein of interest is fused to a tandem affinity tag consisting of, in order, a calmodulin-binding peptide (CBP) and two repeats of the protein A domain. These two tags are separated by a recognition site for the TEV protease. The sequential purification over IgG Sepharose and a calmodulin resin under native conditions in many cases results in the isolation of pure and active macromolecular complexes. We have modified the original procedure (see Note 1) for the rapid and efficient purification of RNA-editing complexes from *T. brucei* (6).

3.3.1. Expression Plasmids

1. Polymerase chain reaction (PCR) amplify the TAP tag sequence from plasmid pBS1539 (see Note 2) using primers 5'-GGAGGATCCATGGAAAAGAG (*Bam*HI site underlined) and 5'-GGAATTCGATATCAATGATCAGGTTGAC TTCCC (*Bcl*II site underlined), double-digest with *Bam*HI/*Bcl*II, and ligate into the *Bam*HI site of the tetracycline-inducible trypanosomal expression plasmid pLew79 ([11]; see Note 3).
2. Determine insertion in sense orientation by sequencing the generated plasmid pLew79-TAP (Fig. 2).
3. For the insertion of genes of interest, pLew79-TAP provides *Hind*III and *Bam*HI sites. Amplify the coding sequence by PCR, introducing *Hind*III and *Bam*HI sites at the 5' and 3' ends of the coding sequence, respectively. Note that the 3' primer should be designed in such a way that the original stop codon is eliminated and an in-frame fusion with the TAP tag is created (see Note 4). For example, to TAP tag the TbREL2 protein, we amplify the gene with primers 5'-ATAAAGCTT ATGTTGCGTCGCCTC (*Hind*III site underlined, start codon italicized) and 5'-GGGAGCCATATCGGATCCTTCGCTAAAGTCAGGAG (*Bam*HI site underlined; introduction of this site eliminated the stop codon).

3.3.2. Transfection into *T. brucei*

The pLew79-TAP plasmid is designed for homologous integration into the rRNA locus in the reverse orientation (11).

1. Linearize 20 µg of plasmid per transfection with *Not*I within its targeting region (in case the gene of interest contains a *Not*I site you need to mutate this site first).
2. Check a 1% aliquot of the digest by agarose gel electrophoresis, and purify the plasmid by phenol extraction and ethanol precipitation (12).
3. Dry and dissolve the plasmid under sterile conditions in 20 µL of H₂O.
4. *T. brucei* cell line 29.13 coexpresses the T7 RNA polymerase and the TET repressor (11) and is suitable for conditional expression of TAP-tagged genes (6). We

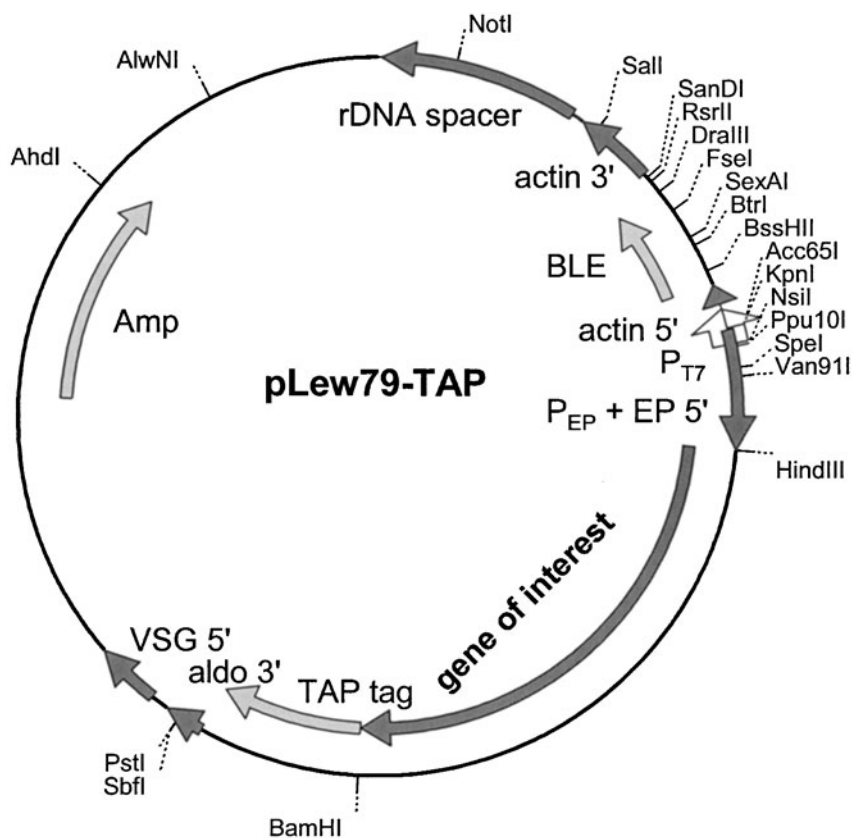


Fig. 2. Plasmid map for pLew79-TAP. P_{T7} , T7 promoter; P_{EP} , EP procyclin promoter. Flanking regions for mRNA processing: EP, EP procyclin; aldo, aldolase. Selectable markers: Amp, β -lactamase; BLE, phleomycin resistance gene.

grow this cell line in SDM-79 medium (13) in the presence of 15 μ g/mL of G418 and 25 μ g of hygromycin B to maintain stable integration of the recombinant genes. For transfection, grow 10 mL of cell line 29.13 to late mid-log phase ($\sim 10 \times 10^6$ cells/mL).

5. Harvest the cells by centrifuging for 10 min at 1300g and 4°C. Wash the cells once in 10 mL of ice-cold CytoMix and centrifuge again.
6. Resuspend the cell pellet in 1 mL of ice-cold CytoMix. Add 20 μ g of the *NotI*-linearized TAP plasmid to an ice-cold 0.2-cm electroporation cuvet, add 0.5 mL of cells in the CytoMix buffer, and mix by careful pipetting.
7. Electroporate the cells with a single pulse (setting for a BTX electroporator: 1600 V, 25 Ω , 50 μ F). Recover the cells in 5 mL of SDM-79.

8. The next day, put the cells under selection with 2.5 $\mu\text{g/mL}$ of phleomycin. It can take up to 1 wk for phleomycin to kill nontransfected cells, and it may be necessary to split the culture during selection (*see Note 5*).

3.3.3. Expression

Once a stable population of phleomycin-resistant parasites has been established, the inducibility of expression of the tagged protein needs to be checked and the induction conditions optimized for maximum recovery of editing complexes.

1. Prepare two 10-mL cultures of recombinant cells (density of $\sim 1\text{--}10^6$ cells/mL), and add 1 $\mu\text{g/mL}$ of tetracycline to one of them (this concentration results in maximal induction).
2. Forty-eight hours later, harvest the cells by centrifuging for 10 min at 1300g.
3. Discard the supernatant, lyse the cell pellet in 50 μL of 2X SDS sample buffer, and shear the genomic DNA by brief sonication.
4. Analyze total lysates by Western blotting (*see Subheading 3.4.2.*). Load the equivalent of $\sim 10^7$ cells/lane. If no specific antibody against the protein of interest is available, the protein A domains in the tagged version allow convenient detection with the PAP reagent (available from Sigma). In most cases, we have not found any detectable expression in the absence of tetracycline (**Fig. 3A**; *see Note 6*).

After expression of the tagged protein is confirmed, optimize the induction conditions.

1. Prepare seven 10-mL cultures and induce with 0, 2 pg/mL, 20 pg/mL, and so on up to 200 ng/mL of tetracycline.
2. After 24 h, harvest 7 mL of culture, collect cells by centrifugation (10 min at 1300g), and replace with fresh medium.
3. Resuspend the cell pellet in 1 mL of 1X PBS, split in two, and pellet the cells by centrifugation.
4. Resuspend one cell pellet in 25 μL of 2X SDS sample buffer and analyze by Western blotting as previously described.
5. Lyse the second cell pellet in IP buffer with 1% Triton X-100, and carry out immunoprecipitations as described in **Subheading 3.2.2.** However, do not coat the magnetic beads with primary antibody. Analyze the immunoprecipitates by Western blotting, using any available antibody directed against an untagged component of the editing complex.
6. Forty-eight hours after induction, repeat **steps 2–5**, adjusting volumes to compensate for any differences in cell density. The results will serve as a guideline for the conditions to be used in large-scale preparations (**Fig. 3B**).

3.3.4. Purification

Once cells are lysed, all steps (with the exception of the TEV digest) should be carried out at 4°C. Precool all buffers. Save aliquots of all lysates, washes, and eluates for subsequent analysis.

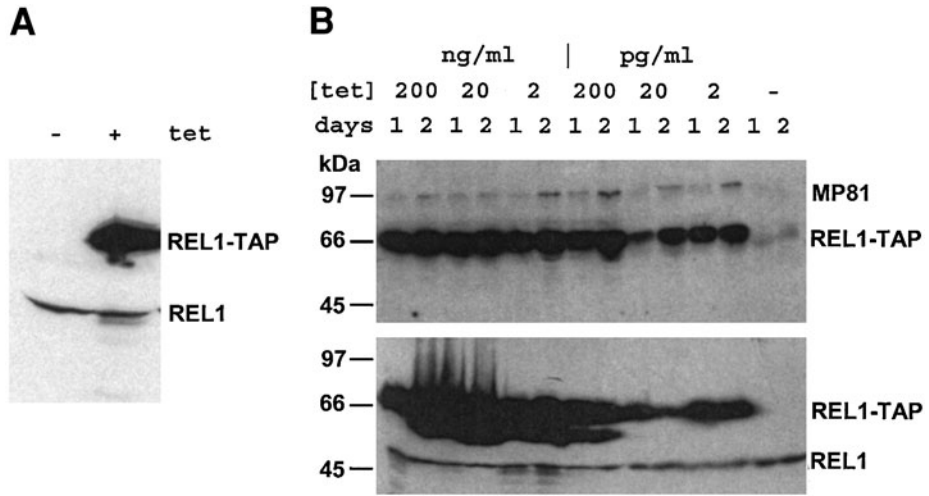


Fig. 3. Analysis of REL1-TAP-expressing cell line by Western blotting. (A) Analysis of total lysates from cells grown in absence (–) or presence (+) of tetracycline (tet). (B) Analysis of immunoprecipitates (top) and total lysates (bottom) of cells grown for 1 or 2 d in presence of various concentrations of tetracycline. The identity of the protein bands is indicated to the right of each panel.

1. Induce 2–4 L of *T. brucei* culture with tetracycline. Cell density and tetracycline concentration depend on the results of the pilot studies. Harvest cells when they reach the mid-log phase ($\sim 20 \times 10^6$ cells/mL). For example, for the REL1-TAP cell line, we use 10 ng/mL of tetracycline, induce at a density of 2.5×10^6 cells/mL, and harvest the cells 48 h later. Depending on the particular protein and experiment, you may wish to use a different concentration and induce for shorter or longer periods of time.
2. Harvest the cells in a JA 8.100 rotor (room temperature, 3500 rpm, 10 min).
3. Wash the pellet in 200 mL of PBS-G.
4. Resuspend the cells in 18 mL of ice-cold IPP150, 1% BSA, and two tablets of complete mini EDTA-free protease inhibitors.
5. Add 2 mL of 10% Triton X-100 and invert the tube several times.
6. Incubate on ice for 20 min or until the cells are completely lysed (monitor lysis under a microscope).
7. Centrifuge the lysate in an SA-600 rotor (4°C, 10,000 rpm, 15 min).
8. In the meantime, add 10 mL of IPP150 and 200 μ L IgG Sepharose beads to a 20-mL disposable Econo-pac column, wash for 5 min at 4°C, and drain by gravity flow.
9. Add cleared lysate to the column, and rotate for 2 h at 4°C.
10. Remove the top plug first, then the bottom plug, and drain by gravity flow.
11. Wash three times with 20 mL of cold IPP150 each time.

12. Wash with 10 mL of TEVCB.
13. Close the bottom of the column, and add 1 mL of TEVCB and 10 μ L of TEV protease (10 U/ μ L). Close the top and mix the beads and protease by carefully flicking the column.
14. Incubate for 2 h at 16°C, flicking the column every 30 min.
15. Remove the top and bottom plugs of the 20-mL column, and recover the eluate by gravity flow.
16. Recover the dead volume with an additional 200 μ L of TEVCB and combine with the first eluate.
17. If particular complexes are to be size selected based on sedimentation, fractionate the eluate on 10–30% glycerol gradients and pool the desired fractions (*see Subheading 3.1.5.*).
18. Measure the total volume of the eluate or fractions. Add 3 vol of CBB and 3 μ L of 1 M CaCl_2 /mL of sample.
19. Add 5 mL of CBB and 200 μ L of calmodulin resin to a 5-mL Econo-Column (or to a 20-mL Econo-Pac column, if you continue from a glycerol gradient and thus have a larger volume), wash for 5 min at 4°C, and drain by gravity flow.
20. Transfer the sample to the column containing the washed calmodulin resin. Rotate for 1 h at 4°C.
21. Drain by gravity flow.
22. Wash with 30 mL of CBB.
23. Elute four 250- μ L fractions with CEB.

Figure 4 shows a representative purification of editing complexes using tagged REL1 as monitored with four MABs against proteins TbREL1, TbMP63, TbMP81, and TbMP42.

3.4. Identification of Protein

3.4.1. Adenylation

In the editing complex the RNA ligase proteins can be identified by their autoadenylation properties in presence of ATP.

1. Incubate the fractions (5–10 μ L) with 2.5 μ Ci of [α - ^{32} P]-ATP in HHE buffer for 15 min at 28°C.
2. Stop the reaction by adding an equal amount of 2X SDS-PAGE sample buffer and separate the proteins by 10% SDS-PAGE gel.
3. Detect the radiolabeled proteins by autoradiography or phosphorimaging. The two RNA-editing ligases TbREL1 and TbREL2 migrate as 50- and 45-kDa bands, respectively (**Fig. 1**).

3.4.2. Western Analysis

Western analysis can conveniently be used to monitor the production and characterization of MABs. In addition, it can be used to monitor the purification

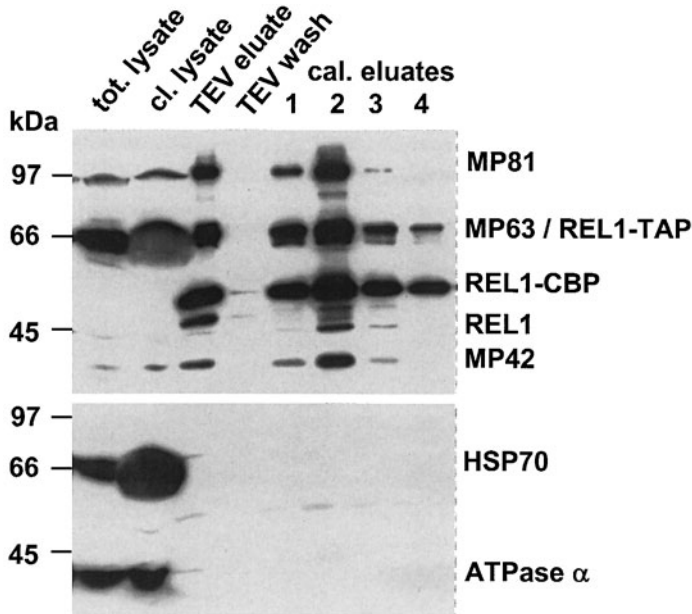


Fig. 4. Monitoring of various steps in TAP protocol by Western blotting. (**Top**) detection of editosome components; (**bottom**) detection of abundant mitochondrial proteins that are not part of editosome. tot., total; cl., cleared; cal., calmodulin.

and characterization of editing complexes. We generated a panel of MAbs that are specifically directed against four different editosome proteins: TbMP81, TbMP63, TbREL1 and TbMP42.

1. Separate the protein fractions to be analyzed by SDS-PAGE and transfer onto a PVDF membrane.
2. Block the membrane with 5% nonfat milk powder in PBST at 4°C overnight, or for 1 h at room temperature.
3. Washed the membrane three times with PBST and incubate with the MAbs (at a dilution of 1:100 for tissue culture supernatants or 1:2000 for purified MAbs in PBST) for 1 h at room temperature with gentle shaking.
4. Wash the filter three times with PBST and incubate for 1 h with antimouse IgG HRP conjugate (Bio-Rad) at a 1:2000 dilution in PBST.
5. Was the filter four times with PBST and develop with ECL or ECL Plus kit (Amersham) per the manufacturer's instructions. Their corresponding MAbs P4D8, P1H3, P3C1, and P3C12 identify proteins TbMP81, TbMP63, TbREL1, and TbMP42, respectively (**Fig. 1**).

3.4.3. Mass Spectrometry

The protein bands are digested in gel with trypsin (**14**) and analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS), by electrospray ionization using an LCQ DECA XP mass spectrometer (Thermo Finnigan).

1. Separate the proteins on a 10% SDS-PAGE gel, and visualize by either Coomassie blue, silver nitrate, or Sypro-Ruby staining. Put the gel on a clean glass plate and excise the protein bands with a new blade, but do not cut extra gel outside the visible band.
2. Slice the gel pieces further into small pieces (approx 1-mm squares) and transfer into 500- μ L tubes.
3. Dehydrate with 200 μ L of acetonitrile (CH_3CN) for 10 min (use gel-loading tips to pipet off liquid and repeat once), and dry in a Speedvac.
4. Add trypsinization solution to the gel pieces (just cover the gel pieces—never use more than 1 μ g of trypsin), and incubate the tubes in ice for 45 min. Once the gel pieces are swollen if needed cover with 50 mM $\text{NH}_4\text{-HCO}_3$ and digest the proteins at 37°C overnight.
5. Following the digestions collect the peptides. Centrifuge the gel pieces for 1 min and collect the supernatant into a fresh tube (the peptides of interest are in the solution now).
6. Further extract the peptides with one change of 20 mM $\text{NH}_4\text{-HCO}_3$, followed by two changes of 5% formic acid in 50% acetonitrile. In each step cover the gel with 10–20 μ L of the solution, vortex briefly, and incubate for 20 min at room temperature. Then spin and collect the liquid.
7. Dry the samples in a Speedvac (at this stage the samples can be stored at –20°C until analysis), and analyze by MS.
8. Separate the peptides by C18 reverse-phase chromatography (10-cm-long, 75- μ m-id column). Use a 30-min linear gradient of 10–60% acetonitrile in 0.4% acetic acid buffer to elute the peptides from the column.
9. Analyze the eluted peptides in-line by mass spectrometer, by switching between MS mode for peptide detection and MS/MS mode for peptide fragmentation and sequence analysis.
10. Analyze the acquired MS/MS data (spectra) with *T. brucei* nucleotide/protein databases using the SEQUEST database search tool for peptide identification. In the case of nucleotide sequences, the search is carried out against all six open reading frames (ORFs). The SEQUEST program compares the theoretical MS/MS spectra from the database with the acquired MS/MS spectra from the samples. Identification of multiple peptides corresponding to an ORF identifies the protein (**4,5**). For example, we identified TbMP99 by identification of 33 peptides corresponding to its gene. Sequence information of a peptide from this protein generated by mass spectrometer is shown in **Fig. 5**.

Several proteins were identified in this way (**15**). Using MAb affinity and TAP-tag affinity pull down of the functional complexes from *T. brucei*, followed

by MS analyses, we showed that at least 16 proteins (TbMP100, TbMP99, TbMP90, TbMP81, TbMP67, TbMP63, TbMP61, mHel61p, TbMP57, TbREL1, TbREL2, TbMP46, TbMP44, TbMP42, TbMP24, and TbMP18) are present in the core of the catalytic complex (6).

3.5. Characterization

Functional roles for the identified proteins can sometimes be predicted based on homology to previously characterized proteins or by computational searches for conserved motifs, using a plethora of tools available on the World Wide Web. For instance, the Conserved Domain Database at NCBI (www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) identifies a ligase motif in TbREL1 and TbREL2. However, these predictions need to be confirmed in vitro as well as in vivo. For functional in vitro studies, we have used assays of recombinant protein function. Functions in vivo are routinely investigated through regulated gene inactivation.

3.5.1. Recombinant Proteins

Small-scale production of recombinant protein using coupled in vitro transcription/translation can yield sufficient amount of active protein for initial functional studies. For instance, we have used this system successfully to confirm the functions of RNA-editing ligases and 3'-terminal uridylyl transferase (16–18) and for protein-protein interaction studies (15; unpublished data). Next we describe the characterization of ligase TbREL2.

3.5.1.1. CLONING

Suitable vectors for protein expression in vitro are, e.g., pSG1 (16) or the pRSET vectors from Invitrogen. These vectors provide a T7 promoter for transcription as well as, in the case of the pRSET vectors, epitope tags (Xpress and 6xHis) that can be used for detection and immunoprecipitation of recombinant proteins. The latter technique was especially useful in obtaining recombinant protein in sufficient purity for functional studies.

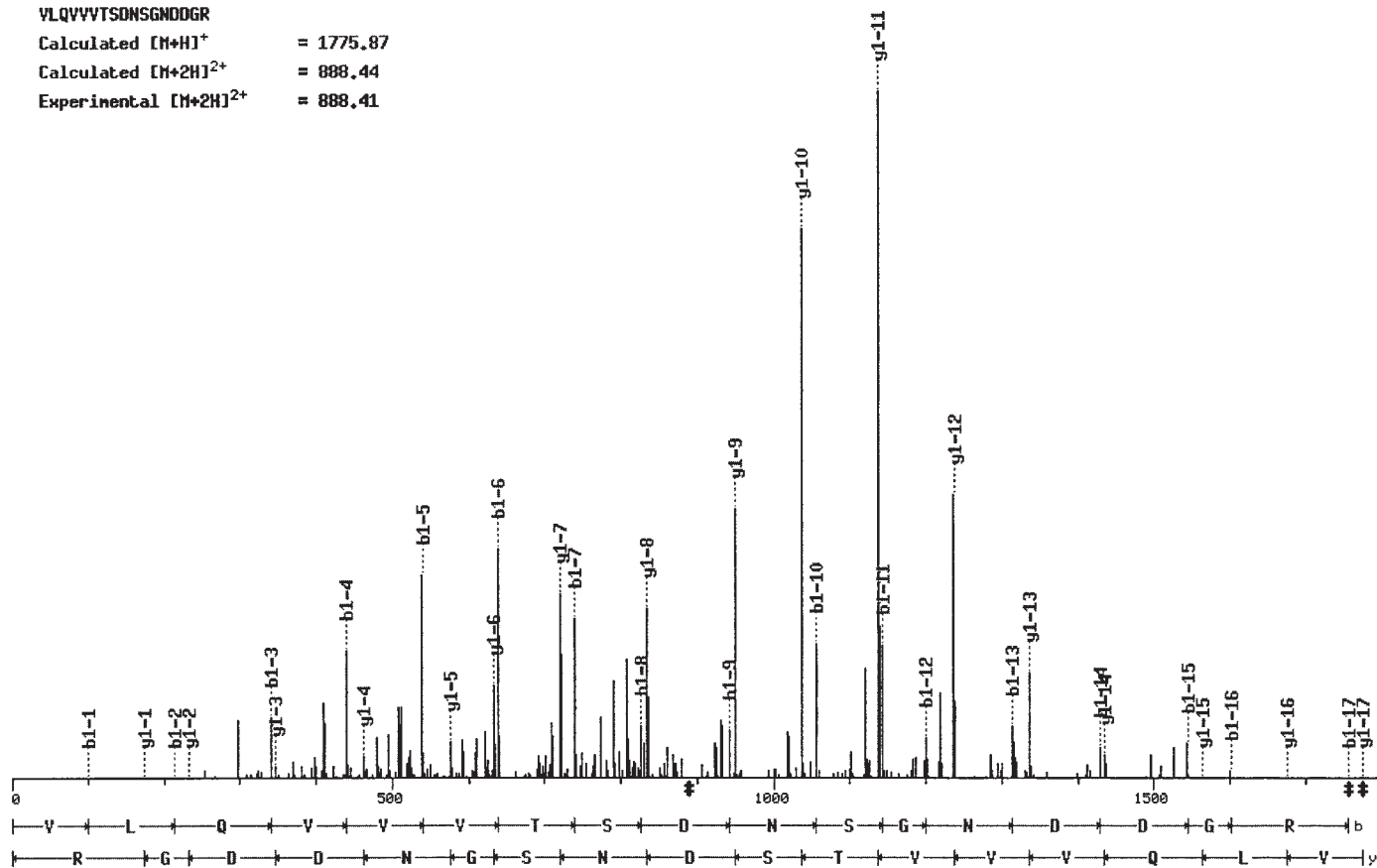
1. Amplify the *REL2* gene from *T. brucei* genomic DNA by PCR (primers 5'-CCG AAT TCA TGC ACC ATC ACC ATC ATG TCG GTG GGG ACG GCA GCA -3' and 5'-CCG GAT CCT CAT TCG CTA AAG TCA G- 3'), digest with *EcoRI* and *BamHI*, and clone into pSG1 vector digested with the same enzymes. The 5' primer introduces a 6xHis tag at the N-terminus that replaces the mitochondrial import signal (amino acids 2–17).
2. Synthesize recombinant proteins in vitro using the TNT cell-free coupled transcription and translation system (Promega) according to the manufacturer's instructions.
3. Immunoprecipitate the recombinant REL2 protein with 6xHis MAb (Clontech). Coat 4×10^7 immunomagnetic beads (Dynabeads M-450; Dynal), with goat anti-

VLQVVVTSONSGNDDGR

Calculated $[M+H]^+$ = 1775.87

Calculated $[M+2H]^{2+}$ = 888.44

Experimental $[M+2H]^{2+}$ = 888.41



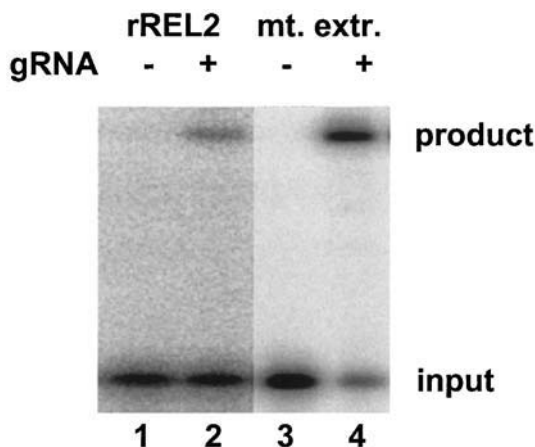


Fig. 6. Ligation assay using recombinant TbREL2 (rREL2, lanes 1 and 2) or mitochondrial extract as positive control (lanes 3 and 4). The requirement for a bridging gRNA demonstrates the specificity of the reaction.

mouse IgG plus 1% BSA, incubate with the MAb diluted 1:100 in IP buffer for 1 h at 4°C with bidirectional mixing, and then wash three times with IP buffer.

4. Incubate the antibody-bound beads with 50 μ L of in vitro-coupled transcription-translation products in 400 μ L of IP buffer containing 1% BSA with bidirectional mixing for 1 h at 4°C.
5. Wash the beads three times with IP buffer and once with HHE buffer. Resuspend the beads in HHE and assay for ligase activity as described in Chapter 13 (**Fig. 6**).

3.5.2. Knockdown Studies

Currently the most straightforward way to inactivate genes in trypanosomes is by way of regulated RNA interference. This strategy entails expression in the cell of double-stranded RNA homologous to the target gene. Although not as “clean” as obtaining actual knockout, or null, mutants, in most cases gene expression can be reduced by 90% and more (“knockdown”) and the regulatable nature of the inactivation allows mutants affecting essential genes to be obtained. For details, we refer the reader to Chapter 4. Our laboratory has used this strategy successfully for the characterization of TbREL2 and TbMP81 (**19**).

Fig. 5. (*see opposite page*) MS analysis of peptide. A CID spectrum of a tryptic peptide from editosome protein TbMP99 is shown. The spectrum matches the predicted spectrum from peptide VLQVVVTSDNSGNDDGR, both from the N- to C-terminus (b-ions) and C- to N-terminus (y-ions).

4. Notes

1. A detailed description of the original procedure as described for yeast is available online at www-db.embl-heidelberg.de/jss/servlet/de.embl.bk.wwwTools.GroupLeftEMBL/ExternalInfo/seraphin/TAP.html.
2. Use of the TAP tag plasmids requires signing a material transfer agreement. Additional information is available at www.uni-frankfurt.de/fb15/mikro/euroscarf/cellzome.html.
3. Plasmid pLew79 facilitates inducible expression through a modified EP promoter, yielding high and moderate expression rates in procyclic and bloodstream stages, respectively (**II**).
4. Some tagged proteins (and associated complexes) may not bind efficiently to the calmodulin resin, perhaps owing to burial of the CBP tag within the protein structure. We do not know in what percentage of cases this problem will occur but inserting an additional spacer region between the protein of interest and the CBP tag may solve it.
5. We have generally found that for expression of TAP-tagged genes and purification of associated complexes, it is not necessary to obtain clonal populations of transfectants. However, if fine-tuning of expression for the particular gene of interest is crucial, such clonal populations can be obtained by limiting dilution.
6. We have found that clones, in general, show tighter regulation than mixed populations. When expression of the tagged version of the protein is detrimental to cell growth or expression of a mutant version of the gene of interest is desired, it is advisable to obtain clonal populations and to pick clones with tight regulation for subsequent studies (*see Note 5*).

Acknowledgments

This work was supported by National Institutes of Health grant AI/GM14102.

References

- 1 Kable, M. L., Heidmann, S., and Stuart, K. (1997) RNA editing: getting U into RNA. *Trends Biochem. Sci.* **22**, 162–166.
- 2 Stuart, K. and Panigrahi, A. K. (2002) RNA editing: complexity and complications. *Mol. Microbiol.* **45**, 591–596.
- 3 Simpson, L., Sbicego, S., and Aphasizhev, R. (2003) Uridine insertion/deletion RNA editing in trypanosome mitochondria: a complex business. *RNA* **9**, 265–276.
- 4 Panigrahi, A. K., Gygi, S., Ernst, N., (2001) Association of two novel proteins, *TbMP52* and *TbMP48*, with the *Trypanosoma brucei* RNA editing complex. *Mol. Cell. Biol.* **21**, 380–389.
- 5 Panigrahi, A. K., Schnauffer, A., Carmean, N., et al. (2001) Four related proteins of the *T. brucei* RNA editing complex. *Mol. Cell. Biol.* **21**, 6833–6840.
- 6 Panigrahi, A. K., Schnauffer, A., Ernst, N., Wang, B., Carmean, N., Salavati, R., and Stuart, K. (2003) Identification of novel components of *Trypanosoma brucei* editosome. *RNA* **9**, 484–492.

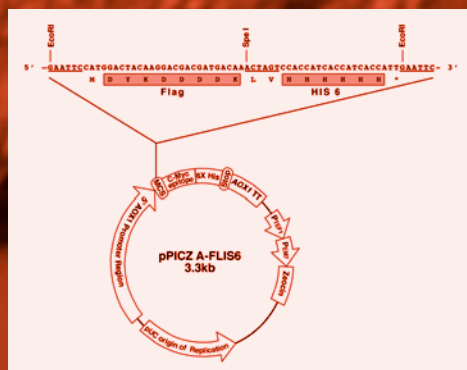
- 7 Stuart, K., Gobright, E., Jenni, L., Milhausen, M., Thomashow, L., and Agabian, N. (1984) The IsTaR 1 serodeme of *Trypanosoma brucei*: development of a new serodeme. *J. Parasitol.* **70**, 747–754.
- 8 Harris, M. E., Moore, D. R., and Hajduk, S. L. (1990) Addition of uridines to edited RNAs in trypanosome mitochondria occurs independently of transcription. *J. Biol. Chem.* **265**, 11,368–11,376.
- 9 Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* **17**, 1030–1032.
- 10 Estévez, A. M., Kempf, T., and Clayton, C. (2001) The exosome of *Trypanosoma brucei*. *EMBO J.* **20**, 3831–3839.
- 11 Wirtz, E., Simone, L., Claudia, O., and Cross, G. A. M. (1999) A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **99**, 89–101.
- 12 Sambrook, J., Fritsch, T., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- 13 Brun, R. and Schonenberger, M. (1979) Cultivation and *in vitro* cloning of procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. *Acta Trop.* **36**, 289–292.
- 14 Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **68**, 850–858.
- 15 Panigrahi, A. K., Allen, T. E., Haynes, P. A., Gygi, S. P., and Stuart, K. (2003) Mass spectrometric analysis of the editosome and other multiprotein complexes in *Trypanosoma brucei*. *J. Am. Soc. Mass Spectrom.* **14**, 728–735.
- 16 Schnauffer, A., Panigrahi, A. K., Panicucci, B., Igo, R. P. Jr., Salavati, R., and Stuart, K. D. (2001) An RNA ligase essential for RNA editing and survival of the bloodstream form of *Trypanosoma brucei*. *Science* **291**, 2159–2162.
- 17 Palazzo, S. S., Panigrahi, A. K., Igo, R. P. Jr., Salavati, R., and Stuart, K. (2003) Kinetoplastid RNA editing ligases: complex association, characterization, and substrate requirements. *Mol. Biochem. Parasitol.* **127**, 161–167.
- 18 Ernst, N., Igo, R. P. Jr., Panicucci, B., Panigrahi, A. K., and Stuart, K. (2003) TbMP57 is a 3′ terminal uridylyl transferase (TUTase) of the *Trypanosoma brucei* editosome. *Mol. Cell* **11**, 1525–1536.
- 19 Drozd, M., Palazzo, S. S., Salavati, R., O’Rear, J., Clayton, C., and Stuart, K. (2002) TbMP81 is required for RNA editing in *Trypanosoma brucei*. *EMBO J.* **21**, 1791–1799.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



Chimeric Templates and Assays Used to Study *Physarum* Cotranscriptional Insertional Editing In Vitro

Elaine M. Byrne

Summary

RNAs made in the mitochondrion of *Physarum polycephalum* are edited relative to their template by the precise addition of nonencoded nucleotides, while they are being synthesized. This insertional editing has been reproduced in vitro during run-on extension of RNAs initiated in vivo, within partially purified mitochondrial transcription elongation complexes (mtTECs), but it does not occur when the mitochondrial polymerase initiates transcription on exogenous cloned DNA. This chapter describes in vitro transcription systems in which mtTEC RNAs are elongated on repositioned parts of the genome or exogenous DNA, in order to investigate how the nontemplated insertions are directed. Restriction enzyme digestion and DNA ligation are used to generate the chimeric templates, and the RNA products are analyzed directly by nuclease dissection (S1 protection followed by RNase T1 digestion) or by reverse transcriptase-polymerase chain reaction (RT-PCR) followed by restriction enzyme analysis or cloning and sequencing.

Key Words: RNA editing; mitochondrion; transcription elongation complexes; run-on transcription; sequence rearrangement; subgenomic circles.

1. Introduction

Many RNAs transcribed from the mitochondrial genome of *Physarum polycephalum* are edited by the precise insertion of single nonencoded nucleotides (mostly cytosine) and dinucleotides (**1**). Edited RNA is generated during run-on transcription in isolated mitochondria (**2**), and from preparations of partially purified mitochondrial transcription elongation complexes (mtTECs), which contain mitochondrial genomes with associated polymerase, nascent transcripts, and other uncharacterized components (**3**). These in vitro systems have demonstrated that the editing is taking place at the growing end of the nascent

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

transcript (4). By contrast, unedited RNA is produced when free RNA polymerase from mitochondrial extracts is used to transcribe “artificial” (tailed or bubble) templates made from cloned DNA bearing the sequence of editable genes (Adam Majewski, E. Byrne, and Jonatha Gott, unpublished data); thus, transcription templates cannot yet be manipulated in traditional ways in order to characterize the *Physarum* editing site determinants.

This chapter describes “intermediate” systems wherein the DNA in mtTECs is rearranged or attached to exogenous DNA, by restriction enzyme digestion and ligation, to generate “chimeric templates” (Figs. 1 and 2). Nascent RNAs begun *in vivo* are then extended across the ligated junctions by allowing run-on transcription, to give RNAs that are chimeric in sequence and/or in terms of the source of the template DNA (Figs. 3 and 4). Specific chimeric transcripts made in the presence of a labeled nucleotide triphosphate (NTP) substrate can be analyzed directly by enzymatic digestion: hybridization to a suitably constructed complementary single-stranded DNA (ssDNA) probe is followed by digestion of unprotected RNA by S1 Nuclease (*see* the schematic in the lower part of Fig. 3), and then protected sections can be fragmented by RNase T1 to detect specific insertions. Alternatively, specific (unlabeled) chimeric transcript sequences can be copied and amplified by reverse transcription followed by polymerase chain reaction (RT-PCR), allowing indirect analysis of individual RNA molecules via cloning and sequencing, or of the sequence pool for a given chimeric transcript by restriction enzyme digestion (*see* the schematic in Fig. 4 for an example).

Thus, the endogenous template can be manipulated somewhat, and exogenous DNA of defined sequence can be used as template by the transcription/editing machinery. These techniques have been used to initiate an examination of the roles of *cis*-acting nucleic acid signals and associated factors in directing the editing process (5), and further studies are proceeding using ligation points closer to editing sites (A. Majewski, A. Rhee, and J. M. Gott, unpublished data). Aspects of these methods might also be applicable to the investigation of other phenomena that occur during transcription, or during other templated polymerization processes such as replication or repair, if, like the *Physarum* transcription-editing process, the coupled systems have not yet been reconstituted from individual components *in vitro*.

2. Materials

1. mtTEC preparation, isolated essentially as described in ref. 3.
2. Commercially available enzymes: restriction enzymes (preferably high concentration), T4 DNA ligase (10 U/ μ L) (Roche, Indianapolis, IN), RNase-free DNase I (10 U/ μ L) (Roche), RNase inhibitor, T7 RNA polymerase or kit with which to make control RNA, S1 Nuclease (Roche), RNase T1, RNase A, RNase H-minus M-MLV RT, polynucleotide kinase, *Taq* DNA polymerase.

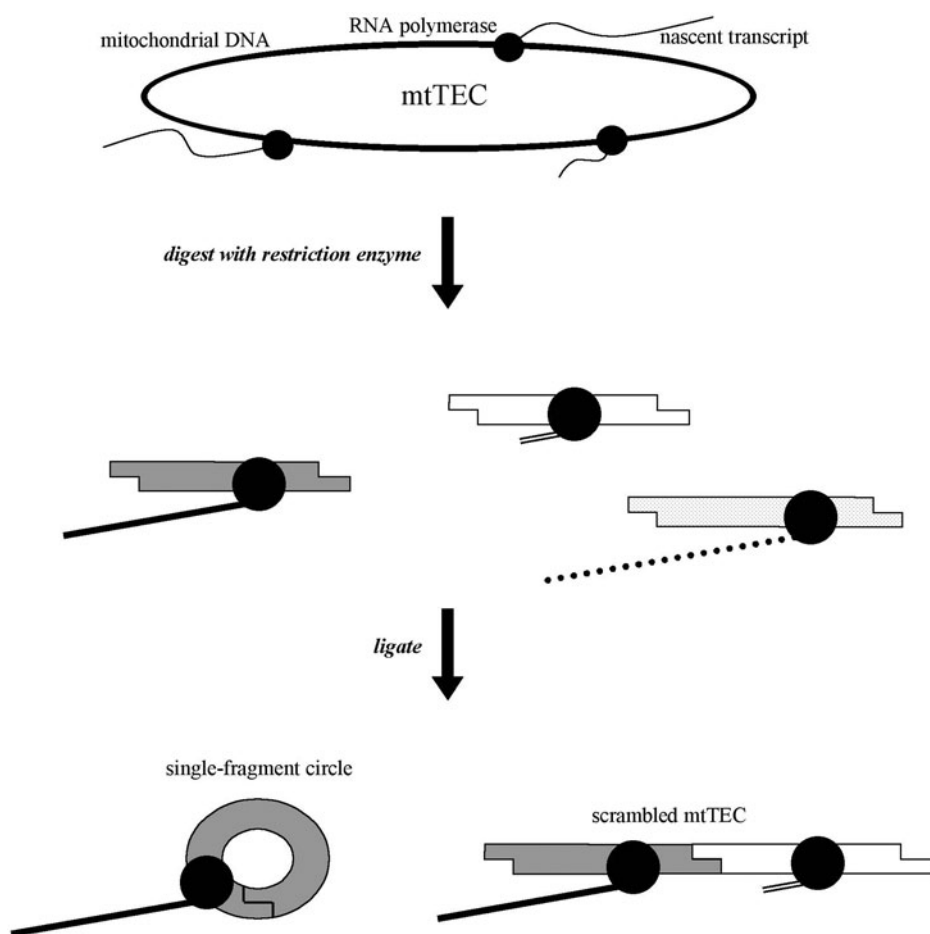


Fig. 1. Generation of chimeric templates in absence of exogenous DNA. Different restriction enzyme fragments and their transcripts are denoted by blocks with different shading and lines of different styles, respectively. Only a selection of the many possible ligation products for the depicted fragments is shown.

3. Oligonucleotide primers.
4. Plasmid vector(s), bacterial strains, and materials for cloning and preparing ssDNA; facilities for sequencing recombinant plasmids.
5. 100 mM buffered solutions of each ribonucleotide triphosphate (NTP), deoxynucleotide triphosphates (dNTPs), [α - ^{32} P]NTP (3000 Ci/mmol, 10 mCi/mL), [γ - ^{32} P]ATP (6000 Ci/mmol, 150 mCi/mL).
6. Glycogen solution (20 mg/mL).
7. TE: 10 mM Tris-HCl, pH 7.5, 1 mM EDTA.

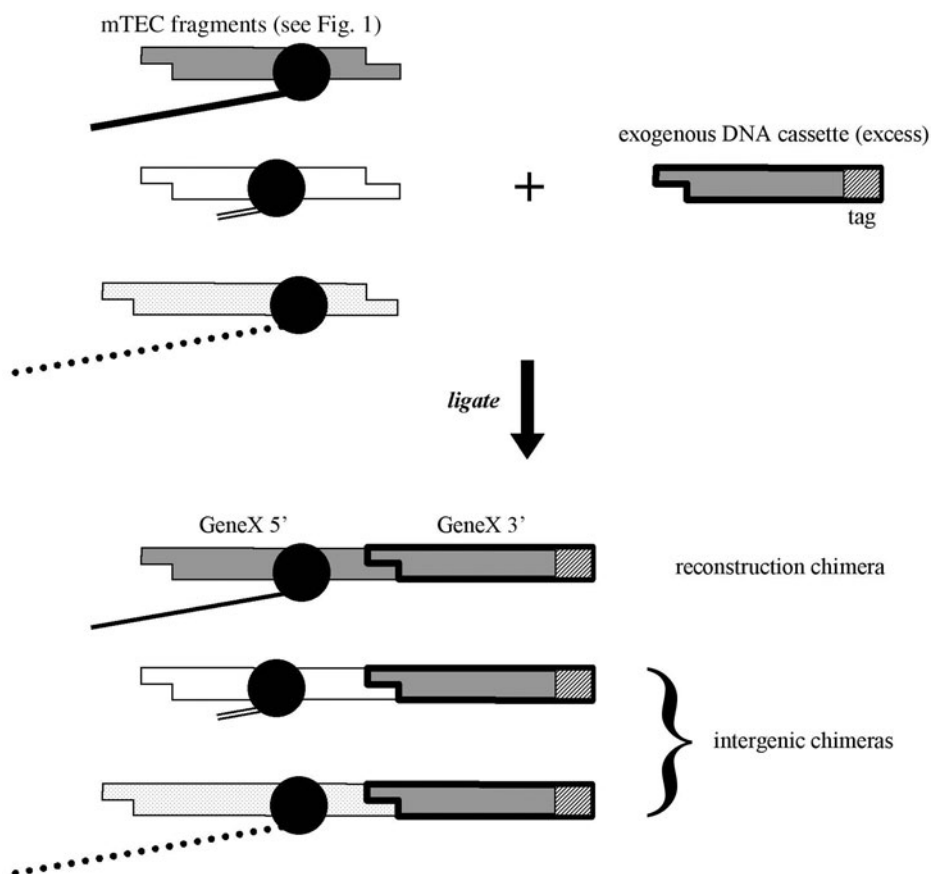


Fig. 2. Additional chimeric templates generated in presence of exogenous DNA. The exogenous cassette (heavy border) shown here has a piece of cloned *Physarum* sequence (GeneX 3') that is a continuation of one of the mtTEC fragments (GeneX 5', gray area), and a 3' tag (hatched area). Again, only a selection of the many possible ligation products is shown. These represent a chimera that "reconstructs" the genomic sequence across the ligation junction between the mtTEC fragment and exogenous DNA, and chimeras that join the exogenous DNA to the mtTEC fragments from other genes, making "intergenic" linkages ("scrambled sequence").

8. Phenol: Tris-saturated. Store at 4°C, stocks at -20°C.
9. Chloroform:isoamyl alcohol ([CIA], 24:1, v/v).
10. BioSpin P30 columns (Bio-Rad, Hercules, CA).
11. 10X Buffer H: 500 mM Tris-HCl, 100 mM MgCl₂, 1 M NaCl, 10 mM dithioerythritol; pH of 1X is 7.5 at 37°C.
12. 10X Transcription stop solution: 100 mM Tris-HCl, pH 7.5, 50 mM EDTA, 1% sodium dodecyl sulfate (SDS). Store at room temperature.

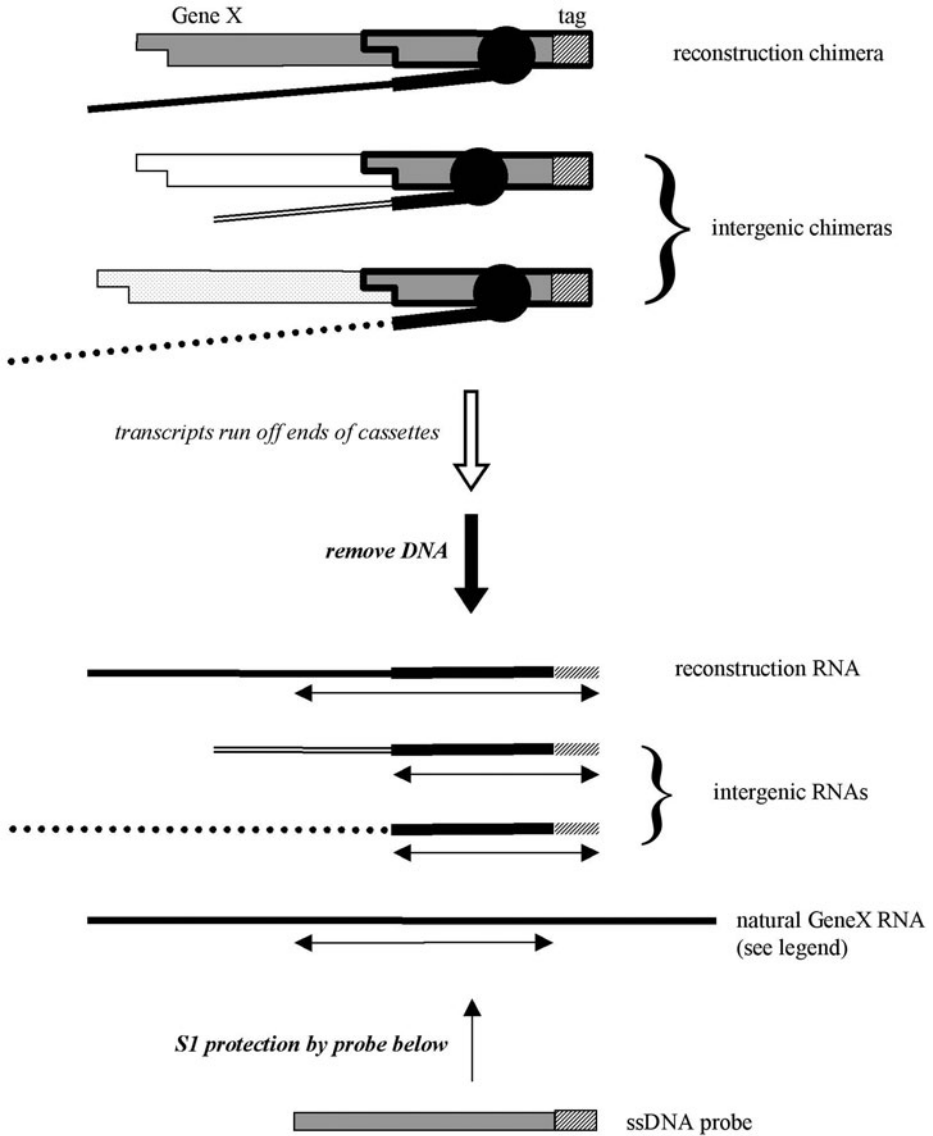


Fig. 3. Run-on transcription of chimeric templates depicted in **Fig. 2**, and S1 nuclease protection analysis. Thick portions of RNA lines represent parts of transcripts templated by the ligated exogenous cassettes. Some labeled natural GeneX RNA may also be present, even if digestion is complete, by run-on transcription from the mtTEC fragment that overlaps the exogenous cassette sequence. S1 probe is antisense to the cassette plus a portion of GeneX upstream of the chimeric junction. Protected regions are indicated by lines with double arrows below RNA.

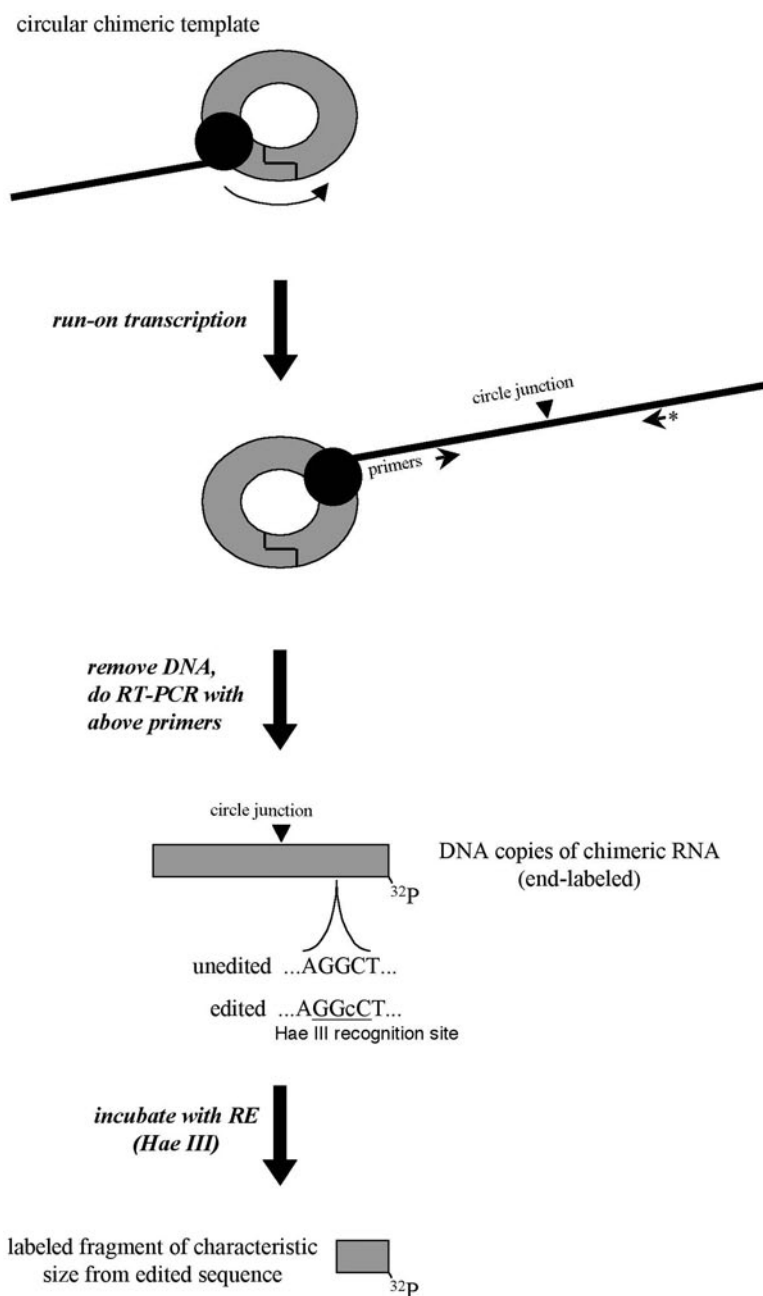


Fig. 4. Run-on transcription from a subgenomic circle (**Fig 1**), amplification of resulting RNA sequence across chimeric junction by RT-PCR, and analysis by restriction enzyme (RE) digestion. In this case the primer downstream of the junction is 5' labeled (asterisk) with ^{32}P . An example is shown of an editing site where insertion of the correct nucleotide, a cytosine (lowercase c), generates an *Hae*III recognition site in the RT-PCR product.

13. S1 hybridization buffer: 40 mM PIPES, pH 6.4, 1 mM EDTA, 0.4 M NaCl, 80% formamide. Store at -20°C .
14. S1 mapping buffer: 0.75 M NaCl, 50 mM sodium acetate, pH 4.5, 4.5 mM ZnSO_4 . Stored at 4°C .
15. S1 stop solution: 2 M sodium acetate, 50 mM EDTA, 50 $\mu\text{g/mL}$ of tRNA; store at -20°C .
16. 10X TAE buffer: 400 mM Tris-acetate, pH 8.0, 10 mM EDTA.
17. 10X TBE buffer: 500 mM Tris-borate, pH 8.3, 10 mM EDTA.
18. Denaturing gel-loading/dye solution: 7 M urea, 1X TBE, 0.25% xylene cyanol, 0.25% bromophenol blue.
19. Elution buffer: 10 mM Tris-HCl, pH 7.5, 0.25 M sodium acetate, 1 mM EDTA, 0.25% SDS.

3. Methods

The following methods outline preparation of DNA cassettes for experiments involving ligation of exogenous DNA to fragmented mtTECs, production of chimeric templates followed by run-on transcription (*see Note 1*), isolation of RNA from these reactions, and analysis of editing in the chimeric RNAs. A number of interdependent parameters must be considered together when planning the details of a chimeric template experiment (e.g., choice of restriction enzymes, design of exogenous DNA cassettes, methods of transcript analysis); these are detailed in **Notes 2–4** and **Table 1**. Recombinant plasmids for exogenous DNA cassette production, control transcripts, and ssDNAs should be prepared in advance using standard molecular biology methods; also *see Notes 5* and **6**.

3.1. Preparation of Exogenous DNA Cassettes (*see Note 7*)

Short cassettes can be made by annealing two suitably designed oligonucleotides, where one has a short unpaired terminus to act as a sticky end for ligation. The nonligating end should be blunt (and preferably without a 5' phosphate). Longer cassettes are most easily obtained by PCR amplification from suitably designed recombinant plasmids: if the cassette is to contain a section of mitochondrial sequence, it should include a downstream 3' terminal "tag" of a sequence not found in the mitochondrial genome (*see Fig. 2*), to aid in the isolation/analysis of chimeric RNAs. The sticky end for ligation is generated by restriction enzyme digestion (*see Note 8*); it is best to then purify the cassette (by, e.g., agarose gel electrophoresis) to avoid the other restriction fragment(s) competing in the subsequent ligation to mtTECs.

3.2. Generation of Chimeric Templates and Allowance of Run-On Transcription

See Notes 9 and **10** before planning details of the following series of reactions.

Table 1
Options for Analysis of Editing in Chimeric Transcripts (see Note 4)^a

Direct (S1-T1 analysis of labeled RNA)	RT-PCR →	
	Restriction enzyme analysis	Cloning and sequencing
+ Visualizes RNA made in vitro, little complication from sequences made in vivo	– Uncertain as to where in vitro transcription commences upstream of the ligation junction	
+ Can quantitate efficiency of editing at individual sites	+ Can quantitate efficiency of editing at individual sites . . . – . . . within limits of above constraint	– Quantitation subject to sampling bias at low numbers; cloning biases possible
– Identity or position of extra nucleotides within T1 fragments not checked by single- dimension electrophoresis	– Distinguishes between only two states at a time (unedited vs edited here)	+ Complete sequence information on subset of individual molecules; can detect misediting, determine site status combinations
– Resolution of editing status of given site determined by distribution of T1 fragment sizes	– Number of assayable sites limited by sequence context	+ All editing sites examined
– May be difficult to get enough signal for T1 analysis of specific RNA	+ Can analyze very small amounts of specific RNA	

^a+, Advantages; –, disadvantages.

3.2.1. Restriction Enzyme Digestion of mtTECs

1. Thaw an aliquot of the mtTEC preparation on ice and remove a volume that contains 5 µg of protein (see **Note 11**).
2. Calculate the amount of water needed for a final digestion volume of 35 µL: from 35 µL subtract the volume of the mtTEC sample (see **step 1**), buffer stock (3.5 µL, see **step 3**) and restriction enzyme(s) (see **step 4**).
3. Combine the water with 3.5 µL of 10X buffer H, and then mix into the mtTEC sample.
4. Mix in 15 U of each restriction enzyme, and incubate at 30°C for 15–20 min.

3.2.2. Ligation Reactions

1. Bring the reaction to 40 μL (*see Note 9d*) by adding in the following order ATP to 500 μM , 2 pmol of exogenous DNA where appropriate (*see Note 12*), and 1.5 U of T4 DNA ligase (the stock enzyme can be diluted in 10 mM Tris-HCl, pH 7.5, as necessary).
2. Incubate at 16°C for 30 min.

3.2.3. Run-On Transcription

Conditions for the run-on transcription reactions are similar to those described in **ref. 3** except for higher levels of Tris (approx 40 mM, pH 7.8) and NaCl (approx 70 mM), which are contributed by the restriction enzyme digestion buffer in addition to the mtTEC buffer (*see Notes 9b* and **13**). The following protocol describes an unlabeled transcription with all nucleotides at the same concentration; to make labeled RNA for direct analysis, and/or to influence editing efficiencies via nucleotide concentrations, the conditions can be altered according to the considerations in **Notes 14** and **15**.

1. Bring the volume to 50 μL with the addition of bovine serum albumin to 100 $\mu\text{g}/\text{mL}$ final; DTT to 2.5 mM final; CTP, GTP, and UTP to 500 μM final; and extra ATP to 200 μM (in addition to ATP from ligation reaction; *see Note 15c*). If the final concentration of free magnesium would be significantly less than 8 mM (taking into account free magnesium from the mtTEC preparation, buffer added for the restriction enzyme digestion reaction, and chelation by the nucleotides), supplement by adding MgCl_2 to this concentration. It is important that the nucleotides be added together and last to the reaction (usually as part of a mix with all the other ingredients), so that each individual nucleotide is incorporated under the same conditions and labeling is uniform.
2. Incubate at 30°C for 30 min.
3. Stop the transcription reactions by adding 1/9 vol of 10X transcription stop solution.

3.3. Isolation of RNA

1. Extract the nucleic acids once with phenol, then once with CIA.
2. Precipitate by adding 1 μL of glycogen (carrier), 0.1 vol of 4 M ammonium acetate, and 2.5 vol of ethanol, cooling on dry ice until the mixture is viscous (or at -20°C for at least 2 h). Spin down the nucleic acids for 20 min in an Eppendorf centrifuge at 12,000 rpm, 4°C. Pipet off the supernatant, and then wash the pellet by adding 500 μL of 70% ethanol and vortexing gently. Spin again for 5 min as before, pipet off the liquid, and air-dry for approx 15 min at room temperature.
3. Resuspend the pellets in 20 μL of TE on ice for a minimum of 15 min. A small amount (e.g., 5%) can be taken as a pre-DNase sample for possible use in control PCR reactions.
4. Digest the DNA with 20 U of DNase I in a final concentration of 10 mM Tris (pH 7.5), 1 mM EDTA, 6.25 mM DTT, 12.5 mM magnesium acetate, and

0.25 U/ μ L of RNase inhibitor, in a final volume of 40 μ L. Incubate for 30 min on ice, then 30 min at 30°C.

5. Pass the reaction through a BioSpin P30 column according to the manufacturer's instructions, to remove DNA digestion products, which might interfere with subsequent S1 nuclease digestion or RT-PCR procedures.
6. To the effluent add SDS (to 0.6% final), acetic acid (to 12 mM), and EDTA (to 30 mM), in a final volume of 100 μ L, to stop any residual DNase activity.
7. Extract the RNA once with a 1:1 mix of phenol:CIA, then once with CIA. If the RNA is labeled, go directly to **Subheading 3.4.1.1.1**. Otherwise, ethanol precipitate as in **step 2**.
8. For RNA that is to be subjected to reverse transcription (*see Subheading 3.4.2.1.*), resuspend as in **step 3**, DNase treat for a second time by repeating **steps 4–7**, and resuspend as in **step 3**.

3.4. Analysis of Chimeric RNAs

3.4.1. Direct Analysis of Labeled RNA

Sections of metabolically labeled chimeric RNAs of interest are isolated by hybridizing them to complementary ssDNA probes (*see Note 16* and schematic in **Fig. 3**), then incubating with S1 nuclease, which digests the unprotected RNA. Labeled control RNAs (edited and unedited) should be made in vitro and S1 protected in separate reactions (*see Note 17*). Polyacrylamide gel electrophoresis (PAGE) is used to resolve the products, allowing one to determine whether a chimeric fragment has been produced, to get an initial idea of whether it is edited (from its mobility relative to edited and unedited control RNA fragments), and to purify it. RNase T1, which cleaves after G residues, is then used to digest the protected sequence into small defined fragments, some of which will encompass one, or occasionally more, insertion sites (*see Note 18*). High-percentage PAGE can be used to separate the T1 fragments at single-nucleotide resolution, allowing the efficiency of insertion at many sites to be measured (also *see Note 4a*).

3.4.1.1. S1 NUCLEASE PROTECTION

S1 nuclease protection experiments are done under conditions that allow individual *Physarum* mtRNA sequences to hybridize efficiently and stably enough for precise and efficient S1 nuclease protection, despite high A/T content, single-nucleotide bulges owing to partial editing, and the complexity of the mtRNA sequence mixture (**6**); *see Notes 19* (important) and **20**. Ethanol precipitations and 70% ethanol washes are carried out as in **Subheading 3.3., step 2**, with modifications as noted; pellets are resuspended for at least 15 min, on ice, unless otherwise indicated.

3.4.1.1.1. Run-On RNA

1. Mix RNA from **step 7** of **Subheading 3.3.** with 4 μg of the appropriate antisense ssDNA, ethanol precipitate, wash twice with 70% ethanol, and dry the pellet for approx 5 min.
2. Resuspend the pellet in 20 μL of hybridization buffer for 15 min at room temperature. Heat at 85°C for 10 min to denature nucleic acids, spin briefly, and incubate at 37°C for a minimum of 16 h to hybridize.
3. At room temperature, spin briefly, add S1 nuclease (usually 200–800 U, as determined in advance; *see Note 19*) in 200 μL of ice-cold S1 mapping buffer, and incubate at 26°C for 1.5 h.
4. Stop the S1 reaction by adding 54 μL of S1 stop solution, extract once with a 1:1 mix of phenol:CIA, then once with CIA. Ethanol precipitate without adding further salt; the pellet does not need to be dried. Resuspend the pellet in 100 μL of TE.
5. Add 4 μg of the ssDNA and ethanol precipitate, without additional salt; glycogen is not needed unless the pellet in **step 4** is very small. Wash twice with 70% ethanol, leaving 5 min on ice each time, and dry the pellet for approx 5 min.
6. Resuspend and denature the nucleic acids as in **step 2**, and hybridize at 37°C for 2 h.
7. Perform S1 nuclease digestion as in **step 3**.
8. Stop the reaction, extract the nucleic acids, and ethanol precipitate as in **step 4**, but wash three times with 70% ethanol, and then dry the pellet for approx 5 min. Count the radioactive signal and resuspend in a small volume of TE (e.g., 20 μL , in case the RNA is ready for final electrophoresis). At this point, samples of the RNA (e.g., 5–10%) can be examined by denaturing PAGE (*see Subheading 3.4.1.1.3.*). If the expected S1 fragment(s) can be resolved above background, the third round of S1 protection (**steps 9–12**) can be omitted.
9. Add 4 μg of the ssDNA and ethanol precipitate, without additional salt; glycogen is not needed unless the pellet in **step 8** is very small. Wash twice with 70% ethanol, leaving 5 min on ice each time, and dry the pellet for approx 5 min.
10. Resuspend and denature the nucleic acids as in **step 2**, and hybridize at 37°C for 2 h.
11. Perform S1 nuclease digestion as in **step 3**.
12. Stop the reaction, extract the nucleic acids, and ethanol precipitate as in **step 4**, but wash three times with 70% ethanol, and then dry the pellet for approx 5 min. Count the radioactive signal and resuspend in a small volume of TE (e.g., 20 μL).

3.4.1.1.2. Control RNA

1. Mix labeled control RNA (*see Note 21*) with 1 μg of the antisense ssDNA. Then ethanol precipitate and dry the pellet for approx 5 min.
2. Resuspend the pellet in 20 μL of hybridization buffer for 15 min at room temperature. Heat at 85°C for 10 min, spin briefly, and incubate at 37°C for 2 h.
3. At room temperature, spin briefly, add S1 nuclease (same amount as in **Subheading 3.4.1.1.1., step 3**) in 200 μL of ice-cold S1 mapping buffer, and incubate at 26°C for 1.5 h.

4. Stop the S1 reaction by adding 54 μL of S1 stop solution, extract once with a 1:1 mix of phenol:CIA, then once with CIA. Ethanol precipitate without adding further salt, wash three times with 70% ethanol, dry the pellet for approx 5 min, count the radioactive signal, and resuspend in TE.

3.4.1.1.3. Separation and Purification of S1 Fragments

1. Subject the S1 nuclease-protected RNAs to denaturing PAGE and autoradiography (*see Note 22*). Generally 4 or 5% acrylamide gels are suitable.
2. Excise the bands of interest from the gel and extract the RNA fragments by soaking in elution buffer at room temperature for several hours (to overnight), preferably twice. Extract the eluent once with phenol, once with CIA, and then ethanol precipitate without additional salt. As a precaution against the presence of residual SDS, the RNA can be reprecipitated, and three 70% ethanol washes should be performed. Count the radioactive signal from the dried RNA pellet and resuspend in a small volume of TE (as little as 5 μL); if there is more than one run-on sample, it is best to resuspend all at the same number of counts per minute per microliter.

3.4.1.2. RNase T1 ANALYSIS

1. Denature the gel-purified S1 nuclease-protected RNA fragments (*see Note 23*) in 5 μL of TE by incubating at 95°C for 2 min. Then cool immediately on ice and spin briefly.
2. Add 1 μL of 10 U/ μL RNase T1 and incubate for 45 min at 37°C.
3. Add 7 μL of denaturing gel-loading/dye solution and heat at 85°C for 2 min.
4. Separate digestion products by denaturing PAGE (20% acrylamide is generally suitable). Autoradiography may take from overnight to several weeks.

3.4.2. Indirect Analysis by RT-PCR

RT-PCR is performed using pairs of primers that will not generate a product from the naturally occurring mtTEC nucleic acids, at least within the distance constraints over which RT and PCR operate (*see Note 24*). Before performing the RT reactions, it may be useful to check for the presence of chimeric DNA by first performing PCR (as in **Subheading 3.4.2.3.**) on the pre-DNase sample (*see Subheading 3.3., step 3*).

3.4.2.1. REVERSE TRANSCRIPTION, PURIFICATION OF cDNA

1. Anneal a sample of the RNA (e.g., 2 μL , i.e., 10%) from **Subheading 3.3., step 9** with 2.5 pmol of RT primer in 10 μL by heating at 70°C for 10 min, then placing on ice for 10 min; spin down any condensation. If RT-minus reactions are to be done as controls, scale-up twofold and split after annealing.
2. For a final reaction volume of 20 μL , add dNTPs to 0.5 mM, 100 U of RNase H-minus M-MLV RT, and reaction buffer (supplied by the manufacturer).

3. Incubate at 48°C for 60 min (preferably in an oven to minimize changes in concentrations owing to condensation).
4. Digest the RNA (*see Note 25*): Add 80 μ L of TE and heat at 95°C for 5 min. Then immediately add 1 μ g of RNase A, and incubate at 37°C for 15 min.
5. Extract the cDNA once with phenol, then once with CIA. Ethanol precipitate as in **Subheading 3.3., step 2**, but use only 2 vol of ethanol, and then resuspend the pellets in 25 μ L of TE.

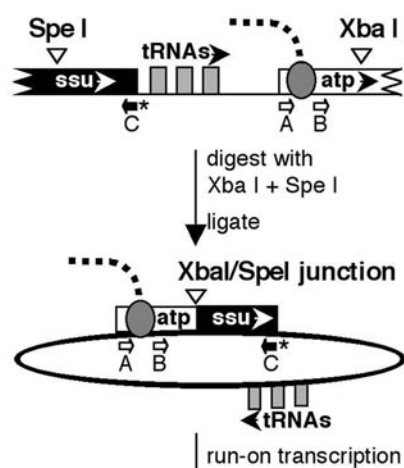
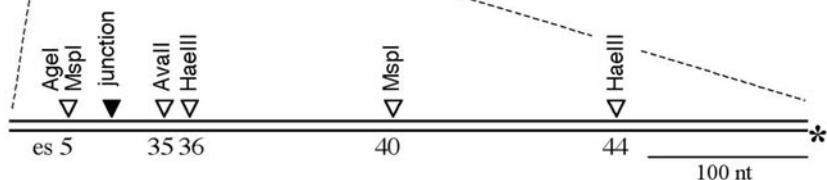
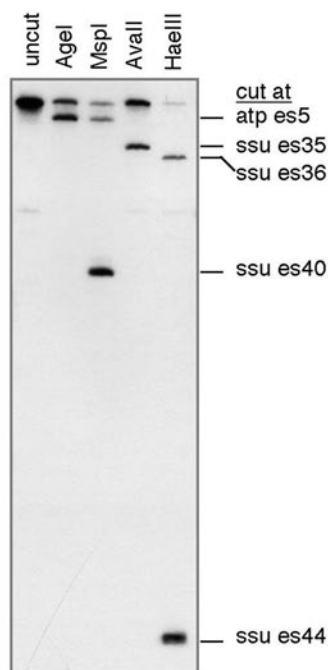
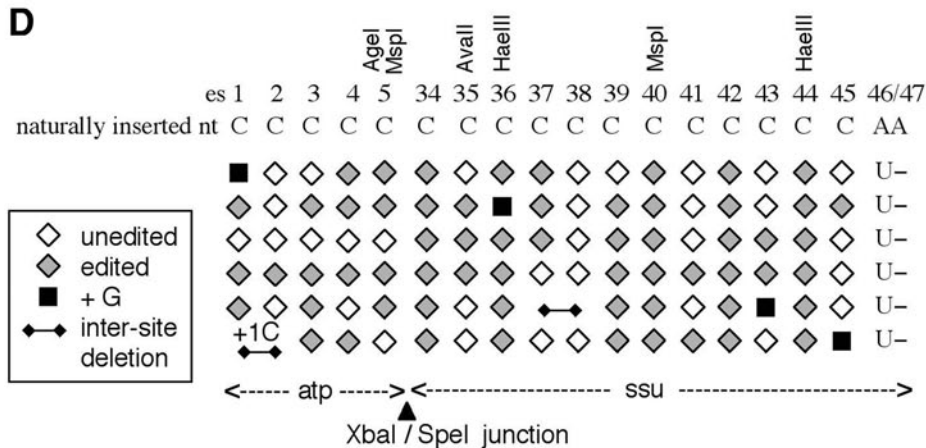
3.4.2.2. PCR AND CLONING

1. Using, e.g., 5 μ L of the cDNA as template, perform PCR over 30 cycles (1 min at 94°C, 1 min at 50°C, and 2 min at 72°C) with a 7-min final extension at 72°C, using 2.5 U of *Taq* DNA polymerase/100 μ L in buffer supplied by the manufacturer, with 200 μ M dNTPs and 100 pmol of each primer (*see Note 26*).
2. Run samples (e.g., 5–10%) on an agarose gel with size markers in 1X TAE, stain with ethidium bromide, and look for bands of mobility consistent with the expected chimeric sequence.
3. Products of the expected size can be gel purified from a preparative gel if necessary (*see Note 27*). Otherwise, clean up the DNA by organic extraction and ethanol precipitation, using a spin column or commercially available kit.
4. Clone into any convenient vector using standard molecular biology techniques, and sequence. **Figure 5D** provides an example of sequence-derived data.

3.4.2.3. HOT PCR AND RESTRICTION ENZYME ANALYSIS

At some editing sites, nucleotide addition generates recognition sequences for restriction enzymes, within the cDNA molecules, which are not present in the original mitochondrial DNA sequence. The extent of correct editing at these sites can be estimated by incubating end-labeled RT-PCR products (*see Note 28*) with the cognate restriction enzymes, followed by PAGE (7). **Figure 5B,C** shows an example of this type of analysis.

1. Using, e.g., 5 μ L of the cDNA as template, perform PCR in a 50- μ L reaction over 30 cycles (1 min at 94°C, 1 min at 50°C, and 2 min at 72°C) with a 7-min final extension, using 1.25 U of *Taq* DNA polymerase in buffer supplied by the manufacturer, with 200 μ M dNTPs, 25 pmol of the unlabeled primer, and 12.5 pmol total of a [^{32} P]-end-labeled primer.
2. Examine the samples (e.g., 5–10%) by denaturing PAGE (4%, unless the expected product is very small) and autoradiography, looking for bands of mobility consistent with the expected chimeric sequence.
3. Gel purify the chimeric sequence product on a native acrylamide or agarose gel if necessary (*see Note 27*).
4. Incubate samples of each PCR product with a series of restriction enzymes that are diagnostic for editing sites and, if possible, one or two enzymes that cut at recognition sites not affected by editing (to confirm the identity of the DNA).

A**B****C****D**

PCR products can generally be digested without cleanup, provided they do not constitute more than, e.g., 10% of the reaction volume.

5. Examine all or a portion of each digest by denaturing PAGE (*see Note 29*) and autoradiography. Look for bands of mobility consistent with cleavage at the editing sites (bearing in mind which terminus is labeled).

4. Notes

1. Somewhat surprisingly, the DNA present in active mtTECs can be efficiently digested by reaction enzymes (5). Ligation can then occur between fragment ends that were not adjacent in the genome (**Fig. 1**). Intramolecular ligation should be kinetically favored, forming single-fragment subgenomic circles (lower left of **Fig. 1**). On run-on transcription (top of **Fig. 4**), RNA polymerase enters the upstream end of the fragment after passing the ligation junction. If the fragment is from the interior of a gene or cotranscribed unit, the polymerase thus revisits the upstream end of the fragment, reiterating the sequence of the fragment in the transcript. In the absence of termination, multiple repeats of the fragment sequence could potentially be transcribed as the polymerase passes the junction repeatedly. Some ligation may also occur between different pieces of the fragmented mitochondrial genome, particularly at higher concentrations of mtTEC, making scrambled mtTEC-mtTEC intermolecular chimeras (lower right of **Fig. 1**; some could also form multifragment circles [not shown]); however, although these molecules have been detected by PCR, the levels of RNA produced are generally very low, often below the threshold of detection. Note that head-to-tail intermolecular ligation

Fig. 5. (*see opposite page*) Example of experiment involving mtTEC rearrangement and analysis of chimeric transcript (from an intramolecular circle) via RT-PCR. (A) The 3' part of the *ssu* gene and 5' part of the *atp* gene are located at the ends of one of the fragments that results from digestion of mtTEC with *Xba*I and *Spe*I. Fusion of the genes by ligation of these ends, followed by run-on transcription, results in a chimeric transcript that contains *atp* editing sites (es) 1–5 and *ssu* es 34–47. Primers (A–C) used for RT-PCR are shown as arrows below the RNA. (B) RT-PCR product made with primer B and end-labeled primer C, with restriction enzyme recognition sites generated by correct insertion (above) at editing sites listed below. (C) PAGE of restriction digests used to assess extent of correct editing at insertion sites depicted in (B). Positions of labeled fragments representing cleavage at edited sites are indicated on the right. (D) Patterns of correct and incorrect (*see ref. 8*) editing in six clones made using primers A and C (*see [A]*). Cytosine insertion sites are indicated by symbols: correctly edited, grey diamonds; unedited, open diamonds; G misinsertion, solid squares; lines with small diamonds at either end represent deletions encompassing regions between adjacent sites, and +1C indicates that the clone has one more cytosines than the genomic sequence in the region of the deletion. Insertions at the dinucleotide editing site (es46/47) are indicated by the appropriate capital letters, while a dash indicates no insertion.

tion between two copies of the same fragment will generate the same junction as an intramolecular circle. Adding exogenous DNA to the ligation reaction will result in chimeras between mtTEC and exogenous DNA (**Fig. 2**), at the expense of some of the mtTEC-mtTEC pairings. Some of the mtTEC-exogenous DNA ligation products will regenerate the original genomic sequence across the ligation point; these products are referred to as “reconstruction chimeras.” The remaining scrambled sequence combinations are named “intergenic chimeras” here.

2. Chimeric transcripts from high-concentration chimeric templates driven by highly transcribed upstream genes are most likely to be detectable. The downstream RNA pool from exogenous DNA ligated to multiple mtTEC genes that drive transcription also falls into this category.
3. Restriction enzymes are chosen with the following considerations in mind:
 - a. Enzymes that make staggered cuts in the duplex are preferable, since sticky ends are more efficiently ligated.
 - b. The surprising persistence of ligated *Xba*1 ends in the continued presence of *Xba*1 has been utilized to analyze transcripts from a circularized *Xba*1 fragment (5). However, it is preferable to use pairs of nonisoschizomers to generate compatible sticky ends on the fragments targeted for ligation, if possible. If a single restriction enzyme is used and does not allow sufficient ligation product to accumulate, it may be possible to deplete the enzyme before ligation, using a BioSpin column.
 - c. Ligation of exogenous DNA should be more efficient if the targeted endogenous mtTEC fragment cannot circularize; in this case it might be helpful to include an additional restriction enzyme that generates an incompatible upstream end.
4. Analysis of editing in the chimeric transcripts can be performed (A) directly on labeled RNA by S1 nuclease protection followed by T1 digestion, or indirectly by (B) making double-stranded DNA copies via RT-PCR, followed by (B1) restriction enzyme analysis or (B2) cloning and sequencing. If more than one technique can be applied for a given chimeric template experiment, it may be advantageous to do so because the different methods provide overlapping sets of information that complement each other in some respects. The advantages and disadvantages of each method are shown in **Table 1**. The sequence of an individual RNA molecule does not reveal the point at which in vivo synthesis gives way to run-on transcription upstream of the chimeric junction; however, since editing is highly efficient and accurate in vivo (2) (cf. in vitro [3,5,8]), any unedited or incorrectly edited sites, and downstream RNA, are known to have been made in vitro. If direct analysis is to be performed:
 - a. Information about the identity and exact position of nucleotide changes from genomic sequence within the individual T1 fragments can often be obtained by fingerprinting and/or secondary analysis (2,6,9).
 - b. See **Note 16a** and **b**.
5. General technical considerations: To avoid nonspecific degradation of RNA, all reagents should be of high purity and handled with gloves. Reaction materials

and intermediates should be kept on ice between incubations and stored at -20°C unless otherwise indicated.

6. Denaturing gels are made with 19:1 acrylamide:bisacrylamide, 7 M urea, and 1X TBE. Samples and labeled size markers are denatured by mixing with an equal volume of denaturing gel-loading/dye solution and heating at 85°C for 2 min, then electrophoresed using 1X TBE. The bands are then visualized by autoradiography; use of a high-sensitivity film and intensification screen is recommended for S1 nuclease and RNase T1 fragments of run-on transcripts.
7. Ligated exogenous DNA does not support editing, at least in sequences located approx 24 bp and farther downstream of the point of attachment to mtTEC DNA as tested in **ref. 5**. However, it can be useful in facilitating the analysis of editing at mtTEC sites that are located very close to the point at which restriction enzyme cleavage is used to remove downstream endogenous DNA: there, the ligated exogenous sequence can act as a tag to enable RT-PCR, or to overcome difficulties in direct analysis of sequences located very close to the 3' end of an RNA (A. Majewski and J. M. Gott, unpublished data). Furthermore, it would be interesting to observe whether editing can occur at exogenous DNA sites that are located extremely close to the ligated junction.
8. If not using a preexisting restriction enzyme site, an appropriate recognition sequence can be engineered either when making the plasmid insert or when amplifying the cassette, by inclusion at the 5' end of the upstream primer. The opposite end of the cassette should be kept blunt.
9. When planning the details of the cleavage, ligation, and transcription series, the following considerations apply:
 - a. Concentration of glycerol (contributed by the mtTEC preparation and added enzymes stocks) should not exceed approx 10%, in order to reduce the chance of restriction enzyme star activity; concentrated restriction enzyme stocks are useful.
 - b. Incubation of mtTEC with *Xba*1 in the absence of added salt seems to cause degradation of the mitochondrial DNA, possibly via a nuclease copurifying with the mtTEC; therefore, a buffer recommended for use with *Xba*1 ("Buffer H") has been added to the digestion mixture. When using other restriction enzymes, test digestions should first be carried out, monitoring the resulting fragment patterns by extracting the DNA and visualizing it via agarose gel electrophoresis.
 - c. For a given quantity of mtTEC protein, reaction volumes could be increased somewhat if required, such as to accommodate a dilute mtTEC preparation, or large total volumes of restriction enzymes or label.
 - d. If intermolecular ligation products are sought, it is best to keep the restriction enzyme and ligation reaction volumes to a minimum; conversely, formation of intramolecular circles will be unaffected or facilitated by lower DNA concentration.
 - e. If making labeled transcripts for direct analysis, reactions can be scaled up by a factor of 2 or more; the nucleic acid from 2X scaled-up reactions can be processed as in **Subheading 3.3**, without modifying the protocol.

- f. Controls can be included in which the restriction enzymes, exogenous DNA, ligase, or run-on transcription step are omitted.
10. If reactions are sufficiently large, samples of DNA may be extracted after each stage to estimate the efficiency of cleavage and the fraction of a given endogenous fragment that is present in the desired ligation product. The DNA is inspected by agarose gel electrophoresis (staining with ethidium bromide, and performing Southern blot analysis if necessary). Bands representing ligation of exogenous DNA to one and/or both ends of the endogenous fragments should usually be visible above background. Intramolecular circles may be resolvable from their linear precursors, depending on their size, topology, and electrophoresis conditions; they can form a significant proportion of the ligation product pool for fragments with self-compatible ends even in the presence of excess exogenous DNA. A background smear of variable intensity may also be present, owing to random ligation of some of the mtTEC fragment pool, as well as any contaminating nuclear DNA and its ligated derivatives, and any sheared mtTEC DNA. Ligation products may also be detected by applying the PCR protocol in **Subheading 3.4.2.3.** to the nucleic acid extracted in **Subheading 3.3., step 3.**
 11. An example would be 22.5 μL of a 0.22 $\mu\text{g}/\mu\text{L}$ prep. A range of 2–10 μg of protein has been used successfully in this protocol. mtTEC preparations may vary considerably in their protein and nucleic acid concentrations. The preparations used in chimeric template experiments typically contained 10% glycerol, 10 mM Tris-HCl (pH 8.0), 5 mM MgCl_2 , 0.5 mM EDTA, and 1 mM DTT; because of the variations in volume used, the concentrations of these substances varied somewhat in the reactions.
 12. The stated amount aims for an approx 20-fold excess of exogenous cassette over mtTEC fragment ends; however, a reasonable proportion of the mtTEC fragment pool can still be ligated to the exogenous DNA at smaller molar excesses of the latter—some ligation has been observed down to an approx 2-fold excess.
 13. Concentrations of 40–50 mM Tris and 67–90 mM NaCl have been present in previous successful experiments.
 14. To make labeled RNA, an $[\alpha\text{-}^{32}\text{P}]\text{NTP}$ (typically 30 μCi) is included, supplemented with the corresponding unlabeled NTP to, e.g., 5 μM (see also **Note 15**). A 10-min 500 μM chase is added at the end of the reaction. Choice of labeled NTP: (1) if radioactive GTP is used, all T1 fragments will be labeled, with equal intensity; (2) if T1 fragments are to be subjected to secondary analysis, the local sequence will determine which label is most useful.
 15. Nucleotide concentrations:
 - a. Reducing the concentration of a given NTP (to e.g., 5 μM) decreases the efficiency with which it is inserted at editing sites and, conversely, increases the efficiency of editing at sites that immediately precede locations where that nucleotide is conventionally encoded (4).
 - b. There are practical limits to reducing the concentrations of more than one nucleotides at a time, since reduced transcription may lower the labeled RNA signal too far.

- c. Some ATP is consumed during the ligation reaction, but the amount has not been determined; if a defined low concentration is desired during transcription, most of the cold ATP can be removed after ligation by passing the reaction through a BioSpin-30 column (one per standard size reaction) before transcription (the columns should first be preequilibrated with 0.75X mtTEC buffer, 1X buffer H, and 1 mM DTT).
16. ssDNA probes are made using recombinant phagemids that contain an M13 origin of replication, designed according to the following considerations:
 - a. It is generally preferable to protect RNA fragments of ~250–400 nt; shorter fragments are likely to be contaminated by incompletely digested RNA fragments, while longer fragments generally give overly complex RNase T1 digestion patterns.
 - b. The size distribution of the fragments that will be produced on T1 digestion of the S1 products; for resolution by single-dimension electrophoresis, important fragments should differ in size by at least 1 nt from all others in the mix.
 - c. The sequence of the protecting ssDNA should correspond as closely as possible to that of the target chimeric transcript in the region to be protected. It is preferable to use the edited versions of the sequence: although single-nucleotide insertion/deletion differences are largely resistant to nicking by S1 nuclease under the conditions used here, any residual nicking takes place at the bulged bases, so it is better not to have these in the RNA component of the hybrid. Two or more successive mismatches or insertions/deletions relative to the target RNA will constitute a substrate for S1 nicking, and up to approximately half the RNA may be cleaved when a mismatch and insertion occur within 2 nt of each other.
 - d. The ssDNA should be checked against labeled control RNA, by following the S1 nuclease protection protocol in **Subheading 3.4.1.1.2.**, before use with experimental samples.
17. Control RNAs should be labeled with the same [α - ^{32} P]NTP used for the mtTEC transcription, and the full-length transcripts gel purified. Their sequence should correspond as closely as possible to that of the expected chimeric RNA in the region to be S1 protected (*see Note 16c*). These RNAs allow one to check the S1 nuclease activity, ssDNA protection, and RNase T1 digestion reaction, while the digestion products from successful procedures act as mobility markers for electrophoresis. To act as a positive control for S1 nuclease activity, the RNAs should extend for some length beyond the region of complementarity with the protecting ssDNA, at least at one end.
18. In principle, any of the nucleotide-specific endonucleases could be used in this analysis, but T1 is most suitable for *Physarum* mitochondrial sequences because the transcripts are very A/U rich, so that cleavage by A- or U-specific enzymes would in general produce few fragments of distinctive size that could be resolved by PAGE; and because many of the added Cs are inserted next to encoded Cs, so that correct editing at these sites would liberate only mononucleotides on cytosine-specific digestion.
19. A wide variation has been observed among the behaviors of S1 nuclease stocks from different suppliers, and even among different batches of enzyme from the

same supplier. Therefore, we recommend use of the product listed in **Subheading 2.**, and that each new batch of nuclease be titrated to determine the number of units that generates a fragment of predicted size from test RNA, with minimal residual undigested RNA or overdigestion (cleavage within the hybridized region). To do this, samples of labeled control RNA are mixed with unlabeled mtTEC RNA and hybridized to protecting ssDNA, following the protocol of **Subheading 3.4.1.1.1., steps 1–4**, then examined as in **Subheading 3.4.1.1.3., step 1**.

20. Three rounds of S1 nuclease protection are usually necessary to obtain a reasonably clean fragment from less abundant individual chimeric transcripts (e.g., the reconstruction chimera in **Fig. 3**). Two rounds may be sufficient for more abundant RNAs, such as the pool of transcripts from exogenous DNA cotranscribed from multiple upstream genes (intergenic chimeras in **Fig. 3**). For the first round of S1 protection, the probe and target must interact in a complex RNA mixture, necessitating a long hybridization incubation; subsequently, the amount of nonspecific RNA is greatly reduced, so less time is needed for pairing between probe and target (however, the hybridizations can be left overnight without harm). Details of the ethanol precipitation, wash, and resuspension steps are designed to minimize loss and degradation of RNA and accumulation of salt. These considerations are especially important for steps subsequent to the first S1 nuclease digestion, when the total quantity of intact RNA has been greatly reduced and the solution is high in salt. The quantity of labeled RNA present at each stage can be estimated by counting the dry pellets using a scintillation counter (although there may be significant coprecipitation of unincorporated label in the initial stages).
21. A substantial amount of signal can be used, such as 50,000 cpm, in order to provide ample protected fragment for subsequent RNase T1 analysis experiments.
22. Care should be taken to avoid loading too much of the control RNA samples onto gels that contain the products from run-on transcription experiments. Bands from run-on transcripts may be substantially less intense than predicted from the radioactivity in the final S1 protection pellet, because there may be a significant background of partially digested RNA and other byproducts, as well as residual unincorporated label; an overnight exposure is usually required.
23. Aim to digest an equivalent amount of radioactivity from each sample and control, so that corresponding bands in different lanes have similar intensity.
24. Thus, e.g., the primers that amplify the subgenomic circle transcripts in **Figs. 4** and **5** are directed away from each other in natural sequences, and a reconstruction chimera is amplified via the use of a primer from upstream of the chimeric junction in conjunction with one from the terminal tag sequence. A “nested” downstream primer can be used in the PCR step in order to decrease the probability of amplifying products based on spurious priming during reverse transcription.
25. Since use of an RNase H-minus RT is recommended, this step is included to prevent any spurious priming by RNA in the subsequent PCR reaction; the nucleic acids are first heated to free the RNA from hybrids with DNA.
26. Termini that are to be ligated directly to the vector should contain a 5′ phosphate—use primers with 5′ phosphates in the PCR reaction.

27. Gel purification is indicated if there is a significant background of higher and/or lower molecular weight products, which could reduce the efficiency of cloning chimeric sequences, or complicate the restriction enzyme analysis of **Subheading 3.4.2.3**. However, it might be worth investigating whether background bands represent novel sequence products of the transcription/editing process, such as deletions between editing sites, as have been found in the *Physarum* system ([8]; see also Fig. 5D).
28. Radioactively labeled PCR products allow for more sensitive restriction enzyme analysis and can be generated by having one of the primers 5'-end labeled with ^{32}P (using polynucleotide kinase and [γ - ^{32}P]ATP). It is probably best to have the unlabeled primer present in, e.g., twofold excess in order to minimize the chances of ending up with residual labeled ssDNA, which could generate false negative signals during restriction enzyme analyses.
29. Choose an acrylamide percentage suitable for resolving the fragments that will result if all sites are edited; if the range of fragment sizes is very large, it may be beneficial to use more than one percentage gel. If necessary, precipitate the DNA and resuspend in a smaller volume before loading.

Acknowledgments

The protocols herein were developed in collaboration with Dr. Jonatha M. Gott in the Center for RNA Molecular Biology at Case Western Reserve University School of Medicine, supported by National Institutes of Health grant GM54663.

I also wish to acknowledge helpful suggestions and comments from Adam Majewski; Yu-Wei Cheng; Amy Rhee; Olga Kourennaia; and Drs. Linda Visomirski-Robic, Tim Nilsen, Pat Maroney, Donal Luse, Pieter de Haseth, David Setzer, Jo Ann Wise, Eric Christian, David McPheeters, and Piet de Boer.

References

1. Gott, J. M. (2001) RNA editing in *Physarum polycephalum*, in *RNA Editing: Frontiers in Molecular Biology* (Bass, B., ed.), Oxford University Press, Oxford, UK, pp. 20–37.
2. Visomirski-Robic, L. M. and Gott, J. M. (1995) Accurate and efficient insertional RNA editing in isolated *Physarum* mitochondria. *RNA* **1**, 681–691.
3. Cheng, Y.-W. and Gott, J. M. (2000) Transcription and RNA editing in a soluble in vitro system from *Physarum* mitochondria. *Nucleic Acids Res.* **28**, 3695–3701.
4. Cheng, Y.-W., Visomirski-Robic, L. M., and Gott, J. M. (2001) Non-templated addition of nucleotides to the 3' end of nascent RNA during RNA editing in *Physarum*. *EMBO J.* **20**, 1405–1414.
5. Byrne, E. M. and Gott, J. M. (2002) Cotranscriptional editing of *Physarum* mitochondrial RNA requires local features of the native template. *RNA* **8**, 1174–1185.

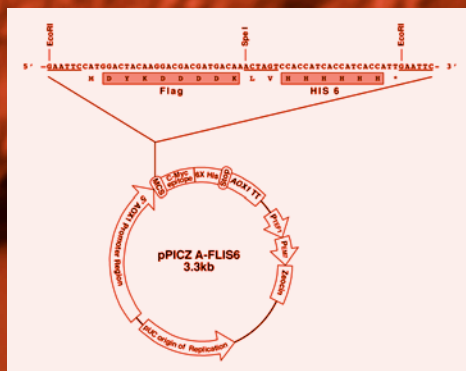
- 6 Visomirski-Robic, L. M. and Gott, J. M. (1997) Insertional editing of nascent mitochondrial RNAs in *Physarum*. *Proc. Natl. Acad. Sci. USA* **94**, 4324–4329.
- 7 Rundquist, B. A. and Gott, J. M. (1995) RNA editing of the *col* mRNA throughout the life cycle of *Physarum polycephalum*. *Mol. Gen. Genet.* **247**, 306–311.
- 8 Byrne, E. M., Stout, A., and Gott, J. M. (2002) Editing site recognition and nucleotide insertion are separable processes in *Physarum* mitochondria. *EMBO J.* **21**, 6154–6161.
9. Visomirski-Robic, L. M. and Gott, J. M. (1997) Insertional editing in isolated *Physarum* mitochondria is linked to RNA synthesis. *RNA* **3**, 821–837.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



Methods for Analysis of Mitochondrial tRNA Editing in *Acanthamoeba castellanii*

Amanda J. Lohan and Michael W. Gray

Summary

We present here methodology for assaying 5'-terminal editing of mitochondrial tRNAs in the amoeboid protist *Acanthamoeba castellanii*. This type of editing involves replacement of one or more nucleotides within the first three positions at the 5' end of a tRNA substrate. The assay procedure involves RNA ligase-mediated joining of the 5' and 3' ends of a tRNA, use of the resulting circularized tRNA as template for cDNA synthesis primed by tRNA-specific primers over a region that encompasses the ligated 5' and 3' halves of the acceptor stem, amplification of cDNA via polymerase chain reaction, and cloning and sequencing of double-stranded cDNA product. This approach has the advantage that it simultaneously reveals potential editing events on the 5' and 3' side of an acceptor stem, as well as serves to identify mature tRNAs (characterized by having a 3'-CCA_{OH} motif) and partially processed intermediates.

Key Words: *Acanthamoeba castellanii*; mitochondria; mitochondrial isolation; RNA ligase; reverse transcription polymerase chain reaction; tRNA circularization; tRNA editing.

1. Introduction

Among the myriad definitions and categorizations of RNA editing in the literature, perhaps the most pertinent with respect to this chapter is “any programmed alteration of RNA primary structure to generate a sequence that could have been directly encoded at the DNA (gene) level” (1). Various forms of mitochondrial tRNA editing have been reported for a wide variety of organisms (2–4), including C-to-U substitution editing in plants (5) and animals (6); C or U insertional editing in myxomycete protozoa (7,8); a type of template-independent editing observed at the 3' ends of mitochondrial tRNAs in many

metazoa (9–14); and the template-dependent processes evidenced as 5′ editing in the mitochondria of amoeboid protozoa (15–18) and several chytridiomycete fungi (19,20), as well as the template-directed 3′ mitochondrial tRNA editing recently characterized in centipede (21).

This chapter focuses on terminal editing events: ones localized to the acceptor stem and involving nucleotide substitution. The process that we address (which is the most frequently observed and widely distributed type of tRNA editing) involves replacement of nucleotides within the 5′ or 3′ half of the tRNA acceptor stem in either a templated or a template-independent fashion. We describe in detail a method that can delineate 5′-terminal tRNA-editing events from those occurring on the 3′ side of the acceptor stem.

In the amoeboid protist *Acanthamoeba castellanii*, the mitochondrial genome encodes 15 bona fide tRNA species (22). Secondary structure modeling predicts that 12 of these 15 tRNAs should have one or more mismatches in the first three base pairs of their acceptor stems (**Fig. 1**), a situation incompatible with normal tRNA function. Investigation of this apparent anomaly revealed that the inferred acceptor stem mis-matches are, in fact, “corrected” by a 5′-terminal editing process (15). A nucleotide incorporation activity (putative tRNA-editing activity) has been partially purified from *A. castellanii* mitochondria and shown to be capable of excising the first three nucleotides at the 5′ end of the acceptor stem, prior to catalyzing their replacement utilizing the corresponding 3′ half as a *cis*-acting template guide (23) (**Fig. 1**). As a consequence of this editing, mismatched acceptor stem positions are converted to canonical Watson-Crick base pairs (most typically G:C).

Initially, Lonergan and Gray (15) used oligonucleotide-primed reverse transcriptase (RT) sequencing to identify 5′ editing in a subset of *Acanthamoeba* mitochondrial tRNAs. Applying a different method (detailed in this chapter), Price and Gray (1) were able to confirm the previously characterized tRNA edits and expand the study to encompass the remaining predicted mitochondrial tRNA edits. The latter methodology capitalizes on the ability of T4 RNA ligase to catalyze the joining of a 5′-phosphoryl moiety to a 3′-hydroxyl group between or within single-stranded regions of RNA or DNA (24). Because mature tRNA molecules are characterized by 5′-monophosphate and 3′-hydroxyl termini as well as a four-nucleotide 3′-terminal extension, they are ideal substrates for intramolecular ligation mediated by RNA ligase (**Fig. 2**). Use of such a circularized template for cDNA synthesis offers several advantages compared with other methods for analysis of terminal editing in tRNA. Primarily, cDNA synthesis on a circularized tRNA template spans the entire region of the acceptor stem, thus allowing simultaneous determination of 5′ and 3′ acceptor stem sequences. Depending on where the forward and reverse primers are situated within the tRNA sequence, regions of the tRNA bounding

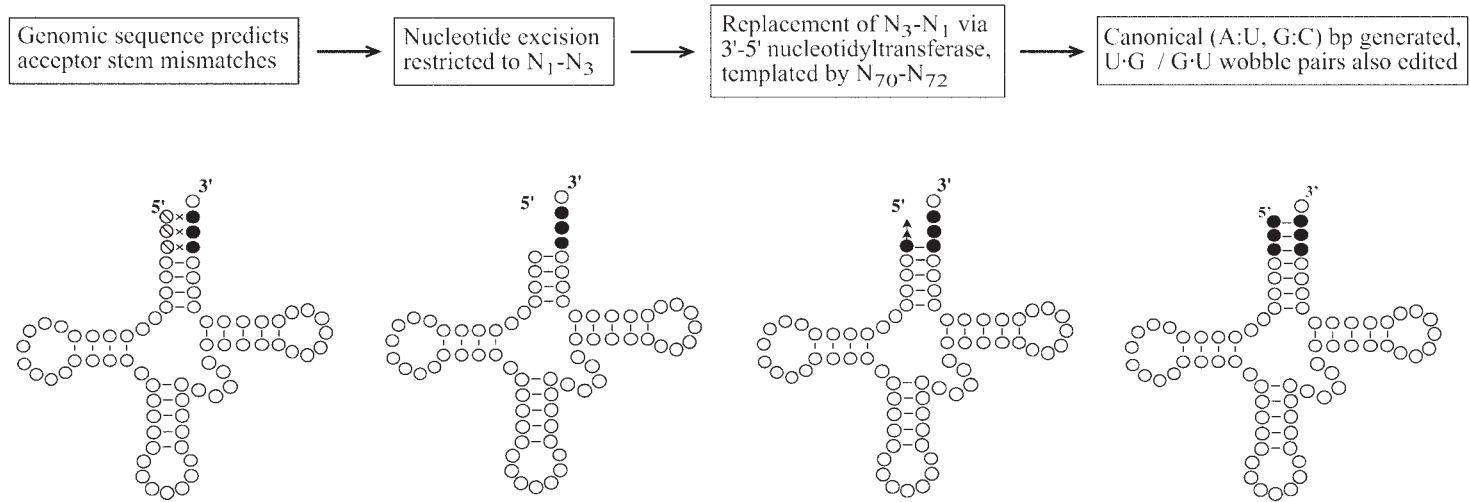


Fig. 1. Schematic summary of proposed mechanism of 5' tRNA editing in *A. castellanii* mitochondria. Secondary structure modeling predicts that 12 of 15 mtDNA-encoded tRNAs should have one or more mismatches in the first three base pairs of their acceptor stems. The tRNA-editing activity in *Acanthamoeba* mitochondria is capable of excising the first three nucleotides at the 5' end of the acceptor stem, prior to catalyzing their replacement utilizing the corresponding 3' half of the stem as a *cis*-acting template guide.

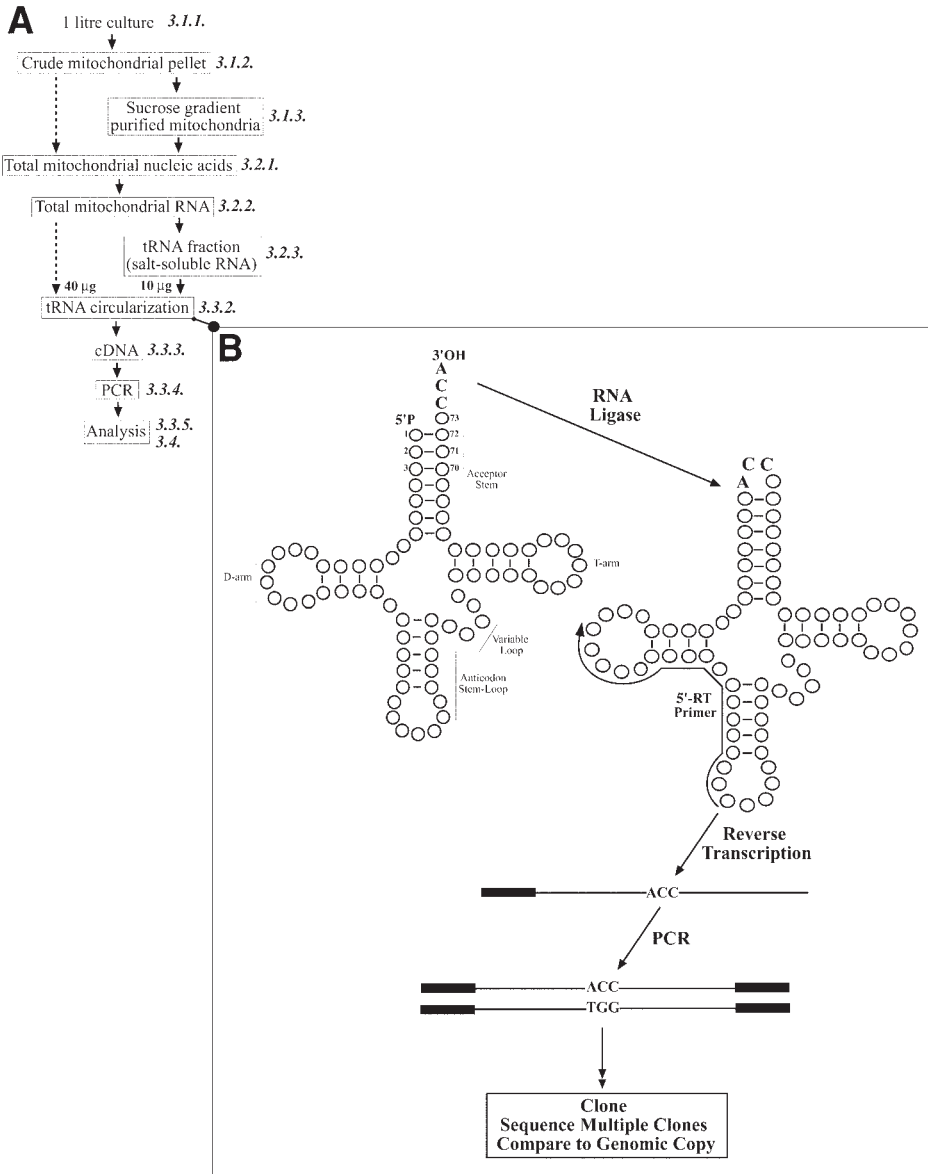


Fig. 2. Overview of strategy for analysis of 5'- or 3'-terminal tRNA editing. (A) Flow chart overview of methods outlined in **Subheading 3.** (see numerals to right of boxes) for analysis of 5' tRNA editing in highly purified *Acanthamoeba* mitochondria. Solid arrows refer to methodology described in detail. Dashed arrows indicate alternate steps or abbreviated methodologies applied to other organisms or when material is not abundant. (B) Illustration of steps involved in tRNA circularization, cDNA synthesis, and polymerase chain reaction (PCR).

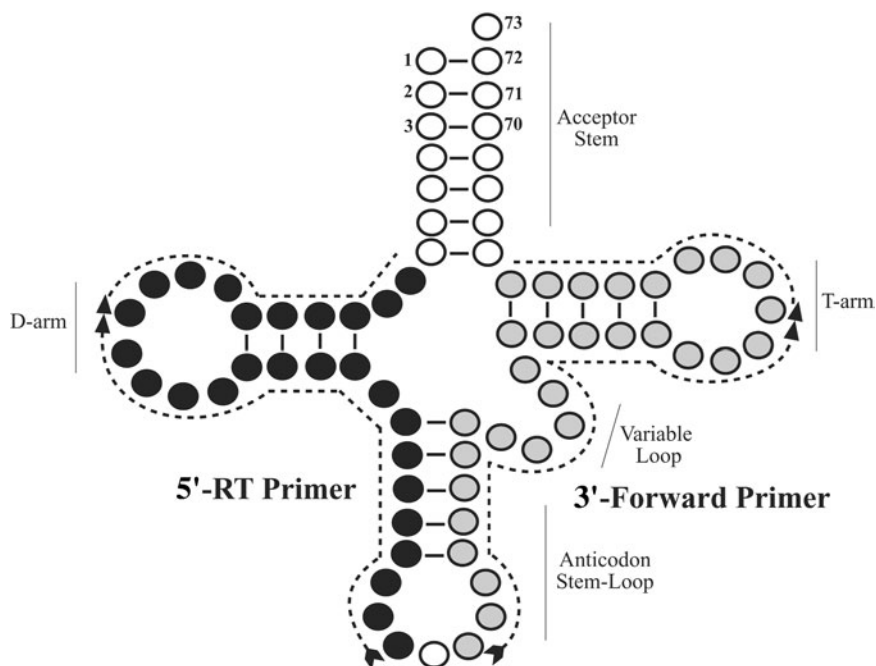


Fig. 3. Illustration of optimal regions within a tRNA for primer placement. The dotted line and shaded circles delineate the regions of primer placement for cDNA synthesis (RT primer) and PCR (RT and 3'-forward primers). Solid circles indicate the 5'-RT primer positioning, and shaded circles indicate the 3'-forward primer. Arrows denote the direction of the primer. Primers may lie within but not necessarily encompass the entire stretch of sequence indicated in the figure.

the acceptor stem may also be analyzed (**Figs. 2 and 3**). A further benefit of the circularization method is that it reveals the presence of the 3'-CCA_{OH} motif, which mature tRNAs require for aminoacylation. Unlike their eubacterial counterparts, mitochondrial tRNAs generally do not encode the -CCA_{OH} extension; rather, this sequence is added posttranscriptionally in mitochondria, and its presence therefore indicates that the corresponding gene is functional.

The procedures outlined here describe in detail the preparation of *Acanthamoeba* mitochondrial RNA and tRNA fractions (**Subheadings 3.1.1.–3.2.3.**), RT polymerase chain reaction (RT-PCR) (**Subheadings 3.3.1.–3.4.**), and expected results. Analysis of *Acanthamoeba* mitochondrial tRNA^{Phe} is presented as an example of 5'-terminal editing in this system.

2. Materials

1. Neff base medium (pH 7.0), 1 L: 0.75% yeast extract, 0.75% proteose peptone, 2 mM KH₂PO₄, 1 mM MgSO₄. Adjust the volume to 890 mL and autoclave. Sub-

sequent to sterilization and just prior to inoculation, add aseptically 100 mL of 15% glucose, 10 mL of 10 mM ferric citrate, 50 μ L of 1 M CaCl_2 , 100 μ L of 10 mg/mL thiamine, 100 μ L of 2 mg/mL D-biotin, and 2 μ L of 0.5 mg/mL vitamin B₁₂ to the 890 mL of autoclaved Neff base medium. Either a 4-L Erlenmeyer or 2.8-L Fernbach flask should be used.

2. Phosphate-buffered saline (PBS) (pH 7.4): 4.3 mM Na_2HPO_4 , 1.4 mM KH_2HPO_4 , 137 mM NaCl, 2.7 mM KCl (1 L of PBS is required for every liter of culture prepped). Autoclave and store at 4°C.
3. 10% Bovine serum albumin (BSA): (a) BSA fraction V, (b) BSA fraction V essentially fatty acid free—used only for the gradient solutions (may be stored at -20°C).
4. 100 mM (1.74 mg/mL) Phenylmethylsulfonyl fluoride (PMSF). Dissolve in isopropanol, and store at -20°C.
5. Homogenization buffer: 10 mM Tris-HCl (pH 7.6), 1 mM ZnCl_2 , 0.25 M sucrose, 1 mM dithiothreitol (DTT), 0.1% BSA—from sterile stock solutions. Prepare fresh on the day of use, with addition of BSA and DTT just prior to use.
6. Mitochondrial wash buffer: 10 mM Tris-HCl (pH 7.6), 10 mM EDTA, 0.35 M sucrose, 1 mM DTT, 0.1% BSA—from sterile stock solutions. Prepare fresh on the day of use, with addition of BSA and DTT just prior to use.
7. Sucrose gradient solutions: 1.55 M sucrose and 1.3 M sucrose in 50 mM Tris-HCl (pH 8.0), 3 mM EDTA (pH 8.0), 1 mM DTT, 0.1% BSA (fatty acid-free stock)—from sterile stock solutions. Prepare fresh on the day of use, with addition of BSA and DTT just prior to use.
8. Pre-sucrose gradient buffer: 50 mM Tris-HCl (pH 8.0), 3 mM EDTA (pH 8.0), 0.25 M sucrose, 1 mM DTT, 0.1% BSA—from sterile stock solutions. Prepare fresh on the day of use, with addition of the BSA and DTT just prior to use.
9. Post-sucrose gradient/storage buffer: 20 mM Tris-HCl (pH 8.0), 0.5 mM EDTA, 0.25 M sucrose, 15% (v/v) glycerol, 1 mM DTT, 1 mM PMSF—from sterile stock solutions. Prepare fresh on the day of use, with addition of DTT and PMSF just prior to use.
10. Phenol-cresol mix: 500 g of phenol crystals dissolved in 55 mL of distilled water (dH_2O) (alternatively, 500 mL liquified phenol), 0.5 g of 8-hydroxyquinoline. Mix and equilibrate against 10–50 mM Tris-HCl (pH 7.6).
11. TE: 10 mM Tris-HCl (pH 7.6), 1 mM EDTA (pH 8.0).
12. DNase I buffer (10X): 400 mM Tris-HCl (pH 7.6), 60 mM MgCl_2 .
13. Dounce tissue grinder (15 or 40 mL) (Wheaton).
14. Avian myeloblastosis virus (AMV) RT and supplied buffer.
15. tRNA-specific oligonucleotide primers: 1 and 10 pmol stocks of 5'-RT primer, 10 pmol stock of 3' forward primer.
16. Motorized or battery-operated pipet controller.

3. Methods

The methods outlined cover the growth and harvesting of purified mitochondria from the amoeboid protozoon *A. castellanii*. Procedures for isolating both total mitochondrial RNA and a tRNA-enriched fraction are given. Condi-

tions for circularization of tRNA species and subsequent RT-PCR conditions for generation of DNA sequences for cloning are also provided.

3.1. Growth of *A. castellanii* and Extraction of Mitochondria

This section includes a detailed description of the growth parameters for a 1-L culture of *A. castellanii*, coupled with protocols for isolation of both crude and sucrose gradient-purified mitochondria.

3.1.1. Growth of *A. castellanii*

Five-milliliter starter cultures are maintained on a shaking platform in 50-mL tissue culture tubes at room temperature. A 4 d starter culture is used to inoculate the 1-L culture. The required growth conditions are 30°C with moderate shaking until an OD₅₅₀ reading of 0.9–1.0 is achieved, typically 4 to 5 d. The doubling time of a healthy culture of *Acanthamoeba* trophozoites is approx 18–20 h (see **Note 1**).

3.1.2. Isolation of Crude Mitochondrial Pellet

Although the mitochondrial isolation protocol presented is specific for *Acanthamoeba*, we have used it, with minor modification (chiefly with respect to the cell disruption technique), for other unicellular eukaryotes (see **Note 2**). All manipulations are carried out on ice, all buffers and tubes are stored on ice, and all centrifugation steps should be at 4°C. With the exception of PBS, all buffers for mitochondrial isolation (see **Subheading 2, items 5–8**) are prepared fresh on the day of use. DTT and BSA or PMSF are added to the solutions just prior to use.

1. Divide the culture between two chilled 750-mL plastic centrifuge bottles and centrifuge at 900g for 5 min. A swinging-bucket rotor is preferable but not an absolute requirement.
2. Wash the cell pellets twice with 0.5 vol of cold PBS, and centrifuge for 5 min at 900g after each wash. Do not vortex; resuspend the pellets by either gentle swirling or flushing the buffer over the pellet using a pipet.
3. Resuspend each washed pellet in 25 mL of homogenization buffer (5 mL of buffer is required for every 100 mL of initial culture).
4. Disrupt the cells in a chilled Dounce tissue grinder, 35 strokes with pestle A followed by five strokes with pestle B. A stroke is a single down-up movement. Greater than 90% cell breakage is achieved with this method (see **Note 2**).
5. Use low-speed centrifugation (900g for 10 min) to sediment the nuclei and any intact cells. A swinging bucket rotor is preferable but not an absolute requirement.
6. Divide the supernatant from the low-speed spin between two cooled 30-mL Corex tubes. When transferring the supernatant be sure to avoid touching the pellet-supernatant interface. A small volume (<1 mL) may be left at the interface in

order to avoid any carryover contamination. Discard the gray-brown pellets (nuclei, large membrane fragments and intact cells).

7. Centrifuge the postnuclear supernatants from **step 6** for a further 20 min at 9000g to yield the crude mitochondrial pellet (typically caramel-brown). Discard the supernatant (crude cytoplasmic fraction) from this spin. Use a fixed-angle rotor for this and all subsequent centrifugations.
8. Wash each of the mitochondrial pellets with 10 mL of wash buffer. For efficient washing, resuspend the pellet in the wash buffer by repeated pipetting using a motorized pipet controller (*see Note 3*). Resediment the mitochondria with a 10-min centrifugation at 9000g.
9. Repeat the wash as in **step 8**, but increase centrifugation time to 15 min. Final pellets constitute the crude mitochondrial yield. The crude mitochondria can be either further purified by sucrose gradient centrifugation (*see Subheading 3.1.3.*) or directly used for RNA extraction (*see Subheadings 3.2. and 3.2.1. and Note 4*).

3.1.3. Purification of Mitochondria on Sucrose Gradients

The mitochondria are further purified by ultracentrifugation on a two-step 1.3 *M*, 1.55 *M* sucrose gradient. Each sucrose gradient consists of a 7.5-mL lower layer of 1.55 *M* sucrose, onto which is carefully layered 15 mL of 1.3 *M* sucrose solution. Both sucrose solutions are prepared in the buffer outlined in **Subheading 2., item 7**. The following protocol is specifically for use with a Beckman SW25.1 swinging-bucket rotor and 1 × 3 in. (25 × 76 mm) centrifuge tubes. The rotor requires three tubes; therefore, three step gradients must be prepared for each run.

1. Pipet 7.5 mL of 1.55 *M* sucrose into each centrifuge tube. To add the upper sucrose layer, place the pipet tip on the side of the tube just above the lower 1.55 *M* layer, and with the motorized pipet at the lowest setting, carefully layer on the sucrose mix. A correctly poured gradient will have a visible sharp interface.
2. Resuspend the pellets from **Subheading 3.1.2., step 9** in 1.5 mL of pre-sucrose gradient buffer by repeated pipetting (employing the automatic pipet aid) (*see Note 3*). The suspension should have no visible clumps. Each gradient will hold 3–5 mL of applied material.
3. Pool the two pellet suspensions and load onto a single gradient. Gently layer the mitochondrial suspension onto the gradient, taking care to avoid mixing with the 1.3 *M* shelf. To act as balance tubes, blind load the remaining two gradients with an equivalent volume of pregradient buffer.
4. Centrifuge the gradients at 4°C for 1 h at 22,500 rpm. The brake should be on the gentlest setting allowed by the manufacturer. The mitochondria should form a compact caramel-brown band at the interface of the 1.3 and 1.55 *M* layers. A cloudy lipid band may be visible near the top of the gradient, at the interface between the gradient and the applied sample. A small pellet may be present, if there was carryover nuclei contamination. A large pellet with little material at the

1.3–1.55 *M* interface is indicative of loss of mitochondrial integrity and/or the loading of incompletely resuspended pellets.

5. Remove the bulk of the upper sucrose layer by aspiration, being especially careful to avoid disturbing the mitochondrial band.
6. Using an 18-gage needle (bent at a 90° angle) attached to a small syringe, slowly recover the mitochondrial band. The bent syringe allows the needle to be placed across the top of the mitochondrial band. Once the needle tip is in the band, pull the contents into the syringe. The mitochondrial material has a viscous (clotlike) consistency, so that it should move cleanly into the syringe.
7. Take note of the volume, generally 3 to 4 mL, and transfer to a chilled 15-mL snap-cap culture tube.
8. Dilute the mitochondria slowly (over a 15- to 20-min period) with 2 vol of post-sucrose gradient/storage buffer. This introduces both glycerol as a cryoprotectant and PMSF and DTT as protease inhibitors.
9. Centrifuge at 12,000*g* for 30 min (4°C).
10. Decant the supernatant and discard. The mitochondrial pellet should be firm, caramel-brown, and 0.6 to 0.7 g in weight (typical yield from a 1-L culture). The pellet may be frozen in liquid nitrogen and stored at –70°C for periods in excess of 18 mo.

3.2. Preparation of Mitochondrial Total RNA and tRNA Fractions

The following procedures are for the preparation of mitochondrial nucleic acids, and subsequent manipulations to yield both total RNA and tRNA fractions. Volumes are small enough that most centrifugations may be carried out in microfuges. Wear disposable gloves throughout all procedures; they are of critical importance in reducing the chance of RNase contamination.

3.2.1. Total Mitochondrial Nucleic Acids

1. Resuspend the crude mitochondrial pellet from **Subheading 3.1.3., step 10** in 2 mL of chilled TE.
2. Lyse the mitochondria by sequential addition of sodium dodecyl sulfate to 1%, NaOAc (pH 5.2) to 0.3 *M*, and an equal volume of phenol-cresol mix (*see Note 5*).
3. Vortex the lysis mix briefly, place on ice for 5 min, and centrifuge at 10,000*g* for 10 min (4°C). Since RNA is the final goal, possible shearing of DNA by vortexing is not a concern here.
4. Transfer the aqueous phase to clean tubes, add NaOAc (pH 5.2) to 0.3 *M*, and reextract with an equal volume of phenol-cresol. If material is still visible at the interface, additional phenol-cresol extractions should be carried out until the interface is clear.
5. Precipitate the nucleic acids from the aqueous phase with 2.5 vol of 100% ethanol (1 h; –70 or –20°C).
6. Centrifuge at 12,000*g* for 15 min.
7. Wash the final pellet once in 70–80% ethanol, dry, and redissolve in 200 µL of nuclease-free dH₂O.

3.2.2. DNase I Treatment to Yield Total RNA Fraction

Because of the acidic condition of the phenolic extractions in **Subheading 3.2.1., steps 3 and 4**, the RNA preparation is minimally contaminated with DNA.

1. In a total reaction volume of 300 μL , mix 180 μL of mitochondrial nucleic acids from **Subheading 3.2.1., step 7** with 150–200 U of RNase-free DNase I.
2. Incubate the DNase reaction at 37°C for 30 min.
3. Add NaOAc (pH 5.2) to 0.3 *M* and extract once with an equal volume of phenol-cresol.
4. Recover the RNA by two rounds of ethanol precipitation: to the aqueous phase add 2.5 vol of ethanol (1 h, –70 or –20°C). On addition of the ethanol, the precipitating RNA will turn the solution milky. Two rounds of precipitation are required in order to ensure complete removal of the phenol.
5. Centrifuge for 15 min at 12,000*g* (4°C).
6. Decant the alcohol, redissolve the pellet in 50 μL of nuclease-free dH_2O , and add 125 μL of 100% ethanol (1 h, –70 or –20°C).
7. Centrifuge at 12,000*g* for 15 min.
8. Wash the final pellet with 70–80% ethanol and dry. Do not overdry the RNA pellet because this will substantially reduce its subsequent solubility.
9. Redissolve the RNA in 100 μL of nuclease-free dH_2O ; the concentration should be approx 5 $\mu\text{g}/\mu\text{L}$. Estimated yield is 300–500 μg of total RNA per 0.4–0.6 g of sucrose gradient-purified mitochondria.
10. Assess the purity of the RNA preparation (*see Note 6*) by standard methods (25): agarose or polyacrylamide gel electrophoresis (PAGE), or ultraviolet absorbance spectrum (200–300 nm).

3.2.3. Salt Fractionation to Recover tRNA Fraction

1. To the mitochondrial RNA from **Subheading 3.2.2., step 9**, add NaCl to a concentration of 1.2 *M*. For salt precipitation to work effectively, the concentration must be minimally 4 to 5 $\mu\text{g}/\mu\text{L}$.
2. Store the RNA-salt solution overnight at 4°C (or not less than 8 h).
3. Centrifuge for 10 min at 12,000*g* (4°C).
4. Draw off the supernatant; this constitutes the bulk of the salt-soluble (mostly transfer) RNAs (=“sRNA”). Store at –70 or –20°C.
5. Resuspend the pellet in 20 μL of nuclease-free dH_2O and repeat fractionation (solution adjusted to 1.2 *M* NaCl, 4°C overnight or not less than 8 h).
6. Once again, centrifuge the RNA-salt mixture at 10,000*g* for 10 min. The pellet constitutes the salt-insoluble (mostly ribosomal and messenger) RNAs (=“iRNA”).
7. Combine the supernatants from **steps 4 and 6** and precipitate sRNA with 2.5 vol of ethanol (1 h, –70°C).
8. Centrifuge for 15 min at 12,000*g* (4°C). Wash the pellet with 80% ethanol, dry, and dissolve in 200 μL of nuclease-free dH_2O .

9. Measure the absorbance at 260 nm and calculate the concentration, assuming that 40 $\mu\text{g/mL}$ of RNA solution gives an A_{260} of 1.0.
10. Add 0.1 vol of 3 M NaOAc (pH 5.2) and extract once with phenol-cresol.
11. To the aqueous phase add 2.5 vol of ethanol (1 h, -70°C), centrifuge for 15 min at 12,000g (4°C).
12. Decant the ethanol, resuspend the pellet in 50 μL of nuclease-free dH_2O , and add 125 μL of 100% ethanol; store for 1 h at -70 or -20°C .
13. Centrifuge for 15 min at 12,000g (4°C).
14. Wash the final pellet once in 70–80% ethanol and dry.
15. Store the sRNA at -70°C at a concentration of 1 $\mu\text{g}/\mu\text{L}$ (see **Note 7**).

3.3. Circularization and Generation of tRNA Clones

This section outlines the conditions required for intra-molecular ligation of individual tRNAs to yield circular products that serve as substrates for cDNA synthesis. The cDNA is then subjected to PCR and the final product visualized. For a schematic of the methods in this section, refer to **Fig. 2B**.

3.3.1. Selection of Primers

Two primers appropriately flanking the acceptor stem are required for each tRNA analyzed (see **Fig. 3**). The primers should be minimally 18 nt long and situated within the regions indicated by the dotted line and shaded circles in **Fig. 3**.

3.3.2. Ligation Conditions for tRNA Circularization

Ligation conditions are a combination of those found in **refs. 1** and **10**, with current adjustments.

1. Denature 10 μg of sRNA (see **Subheading 3.2.3., step 12**) or 40 μg of total RNA (see **Subheading 3.2.1., step 5**) by heating to 90°C for 5 min. Cool on ice and centrifuge briefly (see **Note 8**).
2. Include several controls during the first use of this protocol:
 - a. RNase-treated DNA.
 - b. DNA no RNase treatment.
 - c. Unligated RNA.Both (a) and (b) should be treated under the same conditions as the sRNA in **step 1**, and (c) should be treated as in **Subheading 3.3.3.** and following.
3. The ligation conditions for a 20- μL total reaction volume are 50 mM HEPES (pH 7.5); 15 mM MgCl_2 ; 10% DMSO; 0.1 mg/mL of nuclease-free commercial BSA; 3.3 mM DTT; 100 μM ATP; 15 U of T4 RNA ligase; and the denatured, cooled RNA. Allow the ligation to proceed for 8–10 h at 37°C .
4. Subsequent to ligation, dilute the mix to 50 μL , add 0.1 vol of 3 M NaOAc (pH 5.2), and extract once with an equal volume of phenol-cresol.

5. The circularized tRNAs are precipitated twice. Add 0.1 vol of 3 M NaOAc (pH 5.2) and 2.5 vol of 100% ethanol (-70°C for 1 h). Centrifuge for 15 min at 12,000g (4°C).
6. Decant the ethanol, redissolve the pellet in 50 μL of nuclease-free dH_2O , and add 125 μL of 100% ethanol; store for 1 h at -70 or -20°C .
7. Centrifuge at 12,000g for 15 min.
8. Resuspend the pellet in 10 μL of nuclease-free dH_2O ; the final concentration should be approx 1 $\mu\text{g}/\mu\text{L}$ if sRNA is used, or approx 4 $\mu\text{g}/\mu\text{L}$ if total RNA is the substrate.
9. Store the circularized RNA solution at -70°C .

3.3.3. cDNA Synthesis From Circularized Templates

1. In a total volume of 21.5 μL , combine 1 μL of circularized RNA solution with 1 pmol of 5'-RT primer with dH_2O added to bring the mix up to the required 21.5 μL .
2. Heat the tRNA-primer mix to 90°C for 3 min, allow to cool to room temperature over a 15- to 20-min period, and finally store on ice for a further 15 min.
3. To the cooled primer-tRNA mix, add 0.1 vol of AMV buffer, 1.8 μL of 1 mM deoxynucleotide triphosphate (dNTP) mix (60 μM final concentration each of dATP, dCTP, dGTP, and dTTP), 15 U of AMV RT, and dH_2O , to a final reaction volume of 30 μL (see **Note 9**).
4. Allow cDNA synthesis to proceed at 45°C for 45 min.
5. Store the resultant cDNA at -20°C .

3.3.4. Amplification of Transcribed Circularized tRNAs

1. For a 50- μL total volume, combine 1.0 μL of cDNA, 10 pmol each of the 5'-RT and 3'-forward tRNA-specific primers, 0.1 vol of PCR buffer, 0.2 mg/mL of commercial BSA or gelatin, 2.5 U of *Taq* DNA polymerase, and 250 μM each of the four dNTPs.
2. Cycle parameters are denaturation (94°C , 3 min), followed by 30–35 cycles of denaturation (94°C , 40 s), annealing ($50/55^{\circ}\text{C}$, 40 s), extension (72°C , 60 s), and a final 12-min extension step at 72°C . To prevent aberrant annealing or primer extension during the initial heating to 94°C , preheat the block to 94°C prior to introducing the reaction tubes.

3.3.5. Visualization, Cloning, and Sequencing of Reverse-Transcribed Circularized tRNAs

PCR products from **step 2** in **Subheading 3.3.4.** may be visualized using standard agarose or PAGE techniques (25). Cloning of the PCR products is recommended for sequencing by automated means and in order that multiple clones may be generated to quantify editing. Invitrogen's TOPO TA CloningTM technology is one of the quicker, more reliable, and more efficient means of cloning

the PCR products for analysis. Several clones from each circularized tRNA should be analyzed. Sequencing may be carried out with either the specific tRNA primers or preferably vector primers if TOPO TA Cloning is utilized.

3.4. Analysis

The method outlined in **Subheadings 3.3.1–3.3.5.** allows the identification of both 5'- and 3'-terminal editing events in tRNAs. Depending on the original positioning of the primers, sequences outside of the acceptor stem may also be included in the analysis (**Fig. 2B**).

Special attention should be paid to the presence of a 3'-CCA_{OH} tail when comparing the sequences of the "corrected" cDNA and the corresponding genomic copy. The presence of 3'-CCA_{OH} signifies functionality, since most mitochondrial tRNAs require posttranscriptional addition of this motif. Absence of the expected CCA sequence may be indicative of an editing intermediate. The sequencing of multiple clones from the same circularization reaction may allow identification of partially edited intermediates. In the case of *Acanthamoeba*, no editing intermediates—i.e., sequences lacking the CCA tail or incompletely edited—were found for any of the 15 bona fide tRNAs analyzed (**1**). Refer to **Fig. 4** for a detailed depiction of the results for 5' editing of *Acanthamoeba* mitochondrial tRNA^{Phe}.

Analysis of the characteristics of the editing besides its position on the 5' or 3' side of the acceptor stem can provide insights into the possible mechanism. A surprising finding in the case of the mitochondrial 5' editing described in *Acanthamoeba* was that wobble pairs G·U and U·G are edited to A:U and C:G canonical pairings, respectively (**1**), emphasizing the templated nature of the nucleotide replacement. In the case of centipede mitochondria, the 3' editing activity replaces the U·G pair in tRNA^{Gln} with a Watson-Crick U:A pair, leading us to conclude that this editing, too, is template dependent, using the 5' side of the acceptor stem as a guide.

4. Notes

1. *Acanthamoeba* cultures are easily contaminated so a strict aseptic technique must be observed in all additions to the basal medium. A healthy culture will have a slightly granular appearance. Individual *Acanthamoeba* cells are readily identified by conventional light microscopy (×40 magnification). A healthy cell will have slender hairlike projections (acanthopodia), prominent vacuoles, and a prominent nucleolus. No staining is required to visualize any of these features of a growing cell. The scaling up or down of culture volumes is easily achieved with the proviso that the medium should be no deeper than 4–5 cm in the vessel used, to permit optimal air exchange during growth.

4. If the mitochondrial isolation protocol is used with a different organism and where sucrose gradient conditions are not optimized or mitochondrial yields are low (<0.2 g), the crude mitochondrial pellet may be used for RNA isolation. The pellet may be stored at -70°C if it is first resuspended in 2 mL of postsucrose/storage buffer. While keeping the suspension on ice, slowly add a further 4 mL of buffer over a 10-min period (1 mL every 2 min), and centrifuge for 30 min at $12,000g$ (4°C). The pellet may then be snap frozen in liquid nitrogen and stored for lengthy periods at -70°C .
5. The phenol-cresol mix used for phenolic extraction is based on protocols outlined in **refs. 26** and **27**. The 8-hydroxyquinoline acts as both an antioxidant and a mild RNase inhibitor. Addition of the salt (0.1 vol of 3 *M* NaOAc [pH 5.2]) to the mitochondrial nucleic acid mix prior to phenol extraction acidifies the phenolic lysate. Under acidic conditions, the DNA will collect at the organic-aqueous interface, whereas the RNA will remain in the aqueous phase. An alternate method for nucleic acid extraction involves the use of TRIzol Reagent™, a phenol/guanidine isothiocyanate mix (Gibco-BRL, Gaithersburg, MD).
6. To assess RNA quality, using standard methodology, a small amount of the RNA may be electrophoresed on a 3 to 4% agarose or 10% polyacrylamide gel (**25**). Intact RNA will band reproducibly; intense bands representative of the rRNAs with minimal background smearing are a good sign that the RNA is undegraded. tRNA and other small RNA species will appear as an amorphous blob at the leading edge of the gel. Lack of definition and excessive smearing are usually indicative of degradation of the RNA sample. As an additional quality control, an ultraviolet absorbance spectrum of the RNA should be determined over a range of 200–300 nm. A dilution of 1:200 or 1:500 should be sufficient; a pure sample should generate the characteristic skewed bell-shaped curve, with an ideal A_{260}/A_{280} ratio of approx 2 for pure RNA (**25**).
7. If the RNA is to be used immediately it should be redissolved in dH_2O and stored at -70°C . For longer-term storage, the RNA should be maintained in 70% ethanol.
8. Where salt fractionation is not carried out, total RNA may be substituted for tRNA (sRNA; *see Subheading 3.2.3.*). Because only 10–20% of the total RNA comprises tRNA, 40 μg of total RNA should be substituted for the recommended 10 μg of sRNA, in order to ensure that there is sufficient target in the circularization reaction.
9. The advantage of AMV over other commonly used RTs (e.g., Moloney murine leukemia virus) is that its optimum temperature is 42°C , which helps to reduce possible tRNA secondary structure problems.

Acknowledgments

This work was supported by grant MOP-4124 from the Canadian Institutes of Health Research. We also gratefully acknowledge salary support from the Canada Research Chairs Program and the Canadian Institute for Advanced Research (Program in Evolutionary Biology).

References

- 1 Price, D. H. and Gray, M. W. (1999) Confirmation of predicted edits and demonstration of unpredicted edits in *Acanthamoeba castellanii* mitochondrial tRNAs. *Curr. Genet.* **35**, 23–29.
- 2 Price, D. H. and Gray, M. W. (1998) Editing of tRNA, in *Modification and Editing of RNA* (Grosjean, H. and Benne, R., eds.), ASM Press, Washington, DC, pp. 289–305.
- 3 Gray, M. W. (2001) Speculations on the origin and evolution of RNA editing, in *RNA Editing* (Bass, B. L., ed.), Oxford University Press, Oxford, UK, pp. 160–184.
- 4 Hopper, A. K. and Phizicky, E. M. (2003) tRNA transfers to the limelight. *Genes Dev.* **17**, 162–180.
- 5 Fey, J., Weil, J. H., Tomita, K., Cosset, A., Dietrich, A., Small, I., and Maréchal-Drouard, L. (2002) Role of editing in plant mitochondrial transfer RNAs. *Gene* **286**, 21–24.
- 6 Janke, A. and Pääbo, S. (1993) Editing of a tRNA anticodon in marsupial mitochondria changes its codon recognition. *Nucleic Acids Res.* **21**, 1523–1525.
- 7 Miller, D., Mahendran, R., Spottswood, M., Costandy, H., Wang, S., Ling, M.-L., and Yang, N. (1993) Insertional editing in mitochondria of *Physarum*. *Semin. Cell Biol.* **4**, 261–266.
- 8 Antes, T., Costandy, H., Mahendran, R., Spottswood, M., and Miller, D. (1998) Insertional editing of mitochondrial tRNAs of *Physarum polycephalum* and *Didymium nigripes*. *Mol. Cell. Biol.* **18**, 7521–7527.
- 9 Yokobori, S.-I. and Pääbo, S. (1995) tRNA editing in metazoans. *Nature* **377**, 490.
- 10 Yokobori, S.-I. and Pääbo, S. (1995) Transfer RNA editing in land snail mitochondria. *Proc. Natl Acad. Sci. USA* **92**, 10,432–10,435.
- 11 Tomita, K., Ueda, T., and Watanabe, K. (1996) RNA editing in the acceptor stem of squid mitochondrial tRNA^{Tyr}. *Nucleic Acids Res.* **24**, 4987–4991.
- 12 Yokobori, S.-I. and Pääbo, S. (1997) Polyadenylation creates the discriminator nucleotide of chicken mitochondrial tRNA^{Tyr}. *J. Mol. Biol.* **265**, 95–99.
- 13 Hatzoglou, E., Rodakis, G. C., and Lecanidou, R. (1995) Complete sequence and gene organization of the mitochondrial genome of the land snail *Albinaria coerulea*. *Genetics* **140**, 1353–1366.
- 14 Yamazaki, N., Ueshima, R., Terrett, J. A., et al. (1997) Evolution of pulmonate gastropod mitochondrial genomes: comparisons of gene organizations of Euhadra, Cepaea and Albinaria and implications of unusual tRNA secondary structures. *Genetics* **145**, 749–758.
- 15 Lonergan, K. M. and Gray, M. W. (1993) Editing of transfer RNAs in *Acanthamoeba castellanii*. *Science* **259**, 812–816.
- 16 Lonergan, K. M. and Gray, M. W. (1993) Predicted editing of additional transfer RNAs in *Acanthamoeba castellanii* mitochondria. *Nucleic Acids Res.* **21**, 4402.
- 17 Cole, R. A. and Williams, K. L. (1994) The *Dictyostelium discoideum* mitochondrial genome: a primordial system using the universal code and encoding hydrophilic proteins atypical of metazoan mitochondrial DNA. *J. Mol. Evol.* **39**, 579–588.

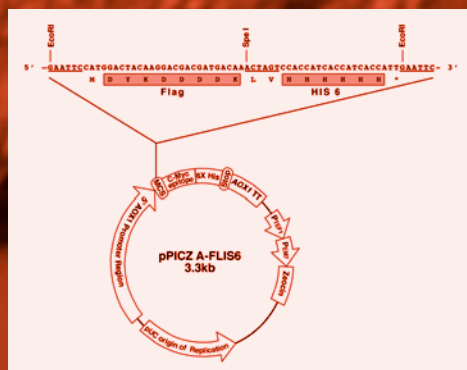
- 18 Angata, K., Kuroe, K., Yanagisawa, K., and Tanaka, Y. (1995) Codon usage, genetic code and phylogeny of *Dictyostelium discoideum* mitochondrial DNA as deduced from a 7.3-kb region. *Curr. Genet.* **27**, 249–256.
- 19 Laforest, M.-J., Roewer, I., and Lang, B. F. (1997) Mitochondrial tRNAs in the lower fungus *Spizellomyces punctatus*: tRNA editing and UAG “stop” codons recognized as leucine. *Nucleic Acids Res.* **25**, 626–632.
- 20 Forget, L., Ustinova, J., Wang, Z., Huss, V. A. R., and Lang, B. F. (2002) *Hyaloraphidium curvatum*: a linear mitochondrial genome, tRNA editing, and an evolutionary link to lower fungi. *Mol. Biol. Evol.* **19**, 310–319.
- 21 Lavrov, D. V., Brown, W. M., and Boore, J. L. (2000) A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus*. *Proc. Natl. Acad. Sci. USA* **97**, 13,738–13,742.
- 22 Burger, G., Plante, I., Lonergan, K. M., and Gray, M. W. (1995) The mitochondrial DNA of the amoeboid protozoon, *Acanthamoeba castellanii*: complete sequence, gene content and genome organization. *J. Mol. Biol.* **245**, 522–537.
- 23 Price, D. H. and Gray, M. W. (1999) A novel nucleotide incorporation activity implicated in the editing of mitochondrial transfer RNAs in *Acanthamoeba castellanii*. *RNA* **5**, 302–317.
- 24 Sugino, A., Snopek, T. J., and Cozzarelli, N. R. (1977) Bacteriophage T4 RNA ligase: reaction intermediates and interaction of substrates. *J. Biol. Chem.* **252**, 1732–1737.
- 25 Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- 26 Spencer, D. F., Gray, M. W., and Schnare, M. N. (1992) The isolation of wheat mitochondrial DNA and RNA, in *Seed Analysis*, vol. 14 in *Modern Methods of Plant Analysis, New Series* (Linskens, H. F. and Jackson, J. F., eds.), Springer-Verlag, Berlin, pp. 347–360.
- 27 Parish, J. H. and Kirby, K. S. (1966) Reagents which reduce interactions between ribosomal RNA and rapidly labelled RNA from rat liver. *Biochim. Biophys. Acta* **129**, 554–562.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



In Vitro RNA Editing Systems From Higher Plant Chloroplasts

Tetsuro Hirose, Tetsuya Miyamoto, Junichi Obokata,
and Masahiro Sugiura

Summary

RNA editing in higher plant chloroplasts involves C-to-U conversion at ~30 specific sites. In vitro systems supporting accurate editing have been developed from tobacco and pea chloroplasts. mRNA substrates labeled with ^{32}P at C residues to be edited provide sensitive detection of editing activity in vitro. The present systems allow the rapid identification of *cis*-elements using mutated mRNA substrates and *trans*-acting factors by ultraviolet crosslinking.

Key Words: Chloroplast; *cis*-acting element; in vitro system; pea; RNA editing; RNA ligation; site-specific labeling; tobacco; *trans*-acting factor.

1. Introduction

In chloroplasts of higher plants, the genetic information encoded in the chloroplast genome is altered by a posttranscriptional phenomenon termed *RNA editing*. RNA editing observed in higher plant chloroplasts is site-specific C $\rightarrow$ U conversions that usually restore codons for conserved amino acids including initiation and termination codons (1). It has now been accepted that this phenomenon commonly exists in chloroplasts of all major land plant phyla (2). In tobacco and maize chloroplasts, total numbers of RNA-editing sites have been identified to be 34 and 27 sites, respectively (3–6). However, no apparent conserved sequence has been found in the flanking sequences of editing sites. With respect to the mechanism of RNA editing in chloroplasts, three important questions have arisen: What is the catalytic mechanism? How does the editing machinery recognize the site to be edited? Is the site recognized by protein or

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

“guide RNA”? The in vitro RNA-editing system provides a powerful tool to address these questions biochemically (7,8).

2. Materials

2.1. Preparation of Chloroplast Extracts From Tobacco and Pea

1. Tobacco (*Nicotiana tabacum* var. Bright Yellow 4) leaves.
2. Pea (*Pisum sativum* var. Alaska) leaves.
3. 5X MCB1 buffer: 1.5 M mannitol, 250 mM HEPES-NaOH, pH 8.0, 10 mM EDTA, 25 mM β -mercaptoethanol.
4. 1X MCB1 buffer with 0.6% polyvinylpyrrolidone (PVP) and 0.1% bovine serum albumin (BSA).
5. 1X MCB1 buffer with 0.1% BSA.
6. 10 and 80% Percol (Amersham, Piscataway, NJ) in 1X MCB1.
7. 1X MCB2 buffer: 0.32 M mannitol; 50 mM HEPES-NaOH, pH 8.0, 2 mM EDTA.
8. Buffer Ex: 30 mM HEPES-KOH, pH 7.7, 10 mM Mg acetate, 2 M KCl, 0.2% Triton X-100, 2 mM dithiothreitol (DTT).
9. Buffer A: 30 mM HEPES-KOH, pH 7.7; 3 mM Mg acetate, 45 mM K acetate, 30 mM NH_4 acetate, 2 mM DTT, 10% glycerol.
10. Cotton gauze (four layers).
11. Polytoron (Kinematica).
12. Microdialyzer (Bio-Tech).
13. Filter cups (mol wt cutoff: 12,000; Bio-Tech).

2.2. In Vitro RNA Editing Reaction and Ultraviolet Crosslinking

1. Plasmid constructs containing the 5'-flanking sequence of the editing site.
2. Ribooligonucleotides corresponding to the editing site, its 3' flanking region (9 nt), and artificial 25-nt sequence (see **Note 1**).
3. Deoxyoligonucleotide (for bridging two RNAs).
4. T3 Megascript RNA synthesis kit (Ambion).
5. Phenol:chloroform:isoamyl alcohol (PCA) (25:24:1).
6. T4 polynucleotide kinase (TaKaRa).
7. T4 DNA ligase (TaKaRa).
8. 10X T4 DNA ligase buffer (TaKaRa).
9. Gel-loading buffer.
10. RNA elution buffer (Ambion).
11. 5X mineral mix: 75 mM HEPES-KOH, pH 7.7, 7.5 mM Mg acetate, 112.5 mM K acetate, 75 mM NH_4 acetate, 5% polyethylene glycol 6000.
12. 100 mM ATP (Boehringer Mannheim).
13. Proteinase inhibitor mixture (CompleteTM; Boehringer Mannheim).
14. RNase inhibitor (TaKaRa).
15. RNase-free water.
16. Nuclease P1 (Sigma, St. Louis, MO).
17. Cellulose thin-layer chromatography (TLC) plate (20 $\times$ 20 cm; Kodak).

18. Bioimaging analyzer BAS2000 (Fuji).
19. RNase A.
20. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer.

3. Methods

The following protocol comprises growth of tobacco and pea plants; preparation of intact chloroplasts, chloroplast extracts, and substrate RNAs, *in vitro* RNA editing reaction and TLC; and ultraviolet (UV) crosslinking to detect putative *trans*-acting factors.

3.1. Growth of Tobacco and Pea Plants

To obtain active chloroplast extracts, the plant stage to be harvested is extremely important (*see Note 2*). To induce germination, tobacco seeds (0.1–0.15 g; they should be dried and stored at 4°C) are put in water at room temperature for 1 d. Seeds are sown on soil in three plastic trays (50 × 30 cm) and wrapped with Saran Wrap until >80% of the seeds are germinated. Plants are grown in a growth chamber (16 h light/8 h dark, 28°C, 8000 lux) until harvest. After approx 6 wk, the stage III leaves (about 10 cm) are harvested. Pea plants are grown for 2 wk under the same conditions. Pea chloroplasts are prepared from young leaves from the upper part of plants.

3.2. Preparation of Intact Chloroplasts

Intact chloroplasts are prepared according to the standard method (9). The same procedure is applied for chloroplast preparation from tobacco and pea leaves.

1. Cut 100–150 g of tobacco leaves (stage III) into about 1 × 1 cm pieces.
2. Collect half of the leaf pieces in a glass tube (diameter: 6.5 cm; depth: 24 cm), add 300 mL of MCB1 + 0.6% PVP + 0.1% BSA, and homogenize by Polytron for 10 s. Repeat with another half of the leaf pieces.
3. Filter the homogenate through four layers of cotton gauze in a cold room.
4. Divide the filtered homogenate into four plastic centrifugation tubes (6 × 15 cm) and centrifuge for 1 s at 1500g with a BH-9 rotor (Tomy).
5. Gently suspend the resultant pellets in MCB1 + 0.1% BSA (3 mL/tube).
6. Overlay each 3-mL chloroplast suspension on a 10–80% Percol gradient (34 mL each in four plastic centrifugation tubes [2.5 × 8.8 cm]).
7. Centrifuge at 6000g for 30 min with a TS-2 rotor with the brake off (Tomy).
8. Two green bands appear. Carefully collect the intact chloroplasts (the lower band) using a glass pipet (*see Note 3*).
9. Wash the chloroplasts twice with 10–20 mL of MCB2 by centrifuging at 1000g for 1 min (40 s for the second washing).

3.3. Preparation of Chloroplast Extracts

Isolated chloroplasts are lysed by nonionic detergent (0.2% Triton X-100) and extracted with salt with a high concentration (2 M KCl). The S60 fraction is prepared by ultracentrifugation.

1. Gently suspend the chloroplast pellets in 300–400 μ L of cold buffer Ex immediately after the second washing.
2. Put a tube on ice for 15 min and occasionally gently shake the tube.
3. Transfer the suspension to an ultracentrifuge open-top tube (11 $\times$ 34 mm).
4. Ultracentrifuge at 13,000 rpm (6000g) at 2°C for 10 min (TLA-100.2).
5. Recover the supernatant in a new ultracentrifuge tube.
6. Ultracentrifuge at 42,000 rpm (60,000g) at 2°C for 120 min (*see Note 4*).
7. Recover the supernatant (S60) in an Eppendorf tube.
8. Transfer the S60 fraction (about 300 μ L) into a Microdialyzer cup (up to 150 μ L in one tube; Bio-Tech).
9. Dialyze against 300 mL of buffer A using a microdialyzer for 3.5 h in a cold room. Exchange the buffer and dialyze further for 3.5 h.
10. Divide the dialyzed S60 fraction into 52- μ L aliquots and freeze in liquid nitrogen. Store the aliquots at -70°C . Extracts are stable for at least for 6 mo.

3.4. Preparation of Substrate RNAs

The substrate is an RNA molecule that is labeled with ^{32}P at the 5' phosphate of the specific cytidine to be edited. This labeled RNA is prepared from two RNA pieces by an enzymatic ligation method (**Fig. 1**; [10]). The upstream RNA is synthesized by in vitro transcription with T3 RNA polymerase. Because it is difficult to obtain substantial amounts of transcripts with the cytidine at the 5' end by in vitro transcription with T7, T3, or SP6 RNA polymerase, the downstream RNA containing the editing site is chemically synthesized (*see Note 1*). Two RNA pieces are ligated with T4 DNA ligase in the presence of the bridge deoxyoligonucleotide.

3.4.1. In Vitro Transcription of Upstream RNAs

1. Linearize the plasmid clone containing the upstream part with an appropriate restriction enzyme for the template of in vitro transcription. The transcript produced from this template should end at a nucleotide upstream of the editing site (**Fig. 1**). PCA extract the linearized plasmid followed by ethanol precipitation.
2. Synthesize the upstream RNA by using a Megascript T3 RNA synthesis kit according to the manufacturer's protocol.
3. Purify the upstream RNA by passing through a Sephadex G50 spin column. If necessary, is further purify by PAGE.

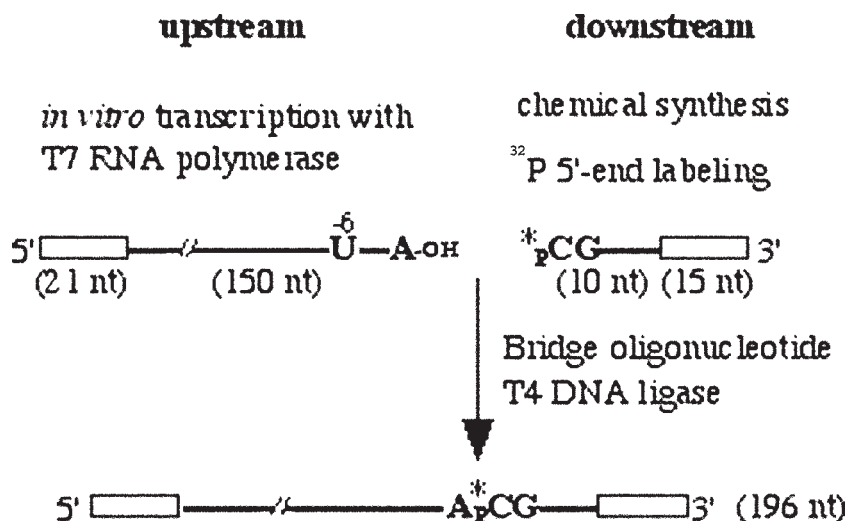


Fig. 1. Synthesis of the substrate RNA specifically labeled at the editing site. The upstream RNA (150 nt preceding the editing site of tobacco *psbL* mRNA with a 5' extension of a 21-nt sequence from pBluescript II) and the downstream RNA (10 nt downstream from the editing site in *psbL* mRNA with the 3' extension of a 15-nt sequence from the KS primer) are ligated with T4 DNA ligase in the presence of the bridge deoxyoligonucleotide. Boxes represent extensions.

3.4.2. Preparation of Downstream Ribooligonucleotide

For this protocol, a synthetic ribooligonucleotide will need to be ordered from an appropriate chemical company (see **Note 1**).

1. Dissolve the ribooligonucleotide in RNase-free ddH₂O and determine the concentration by measuring A_{260} . It can be stored at -20°C . Deprotection of 2' groups might be necessary before dissolving.
2. Label the 5' end of the ribooligonucleotide with $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ and polynucleotide kinase in a 50- μL reaction mixture containing 100 pmol of the ribooligonucleotide, 5 μL of 10X buffer, 10 μL of $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ (185 TBq/mmol), and 2 μL (20 U) of T4 polynucleotide kinase.
3. Incubate for 30 min at 37°C , followed by inactivation of polynucleotide kinase with incubation for 10 min at 65°C .
4. Extract with PCA, and remove unreacted mononucleotides by passing through a Sephadex G25 spin column.
5. Precipitate the labeled ribooligonucleotide with ethanol and dissolve the precipitate in RNase-free water.

3.4.3. RNA Ligation

1. Synthesize the bridge deoxyoligonucleotide possessing at least 20mers complementary to each upstream and downstream of the ligation site.
2. Dissolve the deoxyoligonucleotide in RNase-free water up to a concentration of 200 μM . It can be stored at -20°C .
3. Ligate two RNA pieces with T4 DNA ligase in the presence of the bridge deoxyoligonucleotide. Mix 300 pmol of the in vitro-transcribed upstream RNA, 100 pmol of the ^{32}P -labeled downstream ribooligonucleotide, 200 pmol of the bridge deoxynucleotide, and 5 μL of 10X T4 DNA ligase buffer, and add RNase-free water up to 49.8 μL .
4. Heat the tube at 94°C for 1 min and gradually cool down to room temperature on the bench.
5. After about 2 h (until the heat block is cooled to room temperature), briefly spin the tube and incubate at 25°C for 5 min.
6. Add 0.2 μL of T4 DNA ligase (final 1 U/ μL), and incubate at 25°C overnight.
7. Precipitate the ligated RNA with ethanol. Dissolve the precipitate in 20 μL of the loading buffer.
8. Take 0.5 μL of ligated RNA to measure the radioactivity in a scintillation counter.
9. Denature the ligated RNA at 94°C for 3 min and immediately cool down. Then apply to a 5% PAGE containing 7 M urea (10×10 cm gel).
10. After the ligated RNA is separated properly, stop the electrophoresis and expose the gel to X-ray film for 3 min.
11. Cut the band corresponding to the ligated RNA, and immerse the gel piece in elution buffer at 37°C overnight.
12. Precipitate the RNA with ethanol and dissolve the precipitate in RNase-free water. Calculate the amount of yielded substrate RNA by comparing the radioactivity of the eluted RNA relative to that of the total downstream ribooligonucleotide applied (measured in **step 8**).
13. Adjust the RNA concentration to 50 fmol/ μL with RNase-free water. It can be stored at -20°C up to 1 wk.

3.5. In Vitro RNA-Editing Reaction

The in vitro RNA-editing reaction is carried out using a site-specifically labeled RNA in the chloroplast S60 extract. RNA editing is monitored by observing the alteration in the labeled cytidine to uridine on TLC of 5'-mononucleotides produced by digestion of RNAs with nuclease P1 (*see Fig. 2*).

3.5.1 In Vitro RNA-Editing Reaction

1. Prepare the reaction mixture containing 5 μL of 5X mineral mix, 1 μL of 75 mM ATP, 2 μL of RNase inhibitor (130 U/ μL), 0.5 μL of Complete (50X concentration: 1 tablet/mL), and 3 μL of RNase-free water in a 1.5-mL tube (*see Note 5*). Set a negative control without adding the chloroplast extract.
2. Add 1 μL of RNA substrate (50 fmol) (*see Note 6*).

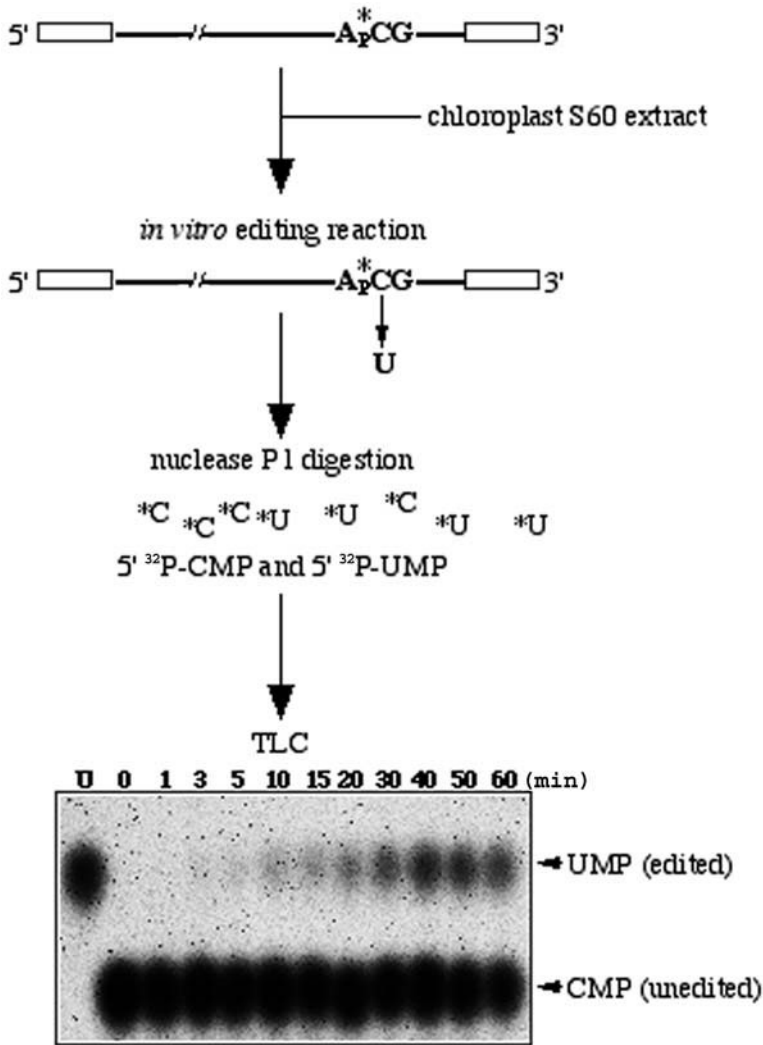


Fig. 2. Schematic diagram of *in vitro* RNA editing reaction. *psbL* mRNA specifically labeled at the editing site (C) is incubated with a chloroplast extract for 0–60 min. RNAs are digested to 5′-mononucleotides containing ³²P-CMP and ³²P-UMP with nuclease P1 and separated by cellulose TLC.

3. Add 12.5 μL of chloroplast S60 extract.
4. Incubate at 28°C for 40 min (*see Note 7*).
5. Extract with PCA, and pass the aqueous phase through a Sephadex G50 spin column.
6. Precipitate the RNA with ethanol.

7. For nuclease P1 digestion, dissolve the RNA in 8.5 μL of RNase-free water and mix with 0.5 μL of 3 M NH_4Ac and 1 μL of nuclease P1 (1 U/ μL).
8. Incubate at 37°C for 3 h.

3.5.2. Thin-Layer Chromatography

1. Spot about 500 cpm of digested RNA on a Cellulose TLC plate and dry. Spot ^{32}P -UMP (nuclease P1 digestion products of appropriate [α - ^{32}P] UTP-labeled transcripts) as a control.
2. Develop with the solvent containing isopropanol, HCl, and water (70:15:15) at room temperature overnight (or for 8 h) in a TLC glass box.
3. Dry the plate and expose it to an imaging plate overnight.
4. Visualize separated spots by using a bioimaging analyzer BAS2000 and quantify the ratio between the intensity of 5'CMP and 5'UMP spots (see **Fig. 2** and **Note 8**).

3.6. UV Crosslinking

The in vitro RNA-editing system can be utilized to detect *trans*-acting factors. UV crosslinking with the RNA containing a site-specific labeling is a powerful tool to detect protein(s) that interact with the specific site during the in vitro RNA-editing reaction (7,8) (see **Note 9**).

1. Incubate the site specifically labeled RNA (50 fmol) under in vitro editing reaction conditions as described in **Subheading 3.5.1**. at 28°C for 15 min.
2. Put the reaction tube on ice with the cap opened.
3. UV irradiate (254 nm) in a Funacrosslinker at a setting of 0.9 J/cm².
4. Repeat UV irradiation with the same setting.
5. Add RNase A (50 $\mu\text{g}/\text{mL}$ final) and incubate at 37°C for 15 min.
6. Precipitate the proteins with cold acetone (50% final) and dissolve in the SDS-PAGE sample buffer.
7. Heat the protein samples to 95°C for 5 min and apply to 12 or 15% PAGE with 0.1% SDS.
8. Dry the gel and expose to an imaging plate.
9. Visualize the separated ^{32}P -labeled proteins using a bioimaging analyzer BAS2000 (**Fig. 5**).

4. Notes

1. For accurate and efficient RNA-editing reactions, the downstream sequence of the editing site does not need to be extended further, at least in the case of *psbL*, *ndhB*, *petB*, and *psbF* mRNAs. Only nine nucleotides after the editing site are sufficient for the editing. The following 25 nt of an artificial sequence is for annealing of the primer for confirming sequences of reverse transcriptase polymerase chain reaction (RT-PCR) products (**Fig. 3**).
2. Various biochemical activities in chloroplasts are altered during plant growth. To obtain active chloroplast extracts, the optimal leaf stage to be harvested should be

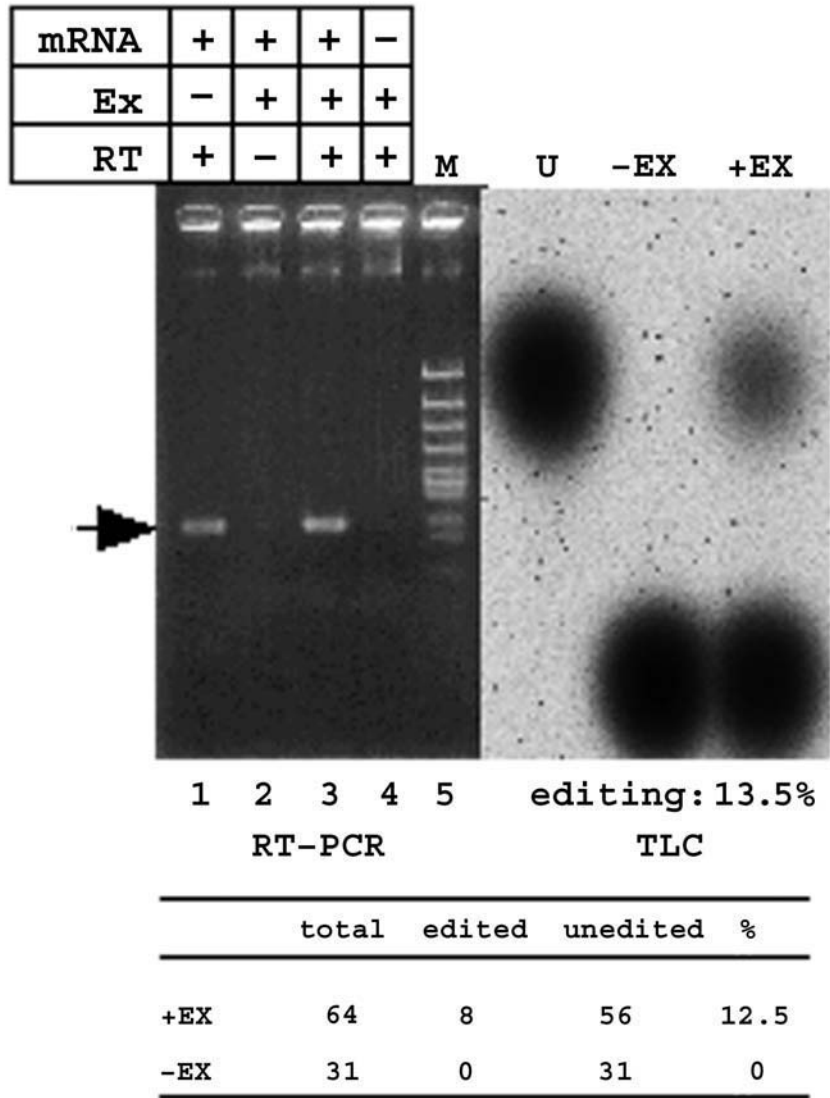


Fig. 3. Confirmation of in vitro RNA editing reaction by sequencing RT-PCR products. The *psbL* mRNA incubated with or without a chloroplast extract was subjected to RT-PCR. The primer corresponding to the 3' extension of 15 nt in Fig. 1 was used for the reverse transcription. The arrow indicates PCR products derived from mRNA substrate. The products were cloned and sequenced. The table shows the numbers of cDNA clones possessing T (edited: 12.5%) and C (unedited) at the editing site. The efficiency of the in vitro editing reaction (13.5% calculated from spots on TLC) is equivalent to the population of edited cDNA clones above. The sequencing result of the control experiment (incubation without extract) is also shown.

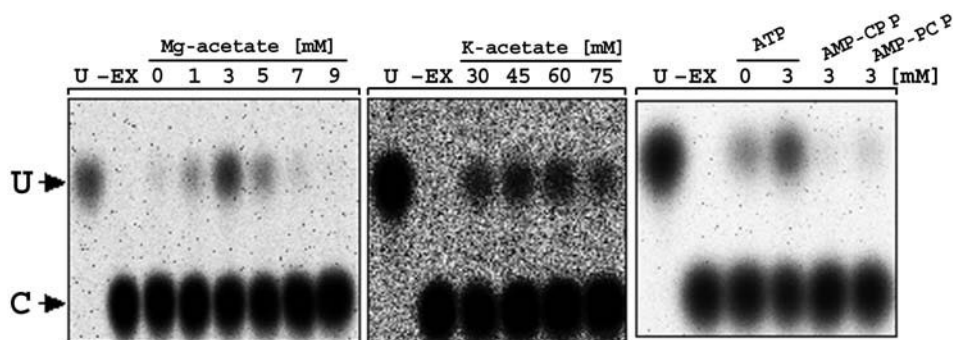


Fig. 4. Influence of concentrations of Mg^{2+} , K^+ , and ATP on in vitro editing reaction of *psbL* mRNA.

carefully determined at the beginning. Stage III (10 cm) with low RNase activity (*II*) needs to be determined according not only to leaf size but also the softness and color.

3. The ratio between the two bands depends on leaf conditions (e.g., starch contents) and leaf homogenization (*see Subheading 3.2.2.*). The amount of extraction buffer (buffer EX) added (*see Subheading 3.3.1.*) needs to be determined according to the yield of intact chloroplasts at this step.
4. The S60 fractionation is not a critical step. We found that rapid preparation of extracts generally yields better results; hence, the ultracentrifugation step can be omitted.
5. Among various conditions, Mg^{2+} concentration and ATP concentration are important factors for the in vitro editing reaction (**Fig. 4**). The conditions previously described were optimized to obtain the highest activity for the editing of tobacco chloroplast *psbL* mRNA. The optimal concentration of these compounds should be checked at the beginning for other mRNA substrates.
6. Site-specifically labeled RNAs of <1 fmol can be detected. To increase the sensitivity of RNA competition experiments, we recommend reducing the amount of RNA substrates.
7. The editing reaction of *petB* mRNA continues up to 2 h. It is possible that the half-life of editing activity in extracts differs in editing sites and/or extract conditions.
8. Chloroplasts contain an endonuclease activity that cleaves the 3' phosphodiester bond of pyrimidine nucleotides (our unpublished observation). If the editing site follows uridine (-NpU³²pCpN-), pU*p would be produced with nuclease P1 digestion. As pU*p comigrates with *pU, RNA-editing activity in vitro needs to be verified by sequencing RT-PCR products.
9. When no specific crosslinked protein is detected reproducibly, the position of the labeled nucleotide needs to be altered from the editing site (+1) to -6 or more

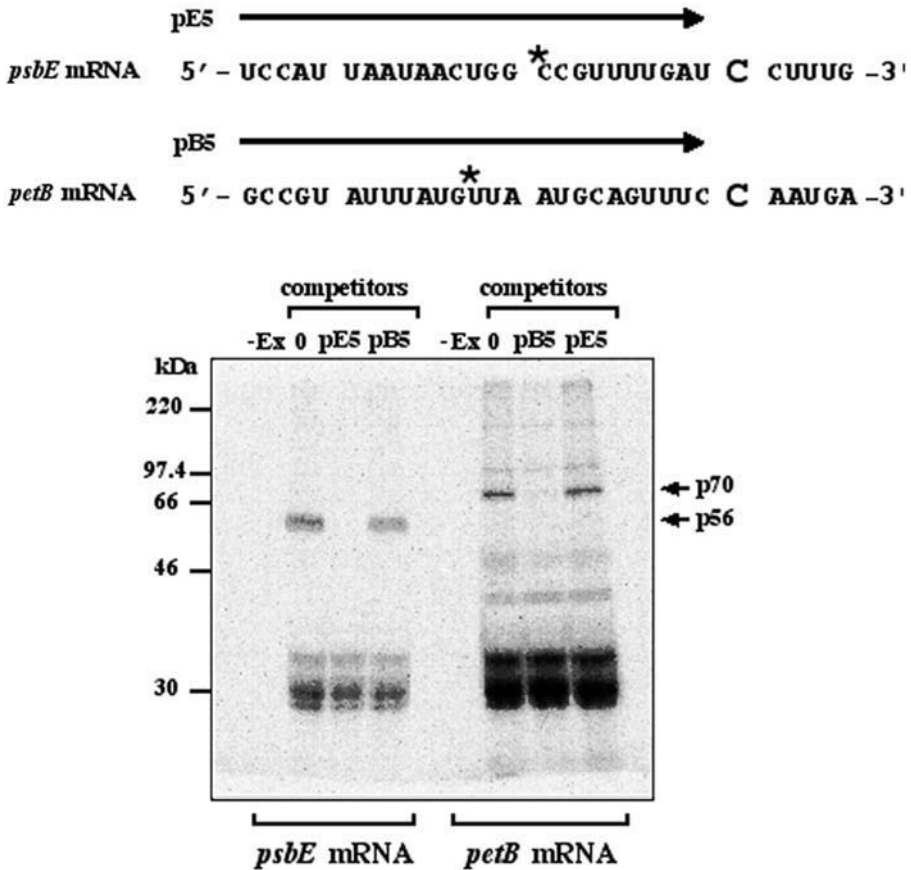


Fig. 5. Detection of tobacco chloroplast proteins bound to *cis*-acting elements for editing of *psbE* and *petB* mRNAs by UV crosslinking. The *psbE* and *petB* mRNAs with ^{32}P labeling at the middle (asterisks) of the *cis*-acting elements were used as substrates. Each mRNA of 10 fmol was incubated under the in vitro editing reaction followed by UV irradiation (254 nm). RNA was digested with RNase A and the crosslinked ^{32}P -labeled protein was separated by SDS-PAGE. Unlabeled RNA competitors (1 pmol) were added as indicated. p56 and p70 specifically bind to the *cis*-element of *psbE* mRNA and *petB* mRNA, respectively.

upstream positions. To investigate the binding specificity of crosslinked proteins, excess amounts of unlabeled competitor RNAs are added. As we have published (7,8), site-specific RNA-binding proteins interact with the upstream *cis*-acting elements of respective RNA-editing sites (Fig. 3).

Acknowledgment

This work was supported in part by a Grant-in-Aid for Scientific Research in Priority Areas C (no. 13201001) from the Ministry of Education, Science, Sports and Culture.

References

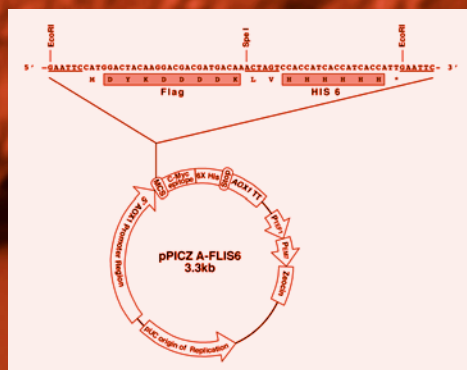
1. Bock, R. (2001) RNA editing in plant mitochondria and chloroplasts, in *RNA Editing* (Bass, B. L., ed.), Oxford University Press, Oxford, UK, pp. 38–60.
2. Freyer, R., Kiefer-Meyer, M.-C., and Kössel, H. (1997) Occurrence of plastid RNA editing in all major lineages of land plants. *Proc. Natl. Acad. Sci. USA* **94**, 6285–6290.
3. Hirose, T., Kusumegi, T., Tsudzuki, T., and Sugiura, M. (1999) RNA editing sites in tobacco chloroplast transcripts: editing as a possible regulator of chloroplast RNA polymerase activity. *Mol. Gen. Genet.* **262**, 462–467.
4. Sasaki, T., Yukawa, Y., Miyamoto, T., Obokata, J., and Sugiura, M. (2003) Identification of RNA editing sites in chloroplast transcripts from the maternal and paternal progenitors of tobacco (*Nicotiana tabacum*): comparative analysis shows the involvement of distinct *trans*-factors for *ndhB* editing. *Mol. Biol. Evol.* **20**, 1028–1035.
5. Maier, R. M., Neckermann, K., Igloi, G. L., and Kössel, H. (1995) Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. *J. Mol. Biol.* **251**, 614–628.
6. Bock, R., Albertazzi, F., Freyer, R., Fuchs, S., Ruf, S., Zeltz, P., and Maier, R. M. (1997) Transcript editing in chloroplasts of higher plants, in *Eukaryotism and Symbiosis* (Schenk, H. E. A., Herrman, R., Jeon, K. W., Müller, N. E., and Schwemmler, W., eds.), Springer, Berlin, pp. 123–137.
7. Hirose, T. and Sugiura, M. (2001) Involvement of a site-specific *trans*-acting factor and a common RNA-binding protein in the editing of chloroplast mRNAs: development of a chloroplast *in vitro* RNA editing system. *EMBO J.* **20**, 1144–1152.
8. Miyamoto, T., Obokata, J., and Sugiura, M. (2002) Recognition of RNA editing sites is directed by unique proteins in chloroplasts: biochemical identification of *cis*-acting elements and *trans*-acting factors involved in RNA editing in tobacco and pea chloroplasts. *Mol. Cell. Biol.* **22**, 6726–6734.
9. Bartlett, S. G., Grossman, A. R., and Chua, N.-H. (1982) *In vitro* synthesis and uptake of cytoplasmically-synthesized chloroplast proteins, in *Methods in Chloroplast Molecular Biology* (Edelman, M., Hallick, R. B., and Chua, N.-H., eds.), Elsevier Biomedical, Amsterdam, pp. 1081–1091.
10. Moore, M. J. and Sharp, P. A. (1992) Site-specific modification of pre-mRNA: the 2'-hydroxyl groups at the splice sites. *Science* **256**, 992–997.
11. Hirose, T. and Sugiura, M. (1996) *Cis*-acting elements and *trans*-acting factors for accurate translation of chloroplast *psbA* mRNAs: development of an *in vitro* translation system from tobacco chloroplasts. *EMBO J.* **15**, 1687–1695.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



Studying RNA Editing in Transgenic Chloroplasts of Higher Plants

Ralph Bock

Summary

RNA editing in plant cell organelles is a posttranscriptional process that changes the identity of individual nucleotides by pyrimidine transitions. C-to-U conversions are the predominantly found modifications in both mitochondrial and chloroplast transcripts of vascular plants. Research on functional, mechanistic, and evolutionary aspects of plant RNA editing has benefited immensely from the development of transgenic technologies for higher plant chloroplasts. This chapter describes methods for the generation of plants with stably transformed chloroplast genomes and their use for studying RNA editing processes in vivo.

Key Words: RNA editing; plastid; chloroplast; tobacco; *Nicotiana tabacum*; biolistic transformation; particle gun; transplastomic plant.

1. Introduction

Conversional RNA editing by pyrimidine transitions occurs in both DNA-containing organelles of plant cells: mitochondria and plastids (chloroplasts). Editing posttranscriptionally alters single-nucleotide positions in plastid and mitochondrial transcripts by cytidine-to-uridine or uridine-to-cytidine conversions at highly specific sites (reviewed in **refs. 1–3**). The vast majority of editing sites is located within coding regions and is functionally significant in that the editing results in changes in the coding properties of the affected triplets. Typically, these alterations result in the restoration of codons for phylogenetically conserved amino acid residues. Editing is an early posttranscriptional event and an essential processing step in the maturation of organellar transcripts in that the changes in the coding properties of the mRNA introduced by edit-

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

ing can be essential for normal protein function (4). While being apparently unrelated to animal RNA-editing systems, the editing processes in higher plant mitochondria and chloroplasts share many similar features, indicating that they might have originated from common evolutionary roots (for a review, *see* **ref. 3**). In both organelles, editing appears to be an evolutionarily late acquisition because it is restricted to higher plants (tracheophytes) and some derived bryophyte lineages, but does not occur in algae and primitive bryophytes.

The molecular mechanisms of RNA-editing processes in plant cell organelles are still largely unknown. For a long time, research on editing in plants has been hampered by a lack of suitable experimental systems for testing mutated substrate RNAs or purifying *trans*-acting editing factors. At least for chloroplasts, recent methodological progress has helped to largely overcome these problems: the development of robust technologies for the genetic transformation of higher plant chloroplasts (5–7) has facilitated the study of plastid RNA editing in an *in vivo* system. The more recent development of an *in vitro* RNA-editing system for chloroplasts (8,9) marks another major stride forward and provides a promising tool toward the biochemical purification of editing factors.

In this chapter, key techniques involved in the *in vivo* analysis of editing by generating plants with transgenic chloroplasts are described. Emphasis is placed on the chloroplast transformation technology for the model plant tobacco and the tissue culture and selection procedures required to obtain stable transplastomic plants (=plants with transformed plastid genomes).

2. Materials

All chemicals, including plant hormones and antibiotics, must be highly pure and certified for plant cell and tissue culture.

2.1. Plant Material and Growth Conditions

Young leaves from tobacco plants (*Nicotiana tabacum*) grown under aseptic conditions are required for chloroplast transformation experiments. Sterile tobacco plants can most easily be produced by surface-sterilizing seeds. To this end, 400 μL of ethanol and 400 μL of 6% bleach are added to a sample of 50–100 seeds in an Eppendorf tube. Following vigorous shaking for 2 to 3 min, the liquid is removed with a sterile pipet, and the seeds are washed at least four times with 1.5 mL of sterile distilled water (each time removing the liquid as completely as possible). Seeds are then transferred to sterile containers and germinated on RM medium (*see* **Subheading 2.3., item 1**). Sterile tobacco plants should be grown in boxes (e.g., Magenta boxes) on RM medium and kept in a growth chamber at 25°C under a 16-h light/8-h dark cycle. The light intensity should be between 100 and 200 $\mu\text{mol quanta}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$.

2.2. Stock Solutions for Plant Culture Media

1. 10X Macro (1 L): 19.0 g of KNO_3 , 16.5 g of NH_4NO_3 , 4.4 g of $\text{CaCl}_2 \cdot 2 \text{H}_2\text{O}$, 3.7 g of $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$, 1.7 g of KH_2PO_4 .
2. 100X Micro (100 mL): 169 mg of $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 86 mg of $\text{ZnSO}_4 \cdot 7 \text{H}_2\text{O}$, 62 mg of H_3BO_3 , 8.30 mg of KI, 2.50 mg of $\text{Na}_2\text{MoO}_4 \cdot 2 \text{H}_2\text{O}$, 0.25 mg of $\text{Cu SO}_4 \cdot 5 \text{H}_2\text{O}$, 0.25 mg of $\text{CoCl}_2 \cdot 6 \text{H}_2\text{O}$.
3. FeNaEDTA (1 g/100 mL).
4. Thiamin•HCl (1 mg/mL).
5. 1-Naphthylacetic acid (NAA) (1 mg/mL) (in 0.1 M NaOH).
6. N6-Benzylamino purine (BAP) (1 mg/mL) (in 0.1 M HCl).

2.3. Plant Culture Media

1. RM medium for growth of sterile tobacco plants (1 L): 10X Macro (100 mL); 100X Micro (10 mL); FeNaEDTA (5 mL); 30 g of sucrose, pH 5.6–5.8 (with 0.1 M KOH). The liquid medium is supplemented with 7.4 g/L of agar (certified for plant tissue culture; Duchefa, The Netherlands) (*see Note 1*), autoclaved, and poured into Petri dishes or sterile containers for plant growth.
2. RMOP medium (1 L) for plant regeneration from tissue explants: 10X Macro (100 mL); 100X Micro (10 mL); FeNaEDTA (5 mL); 30 g of sucrose; Thiamin•HCl (1 mL); NAA (0.1 mL); BAP (1 mL); 100 mg of myoinositol, pH 5.8 (0.2 M KOH). After addition of 7.4 g/L of agar (for plant tissue culture; Duchefa) (*see Note 1*), the medium is sterilized by autoclaving and chilled to approx 60°C. Subsequently, antibiotics are added if required, and the medium is poured into Petri dishes (approx 50 mL of medium/Petri dish of 9 cm diameter and 2 cm height). Regeneration of shoots from tissue explants is induced by addition of the synthetic plant hormones NAA (structural analog of auxin) and BAP (cytokinin analog) to the basic RM medium.

2.4. Particle Preparation and β -Glucuronidase Assays

1. Gold particles (0.6 μM , Bio-Rad, Hercules, CA).
2. 100% Ethanol.
3. Spermidine (free base).
4. CaCl_2 .
5. Phosphate buffer: Mix 8.7 mL of 0.2 M KH_2PO_4 , 12.2 mL of 0.2 M K_2HPO_4 , and 20.9 mL of H_2O (no pH adjustment is necessary).
6. $\text{K}_4[\text{Fe}(\text{CN})_6]$.
7. $\text{K}_3[\text{Fe}(\text{CN})_6]$.
8. 5-bromo-4-chloro-3-indolylglucuronide (X-Gluc).
9. dimethylformamide.
10. Triton X-100.

3. Methods

Currently, plastid transformation is routinely available only for a single higher plant species: the model plant tobacco (**6,10**). Although recently trans-

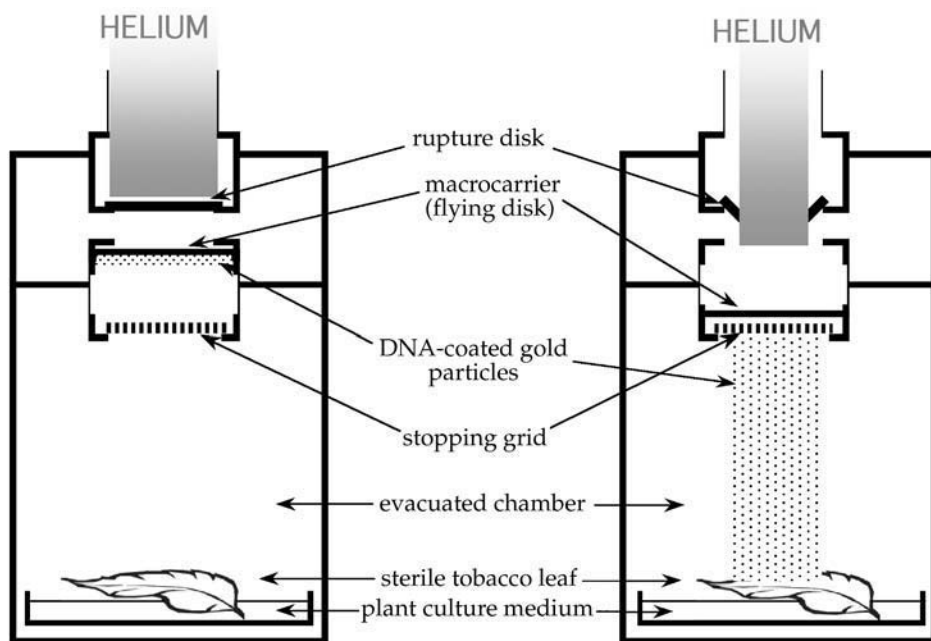


Fig. 1. Biolistic chloroplast transformation with a helium-driven particle gun. A plastic “rupture disk” seals the gas acceleration tube (**left**). The gold particles coated with the transforming plasmid DNA are on the surface of a second plastic disk (macrocarrier). A metal grid (stopping grid) determines the flight distance of the macrocarrier. When the helium pressure reaches the burst pressure of the rupture disk, the disk breaks (**right**), and the helium gas accelerates the flying disk with the particles on its lower side. The flying disk is then stopped by the metal grid, but the DNA-loaded particles fly through the grid and penetrate the cells of the sterile tobacco leaf at the bottom of the chamber (modified from **ref. 32**).

formation protocols have also been developed for potato and tomato plastids (**7,11**), it is highly recommended that at least initially the more efficient and less time-consuming tobacco system be used. Biolistic transformation using a particle gun (**Fig. 1**) appears to be the most effective method of delivering foreign DNA into plastids (**5,12**). Polyethylene glycol treatment of protoplasts provides an alternative to biolistics (**13,14**), but the cell culture and transformation techniques involved are even more demanding and time-consuming than biolistic plastid transformation.

3.1. Plastid Transformation Vector *pRB51*

Tobacco chloroplast transformation technology has been extensively used to map *cis*-acting elements (i.e., RNA sequences) involved in plastid editing site recognition (**15–17**), to define minimum RNA substrates for editing (**18,19**)

and to test for editing of heterologous sites (taken from chloroplast genomes of other species or from plant mitochondrial genomes [4,20]).

To identify *cis*-acting recognition elements or to test mutated editing substrates, editing site-containing DNA segments are placed into a chimeric context and linked to the selectable marker gene *aadA* (a spectinomycin resistance gene encoding an antibiotic-inactivating enzyme). Subsequent transformation of chloroplasts results in introduction of an additional copy of the respective editing site into the chloroplast genome. Such transplastomic lines are useful (1) to test for occurrence and efficiency of editing in the aberrant sequence context and, if editing occurs, (2) to measure competition effects for limiting *trans*-acting site-specific recognition factors (i.e., reduction in the editing efficiency at the endogenous site owing to overexpression of the editing substrate [15,10]).

A suitable vector for testing editing substrates *in vivo* is shown in **Fig. 2**. The plasmid pRB51 contains a chimeric selectable marker gene *aadA* inserted into a noncoding intergenic spacer region (between the *psbE* operon and the *petA* gene; **Fig. 2A,B**) of a cloned plastid DNA fragment. Thus, the *aadA* transgene is flanked by two blocks of chloroplast genome-derived sequence required to mediate transgene integration by homologous recombination (18). Expression of the *aadA* is driven by the strong more or less constitutively expressed Prn promoter taken from the plastid rRNA operon and fused to a synthetic Shine-Dalgarno sequence (6). At its 3' end, the *aadA* coding region is tethered to the 3' untranslated region (3' UTR) from a chloroplast photosystem II gene (*psbA*) that, at the mRNA level, folds into a stable stem loop-type secondary structure and acts as a transcript stabilizing sequence (**Fig. 2C**). In pRB51, the *aadA* gene was further modified by inserting a minilinker between the *aadA* termination codon and the 3' UTR. This minilinker contains a 5' *Xba*I restriction site and a downstream *Bam*HI site. Small fragments containing a chloroplast-editing site can be prepared by polymerase chain reaction (PCR) amplification of the respective genomic sequences introducing a 5' *Xba*I site and a 3' *Bam*HI site with the PCR primers. Any such fragment can then be inserted into the *Xba*I/*Bam*HI minilinker and yields, in a single cloning step, the final plastid transformation vector. Note that, because of their incorporation downstream of the *aadA* stop codon, the editing site-containing sequences will remain untranslated in this context. Instead, they are cotranscribed with the *aadA* and thus can be viewed as a second (RNA gene-like) cistron of a dicistronic operon. The lack of translation will not interfere with editability of the sites since, with the exception of a single site in the *ndhB* transcript (21,22), plastid RNA editing is largely independent of chloroplast translation (23).

3.2. Preparation of Gold Particles for Biolistic Transformation

All steps are done on ice. Water and 100% ethanol must be ice cold!

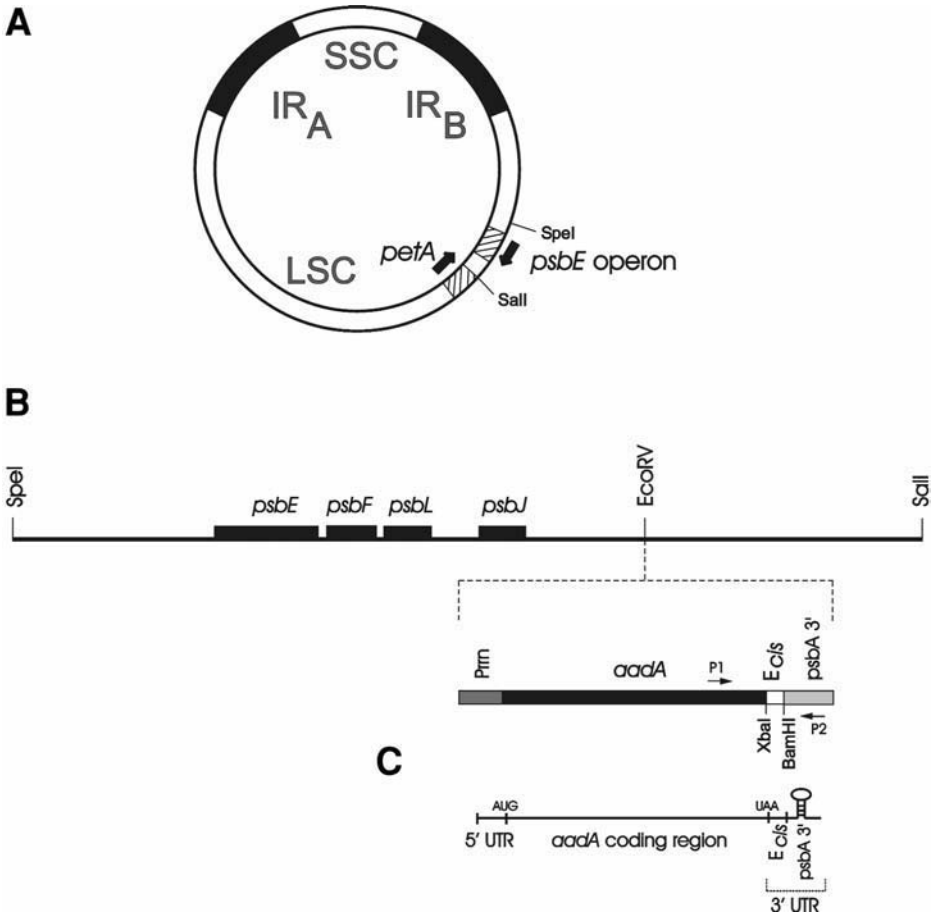


Fig. 2. Physical map of a chloroplast transformation vector suitable for mapping of *cis*-acting sequence elements for recognition of plastid RNA-editing sites (**18**). (**A**) Schematic drawing of tobacco chloroplast genome. The genome contains two large inverted repeat regions (IR_A and IR_B) separated by a large single-copy region (LSC) and a small single-copy regions (SSC). The intergenic spacer between the *psbE* operon and the *petA* gene (striped boxes) serves as a neutral insertion site for transgenes. Note that the *psbE* operon and *petA* are transcribed in opposite directions (arrows). Restriction sites (*SpeI* and *SalI*) used for cloning of this region into a pBluescript vector are also shown. (**B**) Physical map of chloroplast transformation vector pRB51. A chimeric spectinomycin resistance gene (*aadA*) is cloned into a unique *EcoRV* site within the intergenic spacer region between the *psbE* operon and *petA*. A minilinker in between the *aadA* coding region and the transcript-stabilizing *psbA* 3' UTR allows for convenient insertion of PCR-amplified editing site-containing DNA fragments (*E_{cis}*) as *XbaI*/*Bam*HI restriction fragments. (**C**) Structure of chimeric *aadA* transcript as expressed from the transgenic chloroplast genome.

1. Use 2 mg of gold particles (0.6 μm diameter) per 10 shots.
2. Prepare a gold suspension: $n \times 2$ mg in $n \times 100$ μL of 100% EtOH.
3. Vortex for at least 1 min at maximum power.
4. Centrifuge for 1 s at 4000 rpm in a microfuge.
5. Remove the supernatant completely and redissolve the particles in 1 mL of sterile distilled water.
6. Centrifuge for 1 s at 4000 rpm in a microfuge, and remove and discard the supernatant.
7. Resuspend the particles very carefully in $n \times 250$ μL of sterile water.
8. Make 245- μL aliquots (per 10 shots).
9. Add the following components to each aliquot in the order given: 20 μg of DNA (plastid transformation vector; concentration 1 to 2 $\mu\text{g}/\mu\text{L}$), 250 μL of 2.5 M CaCl_2 , 50 μL of 100 mM spermidine. (Vortex briefly immediately after the addition of each component.)
10. Incubate on ice for 10–15 min with brief vortexing every minute.
11. Add 200 μL of 100% EtOH to each sample and briefly mix by vortexing.
12. Centrifuge for 1 s at 3500 rpm in a microfuge, and remove the supernatant completely.
13. Add 600 μL of 100% EtOH, and carefully resuspend the particles by pipetting and vigorous vortexing.
14. Centrifuge for 1 s at 3500 rpm in a microfuge, and remove the supernatant completely.
15. Repeat **steps 13 and 14**.
16. Resuspend the particles very carefully in 65 μL of 100% EtOH by pipetting and vigorous vortexing (*see Note 2*).
17. Use 6 μL per shot, and carefully resuspend the particles immediately before use.

3.3. Biolistic Bombardment of Tobacco Leaves

Use young leaves from tobacco plants (*N. tabacum*) grown under sterile conditions. Place a leaf in a Petri dish on a sterile piece of filter paper on top of a thin layer of RMOP medium without antibiotics. Bombard the abaxial side of the leaf (*see Note 3*) with the plasmid DNA-coated gold particles. For the Bio-Rad PDS1000He biolistic gun (**Fig. 1**), suitable settings for the transformation of tobacco chloroplasts are as follows:

1. Helium pressure at the tank regulator: 1300–1400 psi.
2. Rupture disks: 1100 psi (Bio-Rad).
3. Flying disk assembly: level two from the top.
4. Petri dish holder: level two from the bottom.
5. Vacuum (at time of the shot): 26–28 in. Hg.

A typical chloroplast transformation experiment involves the bombardment of 15–30 leaves per DNA construct.

3.4. Transient Transformation Assays and Histochemical Detection of β -Glucuronidase Activity

The following assay (24) determines the transient transformation frequency and it is important to adjust the settings of the particle gun, optimize the particle preparation procedure, and control for its quality. Although it assays only for nuclear transformation events and, moreover, does not distinguish between transient transformation and stable integration of the transforming DNA, it is reasonable to assume and generally believed that the frequency of transient nuclear transformation correlates with the frequency of chloroplast transformation.

The assay makes use of the β -glucuronidase (GUS) reporter of gene expression. GUS activity can be easily detected by histochemical staining using synthetic substrates, such as X-Gluc. Transiently transformed cells express the enzyme that converts X-Gluc into a blue indigo dye. A large number of plant transformation vectors contain GUS reporter genes under the control of strong constitutive promoters. All of them can also be used for transient transformation assays (*see Note 4*).

A leaf sample bombarded with a GUS-expressing vector is assayed as follows:

1. Mix in an Eppendorf tube 979 μ L of phosphate buffer, 10 μ L of 5 mM $K_4[Fe(CN)_6]$, and 10 μ L of 5 mM $K_3[Fe(CN)_6]$.
2. Dissolve 1 mg of X-Gluc in approx 100 μ L of dimethylformamide.
3. Combine the two solutions, add one drop of the detergent Triton X-100, and mix well.
4. Transfer the bombarded leaf sample into a Petri dish, and distribute the solution onto the abaxial (i.e., the bombarded) side of the leaf.
5. Seal the Petri dish with Parafilm® and incubate overnight at 37°C.

A particle preparation that has worked reasonably well should produce about 10,000 blue spots on the bombarded tobacco leaf. Note that although an individual spot originates from a single GUS-expressing cell, it becomes macroscopically visible only by diffusion of the blue indigo dye into neighboring cells owing to cell lysis caused by the detergent Triton X-100. In addition, it should be borne in mind that only about 2–5% of the cells transiently expressing the marker gene would give rise to stable nuclear transformants (25).

3.5. Selection of Transplastomic Plants

Following biolistic bombardment, the leaves are cut into small pieces (approx 5×5 mm) and placed on the surface of a selective regeneration medium (RMOP containing 500 μ g/mL of spectinomycin). Resistant calli or shoots will appear after 3–6 wk of incubation. Note that not all of the resistant plant lines are true chloroplast transformants: spontaneous spectinomycin resis-

tance can occur through acquisition of specific point mutations in the chloroplast 16S rRNA gene (26). However, such point mutations act in a strictly antibiotic-specific manner: spontaneous spectinomycin-resistant cells are streptomycin sensitive and vice versa. By contrast, the *aadA* selectable marker gene confers resistance to a number of antibiotics of the aminoglycoside type, including spectinomycin and streptomycin. Consequently, when exposed to double selection on plant regeneration medium containing both spectinomycin and streptomycin, tissue samples from spontaneous spectinomycin-resistant lines bleach out, whereas tissue samples from true chloroplast transformants remain green and display growth of callus (10). Alternatively, true chloroplast transformants can be identified by DNA gel blot or PCR analyses testing for the physical presence of the resistance gene in the plastid genome (4,18).

Since plant cells are highly polyploid with respect to their plastid genome, a single tobacco cell may contain up to 10,000 identical copies of its chloroplast DNA (27). For this reason, the primary chloroplast transformant is usually heteroplasmic; that is, it contains a mixture of wild-type and transformed genome copies. To obtain a genetically stable transplastomic plant, residual wild-type genome molecules must be completely eliminated. This is achieved by passing the primary transformant through additional rounds of regeneration on selective medium (10). Usually after two to four rounds, the lines will be homoplasmic and the wild-type genome will no longer be detectable. Homoplasmy is assessed by either DNA gel blot analysis (6) or highly sensitive PCR assays that selectively favor amplification of residual wild-type genomes (18). Verification of homoplasmy is particularly important when the transplastomic lines can be expected to display a mutant phenotype (4) or when the lines are to be used for quantitative analyses of editing efficiency.

Finally, homoplasmic transplastomic shoots regenerated on RMOP medium are transferred to boxes with RM medium. The phytohormone-free RM medium will induce rooting, and the plants can then be taken out of the sterile environment, transferred to the soil (see Note 5), and kept in a greenhouse (see Note 6).

3.6. Analysis of Transplastomic Plants

Since chloroplasts are uniparentally maternally inherited in tobacco (as in the vast majority of angiosperm plant species), plastid transgene localization and homoplasmy can be easily verified genetically by conducting reciprocal crosses. Wild-type and transformed plants are transferred to soil and grown to maturity in a greenhouse. Seed pods are collected from selfed plants and from reciprocal crosses of the transplastomic lines with wild-type plants. Surface-sterilized seeds are then germinated on spectinomycin or streptomycin-containing (500 mg/L) RM medium and analyzed for uniparental inheritance of

the resistance trait. Selfed transformants and crosses with the plastid transformant as maternal parent give rise to uniformly green (i.e., drug-resistant) progeny, whereas seeds collected from wild-type plants and crosses with the plastid transformant as paternal parent yield white (i.e., antibiotic-sensitive) seedlings.

All methods for analyzing RNA editing in transplastomic plants represent standard procedures in plant molecular biology. Simple protocols are available that allow the simultaneous isolation of high-quality DNA and RNA suitable for reverse transcription (e.g., see **ref. 28**). Editing efficiency can be quantitated using a PhosphorImager and published quantitation methods (**12,18,19,30**).

4. Notes

1. The exact amount of agar is dependent on the autoclave used (and its heating and cooling speeds) and must be determined empirically. The solidified medium should be soft enough to allow pushing of leaf pieces 1 to 2 mm into the medium (for optimum medium contact to induce regeneration), but hard enough to stably hold an inserted stem cutting in a vertical position.
2. Coated gold particles tend to aggregate. Careful resuspension of the particles is crucial to the success of the transformation experiment. Clumpy particles hit fewer target cells and will cause severe cell damage on penetration of the leaf surface, thus greatly reducing regeneration rate and transformation frequency.
3. The abaxial (lower) side of the leaf usually has a somewhat softer epidermis and cuticle than the adaxial (upper) leaf side and, thus, is better penetrated by the particles.
4. Vectors for *Agrobacterium*-mediated transformation (e.g., pBI121) are often relatively large plasmids and may give somewhat lower transient transformation frequencies than smaller vectors for biolistics or protoplast transformation. One reasonably small GUS-expressing vector for transient transformation assays is pFF19G (**31**).
5. Residual culture medium should be washed off of the roots of the plants as completely as possible. Because the medium is rich in minerals and also contains sucrose, it would be colonized by soil microbes very rapidly, which in turn might lead to infection of the root tissue.
6. Beware that plants coming from sterile culture containers are extremely sensitive to reduced humidity. To allow for adaptation to greenhouse conditions, they should be kept under a transparent plastic cover for at least a few days, and then the cover is stepwise lifted.

Acknowledgments

I thank Dr. Stephanie Ruf for helpful comments and critical reading of the manuscript. The work on transgenic chloroplasts was supported by grants from the Deutsche Forschungsgemeinschaft.

References

1. Mulligan, R. M., Williams, M. A., and Shanahan, M. T. (1999) RNA editing site recognition in higher plant mitochondria. *J. Hered.* **90**, 338–344.
2. Bock, R. (2000) Sense from nonsense: how the genetic information of chloroplasts is altered by RNA editing. *Biochimie* **82**, 549–557.
3. Bock, R. (2001) RNA editing in plant mitochondria and chloroplasts, in *Frontiers in Molecular Biology: RNA Editing* (Bass, B., ed.), Oxford University Press, pp. 38–60.
4. Bock, R., Kössel, H., and Maliga, P. (1994) Introduction of a heterologous editing site into the tobacco plastid genome: the lack of RNA editing leads to a mutant phenotype. *EMBO J.* **13**, 4623–4628.
5. Svab, Z., Hajdukiewicz, P., and Maliga, P. (1990) Stable transformation of plastids in higher plants. *Proc. Natl. Acad. Sci. USA* **87**, 8526–8530.
6. Svab, Z. and Maliga, P. (1993) High-frequency plastid transformation in tobacco by selection for a chimeric aadA gene. *Proc. Natl. Acad. Sci. USA* **90**, 913–917.
7. Ruf, S., Hermann, M., Berger, I. J., Carrer, H., and Bock, R. (2001) Stable genetic transformation of tomato plastids and expression of a foreign protein in fruit. *Nat. Biotechnol.* **19**, 870–875.
8. Hirose, T. and Sugiura, M. (2001) Involvement of a site-specific trans-acting factor and a common RNA-binding protein in the editing of chloroplast mRNAs: development of a chloroplast in vitro RNA editing system. *EMBO J.* **20**, 1144–1152.
9. Miyamoto, T., Obokata, J., and Sugiura, M. (2002) Recognition of RNA editing sites is directed by unique proteins in chloroplasts: biomedical identification of cis-acting elements and trans-acting factors involved in RNA editing in tobacco and pea chloroplasts. *Mol. Cell. Biol.* **22**, 6726–6734.
10. Bock, R. (2001) Transgenic chloroplasts in basic research and plant biotechnology. *J. Mol. Biol.* **312**, 425–438.
11. Sidorov, V. A., Kasten, D., Pang, S.-Z., Hajdukiewicz, P. T. J., Staub, J. M., and Nehra, N. S. (1999) Stable chloroplast transformation in potato: use of green fluorescent protein as a plastid marker. *Plant J.* **19**, 209–216.
12. Bock, R. (1998) Analysis of RNA editing in plastids. *Methods* **15**, 75–83.
13. Golds, T., Maliga, P., and Koop, H.-U. (1993) Stable plastid transformation in PEG-treated protoplasts of *Nicotiana tabacum*. *Bio/Technology* **11**, 95–97.
14. O'Neill, C., Horvath, G. V., Horvath, E., Dix, P. J., and Medgyesy, P. (1993) Chloroplast transformation in plants: polyethylene glycol (PEG) treatment of protoplasts is an alternative to biolistic delivery systems. *Plant J.* **3**, 729–738.
15. Chaudhuri, S., Carrer, H., and Maliga, P. (1995) Site-specific factor involved in the editing of the psbL mRNA in tobacco plastids. *EMBO J.* **14**, 2951–2957.
16. Bock, R., Hermann, M., and Fuchs, M. (1997) Identification of critical nucleotide positions for plastid RNA editing site recognition. *RNA* **3**, 1194–1200.
17. Hermann, M. and Bock, R. (1999) Transfer of plastid RNA-editing activity to novel sites suggests a critical role for spacing in editing-site recognition. *Proc. Natl. Acad. Sci. USA* **96**, 4856–4861.

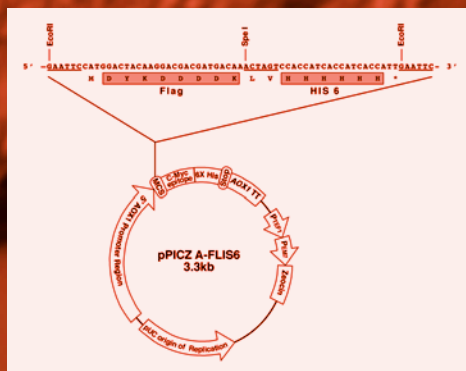
- 18 Bock, R., Hermann, M., and Kössel, H. (1996) In vivo dissection of cis-acting determinants for plastid RNA editing. *EMBO J.* **15**, 5052–5059.
- 19 Chaudhuri, S. and Maliga, P. (1996) Sequences directing C to U editing of the plastid psbL mRNA are located within a 22 nucleotide segment spanning the editing site. *EMBO J.* **15**, 5958–5964.
- 20 Sutton, C. A., Zoubenko, O. V., Hanson, M. R., and Maliga, P. (1995) A plant mitochondrial sequence transcribed in transgenic tobacco chloroplasts is not edited. *Mol. Cell. Biol.* **15**, 1377–1381.
- 21 Karcher, D. and Bock, R. (1998) Site-selective inhibition of plastid RNA editing by heat shock and antibiotics: a role for plastid translation in RNA editing. *Nucleic Acids Res.* **26**, 1185–1190.
- 22 Karcher, D. and Bock, R. (2002) The amino acid sequence of a plastid protein is developmentally regulated by RNA editing. *J. Biol. Chem.* **277**, 5570–5574.
- 23 Zeltz, P., Hess, W. R., Neckermann, K., Börner, T., and Kössel, H. (1993) Editing of the chloroplast rpoB transcript is independent of chloroplast translation and shows different patterns in barley and maize. *EMBO J.* **12**, 4291–4296.
- 24 Stomp, A.-M. (1992) Histochemical localization of (β -glucuronidase, in *GUS Protocols: Using the GUS Gene as a Reporter of Gene Expression* (Gallagher, S. R., ed.), Academic, San Diego, pp. 103–113.
- 25 Klein, T. M., Harper, E. C., Svab, Z., Sanford, J. C., Fromm, M. E., and Maliga, P. (1988) Stable genetic transformation of intact *Nicotiana* cells by the particle bombardment process. *Proc. Natl. Acad. Sci. USA* **85**, 8502–8505.
- 26 Svab, Z. and Maliga, P. (1991) Mutation proximal to the tRNA binding region of the *Nicotiana* plastid 16S rRNA confers resistance to spectinomycin. *Mol. Gen. Genet.* **228**, 316–319.
- 27 Bendich, A. J. (1987) Why do chloroplasts and mitochondria contain so many copies of their genome? *BioEssays* **6**, 279–282.
- 28 Doyle, J. J. and Doyle, J. L. (1990) Isolation of plant DNA from fresh tissue. *Focus* **12**, 13–15.
- 29 Reed, M. L., Peeters, N. M., and Hanson, M. R. (2001) A single alteration 20 nt 5' to an editing target inhibits chloroplast RNA editing in vivo. *Nucleic Acids Res.* **29**, 1507–1513.
- 30 Peeters, N. M. and Hanson, M. R. (2002) Transcript abundance supercedes editing efficiency as a factor in developmental variation of chloroplast gene expression. *RNA* **8**, 497–511.
- 31 Timmermans, M. C. P., Maliga, P., Vieira, J., and Messing, J. (1990) The pFF plasmids: cassettes utilising CaMV sequences for expression of foreign genes in plants. *J. Biotechnol.* **14**, 333–344.
- 32 Hager, M. and Bock, R. (2000) Enslaved bacteria as new hope for plant biotechnologists. *Appl. Microbiol. Biotechnol.* **54**, 302–310.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



Detection and Quantification of Modified Nucleotides in RNA Using Thin-Layer Chromatography

Henri Grosjean, Gérard Keith, and Louis Droogmans

Summary

Identification of a modified nucleotide and its localization within an RNA molecule is a difficult task. Only direct sequencing of purified RNA molecules and high-performance liquid chromatography mass spectrometry analysis of purified RNA fragments allow determination of both the type and location of a given modified nucleotide within an RNA of 50–150 nt in length. The objective of this chapter is to describe in detail a few simple procedures that we have found particularly suited for the detection, localization, and quantification of modified nucleotides within an RNA of known sequence. The methods can also be used to reveal the enzymatic activity of a particular RNA-modifying enzyme *in vitro* or *in vivo*. The procedures are based on the use of radiolabeled RNA (with [^{32}P], [^{14}C], or [^3H]) or [^{32}P]-postlabeled oligonucleotides and two-dimensional thin-layer chromatography of labeled nucleotides on cellulose plates. This chapter provides useful maps of the migration characteristics of 70 modified nucleotides on thin-layer cellulose plates.

Key Words: RNA maturation; posttranscriptional modification; thin-layer chromatography; sequencing; modified nucleotides; 2'-*O*-methylation; nuclease; RNase; kinase; phosphatase.

1. Introduction

The presence of modified ribonucleotides derived from adenosine, guanosine, cytidine and uridine is a hallmark of almost all cellular RNAs of the three phylogenetic domains of life. To date, a total of 96 modified nucleosides to which structures have been assigned have been reported (*1*) (<http://medstat.med.utah.edu/RNAmods> and references therein). However, the pattern of modifications (type and location) depends on the RNA molecule considered as well as on the organism and the organelle they originate from. The largest number

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

of modified nucleosides with the greatest structural diversity (a total of 81) is found in RNAs (2) (www.uni-bayreuth.de/departments/biochemie/trna). Other types of RNA (e.g., small nucleolar RNA, rRNA, mRNA) also contain modified nucleosides (*see* Chapters 20 and 21 and <http://rna.wustl.edu/snoRNAdb>). However, their occurrence and particularly their diversity are lower than in tRNA (3).

In early works, information on the presence and the precise location of a modified nucleotide in a given RNA molecule was obtained by direct sequencing. The very first method used a set of specific RNases to cleave the RNA into oligonucleotides, followed by separation of these fragments by column chromatography and/or electrophoresis. Ultimate identification of modified nucleosides or nucleotides was performed by paper or thin-layer chromatography (TLC), followed by ultraviolet (UV) spectrophotometric determinations. Once identified, they were compared to synthetic modified nucleotides or nucleosides for assignment. This early procedure, called “cold sequencing,” was laborious, time-consuming, and required gram amounts of pure RNA.

Later, using uniformly labeled [^{32}P]-RNA instead of nonlabeled RNA, complete sequencing of an RNA molecule, including identification of modified nucleotides using methods similar to those just described, was performed within a reasonable period of time (months) and with very low amounts of starting material (microgram level). However, this “hot sequencing” technique can be used only with RNA originating from organisms that can be grown in the presence of [^{32}P]-orthophosphate (reviewed in **refs. 4 and 5**).

A major breakthrough in RNA sequence analysis came with the discovery of T4-polynucleotide kinase (PNK). This enzyme allows 5'-end [^{32}P] postlabeling and determination of the corresponding end nucleotide in any RNA fragment resulting from either specific enzymatic digestion or controlled chemical degradation of RNA (6). With this postlabeling technique, sequences can be obtained within only a few weeks of work and from as little as a few micrograms of pure RNA. However, the type of modified nucleotide, especially when its chemical structure is unknown, and its stoichiometric amount within the RNA molecule remained difficult to determine. Nevertheless, the majority of the sequences of naturally occurring tRNAs of various types of cells (more than 500 to date) have been (and are still) obtained using postlabeling sequencing techniques for which various excellent protocols have been described (7–11).

Recently, a promising new development of RNA-sequencing techniques based on a combination of high-performance liquid chromatography (HPLC) and mass spectrometry (MS) analysis has been published (12–15). This approach allows the detection, chemical characterization, and sequence localization of modified nucleotides in a pure fragment of RNA. The technique requires expensive instrumentation and expertise that only few laboratories and

people can acquire. However, HPLC and combined HPLC-MS remain the most powerful tools for the detection, identification, and quantification of modified nucleotides within a pure RNA fragment up to the size of a tRNA or 5S-rRNA (75–150 nt) (*see* for examples **refs. 16–24b**).

Whichever method is used, identification of a modified nucleotide and its localization within an RNA molecule remains a difficult task. Fortunately, in most cases the presence of a given modified nucleotide and the sequence around it in the RNA is expected, and the only information needed is the confirmation of its presence and estimation of its stoichiometric amount.

The objective of this chapter is to describe in detail a few simple procedures that we have found particularly suited for the detection, localization, and quantification of modified nucleotides within an RNA of known sequence. The method can also be used to reveal the enzymatic activity of a particular RNA-modifying enzyme *in vitro* or *in vivo*. The procedure is based on the use of radiolabeled RNA (with [^{32}P], [^{14}C], or [^3H]) or [^{32}P]-postlabeled oligonucleotides and two-dimensional TLC of labeled nucleotides on cellulose plates (2D-TLC), which can be run at low cost in any laboratory. The use of [^{32}P]-postlabeling procedures greatly facilitates the identification of even tiny amounts of modified nucleotides. However, the use of [^{14}C]- or [^3H]-labeling procedures allows the identification of nucleotides, which could be analyzed by alternative techniques as well. This chapter provides useful maps of the migration characteristics of 60 modified nucleotides on thin-layer cellulose plates. These maps complete or simplify those that were already published by one of us (**11,25**) or others (**9,10,26,27**).

2. Materials

2.1. Enzyme Stock Solutions

1. RNase T2 from *Aspergillus oryzae* (cat. no. 18031-13; Gibco-BRL, Gaithersburg, MD): Prepared by dissolving the enzyme in water at a concentration of 5 U/ μL .
2. Nuclease P1 from *Penicillium citrinum* (cat. no. 236225; Roche, Indianapolis, IN) and *Crotalus* snake venom phosphodiesterase (VPD) I (cat. no. 31B217J; Worthington): Dissolve in water at a concentration of 1 mg/mL.
3. Apyrase from potato (*Solanum tuberosum*) (partially purified grade VI from Sigma, St. Louis, MO; cat. no. A-6410): Dissolve in water at 500 U/mL (about 2 mg/mL). Commercial apyrase always has to be tested for possible contamination by phosphatases.
4. PNK (from any company such as USB; Promega, Madison, WI; New England Biolabs; MBI-Fermentas, Hanover, MD; or Roche): This is usually provided in solution at 10 U/ μL in buffer whose composition depends on the company of origin. We prefer recombinant PNK from New England Biolabs (cat. no. 201S) delivered at 10 U/ μL in 50 mM KCl, 10 mM Tris-HCl, pH 7.4, 0.1 mM EDTA, 1 mM dithiothreitol (DTT), 0.1 μM ATP, and 50% glycerol (v/v). The PNK solu-

tion remains active for months or even years when stored at -20°C , depending on how frequently the tube is opened and kept at 4°C during pipetting. Be sure that the working solution contains 50% glycerol because PNK does not support freezing.

5. Commercial T7-RNA polymerase (cat. no. P2075; Promega): This is provided at 15 U/ μL in 20 mM K-phosphate buffer, pH 7.7, 1 mM EDTA, 10 mM DTT, 100 mM NaCl, 0.1% Triton X-100, and 50% (v/v) glycerol together with its TSC5x reaction buffer (200 mM Tris-HCl buffer, pH 7.9, 30 mM MgCl_2 , 10 mM spermidine, 50 mM NaCl).
6. Enzyme solutions: These are aliquoted into 20 to 50- μL samples in Eppendorf tubes, which can be stored at -20°C for years. As a rule, multiple freeze-thaw cycles (especially in phosphate buffer) or exposure to frequent temperature changes is harmful to most enzymes even when glycerol is present.

2.2. Chemicals, Solutions, and Buffers

As a rule, all chemicals are of the utmost quality, and all solutions are made with RNase-free water (either sterilized or treated overnight at 37°C with 0.1% diethylpyrocarbonate [DEPC] and subsequently boiled or autoclaved for 15–20 min to eliminate any trace of DEPC). Eppendorf tubes and tips for pipetting solutions are always autoclaved before use in order to ensure the absence of RNase. Gloves should be worn to avoid contamination by finger RNase (28).

1. Buffers: Prepare a few milliliters of the following solutions: 50 mM NH_4 acetate, pH 4.5 (best buffer for direct RNase T2 digestion and further treatments); 50 mM NH_4 acetate, pH 5.3 (best buffer for direct nuclease P1 treatment); 200 mM NH_4 acetate, pH 5.3 (best buffer for nuclease P1 digestion after RNase T2 digestion and other treatments); 10 mM Na phosphate buffer, pH 7.0. The “labeling buffer” (fivefold concentrated) consists of 100 mM bicine-HCl, pH 8.2 (cat. no. B-3876; Sigma), 50 mM MgCl_2 , 50 mM DTT, 5 mM RNase-free spermidine, pH 8.2 (cat. no. S-0266; Sigma).
2. Stock solution of carrier bulk tRNA: 20 mg of baker’s yeast tRNA (cat. no. R-8759; Sigma) in 1 mL of water.
3. Stock solution of 5′-ATP- Na_2 (at 0.5 mg/mL in water) (SigmaUltra; cat. no. A-7699); neutralize to pH 7.0. The set of ATP, CTP, GTP, and UTP solutions for transcription reactions (Na salts, each at 10 mM, pH 7.0) is from Promega (cat. no. P1221).
4. Homemade mix of the four 5′-monophosphate nucleosides AMP, CMP, GMP, and UMP (Na salts; Sigma A-1752, C-1006, G-8377, U-6375) at a final concentration of 20 mg/mL. Prepare by dissolving 5 mg of each nucleotide in 1 mL of DEPC- H_2O , and adjust the pH to 7.0. This solution can be used as P1-marker mix for TLC as well. However, a homemade mix of P1-markers of bulk tRNA hydrolysate (including modified nucleotides) is preferred. This is prepared by incubating 1 mg of total baker’s yeast tRNA in 250 μL of 50 mM NH_4 acetate buffer (pH 5.3) with 2.5 μg (about 1.2 U) of nuclease P1 for 2 h or overnight in an incubator hood at 37°C . Similarly, homemade T2-marker mix for TLC is prepared by

incubating 1 mg of bulk Baker's yeast tRNA in 250 μ L of 50 mM NH_4 acetate buffer (pH 4.5) with 5 U of RNase T2 for 2 h (or overnight) at 37°C. All solutions are kept at -20°C until used.

5. Radioactive [^{32}P]-labeled nucleotide triphosphates purchased from Amersham (Piscataway, NJ) or ICN. [γ - ^{32}P]-ATP is at 5000 Ci/mmol while [α - ^{32}P]-XTP is at 400 or 800 Ci/mmol. [^{32}P]-orthophosphate is the carrier-free grade.
6. Other compounds: Triethylammonium carbonate buffer (pH 8.0) (cat. no. 7408; Sigma); 25% concentrated ammonia solution (cat. no. 105432; Merck); piperidine (cat. no. 9724; Merck); 1-propanol (cat. no. 100997; Merck); 2-propanol (cat. no. 109634; Merck); isobutyric acid (cat. no. 800472; Merck); 37% concentrated fuming HCl (cat. no. 100317; Merck); NH_4 sulfate (cat. no. 101217; Merck) or better ultrapure NH_4 sulfate (cat. no. 5501UA; Gibco-BRL).

2.3. Thin-Layer Chromatography

1. Solvent A: isobutyric acid:concentrated ammonia:water (66:1:33 [v:v:v]). Mix 330 mL of isobutyric acid with 5 mL of ammonia solution (25%) and 165 mL of water (inclusion of 0.5 M EDTA, pH 8.0, is optional).
2. Solvent B: phosphate buffer: NH_4 sulfate:*n*-propanol (100:60:2 [v:w:v]). Prepare by adding 300 g of (ultrapure) NH_4 sulfate to 500 mL of 100 mM Na phosphate buffer (pH 6.8). Mix well and stir at 50–60°C until all the NH_4 sulfate is dissolved. Allow to cool to room temperature and add 10 mL of 1-propanol. Continue stirring until the cloudy solution becomes clear.
3. Solvent C: isopropanol:concentrated HCl:water (68:18:14 [v:v:v]). Mix 340 mL of 2-propanol, 90 mL of concentrated fuming HCl, and 70 mL of water. All chromatographic solvents are kept at room temperature in a ventilated area. This is particularly important for solvent A, which has a strong pungent and unpleasant odor, and solvent C, which can corrode metallic zones and damage scientific equipment.
4. Glass-coated cellulose thin-layer plates (20 $\times$ 20 cm, 0.1-mm layer) from Merck (cat. no. 105730-001) or Macherey-Nagel (Polygram CEL-300-10; cat. no. 808-013). Plastic-coated cellulose thin-layer plates from Merck (cat. no. 105577-001) or Macherey-Nagel (Polygram CEL-300-10; cat. no. 801-013).
5. Several multiplack chromatographic tanks for 20 $\times$ 20 cm plates, UV lamp with 254-nm filter, UV-blocking eyewear, autoradiography cassettes for PhosphorImager and X-ray films (20 $\times$ 40 cm) from Fuji or Kodak, β -ray safety screen and racks, one water incubator and one incubator hood at 37°C (to prevent the small volume of sample from drying at the bottom of the tube owing to evaporation and condensation of water in the Eppendorf cap), nitrocellulose Millipore filters (0.2 and 0.45 μ m).

3. Methods

The methods described below outline the identification of modified nucleotides within pure RNA or RNA fragments postlabeled with [^{32}P] at their

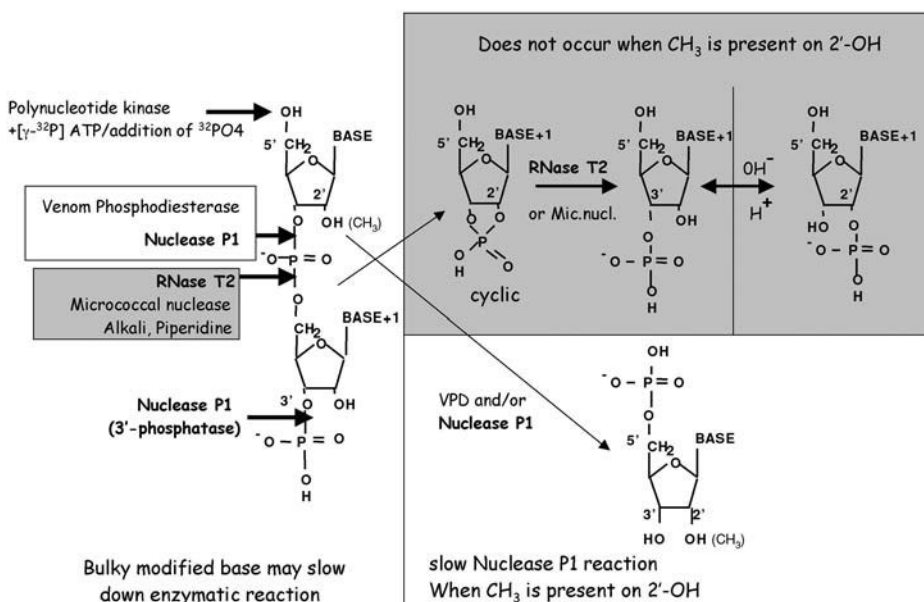


Fig. 1. Various ways to generate monophosphate nucleosides radiolabeled with ^{32}P by postlabeling. For details, see text and **Notes 2–5, 15, and 16**.

5'-ends (**Subheading 3.1.**) uniformly labeled ^{32}P -RNA or ^{32}P -RNA fragments (**Subheading 3.2.**), synthetic RNAs labeled with ^{32}P at a selected position of the RNA molecule (**Subheading 3.3.**), and T7-runoff transcripts generated in the presence of the four triphosphate nucleosides of which only one is radiolabeled with ^{32}P (**Subheading 3.4.**). **Figure 1** illustrates the different cleavage procedures that are used to generate the mononucleotides from purified RNA. In all cases, TLC on cellulose plates is central to the separation and subsequent identification of the modified nucleotides.

As a rule, the RNA molecule should not be too long (maximum length: 150–200 nt) because of difficulties in accurately determining the presence of one or a few modified nucleotides only among the whole RNA sequence. For long RNA, the molecule is cleaved into appropriate smaller fragments and purified by one-dimensional (1D) or 2D electrophoresis or by any other purification technique (for a description and references concerning purification procedures, see **refs. 11, 29, and 30**, and for site-specific cleavages of long RNA, see **refs. 31–34** and references therein).

3.1. Base Composition Analysis by Postlabeling Procedure

3.1.1. Complete Digestion of RNA (or RNA Fragments) into 3'-Monophosphate Nucleosides With RNase T2

1. Lyophilize 1–3 μg of RNA sample ($<0.1 A_{260}$; see **Note 1**) in an Eppendorf tube.
2. Add 10 μL of 50 mM NH_4 acetate buffer, pH 4.5, containing 0.2 U of RNase T2 (see **Note 2**). Mix well and spin down.
3. Incubate the 10- μL reaction mix for 4 h (or overnight in an incubator hood) at 37°C. Keep at –20°C until used. For each experiment described next, we use only 1 μL at a time. Various different samples are usually run in parallel.

3.1.2. [^{32}P]-Labeling of RNase T2 Digest into 3',5-Diphosphate Nucleosides With PNK

1. Lyophilize 1 to 2 μCi of 5'-[γ - ^{32}P]-ATP (5000 Ci/mmol) together with 250–2500 pmol of nonradiolabeled ATP (see **Note 3**). Specific radioactivity will be reduced accordingly.
2. To the lyophilized ATP, add 1 μL of 5X labeling buffer, 1 μL of the RNA hydrolysate obtained in **Subheading 3.1.1.**, 1.5 μL of the PNK solution (1.5 U), and 1.5 μL of DEPC- H_2O . Mix well and spin down.
3. Incubate the 5- μL reaction mix for 30 min to 1 h at 37°C, and then proceed to the next protocol.

3.1.3. Hydrolysis of Excess ATP With Apyrase

1. Just prior to use, dilute the stock solution of apyrase (500 U/mL) 200-fold in DEPC- H_2O (2.5 mU/ μL final).
2. Premix 2 μL (5 mU) of the apyrase dilution with 1 μL of 10 mM Na phosphate buffer, pH 7.0, and 2 μL of 5'-monophosphate nucleosides mix at 20 mg/mL (see **Note 4**).
3. Add this apyrase mix (5 μL) to the mix of 5'-[^{32}P] 3',5'-diphosphate nucleosides obtained in **Subheading 3.1.2**. Mix well and spin down.
4. Incubate the 10- μL reaction mix for 15–30 min at 37°C and proceed to the next protocol.

3.1.4. Removal of 3'-Phosphate of Diphosphate Nucleosides With Nuclease P1

In this protocol, the 3'-phosphatase activity of the nuclease P1 will remove the 3'-phosphate to generate labeled 5'-[^{32}P]-monophosphate nucleosides (see **Note 3**).

1. Prepare a fresh dilution of nuclease P1 at 0.1 mg/mL in 200 mM NH_4 acetate, pH 4.5 (see **Note 5**).

2. Add 5 μL of the diluted nuclease P1 (0.5 μg) to 10 μL of the hydrolysate mix from **Subheading 3.1.3**. Mix well and spin down.
3. Incubate the final 15- μL mix for 2 to 3 h at 37°C. Only a few microliters will be used for the TLC analysis; keep the rest frozen at -20°C.

3.1.5. Analysis of Nucleotides by 2D-TLC

Use commercial TLC plates (20 $\times$ 20 cm) with cellulose (0.1 mm thick) (*see Note 6*). All TLCs are run at room temperature (20–25°C), in closed chromatographic tanks. Solvents A, B, and C (*see Subheading 2.3., items 1–3*) are poured into different tanks so that the height of the solution in the bottom of each tank is 4 to 5 mm. The tanks are kept under a well-ventilated hood, and the solvent is poured in advance to ensure good saturation with solvent vapors. The origin of the deposit is marked on one corner (at 1.5–2 cm from each edge) of the plates with a soft pencil, and the information needed for plate identification is written at the opposite corner. Prior to spotting the radioactive material, about 10–20 μg of the unlabeled marker mix (homemade P1-markers or synthetic mix) is spotted to allow further assignment of the nucleotides under a short-wavelength UV lamp (*see Note 7*). Aliquots of the [^{32}P]-labeled hydrolysate obtained in **Subheading 3.1.4**. (ideally about 20,000–30,000 cpm/Cerenkov counts; *see Note 8*) are then spotted on top of the spot corresponding to the marker mix. These must be applied approx 0.5–1.0 μL at a time to give spots not more than 3 to 4 mm in diameter. The sample is allowed to dry between successive applications.

3.1.5.1. FIRST DIMENSION WITH SOLVENT A

The first dimension is usually developed overnight (15–18 h) with the isobutyric solvent designated A. Once the solvent reaches the top of the plate, some extra time for migration is allowed with this solvent (1 or 2 h) to maximize resolution of the different nucleotides along the migration lane. Several TLC plates (five to seven plates) can be run in parallel in the same multipack chromatographic tank. After migration the plates are withdrawn from the tank and dried thoroughly under a fume hood for several hours (*see Note 9*). Do not dry isobutyric acid and HCl containing solvents together because of the NH_4Cl fume formation.

3.1.5.2. SECOND DIMENSION WITH SOLVENT B (SYSTEM I: A + B)

One of the two plates developed with solvent A is run in the perpendicular direction to the first dimension in a second tank using neutral (pH 6.8) solvent B. This will take about 8–10 h, depending on which commercial plates are used. When solvent B reaches the top, the plate is withdrawn from the tank

and oven dried at 37–50°C in order to avoid slow crystallization and swelling of the NH₄ sulfate as a powder at the surface.

3.1.5.3. SECOND DIMENSION WITH SOLVENT C (SYSTEM II: A + C)

The second of the two dried plates, initially developed with solvent A, is run in the direction perpendicular to the first dimension in a third chromatographic tank using the acidic solvent C (*see Note 10*). This solvent migrates rather slowly and requires 18–20 h to reach the top of the plate. After the run, the plate is withdrawn from the tank and dried thoroughly under a ventilated hood to eliminate acid vapor. The latter darkens X-ray films, corrodes any metallic and other items in the room, and destroys crystals of the costly PhosphorImager screens.

3.1.6. Visual Localization of Nucleotide Markers and Autoradiographic Detection of Labeled Nucleotides

The dried cellulose plates are visualized first in a dark room under UV light with a filter at 254 nm to detect the nucleotide markers. The positions of the radiolabeled nucleotides are obtained after exposure of the plates to autoradiographic X-ray films or PhosphorImager screens. The spots corresponding to the four major canonical 5′-[³²P]-ribonucleotides always give large dark spots on the film, thus limiting resolution of any faint spot corresponding to minor modified nucleotide(s) in their vicinity. This problem is much less serious when [¹⁴C]- or [³H]-labeled nucleotides (or nucleosides) are analyzed by 2D-TLC.

3.1.7. Identification of Modified Nucleotides on Thin-Layer Plates

Detection and identification of the postlabeled [³²P]-modified nucleotides in RNA digests is performed by comparing the autoradiographic and UV patterns against reference maps established in the two TLC systems (I: solvents A + B; II: solvents A + C) described in **Subheadings 3.1.5.2.** and **3.1.5.3.** These reference maps were established by compiling (1) the published localization of modified nucleotides that were found in hundreds of RNAs (mostly tRNAs) sequenced in many laboratories over the last 40 yr, and (2) many published and unpublished results obtained by us in sequencing dozens of tRNAs or by studying tRNA modification in vitro and in vivo. The resulting maps for these two most commonly used chromatographic solvent systems are shown in **Figs. 2–5.** They indicate schematically the relative location of the various 5′-mononucleotides in the 2D-TLC systems. They are grouped into four different sets according to the major nucleotide from which they derive. Altogether, posi-

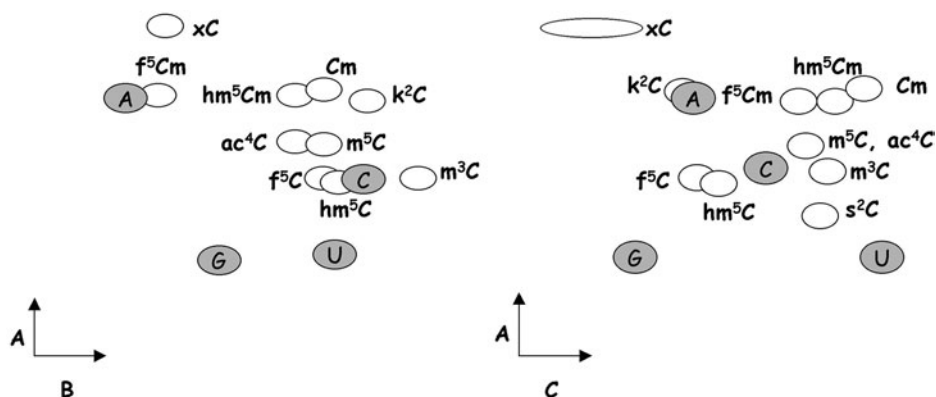


Fig. 3. Diagrams of chromatographic mobility of modified 5'-monophosphate nucleosides derived from cytosine (C) after 2D-TLC. See also text and legend of **Fig. 2**. xC corresponds to the modified cytidine found in the *H. volcanii* tRNA-Ile (anticodon CAU; see **ref. 49**). In solvent system A/C, xC gives a smear, probably because of its instability under acidic conditions (**Note 11**). We lack information for chromatographic mobilities of ac⁴Cm, m⁵Cm present in certain tRNAs and of m⁴C, m⁴Cm present in certain rRNAs. The position of s²C (from tRNA) is shown only in solvent system A/B.

nucleotides in RNA by the postlabeling technique followed by TLC analysis can be found in **refs. 9, 11, 21, and 35**).

3.2. Base Composition Analysis of Uniformly Labeled [³²P]-RNA

An alternative method for detecting modified nucleotides in RNA exists when the study involves microorganisms that can be grown in the presence of highly radioactive [³²P]-orthophosphate. The main advantage of working with uniformly [³²P]-labeled RNA instead of postlabeled RNA is that analysis can be performed directly on hydrolysates of the purified RNA obtained using different types of nucleases. In addition, at variance with the postlabeling technique (see **Note 3**), working with uniformly labeled RNA allows quantitative evaluation of the content of modified nucleotides in RNA. This is particularly interesting when the RNA to be analyzed is extracted from mutant microorganisms that are defective for a given RNA modification enzyme as compared with RNA of the wild-type strain.

Following are detailed only the procedures for the preparation of low-phosphate medium required for efficient incorporation of [³²P] into macromolecules and the protocols for various methods of RNA hydrolysis. The techniques for cell culture, RNA extraction, and purification should be found in other

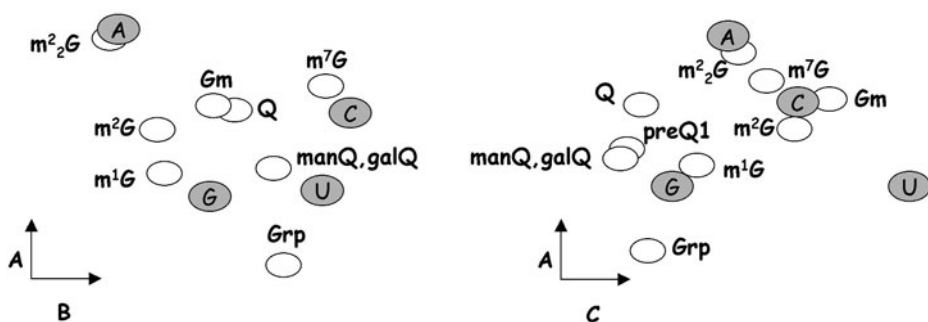


Fig. 4. Diagrams of chromatographic mobility of modified 5'-monophosphate nucleosides derived from guanosine (G) after 2D-TLC. See also text and legend of **Fig. 2**. The chromatographic mobilities of mimG, imG (wyosine or Yt), yW (wybutosine, also called Y), oHyW (hydroxy-Y), oHyW* (Ye), o2yW (peroxy-Y ou Yr), oQ (epoxyqueuosine) from position 37 in the anticodon loop of certain tRNAs, and m²Gm, m²₂Gm from position 10 or 26 in hyperthermophilic archaeal tRNAs and gQ (archaeosine) normally present at position 15 in most archaeal tRNAs are missing. Also those of m^{2,2,7}G and m^{2,7}G present in certain mRNAs and snRNAs are missing. Note that preQo and preQ1 are precursors of Q base (queuine).

specialized method books such as *Current Protocols in Molecular Biology* (Wiley-Interscience) and *Methods in Enzymology* (Academic). The procedures described here apply to RNA labeled with any isotope in addition to [³²P], such as [¹⁴C], [³H] [³⁵S], as well as any selected precursors of modified nucleotides that can be used for incorporation into RNA.

3.2.1. Preparation of Low-Phosphate Medium for Culture of Microorganisms in Presence of [³²P]-Orthophosphate

3.2.1.1. YEAST CULTURE MEDIUM (36):

1. To 100 mL of medium containing 1 g of Bacto yeast extract and 2 g of Bacto peptone, add 1 mL of 1 M MgSO₄ followed by 1 mL of concentrated aqueous ammonia (25%, until the final pH is about 8.5).
2. Gently stir and allow the Mg (NH₄)PO₄ to precipitate at room temperature (or in a cold room at 4°C) for a few hours.
3. Remove the precipitate by centrifugation or filtration through a 0.45-μm Millipore filter.
4. Adjust the pH of the filtrate to 5.8 with HCl and sterilize by autoclaving.
5. To 100 mL of the sterile low-phosphate medium, add 10 mL of a sterile glucose solution at 20% (v/w).

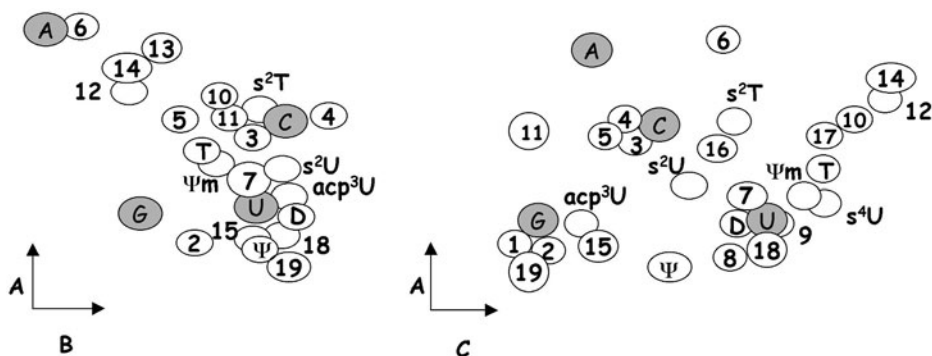


Fig. 5. Diagrams of chromatographic mobility of 5'-monophosphate nucleosides derived from uridine (U) after 2D-TLC. See also text and legend of **Fig. 2**. For clarity of the figure, some U derivatives are designated by numbers: 1 for cmnm^5U (alias cmam^5U); 2 for $\text{cmnm}^5\text{s}^2\text{U}$ (alias $\text{cmam}^5\text{s}^2\text{U}$); 3 for cmnm^5Um ; 4 for mnm^5U (alias mam^5U); 5 for $\text{mnm}^5\text{s}^2\text{U}$ (alias $\text{mam}^5\text{s}^2\text{U}$); 6 for $\text{mnm}^5\text{s}^2\text{Um}$; 7 for $\text{m}^1\psi$; 8 for cmo^5U (also named V); 9 for mo^5U ; 10 for Um ; 11 for mcm^5U ; 12 for $\text{mcm}^5\text{s}^2\text{U}$; 13 for mcm^5Um ; 14 for Tm (m^5Um); 15 for ncm^5U ; 16 for ncm^5Um ; 17 for $\text{m}^1\psi\text{m}$; 18 for cm^5U ; and 19 for $\tau\text{m}^5\text{U}$ and $\tau\text{m}^5\text{s}^2\text{U}$ (taurine derivatives that migrate at the same place in the two solvent systems; see **ref. 102**). Chromatographic mobilities of a given ribonucleotide in both types of solvent systems are not always known. In addition, the possibility exists for misannotation or mispositioning of a few U derivatives of this diagram owing to much confusion in the scientific literature about these U derivatives. Chromatographic mobilities for ho^5U , mcmo^5U (also called mV), chm^5U , mchm^5U , $\text{nm}^5\text{s}^2\text{U}$, $\text{mnm}^5\text{se}^2\text{U}$ present at position 34 of certain tRNAs; s^2Um from tRNAs; and m^3U , m^3Um , $\text{m}^1\text{acp}^3\psi$ present in certain rRNAs are not known yet. Note that $\text{T} = \text{m}^5\text{U}$ and $\text{s}^2\text{T} = \text{s}^2\text{m}^5\text{U}$.

3.2.1.2. *ESCHERICHIA COLI* CULTURE MEDIUM (37)

1. Prepare 100 mL of low-phosphate Bacto peptone (2 g/100 mL) as described in **Subheading 3.2.1., step 1**.
2. Prepare a medium containing 1.5 g/L of KCl, 5.0 g/L of NaCl, 1.0 g/L of NH_4Cl , 2.0 g/L of vitamin-free casamino acids (Difco, Detroit, MI), 100 mL/L of the low-phosphate Bacto peptone, and 12.1 g/L of Tris base in water and autoclave.
3. To every 100 mL of this sterile medium, add 2 mL of sterile 20% (v/w) glucose and 1 mL of sterile 0.1 M MgSO_4 .

3.2.2. Preparation of Uniformly [^{32}P]-Labeled RNA

Since highly radioactive material is handled, experiments (at least at their initial stages) must be performed in a special room, behind plastic screens, using disposable materials for protection against β -rays.

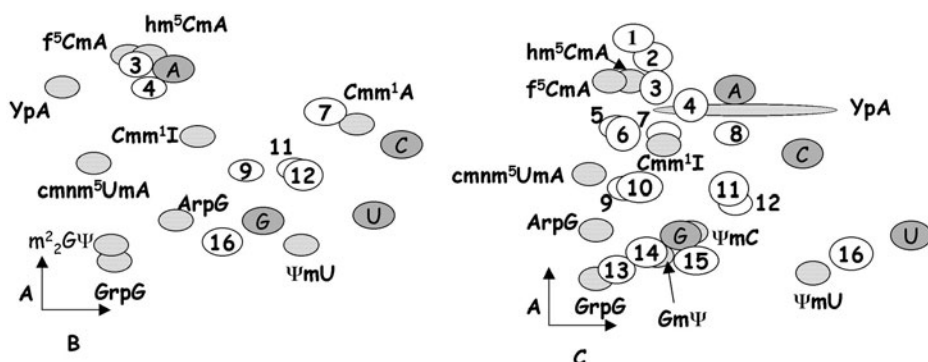


Fig. 6. Diagrams of chromatographic mobility of diphosphate dinucleotides that are resistant to RNase T2 and alkali digestions owing to the presence of a methyl group on the 2'-hydroxyl of the internal ribose (see also Fig. 1 and Notes 2, 5, 15 and 16). Numbers correspond as follows: 1, AmAp; 2, AmCp; 3, CmAp; 4, AmUp; 5, AmGp; 6, GmAp; 7, CmCp; 8, UmAp; 9, CmGp; 10, GmCp; 11, CmUp; 12, UmCp; 13, GmGp; 14, GmUp; 15, UmGp; and 16, UmUp (see also ref. 27). A few diphosphate dinucleotides that contain, in addition to the 2'-methylribose, a modification on one of the two bases are also indicated. In solvent system A/B, not all diphosphate dinucleotides have been identified. The dinucleotide YpA, which is derived from yeast tRNA-Phe recovered from *Xenopus laevis* oocytes (53), gives a smear in solvent A/C because of the instability of the Y derivative under acidic conditions (depurination) (see ref. and Note 12).

Usually, microorganisms are first grown in normal phosphate-containing medium. A small volume (0.1–0.5 mL) of the culture is then transferred into a 50-mL Falcon tube containing 3–5 mL of the low-phosphate medium prepared as described in Subheading 3.2.1. When cells start to grow exponentially, 0.5–1.0 mCi of [^{32} P]-orthophosphate (carrier-free grade) is added. The solution is incubated for several hours at 30°C (yeast) or 37°C (*E. coli*) until most of the [^{32} P]-orthophosphate (60–80%) has been incorporated into the macromolecules. The uptake of [^{32} P]-orthophosphate is measured as follows: an aliquot of the culture (10 μ L) is passed through a nitrocellulose Millipore filter (0.2 μ m). The filter is washed with an isotonic solution, and the radioactivity that remains on the filter is measured by scintillation counting. When most of the [^{32} P] has been incorporated into the cells, 0.5 mL of Na phosphate buffer (pH 7.0) is added to the culture and it is incubated for another 1 or 2 h. This reduces the amount of [^{32} P]-labeled immature RNA in the RNA preparation. At the end of the incubation period, cells are harvested by centrifugation, washed twice with isotonic buffer in order to eliminate most of the unincorporated [^{32} P]-orthophosphate, and transferred into 1.5-mL sure-lock Eppendorf tubes

for further treatments (i.e., cell disruption by vortexing with glass beads, deproteinization with phenol, and precipitation of the nucleic acids [RNA and DNA] by ethanol). Phenol extraction of intact cells will generally allow only the small cellular RNA (e.g., RNA, 5S-RNA, snoRNA) to diffuse out of the permeabilized cells, whereas rRNA and DNA will remain trapped inside the cells together with the denatured proteins (38,39). The uniformly labeled [^{32}P]-nucleic acids are further purified by 1D or 2D polyacrylamide gel electrophoresis (for examples *see* refs. 11 and 29), by adsorption/elution from a small column containing an anion-exchange resin (30), or by affinity chromatography on a small avidin column containing a selected biotinylated oligonucleotide complementary to the RNA of interest (40,41).

3.2.3. Digestions of Uniformly Labeled [^{32}P]-RNA

The purified or semipurified RNA (500–5000 cpm or more), dissolved in DEPC- H_2O (*see* Note 14), is mixed with 1 μL (20 μg) of carrier unlabeled tRNA solution in an Eppendorf tube and lyophilized. A different tube is prepared for each different type of digestion to be tested.

3.2.3.1. COMPLETE DIGESTION WITH RNASE T2

To the lyophilizate obtained in **Subheading 3.2.3.**, 10 μL of 50 mM NH_4 acetate (pH 4.5) containing 0.2 U of RNase T2 is added, mixed well, and spun down. It is then incubated for 4–6 h (or overnight in an incubator hood) at 37°C and kept at –20°C until use. This will produce 3′- [^{32}P]-XMP (where X is any nucleotide, modified or not), except for 2′-*O*-methylated derivatives, for which the corresponding diphosphate dinucleosides (XmpYp) and also 2′-3′-cyclic derivatives of some modified nucleotides will be generated (*see* Note 2).

3.2.3.2. DIGESTION WITH NUCLEASE P1

To the lyophilizate obtained in **Subheading 3.2.3.**, 10 μL of 50 mM NH_4 acetate (pH 5.3) containing 0.25 μg of nuclease P1 is added, mixed well, and spun down. It is then incubated for at least 2 h at 37°C (usually overnight in an incubator hood). This will produce 5′- [^{32}P]-XMP, including 2′-*O*-methylated nucleoside monophosphate (pXm). However, at a lower concentration of nuclease P1 (0.1 $\mu\text{g}/10 \mu\text{L}$) and shorter incubation time (1 h at 37°C), 2′-*O*-methylated dinucleotides (pXmpY) will accumulate because of the slower rate of hydrolysis of phosphodiester bonds when a 2′-*O*-methyl group is present (*see* Note 5).

3.2.3.3. DIGESTION WITH VPD

To the lyophilizate obtained in **Subheading 3.2.3.**, 10 μL of volatile 0.1 *M* triethylammonium carbonate buffer (pH 8.0) or 0.1 *M* Tris-HCl (pH 8.0) con-

taining 50 mU of VPD is added, mixed well, and spun down. It is then incubated for at least 2 h at 37°C. VPD I cuts more easily than nuclease P1 any resistant phosphodiester bond (especially at 2'-*O*-methylated residues), provided that there is a phosphate-free 3'-end. This enzyme allows complete hydrolysis of the RNA into 5'-[³²P]-XMP (see **Note 15**).

3.2.3.4. COMBINED DIGESTIONS WITH NUCLEASE P1 AND VPD

To the nuclease P1 hydrolysate obtained in **Subheading 3.2.3.2.**, 40 µL of 0.1 *M* triethylammonium carbonate buffer (pH 8.0) and 100 mU of VPD is added, mixed well, and spun down. It is then incubated for at least 2 h at 37°C (usually overnight in an incubator hood). In this way, exclusively 5'-monophosphate nucleosides will be produced (see **Notes 5** and **15**).

3.2.3.5. COMPLETE DIGESTION WITH PIPERIDINE

Complete digestion of RNA (20 µg or less) into a mixture of the isomeric forms 2',3' monophosphate nucleosides can be performed under alkaline conditions using piperidine (see **Note 16**). The lyophilized RNA sample (see **Subheading 3.2.3.**) is incubated in 10 µL of 15% piperidine (v/v) for 1 h at 95°C. This type of hydrolysis allows testing for the alkali lability of certain modified nucleotides (see **Note 12**), which then may help to confirm their identities.

3.2.4. TLC Analysis

After digestion, TLC analyses are performed exactly as described in **Subheading 3.1.5.** using either glass or plastic plates coated with cellulose (see **Note 6**). Distribution of 5'-monophosphate nucleosides in the two solvent systems I and II of a nuclease P1 hydrolysate is shown in **Figs. 2–5**, while **Fig. 6** shows the relative positions in solvent system II of the 2'-*O*-methylated dinucleoside derivatives as well as some of the dinucleotides that are resistant or difficult to cleave by nuclease P1 or RNase T2.

Note that the 5'-ribonucleotides generated by nuclease P1 and/or VPD do not migrate at the same rate as their 3'-ribonucleotide counterparts generated by RNase T2 digestion in any of the three solvent systems. However the “relative” migration positions of the different 5'- and 3' derivatives display closely similar global characteristic patterns on the TLC plates. In addition, alkaline treatment of RNA generates a mixture of 2'- and 3'-nucleoside monophosphate that gives characteristic doublets on TLC (**Fig. 7**). Therefore, caution must be taken when interpreting the results from an RNase T2 or alkaline analysis using the reference maps designed for 5'-nucleotides, as in **Figs. 2–5**.

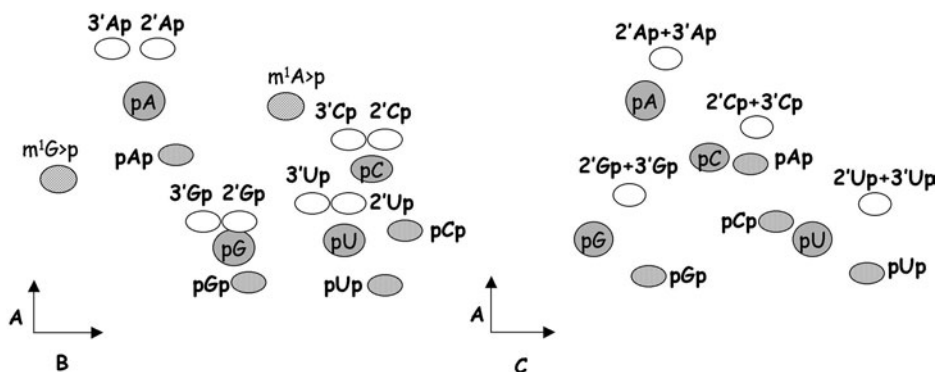


Fig. 7. Effect of phosphate position on chromatographic mobility of monophosphate nucleosides (**Note 16**). pN (N being A, G, C, or A) corresponds to nucleotides with the phosphate group on the 5'-hydroxyl (gray ellipses); 3'- or 2'-Np to the monophosphate nucleosides harboring a phosphate group at position 3' or 2', respectively (open ellipses); and pNp to diphosphate nucleosides harboring one phosphate group on the 5'-hydroxyl and the second one on the 3'-hydroxyl (light gray ellipses). Also shown are the positions for some of the cyclic monophosphate nucleosides (2'-3', designated N>) that often appear on the chromatograms after RNA digestion by RNase T2, especially at low ratio of RNase over the substrate (this happens with other modified nucleotides too). In case of doubt, hydrolyze an aliquot of the same sample with 10 times more RNase T2 and compare the chromatographic profiles.

3.2.5. Evaluation of Stoichiometric Amount of Modified Nucleotides Within an RNA Fragment

If the nucleotide sequence of the uniformly labeled RNA is known, the stoichiometric amount of modified [^{32}P]-XMP can be calculated by measuring the radioactivity found in the corresponding spot of the thin-layer plate, and normalizing this number to the theoretically calculated number of moles of the other nucleotides (modified or not) deduced from the known primary sequence of the RNA fragment. In the case of tRNA, one good reference for the determination of the stoichiometric amount of a given modified nucleotide is the spot corresponding to ribothymidine monophosphate (pT or pm⁵U in **Fig. 5**, except for some tRNAs from higher eukaryotic cells for which nucleotide at position 54 is often pTm or Pm⁵Um). Indeed, ribothymidine corresponds generally to 1 mol/mol tRNA molecule. These quantifications are easily performed after exposure of the TLC plates to a PhosphorImager screen (Image plate) and evaluation of the radioactivity in each individual spot, using ImageQuant

coupled with Excel algorithms. Alternatively, the radioactive compound in each individual spot, as detected after exposure of the TLC plates to X-ray films, can be recovered from the cellulose plates by scraping the cellulose from the glass plate or by cutting off the area of the plastic sheet. After transfer into vials containing scintillation mixture, the radioactivity of the cellulose contained in each individual TLC spot is measured using a scintillation counter (*see Note 18*). Examples of detection of modified nucleotides from uniformly labeled RNA ($[^{32}\text{P}]$, $[^{14}\text{C}]$, or $[^{35}\text{S}]$) extracted from either *E. coli*, *Saccharomyces cerevisiae*, *Halobacterium volcanii*, or *Mycoplasma* can be found in **refs. 42–49** and **101**).

3.3. Identification of Modified Nucleotides in Semisynthetic RNA Radiolabeled at Selected Positions

Naturally occurring, fully modified RNAs isolated from cells are the end products of the complete maturation process and thus cannot usually serve as substrates for testing the activity of a putative RNA-modifying enzyme *in vitro*, except in certain cases of heterologous reactions between enzyme and tRNA from different organisms. However, the development of recombinant DNA and RNA techniques as well as of the chemical and enzymatic synthesis of RNA *in vitro* offers numerous alternatives to obtain synthetic or semisynthetic unmodified or partially modified pre-RNA substrates suitable for the analysis of enzymatic posttranscriptional RNA modifications both *in vitro* and *in vivo* (*see refs. 31, 32, and 50, and references therein*).

The first methodology developed in that direction was based on the earlier work of Bruce and Uhlenbeck (**51**). They showed that it was possible to replace *in vitro* several nucleotides in the anticodon loop of naturally occurring tRNA molecules by any of the four canonical nucleotides. The method involved the use of various nucleases, phosphatases, T4-PNK, and T4-RNA ligase. We originally used this “cut and paste” technology in order to introduce a $[^{32}\text{P}]$ -labeled phosphate 5′ or 3′ adjacent to the nucleoside of interest, simultaneously to anticodon replacement in tRNA (reviewed in **ref. 32**). This site-specific $[^{32}\text{P}]$ -radiolabeled tRNA substrate was then incubated in the presence of an enzymatic preparation or microinjected into *Xenopus* oocytes. After phenol extraction and purification by gel electrophoresis of the RNA, the $[^{32}\text{P}]$ -labeled RNA was further subjected to complete digestion with nuclease P1 or RNase T2 (as described in **Subheadings 3.2.3.1.–3.2.3.5.**) and analyzed by 2D-TLC on cellulose plates (*see Subheadings 3.1.5. and 3.2.4.*). After autoradiography, only the radiolabeled 5′- or 3′- $[^{32}\text{P}]$ -nucleotide and its modified derivative(s) appear as distinct and unambiguous spots on the TLC plates. The radioactivity in the spot corresponding to the modified nucleotide(s) over the sum of the radioactivity found in the different spots on the TLC plate (except for

[^{32}P]-orthophosphate) gives a direct value for the stoichiometric amount of modified nucleotide within the RNA molecule and allows measurement of the degree of advancement of an enzymatic modification reaction *in vitro* or *in vivo* (reviewed in **ref. 32**).

To date, site-specific introduction of a [^{32}P] label (or any other radiolabeled nucleotide) into RNA is much more easily performed using T4-DNA ligase than T4-RNA ligase, provided that the two RNA fragments, of which one is radiolabeled at its 5' end, are brought together using an appropriate splint DNA (**33,34,50**). These "cut-and-paste" methodologies, using either T4-RNA ligase or T4-DNA ligase, have been used successfully to study the enzymatic formation of various modified nucleotides in U2-small nuclear RNA (snRNA) and tRNA, including hypermodified nucleotides located at positions 34 and 37 of the anticodon loop (for examples *see* **refs. 51–57**).

3.4. Identification of Modified Nucleotides in Radiolabeled *In Vitro* Transcripts After Incubation With RNA Modification Enzymes

An alternative method to produce radiolabeled RNA that can be used as substrate for RNA-modifying enzymes *in vitro* or *in vivo* is to use [^{32}P]-radiolabeled *in vitro* runoff transcripts. These RNAs are obtained after transcription of appropriate linearized plasmids or synthetic DNA templates containing a bacteriophage T7 or SP6 promoter upstream of the sequence of interest in the presence of one of the four [α - ^{32}P]-nucleoside triphosphates. Thus, using four independent samples of the same transcript, each being radiolabeled with one of the four [α - ^{32}P]-XTPs, one can in principle identify the nucleotide (modified or not) that is 5' adjacent to a given type of nucleotide in the RNA substrate ("nearest-neighbor analysis"). Indeed, after incubation with the RNA-modifying enzyme, hydrolysis of the recovered RNA by nuclease P1 will generate 5'-monophosphate nucleosides (including [^{32}P]-XMP of the same type as the one used during transcription), while hydrolysis of the same RNA transcript by RNase T2 will generate 3'-monophosphate nucleosides of which only those that were 5'-adjacent to the [^{32}P]-XMP will be radiolabeled (*see* **Fig. 1**). Again, instead of [^{32}P]-label, any other types of radiolabeled ([^{14}C] or [^3H]) nucleotide triphosphate can be used during *in vitro* transcription. In addition, uniformly labeled [^{32}P]-RNA as obtained in **Subheading 3.2**. can be used satisfactorily for detection of the activity of a given modifying enzyme under *in vitro* conditions.

Following are a few examples in which this methodology has been used successfully to detect the activity of several tRNA-modifying enzymes both *in vitro* and *in vivo*, the latter after microinjection of the radiolabeled transcripts into *Xenopus* oocytes (**58–64**). In these studies, TLC analyses using cellulose plates as described in **Subheadings 3.1.5**. and **3.2.4**. were central to the methods used to evaluate and quantify modified nucleotides in RNA.

3.5. RNA Sequencing and Localization of Modified Nucleotides

The recommended procedure to sequence RNAs is the single-hit random labeling-sequencing method (7–11). It allows the sequencing of the primary structure as well as the detection, identification, and localization of most (if not all) modified nucleotides within a pure RNA molecule. It is based on the [^{32}P]-postlabeling and 2D-TLC procedure that we have described. Here, the purified RNA (at least 95% pure, maximum 50–150 nt in length, only micrograms of RNA sample required) is first damaged at high temperature (95°C) under experimental conditions that allow only one phosphodiester bond to be hydrolyzed on average per RNA molecule. The resulting fragments are labeled with [^{32}P] at their 5' end with PNK and [γ - ^{32}P]-ATP as described in **Subheading 3.1.2**. The radiolabeled oligonucleotides are then separated by 1D electrophoresis on a denaturing polyacrylamide gel (50–80 cm in length) or by 2D electrophoresis (at acidic pH in the first dimension and at neutral pH in the second dimension; *see refs. 11 and 29*). After separation of the oligonucleotides at “single-nucleotide resolution,” the oligonucleotide in each of the bands or the spots of the gel is eluted and the 5'-[^{32}P]-end nucleotide of each of the labeled fragments is determined after complete nuclease P1 or VPD digestion followed by 2D-TLC (as described in **Subheadings 3.2.3.2. and 3.2.3.4.**). Since the fragments are separated on the gel according to their length, analysis of their ends depicts not only the nature but also the position of each individual nucleotide in the studied sequence (for details, *see refs. 7, 9, and 10*).

3.6. Use of Reverse Transcriptase to Locate Modified Nucleotides in RNA

Besides the direct sequencing of pure RNA as pointed out in **Subheading 1.**, and the methodologies described in **Subheading 3.**, there is at least one other powerful method to localize certain modified nucleotides in RNA. It is based on the use of reverse transcriptase (usually avian myeloblastosis virus reverse transcriptase [RT] and specific 5'-end-[^{32}P]-labeled oligonucleotide primers (15–20 nt in length) that are complementary to an RNA present in a mixture of unfractionated RNAs (primer extension analysis of RNA; [65]). It is similar to the one described in other chapters in this book for identification of editing sites in RNA molecules (*see, e.g., Chapters 13 and 14*) and characterization of snoRNA (*see Chapter 20*). Here the method takes advantage of the fact that certain nucleotides, when modified at atom positions that interfere with the Watson-Crick pairings, block or slow down (especially when low concentrations of dNTPs are used) the progress of reverse transcription at the nucleotide preceding the modified residue. This is the case for modifications at $N^1\text{-A}$, $N^6\text{-A}$, $N^1\text{-G}$, $N^2\text{-G}$, $N^3\text{-U}$, $N^3\psi$, and $N^3\text{-C}$, such as m^1A , m^1I , m^6_2A , m^1G , m^2G , m^2_2G ,

m^3U , acp^3U , $m^3\psi$, $m^1acp^3\psi$, and “bulky” hypermodified bases, such as wyosine and wybutosine derivatives ms^2t^6A . The existence of such “strong stops” or “pauses” of chain elongation, as identified after autoradiography of 1D-denaturing gel electrophoresis of the *in vitro* transcription products, may only indicate the presence of a putative modified nucleotide. Indeed, a pause or a strong stop can also result from stable secondary structure of the RNA (“false positive”) or any other physical parameter of the RNA molecules that interferes with the normal progression of RT along the RNA molecules. Moreover, depending on the “degree” of modification of a given nucleotide (partially or fully modified) within the RNA population, the arrest of RT is not always easily detectable and appropriate controls always have to be done (for examples *see* **refs. 57 and 65–73**; partly reviewed in **ref. 32**). Reverse transcription has also been used to detect oxidative damage in RNA (**74**).

Localization of a given modified nucleotide within an RNA population by reverse transcription can also be obtained either after introducing a bulky group on a given modified nucleotide, such as *N*-cyclohexyl-*N'*- β -(4-methylmorpholinium) carbodiimide on N^3 of pseudouridine (**75–77a**), or after scission of the phosphodiester backbone of RNA resulting from chemical alteration of m^7G , dihydrouridine (D), or wybutosine of tRNA (**78–82**). In the first case, the progress of the polymerase is impaired by the inability of the template nucleotide to base pair. In the second case, cDNA synthesis stops when the polymerase falls off the template RNA. Here, again, detection of the modified bases within an RNA population will depend on whether the chemical treatment changing the structure of the base and/or inducing selective cleavage of the phosphodiester bond can be performed without affecting the stability of the phosphodiester bonds at other sites of the tRNA molecule. The main advantage of the primer reverse extension assays, as compared to all the other methods that we have described, is that it can be used with completely unfractionated RNA of any length.

3.7. Determination of Modified Deoxynucleotides in DNA

For determination of modified deoxynucleotides in DNA, the postlabeling procedure followed by cellulose TLC as previously described can also be applied with slight accommodations. In this case, the digestion is performed using a mixture of micrococcal nuclease (0.2–0.4 U) and spleen phosphodiesterase (0.5–1.0 μ g), which liberates 3'-monophosphate deoxynucleosides, in 3 μ L of 40 mM Na succinate (pH 6.0) and 20 mM $CaCl_2$ for 4 h at 37°C. This step replaces the RNase T2 hydrolysis used in the case of RNA. The next steps are identical to the ones described for RNA digests, except that for the first solvent we prefer the composition isobutyric acid: NH_4OH :water (66:1.1:28.9 [v/v/v]) when good separation between pdC and pdm⁵C is required for subse-

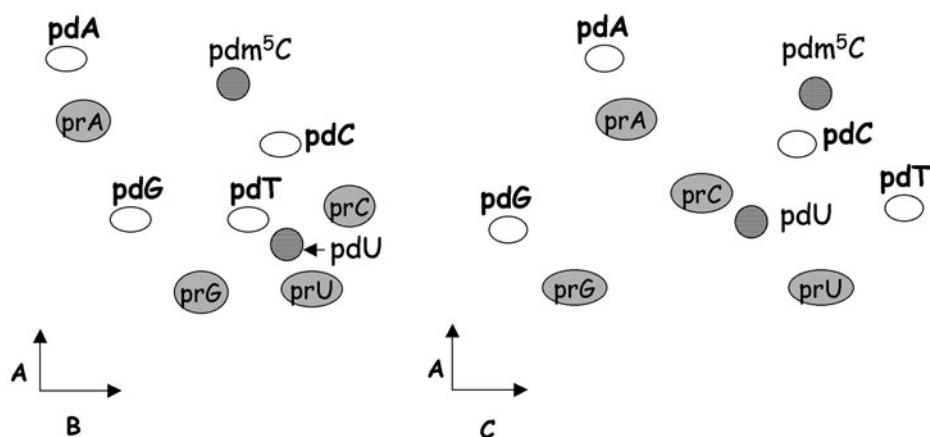


Fig. 8. Comparison of chromatographic mobility of 5'-monophosphate nucleosides derived from *ribonucleotides* (prA, prG, prC, and prU as in **Figs. 2–7**) and corresponding *deoxyderivatives* (pdA, pdG, pdC, and pdT). Positions for the 5'-monophosphate methyldeoxycytosine (pdm⁵C), a naturally occurring modified deoxyribonucleotide in higher eukaryotic DNA and of 5'-monophosphate deoxyuridine (pdU) that occurs spontaneously by deamination of dC, are also shown.

quent accurate quantification of the modified nucleotide. **Figure 8** illustrates the relative migration patterns of the deoxyribonucleotides dA, dC, dG, dT, m⁵dC, and dU of DNA as well as of ribonucleotide rU: m⁵dC is a naturally occurring modified nucleotide in higher eukaryotic DNA (83) while dU is often present because of naturally occurring deamination of dC (84). Only solvent system I (A + B) can be used for separating the four major labeled deoxynucleotides, pdm⁵C, and pdU from the DNA digest. Solvent system II is not recommended for [³²P]-labeled 5' deoxyribonucleotide separations because, unlike nonradioactive DNA, [³²P]-labeled 5'-purine deoxyribonucleotides are very unstable in the acidic solvent C, probably owing to a strong induced radiolytic depurination of these labeled nucleotides (85). There is also slight depurination in the case of ribonucleotides during migration, especially in solvent C, but to a much smaller extent than for deoxyderivatives. Finally, combined digestion with micrococcal nuclease and spleen phosphodiesterase can also be used to detect adduct formation in DNA of organisms that have been exposed to xenobiotic toxic compounds. The method uses a combination of enzymatic treatments to favor formation of the adducts vs the major nucleotides, followed by separation using a TLC procedure on polyethyleneimine-cellulose. This very sensitive method allows the detection of adducts at a level of 10⁻⁹–10⁻¹⁰ mol

adduct/mol of nt, which corresponds to about one adduct per human genome (for more information, see refs. 86–88).

4. Notes

1. As a rule, all RNase-free RNA samples to be subjected to various nuclease digestions have to be free of detergent, phenol, and, as much as possible, salts that can interfere with subsequent enzyme reactions or ruin the quality of the TLC pattern. This is usually obtained by precipitating the RNA by ethanol, washing two to three times with ethanol/water (70% [v/v]), and dissolving in DEPC-treated water. The purity of the RNA sample is of utmost importance for accurate nucleotide sequencing and/or base composition analysis by postlabeling technique as described in **Subheadings 3.1.–3.5.**). Only the primer reverse extension assay (**Subheading 3.6.**) analysis can be performed on unfractionated RNA molecules.
2. RNase T2 (from *A. oryzae*) is a nonspecific ribonuclease that cleaves the phosphodiester bonds of any single-stranded ribonucleic acid to generate 3'-monophosphate nucleosides. It does not cut a dinucleotide in which the 2'-hydroxyl of the ribose moiety is occupied by a methyl (CH₃) group or another group like ribose phosphate, as found in the case of 2'-*O*-ribosylguanosine phosphate (Grp) and 2'-*O*-ribosyladenosine phosphate (Arp) (89,90). RNase T2 activity is also slowed down in the case of certain methylations in the base moiety, such as m¹G, m²₂G, m⁷G, or wybutosine, leading to 2',3'-cyclic monophosphate nucleoside intermediates that could be confused for another modified nucleotide on the TLC plates. To circumvent this problem, one could either use 10 times more enzyme and digest for a longer time (overnight) or treat the digest overnight at pH 1.0 (in 0.1 M HCl) to open the cyclic structure before TLC analysis. However, in the latter condition some modified bases are unstable (see **Note 11**). We found that the enzyme extract prepared from takadiastase powder of *A. oryzae* following the procedure described in **ref. 91**) cleaves RNA more efficiently than the commercial RNase T2. This is probably owing to contamination of RNase T2 by other RNase activities (e.g., RNase T1, RNase A). The enzyme from *A. oryzae* is very easy to prepare and can be stored in water at –20°C for years. Alternatively, a mix of commercial RNase T2 + RNase T1 + RNase A can also be used.
3. Recombinant T4-PNK (PNK from T4-infected *E. coli*) allows transfer of the γ -phosphate of ATP to the 5' end of any 3'-monophosphate nucleoside or nucleic acid fragment bearing a free 5'-hydroxyl group. However, nonrecombinant PNK was reported to have a 3'-phosphatase activity (92), which probably results from contamination of the PNK. Note that the optimal pH for each of these two activities is different, the kinase reaction being optimal between pH 7.0 and 9.0, while the phosphatase activity is optimal between 5.0 and 7.0. When [γ -³²P]-ATP is used, it transfers the [γ -³²P] into the (oligo)nucleotide. More important, PNK does not show equal efficiency of [γ -³²P]-incorporation into each of the different mononucleotides, nor for individual modified nucleotides. Thus, determination of the relative amount of nucleotides (modified or not) will not be possible using

the postlabeling method. Finally, the amount of “cold” ATP to be added to the reaction mixture should correspond roughly to a slight excess of the molar amount of nucleotides of the starting RNA to be labeled. This allows the labeling of all nucleotides and reduces significantly the amount of unused [γ - ^{32}P]-ATP that will have to be destroyed at the next step of the procedure.

4. Apyrase from potato (*S. tuberosum*) is a mixture of ATPase and ADPase that catalyzes the degradation in two steps of ATP into ADP + Pi and ADP into AMP + Pi. Unused [γ - ^{32}P] from ATP in the postlabeling mix will therefore become free [^{32}P]-orthophosphate, which will migrate ahead of nucleotides with the front of the solvents B and C (not A) during chromatography. Commercial apyrase often contains phosphatase activities that are harmful for the labeled nucleotides if very tiny quantities of postlabeled material are handled. This is why phosphate buffer and carrier 5'-monophosphate nucleosides are routinely added to the reaction mix in order to limit dephosphorylation of the [^{32}P]-nucleotides. This trick may serve for any other reaction in which undesired phosphatase activities are suspected. In the procedure described earlier (8), [γ - ^{32}P] from radiolabeled ATP was selectively removed by catalyzing the formation of glucose-6-[^{32}P]phosphate from glucose in the presence of yeast hexokinase. The glucose-6-P migrates almost with the front of solvents C and B, thus far from of all nucleic acid derivatives.
5. Nuclease P1 (from *P. citrinum*) has two activities: as an endonuclease it cleaves RNA and DNA into the corresponding 5'-monophosphate nucleosides, and as a 3'-nucleotidase it removes the phosphate groups at the 3' end of nucleotides (such as nonrecombinant/commercial PNK). However, when 2'-*O*-methylnucleotides or bulky nucleotides (e.g., wye, queuosine derivatives) are present in the RNA molecule, its activities are heavily reduced. As a result, dinucleoside diphosphates of the type pXmpY may be present in the P1-hydrolysate. To circumvent this problem, increasing the amount of enzyme in the reaction mix (10 times) and increasing the incubation time (overnight), or coupling the P1 reaction with VPD (see **Subheading 3.2.3.4.**), generally allows the generation of exclusively 5'-monophosphate nucleosides. However, this drawback can become an advantage when 2'-*O*-methyl dinucleosides, or heavily modified nucleosides, have to be identified by comparing hydrolysates resulting from RNA digestion at low P1/RNA ratio vs digestion at high P1/RNA ratio.
6. It is preferable to use glass plates when large (20 × 20 cm) TLC has to be used because plastic plates are difficult to hold in the rills of the multiplack tanks. However, if previously used glass TLC plates are treated to remove the cellulose layer, the clean glass plates can be used to support the plastic-coated plates in the chromatographic tanks. The plastic plates can be either stuck on the glass plates using small double-sided tape squares or simply fixed by humidifying the glass plates with the chromatographic solvent. One advantage of plastic-coated plates is that they can be cut into smaller sizes (10 × 10 cm or even 6 × 6 cm). When the chromatographic pattern of the modified nucleotides to be analyzed is not too complex, this allows one to perform 2D chromatography in only a few hours for

each dimension, instead of 2 d for the large 20 × 20 cm plates. In some cases, the identification of a given modified nucleotide is even possible by 1D chromatography using solvent B and/or C. Of the commercially available TLC plates, G-1440 Avicel 400 plates from Schleicher & Schuell are highly recommended, but they are unfortunately temporarily unavailable. Those from Macherey-Nagel and Merck are satisfactory, the latter supporting slower solvent migration rates than the former ones. Do not use plates with UV indicator because the latter acquires a strong dark blue when solvent C is used, which hinders detection of marker controls by UV light.

7. The 5'-monophosphate nucleosides mix used as chromatographic marker mix can be either of commercial origin or a homemade enzymatic nuclease P1 digest of bulk tRNA or rRNA of any origin (*see Subheading 2.2.*). When performing TLC of tRNA hydrolysates, it is of course preferable to use markers derived from tRNA hydrolysates, such as the baker's yeast tRNA, because it contains modified nucleotides. Ideally, 10–20 µg of total ribonucleotides should be spotted on each TLC plate. More than 40 µg will produce smears during migration. Use a moderately warm air stream from a hair dryer between successive applications. Small spots are important for good 2D separation of the nucleotides.
8. We suggest using Cerenkov counting that allows quick counting of [³²P] in the absence of scintillation liquid. The yield of Cerenkov counting (tritium channel on the scintillation counter) is as high as 30% compared with scintillation counting. It has the advantage of being a “dry” method and allows further analysis after elution of the radioactive material from the cellulose with water. Counting is done directly by hanging the Eppendorf tube containing the cellulose powder or the piece of cut plastic sheet with radioactive material into the counting flasks.
9. To avoid the influence of isobutyric acid solvent on migration of nucleotides in the second dimension, be sure that all of the acid has evaporated (as tested from its smell) before starting the migration of the second dimension (eventually use a hair dryer with moderately warm air under a ventilated hood). Information about migration of the nucleotides in the first dimension can be obtained by quick inspection of the plates under UV light with a filter at 254 µm. Normally three dark blue spots should be seen corresponding to [GMP + UMP], CMP, and AMP (from bottom to top of the migration line).
10. In both types of chromatographic solvents, AMP, CMP, and UMP appear as dark blue spots while GMP yields a light blue fluorescent spot in solvent C. Normally, 3–10 µg of a nucleotide can be adequately visualized under UV light. Use a marker pen if needed to indicate positions of these spots on the back of the plates, or a soft pencil on the cellulose plates.
11. Caution should be taken when comparing data with the established reference maps. Indeed, one should remember that the maps were generated by compiling data from many independent analyses in various laboratories. Therefore, the true relative position of each spot on the TLC plates in a given experiment could be slightly different, as compared to the reference maps that have to be used as a

“guide.” The majority of the modified nucleotide monophosphates identified in **Figs. 2–5** have never migrated together, because they do not coexist in the same RNA molecule, or in RNAs of the same organism. Conclusive identification of a modified nucleotide should be based on chromatographic comigration with the corresponding UV markers (unfortunately, most modified nucleotide markers are not commercially available). Slight differences in migration of individual spots on the TLC plates can also arise from differences in “room” temperature, incomplete saturation of the tank with vapor, impurities present in the chemicals used to prepare each solvent, the origin of the commercial cellulose layer, and the age of the solvents. As a rule, do not use the solvents more than three to four times. Whatever the exact positions, the pattern of the different spots corresponding to modified nucleotides is always relative to the four major nucleotides as detected by autoradiography or by visual inspection under UV light.

12. Solvent C is strongly acidic (pH below 1.0) and causes depurination of the nucleotides during migration, leading to characteristic [^{32}P] smears ahead of the migrating spots, with the corresponding degradation products (ribosylphosphate and inorganic phosphate) migrating faster than the parent nucleotides. In addition, some modified nucleotides are known to be unstable in contact with acid (pH below 2.0) as well as in alkaline solution (pH 8.0 or higher). Some of them give rise to new stable compounds that were not present in the original RNA samples and give “abnormally” located spots on TLC. Among them are m^1A , which gives m^6A ; m^3C , which gives m^3U ; s^4U , which gives U or C; and ac^4C , which gives C. Some of the hypermodified nucleotides such as wybutosine give more complicated, as-yet unidentified derivatives (*see, e.g., ref. 53*). Other alkali-sensitive nucleotides (pH 9.0 or higher) are D, m^1I , m^7G , sulfur-containing nucleotides such as ms^{216}A or ms^{216}A , Wybutosine derivatives, and most queuosine derivatives, while m^5C and m^4C , e.g., are thermolabile (the list not being exhaustive). Such unstable nucleotides can be identified by comparing 2D chromatograms of RNA hydrolysates that were exposed or not for a few minutes at high temperature (80°C or higher), to high pH (9.0 or higher), low pH (2.0 or lower), or in the presence of H_2O_2 (which selectively oxidizes thiopyrimidine derivatives such as s^2U , s^2C , and s^4U ; [93]).
13. The two chromatographic solvent systems described in this chapter to analyze nucleotide digests are the most commonly used ones but there are others (*see refs. 94 and 95 and references therein*). Alternative methods are based on analyzing the pattern of modified *nucleosides* instead of *nucleotides*. To this end, the nucleosides of an RNA digest obtained by combined digestion with nuclease P1 or T2 and bacterial or intestinal alkaline phosphatase are postlabeled with tritium after periodate oxidation of the 2',3'-hydroxyl group of the ribose and reacted with tritiated borohydride, followed by analysis by TLC (99–100). High resolution of nonlabeled nucleosides by TLC has also been described using special chromatographic solvent systems but requires at least 200 μg of a pure tRNA digest to

detect any minor modified nucleoside by its position on the TLC plate and to establish its UV spectrum (**101**). However, with nucleosides, the method of choice is certainly the HPLC technique coupled with UV and MS (*see refs. 1–24b*). This technique requires at each run a minimum of 20 µg of pure RNA or RNA fragment per test depending on the relative amount of modified nucleotide(s) in the RNA sample (two or three tests should be made). The main advantage of liquid chromatographic techniques is that both the UV spectra and the mass spectra of the modified nucleosides (possibly also detection of the radioactivity) can be recorded “on-line” during chromatography, which allows identification of the chemical structure as well as quantification of the compound(s) of interest.

14. The addition of carrier RNA is important when a very low amount of labeled RNA has to be hydrolyzed. Carrier will protect against contaminating RNase and/or phosphatase activities that could be harmful to the result, especially if the enzyme-to-RNA ratio is high (*see also Note 4*). It also helps migration of nucleotides and serves as control spots on TLC for UV shadowing.
15. VPD from *Crotalus* snake is an exonuclease that efficiently cleaves RNA and DNA into 5'-monophosphate nucleosides. It allows easier cleavage of the 2'-*O*-methyl-containing nucleotide bonds as well as any other bonds adjacent to nucleotides that are resistant to nuclease P1 or RNase T2. Nevertheless, it is slightly slowed down at pseudouridine residues, *N*⁶-aminoacylated nucleosides, and ribose-methylated nucleosides.
16. Piperidine should be preferred over hydroxides (NaOH or KOH) because it is volatile and evaporates completely when spotting on the TLC plate. It will not destroy the cellulose layer, as do hydroxide compounds, and avoids Na⁺ or K⁺ salts, which favor smearing of nucleotides during migration on thin-layer plates. Because the 2'- and 3'-monophosphate nucleosides do not migrate at the same rate in chromatographic system I (A + B), one always obtains two “tandem” spots for each nucleotide after chromatography, one migrating as expected for 3'-phosphate derivatives (as obtained after RNase T2 hydrolysis) and the other migrating nearby (2'-derivative; *see Fig. 7*).

Acknowledgments

We thank Jean-Pierre Waller for advice and improvements on the manuscript. We also thank Juan Alfonzo from Ohio State University, Dolf Hatfield from the National Institutes of Health, and Tom Suzuki from University of Tokyo for advice on establishing the final diagrams in **Figs. 2–5**). This work was supported by research grants from the Centre National de la Recherche Scientifique (CNRS); the Ministère de l'Éducation Nationale, la Recherche Scientifique et de la Technologie (Programme Interdépartemental de Géomicrobiologie des Environnements Extrêmes and Programme de Recherche Fon-

damentale en Microbiologie et Maladies Infectieuses et Parasitaires); the Fonds pour la Recherche Fondamentale Collective, the French Community of Belgium (Actions de Recherches concertées"); and the CNRS. L. D. is a research associate of the Fonds National de la Recherche Scientifique.

References

- 1 Rozenski, J., Crain, P. F., and McCloskey, J. A. (1999) The RNA modification database—1999 update. *Nucleic Acids Res.* **27**, 196–197.
- 2 Sprinzl, M., Horn, C., Brown, M., Ioudovitch, A., and Steinberg, S. (1998) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* **26**, 148–153.
- 3 Limbach, P. A., Crain, F. C., and McCloskey, J. A. (1994) Summary: the modified nucleosides of RNA. *Nucleic Acids Res.* **12**, 2183–2196.
- 4 Brownlee, G. G. (1967) Determination of sequences in RNA, in *Laboratory Techniques in Biochemistry and Molecular Biology* (Work, T. S. and ork, E. E., eds.), North-Holland/American Elsevier, Amsterdam.
- 5 Sanger, F. and Brownlee, G. G. (1967) A two dimensional fractionation method for radioactive nucleotides, in *Methods in Enzymology*, vol. XII (Grossman, L. and Moldave, K., eds.), Academic, New York, pp. 361–381.
- 6 Szekely, M. and Sanger, F. (1969) Use of polynucleotide kinase in fingerprinting non-radioactive nucleic acids. *J. Mol. Biol.* **43**, 607–617.
- 7 Stanley, J. and Vassilenko, S. (1978) A different approach to RNA sequencing. *Nature* **274**, 87–89.
- 8 Silberklang, M. O., Gillum, A. M., and RajBhandary, U. L. (1979) Use of *in vitro* ³²P labeling in the sequence analysis of non-radioactive tRNAs, in *Methods in Enzymology*, vol. LIX (Moldave, K. and Grossman, L., eds.), Academic, New York, pp. 58–109.
- 9 Kuchino, Y., Hanyu, N., and Nishimura, S. (1987) Analysis of modified nucleosides and nucleotide sequence of tRNA, in *Methods in Enzymology*, vol. 155 (Wu, R., ed.), Academic, New York, pp. 379–396.
- 10 Nishimura, S. and Kuchino, Y. (1983) Characterization of modified nucleosides in tRNA, in *Methods of DNA and RNA Sequencing* (Weissman, M. S., ed.), Praeger, New York, pp. 235–255.
- 11 Keith, G. (1990) Nucleic acid chromatographic isolation and sequence methods, in *Chromatography and Modifications of Nucleosides*, vol. 45A, Journal of Chromatography Library Series (Gehrke, C. W. and Kuo, K. C., eds.), Elsevier, The Netherlands, pp. A103–A141.
- 12 McCloskey, J. A. (1990) Constituents of nucleic acids: overview and strategy, in *Methods in Enzymology*, vol. 193 (McCloskey, J. A., ed.), Academic, New York, pp. 796–824.
- 13 Kowalak, J. A., Pomerantz, S. C., Crain, P. F., and McCloskey, J. A. (1993) A novel method for the determination of posttranscriptional modification in RNA by mass spectrometry. *Nucleic Acids Res.* **21**, 4577–4585.

- 14 Rozenski, J. and McCloskey, J. A. (1999) Determination of nearest neighbors in nucleic acids by mass spectrometry. *Anal. Chem.* **71**, 454–459.
15. Crain, P. F. (1998) Detection and structure analysis of modified nucleosides in RNA by mass spectrometry, in *Modification and Editing of RNA* (Grosjean, H. and Benne, R., eds.), ASM Press, Washington, DC, pp. 47–57.
- 16 Buck, M, Connick, M., and Ames, B. N. (1983) Complete analysis of tRNA-modified nucleosides by high-performance liquid chromatography: the 29 modified nucleosides of *S. typhimurium* and *E. coli* tRNA. *Anal. Biochem.* **129**, 1–13.
- 17 Gehrke, C. W. and Kuo, K. C. (1989) Ribonucleoside analysis by reversed-phase high-performance liquid chromatography. *J. Chromatogr.* **471**, 3–36.
18. Gehrke, C. W. and Kuo, K. C. (1990) Ribonucleoside analysis by reverse-phase high performance liquid chromatography, in *Chromatography and Modifications of Nucleosides*, vol. 45A, Journal of Chromatography Library Series (Gehrke, C. W. and Kuo, K. C., eds.), Elsevier, The Netherlands, pp. A3–A71.
19. Pomerantz, S. C. and McCloskey, J. A. (1990) Analysis of RNA hydrolyzates by liquid chromatography-mass spectrometry, in *Methods in Enzymology*, vol. 193 (McCloskey, J. A., ed.), Academic, New York, pp. 796–824.
- 20 Polson, A. G., Crain, P. F., Pomerantz, S. C., McCloskey, J. A., and Bass, B. L. (1991) The mechanism of adenosine to inosine conversion by double-stranded RNA unwinding/modifying activity: a high-performance liquid chromatography-mass spectrometry analysis. *Biochemistry* **30**, 11,507–11,514.
- 21 Xue, H., Glasser, A. L., Desgres, J., and Grosjean, H. (1993) Modified nucleotides in *Bacillus subtilis* tRNA-trp hyperexpressed in a *E. coli*. *Nucleic Acids Res.* **21**, 2479–2486.
- 22 Takeda, N., Nakamura, M., Yoshizumi, H., and Tatematsu, A. (1994) Structural characterization of modified nucleosides in tRNA hydrolysates by frit-fast atom bombardment liquid chromatography/mass spectrometry. *Biol. Mass Spectrom.* **23**, 465–475.
- 23 Dalluge, J. J., Hashizume, T., and McCloskey, J. A. (1996) Quantitative measurement of dihydrouridine in RNA using isotope dilution liquid chromatography-mass spectrometry (LC/MS). *Nucleic Acids Res.* **24**, 3242–3245.
- 24 Tanigushi, H. and Hayashi, N. (1998) A liquid chromatography/electrospray mass spectrometric study on the posttranscriptional modification of tRNA. *Nucleic Acids Res.* **26**, 1481–1486.
- 24a. Kirpekar, F., Douthwaite, S., and Roepstorff, P. (2000) Mapping posttranscriptional modifications in 5S ribosomal RNA by MALDI mass spectrometry. *RNA* **6**, 296–306.
- 24b. Madsen, C. T., Mengel-Jorgensen, J. M., Kirpekar, F., and Douthwaite, S. (2003) Identifying the methyltransferases for m5U747 and m5U1939 in 23S rRNA using MALKDI mass spectrometry. *Nucleic Acids Res.* **31**, 4738–4746.
- 25 Keith, G. (1995) Mobilities of modified ribonucleotides on two-dimensional cellulose thin-layer chromatography. *Biochimie* **77**, 142–144.
26. Nishimura, S. (1979) Chromatographic mobilities of modified nucleotides, in *Transfer RNA, Structure, Properties, and Recognition* (Schimmel, P. R., Söll, D.,

- and Abelson, J. N., eds.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 547–552.
- 27 Hashimoto, S., Sakai, M., and Muramatsu, M. (1975) 2'-O-methylated oligonucleotides in ribosomal 18S and 28S RNA of a mouse hepatoma, MH134. *Biochemistry* **14**, 1956–1964.
 - 28 Holley, R. W., Apgar, J., and Merrill, S. H. (1961) Evidence for the liberation of a nuclease from human fingers. *J. Biol. Chem.* **236**, PC42, PC43.
 - 29 Ikemura, T. (1989) Purification of RNA molecules by gel techniques, in *Methods in Enzymology*, vol. 180 (Moldave, K. and Grossman, L., eds.), Academic, New York, pp. 14–25.
 - 30 Tanner, K. (1990) Purifying RNA by column chromatography, in *Methods in Enzymology*, vol. 180 (Moldave, K. and Grossman, L., eds.), Academic, New York, pp. 25–29.
 - 31 Cedergren, R. and Grosjean, H. (1987) RNA design by *in vitro* RNA recombination and synthesis. *Biochem. Cell. Biol.* **65**, 677–692.
 - 32 Grosjean, H., Motorin, Y., and Morin, A. (1998) RNA-modifying and RNA-editing enzymes: methods for their detection, in *Modification and Editing of RNA* (Grosjean, H. and Benne, R., eds.), ASM Press, Washington, DC, pp. 21–46.
 - 33 Moore, M. J. and Query, C. C. (1998) Uses of site-specifically modified RNAs constructed by RNA ligation, in *RNA-Protein Interactions: A Practical Approach* (Smith, C., ed.), IRL, Oxford, UK, pp. 75–108.
 - 34 Yu, Y. T. and Steitz, J. A. (1997) A new strategy for introducing photoactivable 4-thiouridine (S⁴U) into specific positions in a long RNA molecule. *RNA* **3**, 807–810.
 - 35 Paul, M. S. and Bass, B. L. (1998) Inosine exists in mRNA at tissue-specific levels and is most abundant in brain mRNA. *EMBO J.* **17**, 1120–1127.
 - 36 Warner, J. R. (1991) Labeling of RNA and phosphoproteins in *S. cerevisiae*, in *Methods in Enzymology*, vol. 194 (Guthrie, C. and Fink, G. R., eds.), Academic, New York, pp. 423–428.
 - 37 Landy, A., Abelson, J., Goodman, H. M., and Smith, J. D. (1967) Specific hybridization of tyrosine tRNA with DNA from transducing bacteriophage Phi80 carrying the amber suppressor gene suIII. *J. Mol. Biol.* **29**, 457–471.
 - 38 Monier, R., Stephenson, M. L., and Zamecnick, P. C. (1960) The preparation and some properties of a low molecular weight RNA from Baker's yeast. *Biochem. Biophys. Acta* **43**, 1–8.
 - 39 Brubacker, L. H. and McCorquodale, D. J. (1963) The preparation of amino acid-tRNA from *Escherichia coli* by direct phenol extraction of intact cells. *Biochem. Biophys. Acta.* **76**, 48–53.
 - 40 Tsurui, H., Kumazawa, Y., Sanokawa, R., Watanabe, Y., Kuroda, T., Wada, A., Watanabe, K., and Shirai, T. (1994) Batchwise purification of specific tRNAs by solid-phase DNA probe. *Anal. Biochem.* **221**, 166–172.
 - 41 Mörl, M., Dörner, M., and Pääbo, S. (1994) Direct purification of tRNAs using oligonucleotides coupled to magnetic beads, in *Advances in Biomagnetic*

- Separation* (Uhlén, M., Hornes, E., and Olsik, O., eds.). Eaton, Natik, MA, pp. 107–111.
- 42 Sakano, H., Shimura, Y. and Ozeki, H. (1974) Selective modification of nucleosides of tRNA precursors accumulated in a temperature sensitive mutant of *E. coli*. *FEBS Lett.* **48**, 117–121.
- 43 Nishikura, L. and De Robertis, E.M. (1981) RNA processing in microinjected *Xenopus laevis* oocytes: sequential addition of base modification in a spliced tRNA. *J. Mol. Biol.* **154**, 405–420.
- 44 Sullivan, M. A., Cannon, J. F., Webb, F. H., and Bock, R. M. (1985) Anti-suppressor mutation in *E. coli* defective in biosynthesis of 5-methylaminomethyl-2-thiouridine. *J. Bacteriol.* **161**, 368–376.
- 45 Adanchi, Y., Yamao, F., Muto, A., and Osawa, S. (1989) Codon recognition patterns as deduced from sequences of the complete set of tRNA species in *Mycoplasma capricolum*. *J. Mol. Biol.* **209**, 37–54.
- 46 Laten, H. M., Cramer, J. H., and Rownd, R. H. (1983) Thiolated nucleotides in yeast transfer RNA. *Biochem. Biophys. Acta* **741**, 1–6.
47. Cavaillé, J., Chetouani, F., and Bachellerie, J. P. (1999) The yeast *S. cerevisiae* YDL112w ORF encodes the putative 2'-O-ribose methyltransferase catalyzing the formation of Gm18 in tRNAs. *RNA* **5**, 66–81.
- 48 Pintard, L., Lecoite, F., Bujnicki, J. M., Bonnerot, C., Grosjean, H., and Lapeyre, B. (2002) Trm7p catalyses the formation of two 2'-O-methylriboses in yeast tRNA anticodon loop. *EMBO J.* **21**, 1811–1820.
- 49 Gupta, R. (1984) *Halobacterium volcanii*: identification of 41 tRNAs covering all amino acids and the sequences of 33 class I tRNAs. *J. Biol. Chem.* **259**, 9461–9471.
50. Zimmermann, R. A., Gait, M. J., and Moore, M. J. (1998) Incorporation of modified nucleotides into RNA for studies on RNA structure, function and intermolecular interactions, in *Modification and Editing of RNA* (Grosjean, H. and Benne, R., eds.), ASM Press, Washington, DC, pp. 59–84.
- 51 Bruce, A. G. and Uhlenbeck, O. C. (1982) Enzymatic replacement of the anticodon of yeast phenylalanine tRNA. *Biochemistry* **21**, 855–861.
- 52 Carbon, P., Haumont, E., Fournier, M., de Henau, S., and Grosjean, H. (1983) Site-directed *in vitro* replacement of nucleosides in the anticodon loop of tRNA: application to the study of structural requirements for queuine insertase activity. *EMBO J.* **2**, 1093–1097.
- 53 Droogmans, L. and Grosjean, H. (1987) Enzymatic conversion of guanosine 3'-adjacent to the anticodon of yeast tRNA-phe to N¹-methylguanosine and the Wye nucleoside: dependence on the anticodon sequence. *EMBO J.* **6**, 477–483.
54. Kretz, K. A., Trewyn, R. W., Keith, G., and Grosjean, H. (1990) Site directed replacement of nucleotides in the anticodon loop of tRNA: application to the study of inosine biosynthesis in yeast tRNA-ala, in *Chromatography and Modifications of Nucleosides* vol. 45B, Journal of Chromatography Library Series (Gehrke, C. W. and Kuo, K. C., eds.), Elsevier, The Netherlands, pp. B143–B171.

- 55 Droogmans, L. and Grosjean, L. (1991) 2'-O-methylation and inosine formation in the wobble position of anticodon-substituted tRNA-phe in a homologous yeast *in vitro* system. *Biochimie* **73**, 1021–1025.
- 56 Ma, X., Zhao, X., and Yu, Y. T. (2003) Pseudouridylation (Psi) of U2 snRNA in *S. cerevisiae* is catalyzed by an RNA-independent mechanism. *EMBO J.* **22**, 1889–1897.
- 57 Jackman, J. E., Montange, R. K., Malik, H. S., and Phizicky, E. M. (2003) Identification of the yeast gene encoding the tRNA m¹G methyltransferase responsible for modification at position 9. *RNA* **9**, 574–585.
- 58 Grosjean, H., Droogmans, L., Giegé, R., and Uhlenbeck, O. C. (1990) Guanosine modifications in runoff transcripts of synthetic tRNA-Phe microinjected into *Xenopus laevis* oocytes. *Biochim. Biophys. Acta.* **1050**, 267–273.
- 59 Grosjean, H., Edqvist, J., Sträby, K. B., and Giegé, R. (1996) Enzymatic formation of modified nucleosides in tRNA: dependence on tRNA architecture. *J. Mol. Biol.* **255**, 67–85.
- 60 Jiang, H.-Q., Motorin, Y., Jin, Y.-X., and Grosjean, H. (1997) Pleiotropic effects of intron removal on base modification pattern of yeast tRNA-Phe: an *in vitro* study. *Nucleic Acids Res.* **25**, 2694–2701.
- 61 Morin, A., Auxilien, S., Senger, B., Tewari, R., and Grosjean, H. (1998) Structural requirements for enzymatic formation of threonylcarbamoyladenosine (t⁶A) in tRNA: an *in vivo* study with *Xenopus laevis* oocytes. *RNA* **4**, 24–37.
- 62 Constantinesco, F., Motorin, Y., and Grosjean, H. (1999) Transfer RNA modification enzymes from *Pyrococcus furiosus*: detection of the enzymatic activities *in vitro*. *Nucleic Acids Res.* **27**, 1308–1315.
- 63 Motorin, Y. and Grosjean, H. (1999) Multisite-specific tRNA:m⁵C-methyltransferase (Trm4) in yeast *S. cerevisiae*: identification of the gene and substrate specificity of the enzyme. *RNA* **5**, 1105–1118.
- 64 Alexandrov, A., Martzen, M. R., and Phizicky, E. M. (2002) Two proteins that form a complex are required for 7-methylguanosine modification of yeast tRNA. *RNA* **8**, 1253–1266.
- 65 Boorstein, W. R. and Craig, E. A. (1989) Primer extension analysis of RNA, in *Methods in Enzymology*, vol. 180 (Dahlberg, J. E. and Abelson, J. N., eds.), Academic, New York, pp. 347–369.
- 66 Hagenbüchle, O., Santer, M., and Steitz, J. A. (1978) Conservation of the primary structure at the 3' end of 18S rRNA from eukaryotic cells. *Cell* **13**, 551–563.
- 67 Wittig, B. and Wittig, S. (1978) Reverse transcription of tRNA. *Nucleic Acids Res.* **5**, 1165–1178.
- 68 Youvan, D. C. and Hearst, J. E. (1979) Reverse transcriptase pauses at N²-methylguanine during *in vitro* transcription of *E. coli* 16S ribosomal RNA. *Proc. Natl. Acad. Sci. USA* **76**, 3751–3754.
- 69 Youvan, D. C. and Hearst, J. E. (1981) A sequence from *Drosophila melanogaster* 18S rRNA bearing the conserved hypermodified nucleoside amPsi: analysis by reverse transcription and high performance liquid chromatography. *Nucleic Acids Res.* **9**, 1723–1741.

- 70 Maden, B. E. H., Corbett, M. E., Heeney, P. A., Pugh, K., and Ajuh, P. M. (1995) Classical and novel approaches to the detection and localization of the numerous modified nucleotides in eukaryotic ribosomal RNA. *Biochimie* **77**, 22–29.
- 71 Liu, J., Zhou, W., and Doetsch, P. W. (1995) RNA polymerase bypass at sites of dihydrouracil : implications for transcriptional mutagenesis. *Mol. Cell. Biol.* **15**, 6729–6735.
- 72 Auxilien, S., Keith, G., Le Grice, F. J., and Darlix, J.-L. (1999) Role of post-transcriptional modifications of primer tRNA-Lys₃ in the fidelity and efficacy of plus strand DNA transfer during HIV-1 reverse transcription. *J. Biol. Chem.* **274**, 4412–4420.
- 73 Pintard, L., Bujnicki, J. M., Lapeyre, B., and Bonnerot, C. (2002) MRM2 encodes a novel yeast mitochondrial 21S rRNA methyltransferase. *EMBO J.* **21**, 1139–1147.
- 74 Rhee, Y., Valentine, M. R., and Termini, J. (1995) Oxidative base damage in RNA detected by reverse transcriptase. *Nucleic Acids Res.* **23**, 3275–3282.
- 75 Ho, N. W. Y. and Gilham, P. T. (1971) Reaction of pseudouridine and inosine with *N*-cyclohexyl-*N'*-beta-(4-methylmorpholinium)ethylcarbodiimide. *Biochemistry* **10**, 3651–3657.
- 76 Bakin, A. V. and Ofengand, J. (1993) Four newly located pseudouridylate residues in *E. coli* 23S ribosomal RNA are all at the peptidyl center: analysis by the application of a new sequencing technique. *Biochemistry* **32**, 9754–9762.
77. Bakin, A. V. and Ofengand, J. (1998) Mapping of pseudouridine residues in RNA to nucleotide resolution, in *Methods in Molecular Biology*, vol. 77: Protein Synthesis: Methods and Protocols (Martin, R., ed.), Humana, Totowa, NJ, pp. 297–309.
- 77a. Ofengand, J., Del Campo, M., and Kaya, Y. (2001) Mapping pseudouridines in RNA molecules. *Methods* **25**, 365–373.
- 78 Wintermeyer, W. and Zachau, H. G. (1970) A specific chemical chain scission of tRNA at 7-methylguanosine. *FEBS Lett.* **11**, 160–164.
79. Wintermeyer, W. and Zachau, H. G. (1974) Replacement of odd bases in tRNA by fluorescent dyes, in *Methods in Enzymology*, vol. XX (Grossman, L. and Moldave, K., eds.), Academic, New York, pp. 667–673.
80. Thiebe, R. and Zachau, H. G. (1971) Half-molecules from phenylalanine tRNA's, in *Methods in Enzymology*, vol. XX (Moldave, K. and Grossman, L., eds.), Academic, New York, pp. 178–182.
- 81 Beltchev, B. and Grunberg-Manago, M. (1970) Preparation of pG-fragment from yeast tRNA-Phe by chemical scission at the dihydrouracil and inhibition of yeast tRNA-Phe charging by this fragment when combined with the –CCA half of this tRNA. *FEBS Lett.* **12**, 24–27.
- 82 Cerutti, P., Holt, J. W., and Miller, N. (1968) Detection and determination of 5, 6-dihydrouridine and 4-thiouridine in transfer RNA from different sources. *J. Mol. Biol.* **34**, 505–518.
- 83 Bird, A. (1980) DNA methylation and the CpG frequency in the animal DNA. *Nucleic Acids Res.* **8**, 1499–1505.

- 84 Lindahl, T. and Nyberg, B. (1974) Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry* **13**, 3405–3410.
- 85 Lindahl, T. and Nyberg, B. (1972) Rate of depurination of native deoxyribonucleic acid. *Biochemistry* **11**, 3610–3618.
- 86 Keith, G. and Dirheimer, G. (1995) Postlabeling: a sensitive method for studying DNA adducts and their role in carcinogenesis. *Curr. Opin. Biotechnol.* **6**, 3–11.
87. Dirheimer, G., Pfohl-Leszkowicz, A., and Keith, G. (1999) Methods of DNA-adduct detection : some applications to detect environmental carcinogen exposure, in *Trends in Environmental Mutagenesis* (Sobti, R. C., Obe, G., and Quillardet, A., eds.), Tausco Book Distributors, New Delhi, pp. 23–46.
- 88 Durantont, B., Keith, G., Gossé, F., Bergmann, C., Schleiffer, R., and Raul, F. (1998) Concomitant changes in polyamine pools and DNA methylation during growth inhibition of human colonic cancer cells. *Exp. Cell Res.* **243**, 319–325.
- 89 Glasser, A. L., Desgres, J., Heitzler, J., Gehrke, C. W., and Keith, G. (1991) *O*-Ribosyl-phosphate purine as a constant modified nucleotide located at position 64 in cytoplasmic initiator tRNA-Met of yeasts. *Nucleic Acids Res.* **19**, 5199–5203.
90. Gehrke, C. W., Desgres, J., Keith, G., Gerhardt, K. O., Agris, P. F., Gracz, H., Tempesta, M. S., and Kuo, K. C. (1990) Structural elucidation of nucleosides in nucleic acids, in *Chromatography and Modifications of Nucleoside*, vol. 45A, Journal of Chromatography Library Series (Gehrke, C. W. and Kuo, K. C., eds.), Elsevier, The Netherlands, pp. A159–A223.
- 91 Hiramaru, M., Ushida, T., and Egami, F. (1966) Ribonuclease preparation for the base analysis of polyribonucleotides. *Anal. Biochem.* **17**, 135–142.
- 92 Cameron, V. and Uhlenbeck, O. C. (1977) 3'-Phosphatase activity in T4 polynucleotide kinase. *Biochemistry* **16**, 5120–5126.
- 93 Watanabe, K. (1980) Reactions of 2-thioribothymidine and 4-thiouridine with hydrogen peroxide in transfer RNA from *Thermus thermophilus* and *Escherichia coli* as studied by circular dichroism. *Biochemistry* **19**, 5542–5549.
- 94 Feldmann, H. and Falter, H. (1971) Transfer ribonucleic acid from *Mycoplasma laidlawii* A. *Eur. J. Biochem.* **18**, 573–581.
- 95 Hall, R. H. (1971) *The Modified Nucleosides in Nucleic Acids*. Columbia University Press, New York.
96. Randerath, K., Gupta, R. C., and Randerath, E. (1980) ³H and ³²P derivative methods for base composition and sequence analysis of RNA, in *Methods in Enzymology*, vol. 65 (Grossman, L. and Moldave, K., eds.), Academic, New York, pp. 638–680.
- 97 Gupta, R. C., Randerath, E., and Randerath, K. (1976) A double-labeling procedure for sequence analysis of picomole amounts of nonradioactive RNA fragments. *Nucleic Acids Res.* **3**, 2895–2914.
- 98 Gupta, R. C., Randerath, E., and Randerath, K. (1976) An improved separation procedure for nucleoside monophosphates on polyethyleneimine-(PEI)-cellulose thin layers. *Nucleic Acids Res.* **3**, 2915–2922.

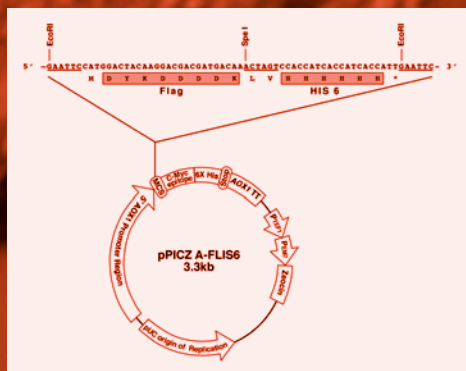
- 99 Gupta, R. C. and Randerath, K. (1979) Rapid readthrough technique for sequencing of RNAs containing modified nucleotides. *Nucleic Acids Res.* **6**, 3443–3458.
100. Randerath, K. and Randerath, E. (1983) Selected postlabeling procedures for base composition and sequence analysis of nucleic acids, in *Methods of DNA and RNA Sequencing* (Weissman, M. S., ed.), Praeger, New York, pp. 169–233.
- 101 Rogg, H., Brambilla, R., Keith, G., and Staehelin, M. (1976) An improved method for the separation and quantitation of the modified nucleosides of transfer RNA. *Nucleic Acids Res.* **3**, 285–295.
102. Suzuki, T., Suzuki, T., Wada, T., Saigo, K., and Watanabe, K. (2002) Taurine as a constituent of mitochondrial tRNAs: new insights into the functions of taurine and human mitochondrial diseases. *EMBO J.* **21**, 6581–6589.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



Functional Characterization of 2'-*O*-Methylation and Pseudouridylation Guide RNAs

Tamás Kiss and Beáta E. Jály

Summary

Pseudouridines and 2'-*O*-methylated nucleotides are ubiquitous constituents of stable cellular RNAs. In eukaryotes, posttranscriptional synthesis of most pseudouridines and 2'-*O*-methylated nucleotides is directed by sequence-specific guide RNAs (gRNAs). In recent years, an enormous number of novel putative modification gRNAs have been identified in a broad variety of organisms by using bioinformatics and large-scale cDNA-sequencing approaches. Understanding of the function of the novel modification gRNAs as well as the functional importance of modified nucleotides requires techniques that support the site-specific detection of 2'-*O*-methylated nucleotides and pseudouridines. Here, we describe rapid, reverse transcription–based methods to map 2'-*O*-methylated nucleotides and pseudouridines in any cellular RNA.

Key Words: Posttranscriptional RNA modification; pseudouridine; 2'-*O*-methylated nucleotide; guide RNAs; small nucleolar RNAs; small Cajal body-specific RNAs; modification mapping; primer extension.

1. Introduction

The correct function of stable cellular RNAs is facilitated by many post-transcriptionally synthesized modified nucleotides (**1–4**). In rRNAs, small nuclear (sn)RNAs, and small nucleolar (sno)RNAs, the site-specific synthesis of the most prevalent modified nucleotides, pseudouridines and 2'-*O*-methylated nucleotides (**Fig. 1**), is directed by sequence-specific guide RNAs (gRNAs) (**Fig. 2**). The box C/D 2'-*O*-methylation and box H/ACA pseudouridylation gRNAs accumulate and function in either the nucleolus or the Cajal body and, accordingly, they are called snoRNAs or small Cajal body-specific RNAs (**5–9**).

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

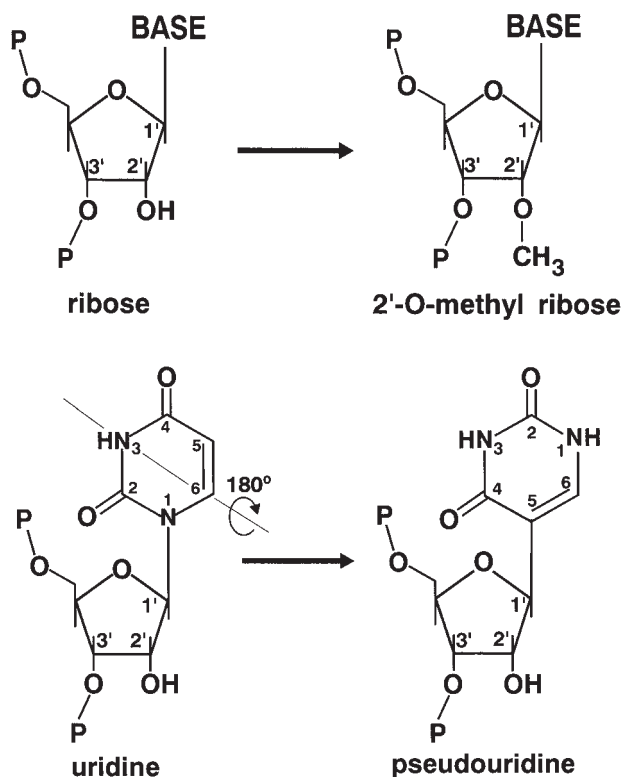


Fig. 1. Methylation of 2'-hydroxyl group of ribonucleotides and conversion of uridine into pseudouridine. After scission of the N1-C1' glycosidic bond of uridine, the uracil base is rotated 180° about its N3-C6 diagonal axis and reattached through the C5 position to the C1' position of ribose.

In recent years, a myriad of putative pseudouridylation and 2'-O-methylation gRNAs have been reported from a broad variety of organisms ranging from *Archaea* to human (10). Since modification gRNAs seem to be one of the most abundant groups of cellular RNAs besides tRNAs, rRNAs, and snRNAs, researchers can anticipate that many more guide RNAs will be identified in the future. Although some of the newly identified gRNAs can be linked to known modification sites in rRNAs or snRNAs based on the presence of specific substrate recognition sequences, scores of the new gRNAs lack reported modification sites in any cellular RNA. This is mainly owing to the fact that only limited information is available on the distribution of modified nucleotides in stable cellular RNAs. Therefore, understanding the function of novel modifi-

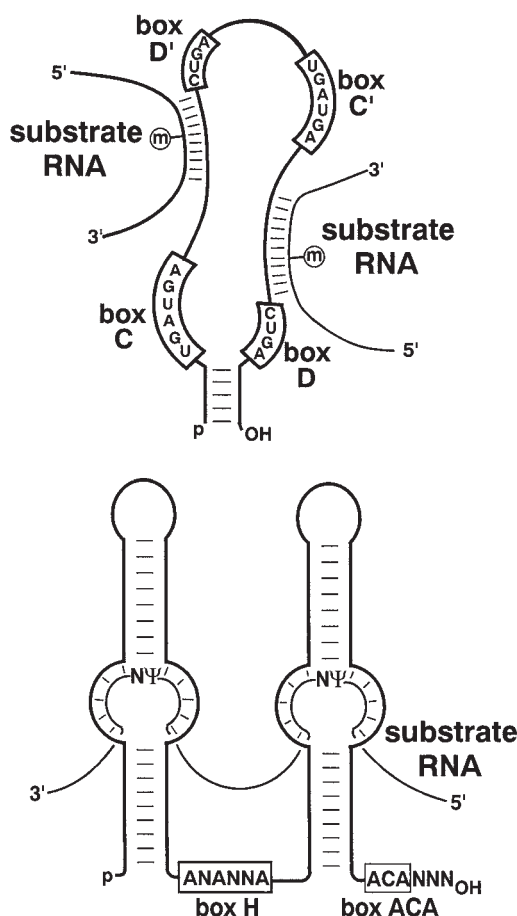


Fig. 2. Selection of 2'-O-methylation and pseudouridylation sites by box C/D and H/ACA modification gRNAs. The 2'-O-methylation gRNAs carry four conserved sequence elements (boxes C, C', D, and D') and form 10- to 21-bp-long helices with their substrate RNAs. This transient interaction places the D or D' box of the gRNA 5 nt upstream of the target residue to be 2'-O-methylated (circled "m") (11,22). The pseudouridylation gRNAs contain the box H (ANANNA) and ACA motifs and fold into the consensus "hairpin-hinge-hairpin-tail" secondary structure (23,24). The distal part of an internal loop in the 5'- and/or 3'-terminal hairpin forms two short helices with the substrate RNA. The target uridine selected for pseudouridylation remains unpaired, and it is located about 15 nt upstream from the H or ACA box of the gRNA (12). The 2'-O-methyl transfer and the uridine-to-pseudouridine conversion reactions are catalyzed by the fibrillarin/Nop1 and dyskerin/Cbf5 proteins that are associated with all box C/D and box H/ACA gRNAs, respectively.

cation gRNAs may necessitate experimental verification of the presence of 2'-*O*-methylated nucleotides and pseudouridines in their predicted substrate RNAs.

In this chapter, we describe simple and straightforward techniques to detect 2'-*O*-methylated nucleotides and pseudouridines in any RNA. These modification mapping experiments can be successfully carried out in any molecular biology laboratory equipped with standard apparatuses, and in combination with genetic or biochemical depletion experiments (11–15), they can provide ultimate evidence for the guide function of new box C/D or box H/ACA RNAs.

2. Materials

1. RNA extraction solution: 4 *M* guanidine thiocyanate, 25 *mM* sodium citrate, pH 7.0, 0.5% sarcosyl, 100 *mM* 2-mercaptoethanol.
2. 3 *M* Sodium acetate, pH 5.0.
3. Water-saturated phenol.
4. Chloroform/isoamyl alcohol (24:1 [v/v]).
5. Isopropanol.
6. 0.3 *M* Sodium acetate, pH 5.0.
7. 70% Ethanol. Store at –20°C.
8. Sterile water.
9. Formamide RNA hydrolysis solution: 0.4 *mM* MgCl₂ in deionized formamide. Store at –20°C.
10. Ethanol. Store at –20°C.
11. *N*-Cyclohexyl-*N'*-β-(4-methylmorpholinium)-ethylcarbodiimide *p*-tosylate (CMC) reaction buffer: 50 *mM* bicine, pH 8.3, 170 *mM* CMC (Sigma, St. Louis, MO), 7 *M* urea, 4 *mM* EDTA. Store at –20°C.
12. Alkali buffer: 50 *mM* Na₂CO₃, pH 10.4.
13. 10X T4 polynucleotide kinase buffer: 500 *mM* Tris-HCl, pH 7.5, 100 *mM* MgCl₂, 50 *mM* dithiothreitol (DTT), 0.5 mg/mL of gelatin. Store at –20°C.
14. [γ-³²P]ATP (5000 Ci/mmol) (Amersham, Piscataway, NJ).
15. T4 polynucleotide kinase.
16. TE: 10 *mM* Tris-HCl, pH 7.5; 1 *mM* EDTA.
17. 7.5 *M* ammonium acetate.
18. Carrier tRNA: 5 μg/μL of *Escherichia coli* tRNA in sterile water. Store at –20°C.
19. Formamide-loading solution: 90% deionized formamide, 10 *mM* EDTA, 0.1% (w/v) bromophenol blue, 0.1% xylene cyanol FF. Store at –20°C.
20. 40% (w/v) Acrylamide/bisacrylamide (19:1) solution. Store at 4°C.
21. 10X TBE running buffer: 1 *M* Tris base, 1 *M* boric acid, 25 *mM* EDTA.
22. 50% (w/v) Ammonium persulfate. Prepare freshly.
23. TEMED. Store at 4°C.
24. Nucleic acid elution solution: 500 *mM* ammonium acetate, 0.1 *mM* EDTA, 0.1% sodium dodecyl sulfate.

25. Carrier glycogen: 10 $\mu\text{g}/\mu\text{L}$ of glycogen in sterile water. Store at -20°C .
26. 10X Reverse transcriptase (RT) buffer: 500 mM Tris-HCl, pH 8.0, 500 mM KCl, 50 mM MgCl_2 , 50 mM DTT, 0.5 mg/mL of gelatin. Store at -20°C .
27. 10 mM Deoxynucleotide triphosphate (dNTP) mix: 10 mM dATP, 10 mM dGTP, 10 mM dCTP, and 10 mM TTP (Amersham). Store at -20°C .
28. 0.04 mM dNTP mix: 0.04 mM dATP, 0.04 mM dGTP, 0.04 mM dCTP, and 0.04 mM TTP (Amersham). Store at -20°C .
29. Ribonuclease inhibitor (RNasin[®]); (Promega, Madison, WI).
30. Avian myeloblastosis virus (AMV) RT (Promega).
31. Phenol/chloroform/isoamyl alcohol (25:24:1 [v/v]) saturated with Tris-HCl, pH 8.0.
32. Alkaline denaturation solution: 2 N NaOH, 2 mM EDTA.
33. 5X Sequencing buffer: 200 mM Tris-HCl, pH 7.5, 100 mM MgCl_2 , 250 mM NaCl. Store at -20°C .
34. 100 mM DTT. Store at -20°C .
35. Chain termination mixes: 80 μM dNTPs in 50 μM NaCl supplemented with 8 μM ddGTP (G mix), ddATP (A mix), ddCTP (C mix), or ddTTP (T mix). Store at -20°C .
36. T7 DNA polymerase, Sequenase Version 2.0 (Amersham).

3. Methods

The classic biochemical approaches used for mapping modified nucleotides are complicated, tedious, and, most important, not applicable to low-abundance RNAs (*16,17*). Here we describe straightforward sequencing methods based on primed reverse transcription to map 2'-O-methylated nucleotides and pseudouridines in any cellular RNA. These methods were originally developed in the laboratories of B. E. Maden (*17*) and J. Ofengand (*18*). The specificity of the reverse transcription-based modification mapping reactions is provided by terminally labeled sequence-specific oligodeoxynucleotide primers, which can selectively base pair to the appropriate test RNA even in a nonfractionated cellular RNA preparation. We present two alternative approaches to determine the position of 2'-O-methylated nucleotides. One of these techniques is based on the fact that 2'-O-methyl groups confer resistance to alkali (*19*). When a partially alkali hydrolyzed RNA is used as a template for primer extension reaction, the 2'-O-methylated nucleotides appear as "gaps" in the ladder of the primer extension products (**Fig. 3**). This method gives good, unambiguous results with abundant cellular RNAs, such as rRNAs and spliceosomal snRNAs. The second method is based on the observation that in the presence of a low concentration of dNTPs, 2'-O-methylated nucleotides interfere with the passage of RT. The enzyme pauses immediately preceding the 2'-O-methylated nucleotides and tends to terminate cDNA synthesis (*17*). Therefore, when

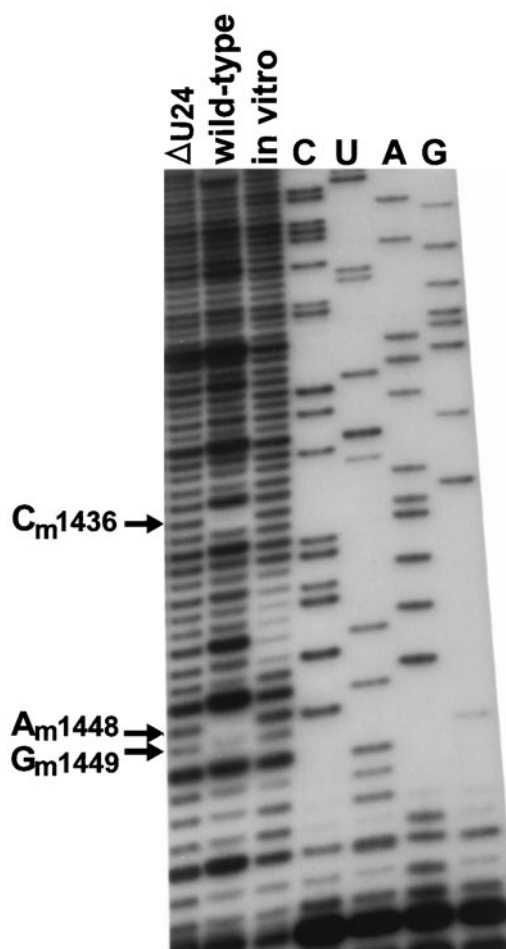


Fig. 3. Primer extension analysis of ribose-methylated nucleotides in yeast 25S rRNA. 2'-*O*-ribose methylation of yeast 25S rRNA at Cm1436, Am1448, and Gm1449 is directed by the U24 box C/D methylation gRNA (*II*). RNA isolated from yeast *Saccharomyces cerevisiae* cells either expressing (wild-type) or lacking U24 snoRNA (Δ U24) was partially hydrolyzed and annealed to a terminally labeled oligodeoxynucleotide primer complementary to nt 1458–1477 of the yeast 25S rRNA. The annealed primer was extended by AMV RT, and the cDNA products were fractionated on a 6% sequencing gel. As a control, in vitro-transcribed 25S rRNA was also partially hydrolyzed and analyzed by primer extension (in vitro). (Reproduced from ref. *II* with permission from Elsevier.)

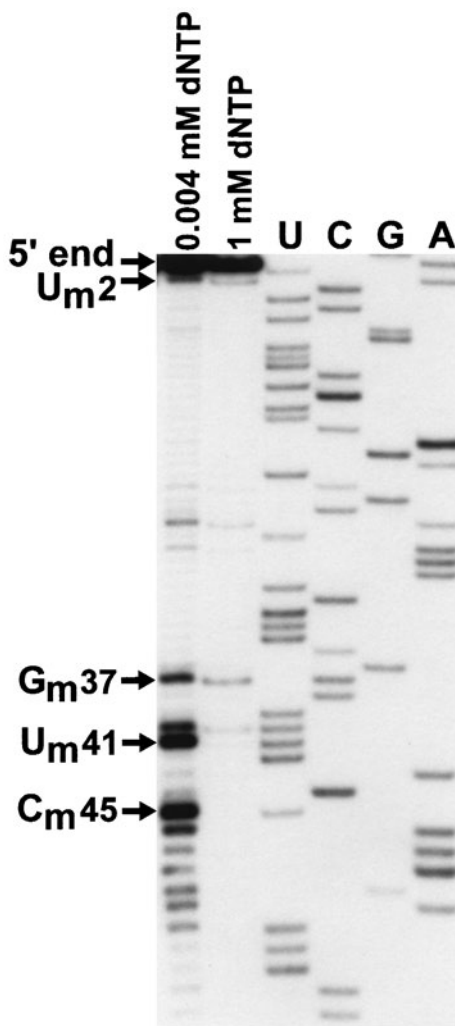


Fig. 4. Primer extension analysis of 2'-O-ribose-methylated nucleotides in mouse U5 spliceosomal snRNA. A ^{32}P -labeled oligodeoxynucleotide primer complementary to the mouse U5 snRNA from C71 to A87 was annealed to mouse cellular RNA and extended by AMV RT in the presence of 1 mM or 0.004 mM dNTPs. Lanes A, G, C, and U represent dideoxy sequencing reactions using the same oligodeoxynucleotide as primer and a recombinant plasmid carrying a cDNA of U5 snRNA as a template. Stop signals representing 2'-O-methylated nucleotides are indicated (25).

primer extension products are separated on a sequencing gel, a stop signal indicates the presence of a 2'-O-methylated ribonucleotide 1 nt downstream from its actual position (Fig. 4) (see Note 1).

This method is recommended for analysis of both low- and high-abundance RNAs. The technique for locating pseudouridines by reverse transcription takes advantage of the selective reactivity of pseudouridine toward carbodiimide (**18**). In the first reaction, uridines and pseudouridines are partially modified by CMC treatment. In the next step, the unstable N₃-CMC-uridine is selectively cleaved by mild alkali treatment. Under the applied conditions, the N₃-CMC-pseudouridine remains stable. The bulky CMC group at the N3 position of pseudouridine arrests RT. Consequently, a stop signal will be present 1 nt downstream of each pseudouridine following electrophoresis of cDNA products (**Fig. 5**). Since the cleavage of CMC-modified uridines is usually incomplete, weak primer extension stops are also visible 1 nt before each uridine.

3.1. RNA Extraction

Most of the commonly used extraction procedures that produce intact RNA are suitable for purification of test RNA. In our laboratory, we routinely use the guanidine thiocyanate-phenol-chloroform purification procedure for RNA extraction from cultured mammalian cells and tissues, yeast cells, and plant protoplasts (**20**) (*see Note 2*).

1. Lyse about 5×10^6 cultured mammalian cells or about 10^8 plant protoplasts in 0.6 mL of RNA extraction solution. Tissues should be homogenized in a glass Teflon homogenizer.
2. Add 40 μ L of 3 M sodium acetate (pH 5.0), 0.6 mL of water-saturated phenol, and 0.2 mL of chloroform/isoamyl alcohol (*see Note 3*).
3. Mix by vortexing for 1 min and incubate on ice for 10 min.
4. Centrifuge in a refrigerated microcentrifuge for 10 min at 13,000 rpm and 4°C.
5. Transfer the upper aqueous phase into a fresh Eppendorf tube, and extract again with 0.3 mL of water-saturated phenol and 0.1 mL of chloroform/isoamyl alcohol.
6. Transfer the aqueous phase into a fresh Eppendorf tube and precipitate the RNA by adding 0.7 mL of isopropanol. Mix well and incubate for 20 min at -20°C.
7. Pellet the RNA by centrifuging for 10 min at 13,000 rpm and 4°C.
8. Remove the supernatant and dissolve the RNA in 0.4 mL of 0.3 M sodium acetate (pH 5.0).
9. Precipitate immediately by adding 0.4 mL of isopropanol. Mix and chill for 20 min at -20°C.
10. Centrifuge for 10 min at 13,000 rpm and 4°C. Rinse the pellet with 1 mL of 70% ethanol and dry under vacuum.
11. Dissolve the purified RNA in sterile water (about 1–5 μ g/ μ L concentration) and store at -20°C.

3.2. Partial Hydrolysis of Test RNA

To produce partially hydrolyzed template RNA for reverse transcription analysis, we hydrolyze RNA samples in deionized formamide supplemented

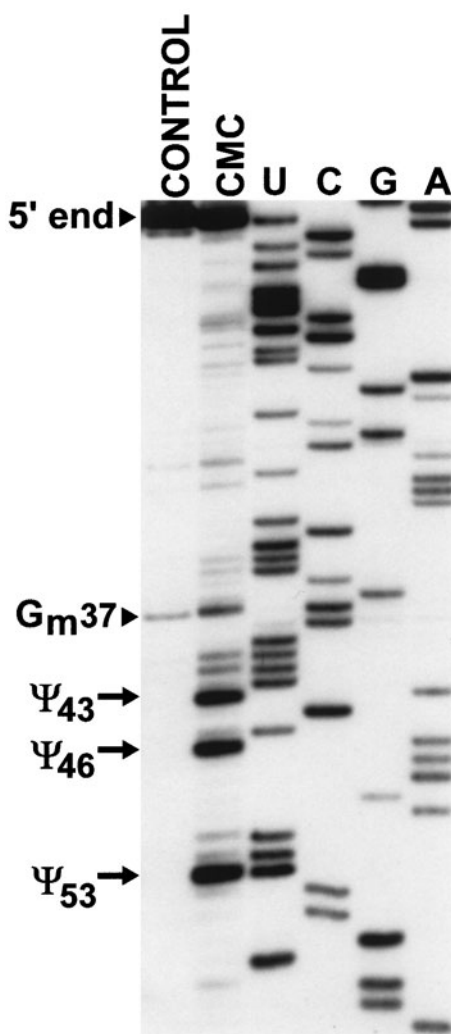


Fig. 5. Primer extension mapping of pseudouridines. Mouse cellular RNA either nontreated (control) or treated with CMC was annealed with a labeled oligodeoxynucleotide primer specific for the U5 spliceosomal snRNA and incubated with AMV RT in the presence of 1 mM dNTPs. For other details, see the legend to **Fig. 4**.

with 0.4 mM MgCl_2 . This procedure hydrolyzes RNA with an evenly distributed cleavage pattern, and, more important, progression of the hydrolysis reaction can be readily controlled.

1. In an Eppendorf tube, mix 5 μL of test RNA (about 20–25 μg) in water and 20 μL of formamide RNA hydrolysis solution.

2. After 5, 10, 15, 20, and 25 min of incubation in boiling water, collect 5-μL samples and pool them into a fresh tube kept on ice.
3. Add 75 μL of 0.3 M sodium acetate (pH 5.0), and precipitate the RNA by adding 400 μL of cold ethanol. Mix well and chill at -70°C for 20 min.
4. Collect the RNA by centrifuging in a microcentrifuge at 13,000 rpm for 15 min at room temperature.
5. Rinse the pellet with 1 mL of 70% cold ethanol and dry under vacuum.
6. Dissolve the partially hydrolyzed RNA in 10 μL of sterile water (about 0.5 μg/μL) and store at -20°C.

3.3. *Modification of Template RNA With CMC*

1. In an Eppendorf tube, lyophilize about 20–25 μg of test RNA to dryness, and dissolve it in 50 μL of CMC reaction buffer.
2. After incubation at 37°C for 20 min, stop the reaction by adding 120 μL of ice-cold 0.3 M sodium acetate (pH 5.0) and precipitate the RNA with 700 μL of cold ethanol.
3. Chill on ice for 10 min, and collect the modified RNA by centrifuging in a microcentrifuge at 13,000 rpm for 10 min at room temperature.
4. Wash the RNA pellet with 1 mL of 70% cold ethanol.
5. After dissolving in 100 μL of 0.3 M sodium acetate (pH 5.0), transfer the RNA to a fresh tube and precipitate with 300 μL of ethanol.
6. Incubate on ice for 10 min, and centrifuge in a microcentrifuge at 13,000 rpm for 10 min at room temperature.
7. Rinse the pellet with 1 mL of 70% ethanol and dry under vacuum.
8. Dissolve the RNA in 80 μL of alkali buffer and incubate for 4 h at 37°C.
9. Terminate the reaction by adding 200 μL of 0.3 M sodium acetate and 900 μL of cold ethanol.
10. Chill on ice for 10 min, and centrifuge at 13,000 rpm for 15 min at room temperature to pellet the RNA.
11. Rinse twice with 70% ethanol, dry, and dissolve the modified RNA in 20 μL of sterile water (about 1 μg/μL final concentration). The CMC-modified RNA either can be used immediately as a substrate in a primer extension reaction or can be stored at -20°C.

3.4. *Labeling and Purification of Oligodeoxynucleotide Primers*

For primer extension, we routinely use 18- to 22-nt-long oligodeoxynucleotide primers. We noticed that the nucleotide sequence rather than the length determines the efficacy of primers. Ideally, the primer hybridizes about 15–40 nt downstream of the putative modification site to be mapped (*see* **Notes 4** and **5**). After 5'-end labeling, the labeled oligodeoxynucleotide primer should be purified by size fractionation on a denaturing polyacrylamide gel in order to remove shorter, incompletely synthesized molecules.

1. For 5'-end labeling, mix together 2.5 μL of 10X T4 polynucleotide kinase buffer, 10 pmol of oligodeoxynucleotide primer, 5 μL of [$\gamma\text{-}^{32}\text{P}$]ATP (5000 Ci/mmol, 10 $\mu\text{Ci}/\mu\text{L}$), and add sterile water to fill the reaction mix up to 24 μL .
2. Add 1.0 μL (about 10 U) of T4 polynucleotide kinase and incubate for 30 min at 37°C.
3. To terminate the reaction, add 20 μL of TE (pH 7.5), 25 μL of 7.5 M ammonium acetate, 5 μL of carrier tRNA, and 150 μL of cold ethanol. Mix well and chill at -70°C for 30 min.
4. Collect the labeled oligodeoxynucleotide by centrifuging at 13,000 rpm for 10 min at room temperature, rinse with 1 mL of 70% cold ethanol, and dry under vacuum.
5. Dissolve the labeled oligodeoxynucleotide primer in 6 μL of formamide-loading solution.
6. Prepare a 20% polyacrylamide-8 M urea sequencing gel (16 $\times$ 16 cm, 0.4-mm spacer, 1-cm comb). For 20 mL of 20% polyacrylamide-8 M urea gel, dissolve 9.6 g of urea in 10 mL of 40% acrylamide/bisacrylamide (19:1) solution. Add 2 mL of 10X TBE buffer and fill up to 20 mL with distilled water. Add 40 μL of 50% ammonium persulfate and 15 μL of TEMED. Cast the gel immediately.
7. Warm up the gel by running at about 15-18 W of constant power for 20 min.
8. Denature the formamide-loading solution containing the labeled oligodeoxynucleotide primer at 95°C for 1 min and load onto the gel.
9. Run the gel at about 15-18 W of constant power, which keeps the gel hot enough to keep oligonucleotides fully denatured. To eliminate the remaining radioactive ATP, let the bromophenol blue tracking dye (migrates with 8-nt-long oligodeoxynucleotides) run out of the gel. The xylene cyanol tracking dye migrates together with 23-nt-long oligodeoxynucleotides.
10. After removing of the upper plate, wrap the gel in Saran Wrap, tape marker strips labeled with fluorescent spots to the Saran Wrap, and expose the gel to autoradiography film for 15-30 s.
11. Localize and cut the labeled full-length oligodeoxynucleotide from the gel.
12. Place the gel slice into an Eppendorf tube and soak in 300 μL of nucleic acid elution solution for at least 6 h at 37°C. Transfer the solution containing the eluted radioactive oligodeoxynucleotide into a fresh tube, and add 5 μL of carrier glycogen (10 $\mu\text{g}/\mu\text{L}$).
13. Precipitate the oligodeoxynucleotide by adding 900 μL of cold ethanol, vortex, chill at -20°C for 20 min, and centrifuge for 10 min at room temperature.
14. Rinse the pellet with 70% cold ethanol, dry, dissolve in 20 μL of sterile water (final concentration of about 0.5 pmol/ μL), and store at -20°C.

3.5. Primer Extension Reaction

1. Mix 1 μL of 10X RT buffer, 1 μL of labeled oligonucleotide primer, and about 5-20 μg of test RNA (either nontreated, CMC modified, or partially hydrolyzed) in 10 μL of sterile water (*see* **Notes 6** and **7**). Denature the RNA by incubating at 95°C for 1 min.

2. Transfer immediately into a 43°C water bath, and let it anneal by incubating for 10 min.
3. In the meantime, prepare the extension mix for each reaction by mixing 4 µL of sterile water, 1 µL of 10X RT buffer, 2 µL of 10 mM dNTP mix, 1 µL of ribonuclease inhibitor (about 40 U), and 1 µL of AMV RT (10 U). For mapping of 2'-O-methylated nucleotides in nontreated RNA, replace the 10 mM dNTP with 0.04 mM dNTP solution.
4. Add the extension mix to the annealed test RNA and incubate at 43°C for 30 min.
5. Stop the reaction by adding 50 µL of TE (pH 7.5), and extract with 70 µL of phenol/chloroform/isoamyl alcohol (pH 8.0) (*see Note 8*).
6. Transfer the upper aqueous phase into a fresh Eppendorf tube and add 7 µL of 3 M sodium acetate (pH 5.0), 1 µL of carrier tRNA (5 µg/µL), and 240 µL of cold ethanol.
7. Chill at -70°C for 20 min and centrifuge for 10 min at room temperature.
8. Wash the pellet with 1 mL of 70% cold ethanol and dry under vacuum.
9. Dissolve the pellet in 10 µL of formamide-loading solution. Samples can be stored at -20°C for a few days, but not longer than 1 wk.

3.6. Sequencing Reactions

To determine the exact position of 2'-O-methyl groups and pseudouridines, dideoxy-sequencing products should be run as molecular markers next to the extension products of the modification mapping reactions. In these sequencing reactions, a recombinant plasmid carrying a cDNA of the RNA of interest is used as a template.

3.6.1. Alkaline Denaturation of Plasmid for Sequencing

1. In an Eppendorf tube, mix 20 µL of supercoiled plasmid carrying the test RNA gene (about 5–10 µg) and 2 µL of alkaline denaturation solution at room temperature.
2. Incubate for 5 min before adding 7 µL of sterile water and 7 µL of 3 M sodium acetate (pH 5.0).
3. Precipitate the denatured DNA by adding 75 µL of cold ethanol. Mix well and incubate at -70°C for 10 min.
4. Collect the DNA by centrifuging in a microcentrifuge for 10 min at room temperature.
5. Rinse the pellet with 1 mL of 70% cold ethanol and dry under vacuum. The dried template DNA can be stored for several weeks at -20°C.

3.6.2. Chain Termination Sequencing Reaction

1. Dissolve denatured template DNA in 7 µL of sterile water, and add 2 µL of 5X sequencing buffer and 1 µL of terminally labeled oligodeoxynucleotide primer used for modification mapping. Mix well, spin briefly, and anneal the template DNA and primer by incubating at 37°C for 5 min.

2. In the meantime, label four Eppendorf tubes (G, A, T, and C), and add 2.5 μL of the appropriate chain termination mix into the labeled tubes.
3. Add 1 μL of 100 mM DTT, 4 μL of TE, and 0.2 μL of T7 DNA polymerase (1 to 2 U) to the annealing reaction. Distribute 3.5 μL of the reaction mix into each tube containing the G, A, T, or C termination mixes.
4. Incubate at 37°C for 5 min and terminate the reaction by adding 4 μL of formamide-loading solution. To obtain sequencing bands with intensity comparable with that of the primer extension products of the modification mapping reactions, the sequencing reactions usually need to be diluted by 5- to 20-fold.

3.7. Polyacrylamide Gel Electrophoresis

1. To analyze primer extension products, use a 6% polyacrylamide sequencing gel (30 $\times$ 40 cm, 0.4-mm spacer) containing 8 M urea. For the 6% polyacrylamide–8 M urea gel, dissolve 36 g of urea in 11.25 mL of 40% acrylamide/bisacrylamide (19:1) solution, 7.5 mL of 10X TBE buffer, and sterile water. The final volume should be 75 mL. Add 200 μL of 50% ammonium persulfate and 50 μL of TEMED. Mix and cast the gel immediately.
2. Warm up the gel by running at 75 W of constant power for 20–30 min.
3. Heat denature the primer extension and sequencing reactions in formamide-loading solution at 95°C for 3 min immediately prior to loading.
4. Load 3 μL of samples and 2 μL of each sequencing reaction and store the rest at –20°C.
5. Run the gel at 75 W of constant power. In a 6% denaturing polyacrylamide gel, the bromophenol blue and xylene cyanol tracking dyes migrate with 26 and 105 nt, respectively.
6. After electrophoresis, fix the gel in 20% ethanol and 10% acetic acid and dry on a gel dryer. To avoid undesired radioactive contamination of the gel during fixation, it is recommended that the nonextended radiolabeled oligodeoxynucleotide primer be allowed to run out into the lower reservoir of the gel electrophoresis apparatus.
7. Expose the gel to autoradiography film for 1–5 d at –70°C with an intensifying screen.

4. Notes

1. In some instances, the RT stop appears opposite to the 2'-O-methylated nucleotide or more frequently, two stop signals are observed: a strong stop preceding a 2'-O-methylated nucleotide and another, usually weaker stop opposite to the 2'-O-methylated nucleotide. The reason for this phenomenon is not yet understood, but apparently it may somehow depend on the actual sequence and/or structural context of the 2'-O-methylated residue.
2. To obtain satisfactory results, modification mapping of low-abundance RNAs extracted from some particular sources (e.g., plant leaves, roots, or cultured cells) might require purification of the test RNA by centrifugation through a CsCl cushion (21).

3. After adding chloroform/isoamyl alcohol, the mixture should turn milky on vortexing, indicating that the aqueous and organic phases are well separated. Normally, about 0.6 mL of aqueous (upper) phase should be recovered by centrifugation. If the two phases fail to separate, add more chloroform/isoamyl alcohol.
4. We have frequently observed that in pseudouridine mappings, the first uridine residue located upstream from the primer gives a strong stop signal, which we call "pseudo-pseudouridine." Therefore, to avoid confusion, it is highly recommended that not just one but rather two uridines occur in the template RNA between the expected pseudouridylation site and the 5' end of the primer-binding region.
5. The low concentration of dNTPs present in the 2'-O-methylation mapping reaction (0.004 mM dNTPs) highly reduces the efficiency of the primer extension reaction and, therefore, does not support mapping of 2'-O-methylated nucleotides located more than 80–100 nt upstream from the primer-binding site. Instead, we have frequently observed that during long runs, 2'-O-methylated nucleotide-induced stops appear in the control primer extension reaction that is performed in the presence of 1 mM dNTPs.
6. The amount of RNA used for a primer extension reaction depends on the cellular abundance of the RNA in question. While utilization of 5 µg of total RNA is sufficient for mapping rRNAs and snRNAs, we usually use 20–30 µg of template RNA to analyze low-abundance RNAs.
7. For each modification mapping reaction, appropriate control primer extension reactions should also be performed using nontreated cellular RNAs or in vitro synthesized test RNAs as templates (see **Figs. 3–5**).
8. When a large amount of template RNA (more than 30 µg) is necessary to obtain detectable extension signals, it is useful to treat the primer extension products with 1 µg of DNase-free RNase A for 30 min at 37°C before phenol extraction.

Acknowledgments

B.E.J. was supported by the French government and the Hungarian Research Foundation (OTKA T029042). This work was supported by grants from Association pour la Recherche contre le Cancer, la Ligue Nationale contre le Cancer, and the Hungarian Research Foundation (OTKA T029042 and T031738).

References

- 1 Agris, P. F. (1996) The importance of being modified: roles of modified nucleosides and Mg²⁺ in RNA structure and function. *Prog. Nucleic Acid Res. Mol. Biol.* **53**, 79–129.
- 2 Charette, M. and Gray, M. W. (2000) Pseudouridine in RNA: what, where, how, and why. *IUBMB Life* **49**, 341–351.
- 3 Decatur, W. A. and Fournier, M. J. (2002) rRNA modifications and ribosome function. *Trends Biochem. Sci.* **27**, 344–351.

- 4 King, T. H., Liu, B., McCully, R. R., and Fournier, M. J. (2003) Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. *Mol. Cell* **11**, 425–435.
- 5 Decatur, W. A. and Fournier, M. J. (2003) RNA-guided nucleotide modification of ribosomal and other RNAs. *J. Biol. Chem.* **278**, 695–698.
- 6 Filipowicz, W. and Pogacic, V. (2002) Biogenesis of small nucleolar ribonucleoproteins. *Curr. Opin. Cell Biol.* **14**, 319–327.
- 7 Kiss, T. (2002) Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell* **109**, 145–148.
- 8 Kiss, T. (2001) Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. *EMBO J.* **20**, 3617–3622.
- 9 Terns, M. P. and Terns, R. M. (2002) Small nucleolar RNAs: versatile trans-acting molecules of ancient evolutionary origin. *Gene Expr.* **10**, 17–39.
- 10 Bachellerie, J. P., Cavaillé, J., and Hüttenhofer, A. (2002) The expanding snoRNA world. *Biochimie* **84**, 775–790.
- 11 Kiss-László, Z., Henry, Y., Bachellerie, J. P., Caizergues-Ferrer, M., and Kiss, T. (1996) Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell* **85**, 1077–1088.
- 12 Ganot, P., Bortolin, M. L., and Kiss, T. (1997) Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell* **89**, 799–809.
- 13 Tycowski, K. T., Smith, C. M., Shu, M. D., and Steitz, J. A. (1996) A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in *Xenopus*. *Proc. Natl. Acad. Sci. USA* **93**, 14,480–14,485.
- 14 Tycowski, K. T., You, Z. H., Graham, P. J., and Steitz, J. A. (1998) Modification of U6 spliceosomal RNA is guided by other small RNAs. *Mol. Cell* **2**, 629–638.
- 15 Jády, B. E. and Kiss, T. (2001) A small nucleolar guide RNA functions both in 2'-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. *EMBO J.* **20**, 541–551.
- 16 Maden, B. E. (1990) The numerous modified nucleotides in eukaryotic ribosomal RNA. *Prog. Nucleic Acid. Res. Mol. Biol.* **39**, 241–303.
- 17 Maden, B. E., Corbett, M. E., Heeney, P. A., Pugh, K., and Ajuh, P. M. (1995) Classical and novel approaches to the detection and localization of the numerous modified nucleotides in eukaryotic ribosomal RNA. *Biochimie* **77**, 22–29.
- 18 Bakin, A. and Ofengand, J. (1993) Four newly located pseudouridylate residues in *Escherichia coli* 23S ribosomal RNA are all at the peptidyltransferase center: analysis by the application of a new sequencing technique. *Biochemistry* **32**, 9754–9762.
- 19 Smith, J. D. and Dunn, D. B. (1959) An additional sugar component of ribonucleic acids. *Biochim. Biophys. Acta* **31**, 573–575.
- 20 Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**, 156–159.
- 21 Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294–5299.

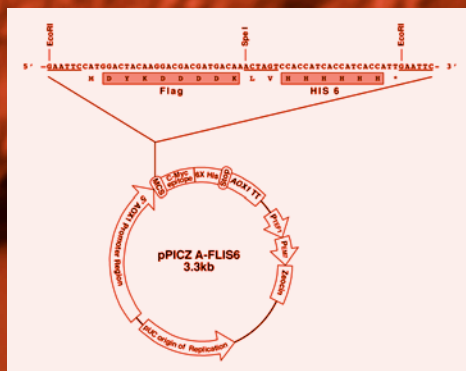
22. Cavaillé, J., Nicoloso, M., and Bachellerie, J. P. (1996) Targeted ribose methylation of RNA in vivo directed by tailored antisense RNA guides. *Nature* **383**, 732–735.
- 23 Balakin, A. G., Smith, L., and Fournier, M. J. (1996) The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions. *Cell* **86**, 823–834.
- 24 Ganot, P., Caizergues-Ferrer, M., and Kiss, T. (1997) The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. *Genes Dev.* **11**, 941–956.
25. Massenet, S., Mougin, A., and Branlant, C. (1998) Posttranscriptional modification in the U small nuclear RNAs, in *Modification and Editing of RNA* (Grosjean, H. and Benne, R., eds.), ASM Press, Washington DC, pp. 201–227.

RNA Interference, Editing, and Modification

Methods and Protocols

Edited by

Jonatha M. Gott



Experimental RNomics

A Global Approach to Identifying Small Nuclear RNAs and Their Targets in Different Model Organisms

Alexander Hüttenhofer, Jérôme Cavaillé, and Jean-Pierre Bachellerie

Summary

Non-messenger RNAs (nmRNAs) play a wide and essential role in cellular functions. Computational identification of novel nmRNAs in genomes of model organisms is severely restricted owing to their lack of an open reading frame. Hence, we describe experimental approaches for their identification by generating cDNA libraries derived from nmRNAs for which we coined the term *experimental RNomics*. Two different procedures are introduced for cDNA library construction. First, we describe the construction of a general purpose cDNA library from sized RNA fractions. Second, we introduce a more specialized RNomics strategy employing this approach to generate a cDNA library from a specific abundant class of nmRNAs. This is illustrated using as a paradigm the two families of small nucleolar RNAs that guide modification of nucleotides in rRNAs or spliceosomal RNAs small nuclear RNAs (snRNAs) by short antisense elements complementary to the modification site. Following the identification of novel members from the class of small nuclear RNAs by experimental RNomics, we demonstrate how their target sequences in rRNAs or snRNAs can be identified.

Key Words: RNomics; small non-messenger RNA; small nuclear RNA; 2'-O-methylation; pseudouridylation; cDNA library; C-tailing; DNA/RNA-linker.

1. Introduction

Non-messenger RNAs (nmRNAs) do not encode proteins but function directly on the level of RNA in the cell. Over the last few years, the importance of this surprisingly diverse class of molecules has been widely recognized (1–4). nmRNAs have been identified in unexpectedly large numbers, with pre-

From: *Methods in Molecular Biology*, vol. 265:
RNA Interference, Editing, and Modification: Methods and Protocols
Edited by: J. M. Gott © Humana Press Inc., Totowa, NJ

sent estimates in the range of thousands per genome. They play key roles in a variety of fundamental processes in all three domains of life: Eukarya, Bacteria, and Archaea. Their functions include DNA replication and chromosome maintenance, regulation of transcription, processing of RNA (not only RNA cleavage and religation, but also RNA modification and editing), translation and stability of mRNAs, and even stability and translocation of proteins (**4–15**). Many nmRNAs have been discovered fortuitously, suggesting that they merely represent the tip of the iceberg. The vast majority of nmRNAs are relatively small, much shorter than most mRNAs, hence their generic denomination of small nmRNAs (snmRNAs). The highly specific roles of snmRNAs reflect in most cases their ability to bind selectively a small set of proteins as well as their potential to specifically recognize definite RNA targets by sequence-specific base complementarity.

Recently, systematic searches of snmRNAs, dubbed experimental RNomics, have been initiated in various model organisms, by screening cDNA libraries constructed from sized RNA fractions (**16–18**). Since a large number of snmRNAs can be grouped into specific RNA classes (based on size, structural signatures, protein partners, RNA targets, or subcellular location), experimental RNomics can be focused on a particular snmRNA type, using tailor-made cDNA libraries and/or appropriate screening procedures.

In this chapter, we first describe procedures for the construction and screening of a general purpose cDNA library from sized RNA fractions. To illustrate a more specialized RNomics strategy we next use as a paradigm one of the two large families of small nucleolar RNAs (snoRNAs), both associated with a small set of specific proteins. The two snoRNA subtypes, termed C/D and H/ACA snoRNAs, guide two prevalent types of nucleotide modification, 2'-O-methylation and pseudouridylation, respectively, by forming a specific duplex around each target RNA modification site. Cellular RNAs targeted by guide snoRNAs include rRNAs in Eukarya and Archaea, small nuclear RNAs (snRNAs) and archaeal tRNAs, as well as probably a range of still unidentified RNAs (**4,10,11**). Finally, we detail the sequence analysis of resulting clones and their assignment to different snmRNA types, especially by reference to the two major snoRNA families mentioned.

2. Materials

1. Oligonucleotide primers.
 - a. Gibco I-Adaptor: 5'-TCGCGAGCGCCGCGGGGGGGGGGGGGG-3'.
 - b. M13 fsp: 5'-GCTATTACGCCAGCTGGCGAAAGGGGGATGTG-3'.
 - c. M13 rsp: 5'-CCCCAGGCTTTACACTTTATGCTTCCGGCTCG-3'.
 - d. Primer 1: 5'-ATAAAGCGGCCGCGGATCCAA-3'.

- e. Primer 2: 5'-TTGGATCCGCGGCCGCTTTATTNNNNTCAG-3'.
- f. Primer 3: 5'-AATAAAGCGGCCGCGGATCCAANNNNNRTGATGA-3'.
2. TRIzol extraction kit (Gibco-BRL, Eggenstein, Germany).
3. Sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis equipment.
4. Poly(A) polymerase (MBI Fermentas).
5. pSPORT 1 vector (included in the Gibco Superscript® cloning system).
6. pKS vector (Stratagene).
7. Gibco Superscript cloning system (Gibco-BRL).
8. BigDye terminator cycle sequencing reaction kit (PE Applied Biosystems).
9. Lasergene sequence analysis program package.
10. T4 polynucleotide kinase (New England Biolabs).
11. T4 RNA ligase (New England Biolabs).
12. Protein A-Sepharose (Sigma, St. Louis, MO).
13. *Taq* polymerase I (Appligene).
14. Glycogen (20 mg/mL) (Promega, Madison, WI).
15. $^{32/33}\text{P}$ -(γ)-ATP.
16. Hybridization buffer 1: 0.5 M sodium phosphate, pH 7.2; 7% SDS, 1 mM EDTA.
17. Hybridization buffer 2: 1 M sodium phosphate buffer, pH 6.2; 7% SDS.
18. Prehybridization buffer: 1 M sodium phosphate buffer, pH 6.2; 7% SDS.
19. Wash buffer 1: 2X saline sodium citrate (SSC) (20 mM sodium phosphate, pH 7.4; 0.3 M NaCl; 2 mM EDTA), 0.1% SDS.
20. Wash buffer 2: 0.1X SSC, 0.5% SDS.
21. Electrophoresis buffer: 7 M urea in 1Xx TBE (90 mM Tris-HCl, pH 8.3, 90 mM boric acid, 2.5 mM EDTA).
22. 1X Tailing buffer: 50 mM Tris-HCl, pH 8.0, 200 mM NaCl, 10 mM MgCl_2 , 2 mM MnCl_2 , 0.4 mM EDTA, 1 mM dithiothreitol (DTT), 2 mM CTP.
23. 1X Ligation buffer: 50 mM Tris-HCl, pH 7.8, 10 mM MgCl_2 , 10 mM DTT, 1 mM ATP.
24. 1X Kinase buffer: 70 mM Tris-HCl, pH 7.6, 10 mM MgCl_2 , 5 mM DTT, 1 mM ATP.
25. 5X First-strand buffer: 250 mM Tris-HCl, pH 8.3, 375 mM KCl, 15 mM MgCl_2 , 0.1 mM DTT.
26. 1X *Taq*I buffer: 10 mM Tris-HCl, pH 9.0, 50 mM KCl, 0.1% Triton X-100.
27. NET-150 buffer: 50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.05% Nonidet P-40.
28. Nylon membranes (Quiabran Nylon Plus; Quiagen, Germany; or Qbiogene).
29. Bio-Rad semidry blotting apparatus (Trans-blot SD; Bio-Rad, Germany).
30. Stratagene crosslinker.
31. Kodak MS-1 films.
32. Homogenizer One Shot Model (Constant Systems).
33. ddH_2O .

3. Methods

The following methods outline the construction and screening of a cDNA library from sized RNA fractions; the construction of a more specialized cDNA

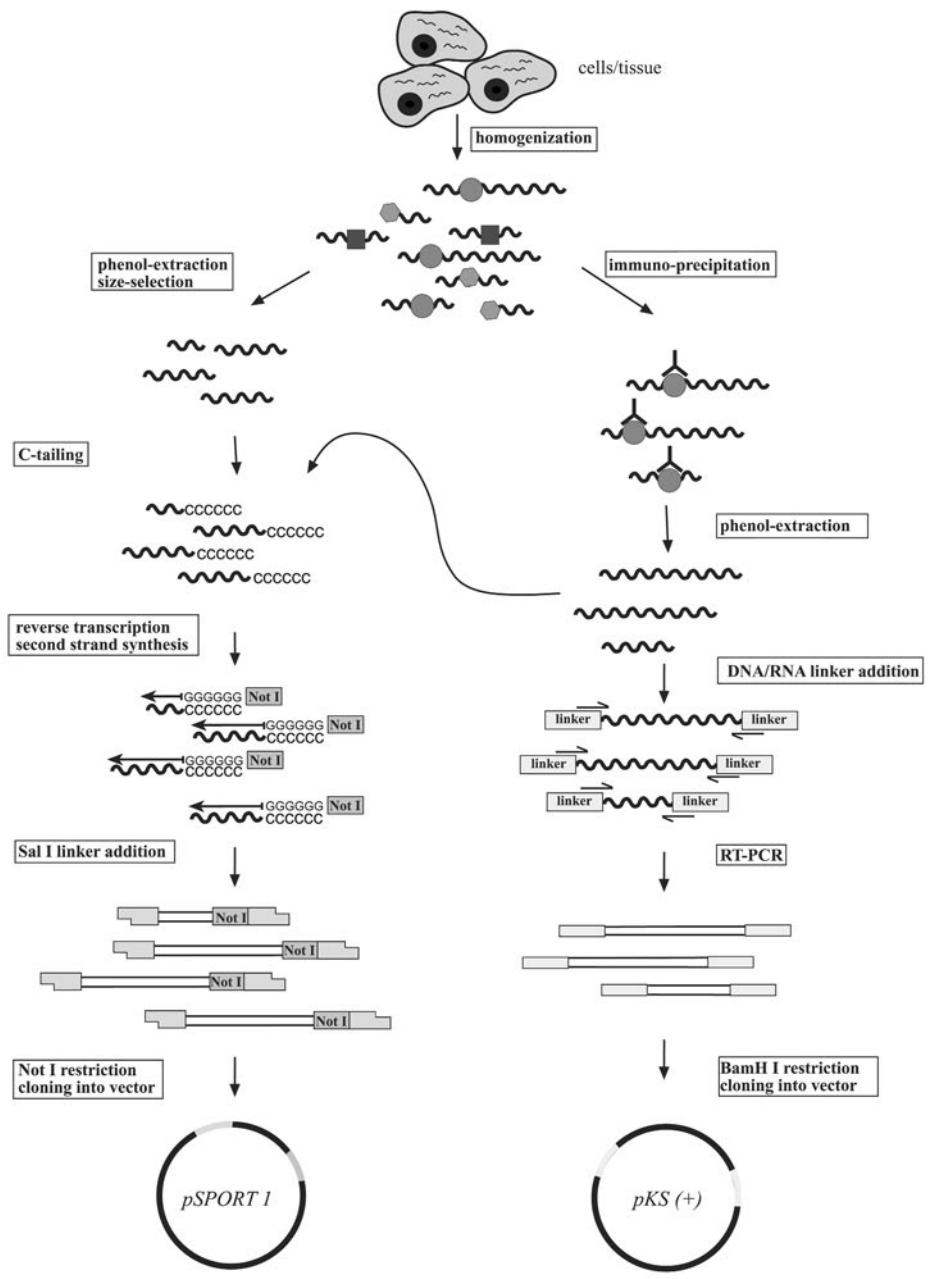
library, starting from an RNA enriched with the snmRNA subtype of interest and/or using specialized oligonucleotide primers for a preferential cDNA synthesis of this snmRNA subtype; and the sequence analysis of clones and assignment of snmRNA types, with emphasis on the two snoRNA subtypes, C/D and H/ACA, including prediction of their cognate modification sites in target RNAs.

As for the construction of the two different cDNA libraries, two different methods for reverse transcription of snmRNAs into cDNAs are applied (**Fig. 1**). In the first method (*see Subheading 3.1.*), 3'-tailing of snmRNAs by CTP and poly(A) polymerase is followed by reverse transcription of snmRNAs by an oligo d(G) primer. As an alternative to C-tailing of RNAs, we also used a "linker-ligation" method for cloning small RNA molecules (*see Subheading 3.2. and Fig. 1*): Two short identical phosphorylated DNA linkers of known sequence are ligated at the 5' and 3' ends of snmRNAs employing T4 RNA ligase. Subsequently, ligated RNAs are transcribed into cDNA by reverse transcriptase polymerase chain reaction (RT-PCR) using oligonucleotides derived from the DNA-linker sequences. The disadvantage of this method is, however, the possibility of oligomerization of DNA linkers and the failure of linker attachment to modified (e.g., 5') termini. The advantage of this method is that often full-length cDNA clones can be obtained, unlike for the C-tailing method.

3.1. Construction and Screening of a cDNA Library

This section describes the isolation of total RNA from the organism of interest, as well as the purification and size selection and the reverse transcription of snmRNAs into cDNA (**Fig. 1**). Since snmRNAs do not contain a poly(A) tail (as do mRNAs), snmRNAs are tailed with CTP and poly(A) polymerase (which also uses CTP as a substrate). After the addition of C-tails, snmRNAs are reverse transcribed employing an oligo d(G) primer and the Gibco Superscript

Fig. 1. (*see facing page*) cDNA library construction strategy for cloning snmRNAs for unbiased, general-purpose cDNA library (**left**) or more specialized library encoding snmRNAs from a certain RNA class, as, e.g., snoRNAs (**right**). Two different ways to generate cDNA library are introduced. For the general-purpose cDNA library, the C-tailing method is shown, and for the specialized cDNA library, a "linker-ligation" protocol is also introduced. The latter procedure starts with a fraction enriched for a specific RNA subclass, such as by immunoprecipitation of a cell lysate with an antibody directed against an RNA-binding protein from that class (indicated by squares, circles, or hexagons). The RNA from the immunoprecipitation reaction can then be either C-tailed (left) or DNA/or RNA linkers can be added followed by RT-PCR (right).



system. This procedure prevents the reverse transcription of small mRNAs (with poly[A] tails) exhibiting sizes similar to noncoding RNAs into cDNA.

3.1.1. Construction of a cDNA Library Encoding snmRNAs

1. Freeze 1 to 2 mg of tissue in liquid nitrogen and prepare total RNA with the TRIzol extraction kit (Gibco-BRL). Because of space constraints, we omit here the detailed description of RNA preparation, which is indicated in the TRIzol protocol (*see* **Note 1**).
2. Fractionate about 200 µg of total RNA on a denaturing 8% polyacrylamide gel (7 M urea, 1X TBE buffer); run the gel (20 × 30 × 0.2 cm) at 400 V for 1.5 h (*see* **Note 2**).
3. Excise the RNA from the gel corresponding to the desired size range (e.g., 50–500 nt). Passively elute the RNA by vigorous shaking in 0.3 M sodium acetate (pH 5.2), 0.1% SDS, and 0.2% phenol at 4°C overnight; extract with phenol/chloroform; and ethanol precipitate the RNA.
4. Tail 5 µg of RNA with CTP using poly(A) polymerase. The C-tailing reaction is carried out in a volume of 50 µL containing 1X tailing buffer and 1.5 U of poly(A) polymerase. Incubate the reaction mix at 37°C for 1.5 h. After incubation, add 20 µL of 3 M sodium acetate (pH 5.2) and 130 µL of ddH₂O. Extract the RNA with phenol/chloroform, precipitate by ethanol, and dissolve in 5 µL of RNase-free ddH₂O.
5. Reverse transcribe the RNA into cDNA using the Gibco I-Adaptor oligonucleotide primer (*important*: do not use the *NotI*-Adaptor supplied by the Gibco Superscript cloning kit, because it contains an oligo-d[T] stretch suitable only for reverse transcription of mRNAs). Carry out first-strand cDNA synthesis (e.g., reverse transcription) in a reaction volume of 15 µL consisting of 5 µL of C-tailed RNA (obtained in **step 4**) and 2 µL of Gibco I oligonucleotide (0.5 µg/µL). Denature the mixture at 90°C for 1 min, immediately transfer to 37°C, and incubate for 10 min. Subsequently, add 4 µL of 5X first-strand buffer, 2 µL of 0.1 M DTT, 1 µL of 10 mM deoxynucleotide triphosphate (dNTP) mix, and 1 µL of ddH₂O (all provided by the Gibco Superscript system) to the reaction mix. Prior to adding 5 µL of avian myeloblastosis virus RT (200 U/µL), incubate the reaction mix at 37°C for 2 min. Carry out first-strand synthesis at 45°C for 1.5 h. Perform second-strand DNA synthesis by adding 30 µL of 5X second-strand buffer, 1 µL of *Escherichia coli* DNA ligase, 4 µL of *E. coli* DNA polymerase I, 1 µL (2 U/µL) of *E. coli* RNase-H, and 3 µL of 10 mM dNTPs (all provided by the Gibco Superscript system) to 15 µL of the first-strand cDNA. Incubate the reaction mix at 16°C for 2 h. Add 2 µL of T4 DNA polymerase (5 U/µL) to the reaction mix, and further incubate for 5 min at 16°C before placing the mixture on ice for 5 min. Stop the reaction by adding 10 µL of 0.5 M EDTA (pH 8.0), followed by one phenol/chloroform extraction and ethanol precipitation. Subsequently, dissolve the pellet in 25 µL of RNase-free ddH₂O.
6. Before cloning into the pSPORT1 vector, add *SalI* linker adapters to the 5' and 3' ends of cDNAs as follows: Add 10 µL of 5X T4 ligase buffer, 10 µL of *SalI*

adapter, and 5 μL of T4 DNA ligase (all provided by the Gibco Superscript system) to 25 μL of cDNA in a total volume of 50 μL . Carry out the reaction at 16°C for 16 h. Extract the DNA with phenol/chloroform, precipitate by ethanol, and dissolve in 41 μL of RNase-free ddH₂O. Subsequently, digest the DNA by the *NotI* restriction enzyme. Carry out the reaction in a total volume of 50 μL containing 5 μL of reaction buffer 1 and 4 μL (10 U/ μL) of *NotI* restriction enzyme for 2 h at 37°C (all provided by the Gibco Superscript system). Electrophoretically separate the restriction fragments on a native 6% polyacrylamide (w/v), 1X TBE gel. Visualize the cDNA by staining with ethidium bromide, subsequently excise from the gel, and passively elute in 0.3 M sodium acetate buffer (pH 7.2) at 4°C overnight. Phenol/chloroform extract the DNA precipitate by ethanol, and dissolve in 15 μL of RNase-free ddH₂O.

7. To clone cDNAs into pSPORT 1 vector, from the 15 μL of cDNA prepared in **step 6**, use 1.5 μL for the ligation reaction. Carry out the ligation reaction using enzyme and buffer provided in the Gibco Superscript system. Incubate the ligation reaction, containing 1X ligation buffer, 50 ng of pSPORT1 vector, and about 100 ng of *NotI*-digested cDNAs, at room temperature for 3 h. Extract the ligation products with phenol/chloroform, precipitate by ethanol, and dissolve in 10 μL of ddH₂O. From this, use 3 μL for transformation into *E. coli* strain DH10B competent cells (Gibco-BRL).
8. Amplify the cDNAs by PCR using primers M13 fsp and M13 rsp (*see Subheading 2., items 1b and 1c*) and spot the PCR products with robots in high-density arrays on filters (**19**).

3.1.2. Sequencing and Sequence Analysis of cDNA Clones

1. Sequence cDNA clones using the M13 rsp reverse primer and the BigDye terminator cycle sequencing reaction kit on an ABI Prism 3700 (Perkin-Elmer) sequencer (*see Note 3*).
2. After exclusion of the most abundant known snmRNAs by filter hybridization screening (*see Subheading 3.1.3.*), sequence additional cDNA clones and compare them with each other using the Lasergene Seqman II program package to identify identical clones.
3. Search the GenBank database (NCBI) for all novel sequences using the BLASTN program. All sequences that have not been annotated in GenBank previously can be treated as potential candidates for novel snmRNAs.

3.1.3. Exclusion of cDNA Clones for the Most Abundant, Known Small RNA Species

1. End label oligonucleotides derived from the sequences identified in **Subheading 3.1.2., step 2** with ³³P-(γ)-ATP and T4 polynucleotide kinase using standard molecular biology techniques (**20**).
2. Hybridize the labeled oligonucleotides to cDNA arrays spotted on filters (*see Subheading 3.1.1., step 8*). Prior to hybridization, pretreat the membrane by washing

it twice in prehybridization buffer. Dot-blot hybridization is performed in hybridization buffer 1 at 53°C for 12 h.

3. Wash the filters twice at room temperature for 15 min in wash buffer 1 followed by another washing step in wash buffer 2 for 1 min at 53°C.
4. Expose to a phosphorimaging screen, and rank intensities of hybridization signals by computer-aided analysis as described in **ref. 21**.

3.1.4. Verification of Expression of Cloned snmRNAs by Northern Blot Analysis

1. Fractionate total RNA prepared from the organism or tissue of interest using the TRIzol method described in **Subheading 3.1.1**, on an 8% denaturing polyacrylamide gel (7 M urea, 1X TBE buffer).
2. Transfer onto nylon membranes using the Bio-Rad semidry blotting apparatus, and immobilize RNAs for 1.5 min using a Stratagene UV crosslinker (1200 J/cm²).
3. Incubate the nylon membranes for 1 h in prehybridization buffer.
4. End label 22- to 24-nt-long oligonucleotides complementary to potential novel RNA species with ³²P-(γ)-ATP and T4 polynucleotide kinase, and hybridize the nylon membranes for 12 h at 58°C in hybridization buffer 2.
5. Wash blots twice at room temperature for 15 min in wash buffer 1, and then at 58°C for 1 min in wash buffer 2.
6. Expose the membranes to Kodak MS-1 film for 1 h to 5 d.

3.2. Construction of a Specialized cDNA Library

Construction of a specialized cDNA library is performed by using RNA isolated from a definite subcellular fraction (22), or enriched in the snmRNA type of interest from a total cellular extract (**Fig. 1**; see **Subheading 3.2.1**). Alternatively, or cumulatively, the enrichment can be achieved at the level of cDNA synthesis (see **Subheading 3.2.2**).

3.2.1. Using an RNA Sample Enriched Through Immunoprecipitation

The approach is illustrated in the case of a cDNA library specific for C/D snoRNAs, using antibodies against fibrillarin, a protein that binds specifically all members of this large snmRNA family. The monoclonal antibody 72B9 (provided by M. Pollard) can be used.

1. Rapidly cut freshly prepared or frozen tissues stored at -80°C into small pieces using a sterile cutter, and resuspend in 8 mL of cold NET-150 buffer.
2. Homogenize the cellular sample using the One Shot Model at 2.0 kbar and clarify the extract by centrifugation (12,000g, 10 min, 4°C).
3. In siliconized RNase-free Eppendorf tubes, incubate 10–50 μL of the supernatant of monoclonal culture with 2.5 mg of swollen protein A-Sepharose (PAS) in 0.5 mL of NET-150 buffer for 2 h at room temperature with gentle agitation.

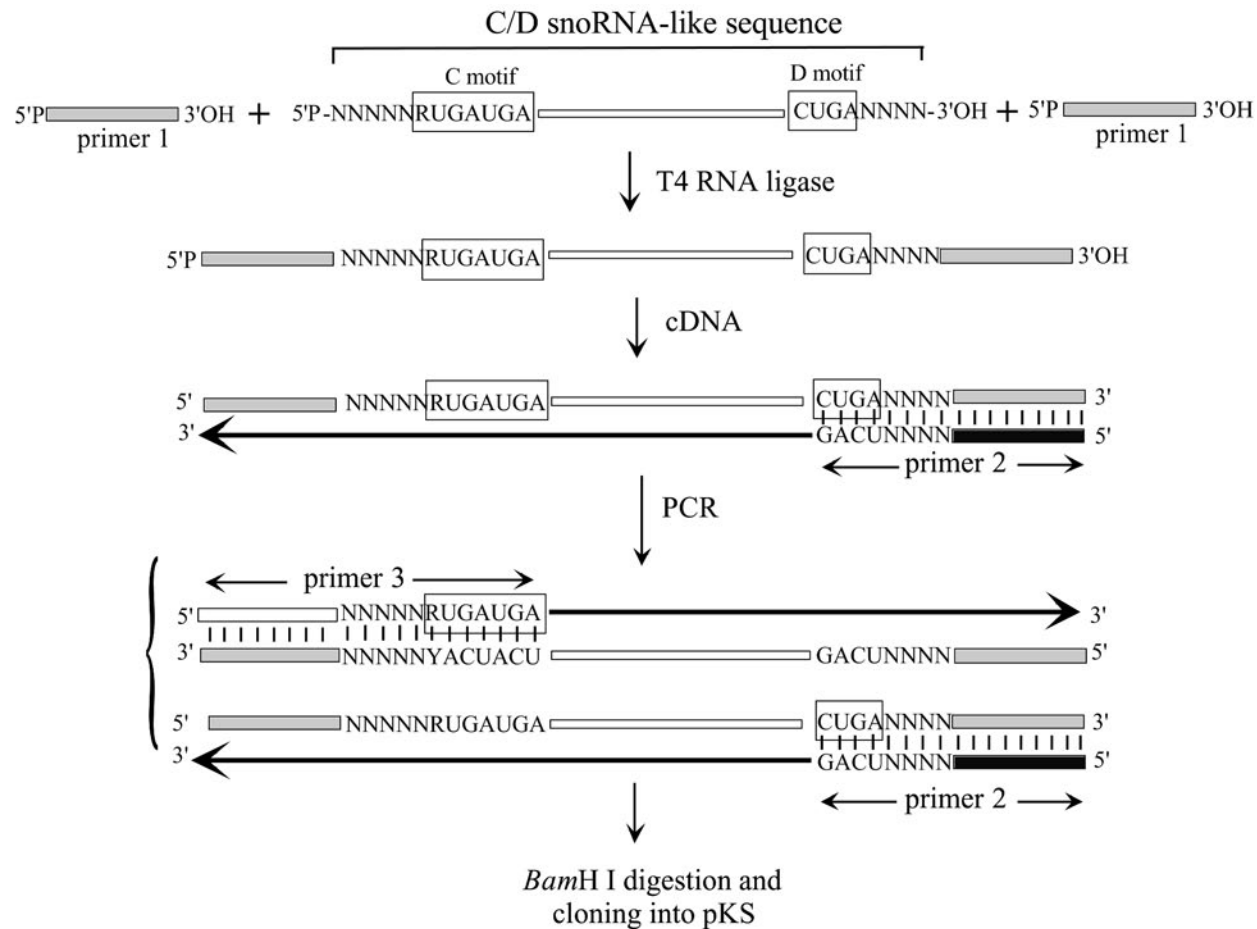
Pellet the PAS-Ig by briefly centrifuging (500g, 20 s), and wash the PAS-Ig two times with 1 mL of NET-150 buffer, before resuspension in 0.5 mL of NET-150.

4. Add 0.5 mL of cell extract supernatant to the 0.5 mL of resuspended PAS-Ig. After 60 min at 4°C with gentle agitation, repellet and wash the PAS-Ig eight times as in **step 3**.
5. To each washed pellet, add 300 μ L of NET-150, 1 μ L of glycogen (20 mg/mL), 30 μ L of SDS (10%), and 300 μ L of water-saturated phenol-chloroform. Incubate for 15 min at 37°C with occasional vortexing.
6. After centrifugation (16,000g, 10 min, at 4°C), recover the aqueous phase, add 40 μ L of 3M Na acetate (pH 5.3), and 1 mL of cold 100% ethanol and chill at -20°C for 30 min. Centrifuge again as previously and wash the pellet with 1 mL of 70% ethanol. Recentrifuge and redissolve the air-dried pellet in 10 μ L of RNase-free water. Store at -20°C.
7. Proceed as in **Subheading 3.1.1**. (starting from **step 4**) or **3.2.2**.

3.2.2. Using an RT-PCR Procedure With Specialized Oligo Primers

Extracted RNAs can be oligonucleotide tagged at both ends by T4 RNA ligase and amplified by RT-PCR (22). Note that the reaction catalyzed by T4 RNA ligase requires RNA substrates with 5' monophosphate and 3' OH termini. In a variant form of this approach, the RT-PCR primers are designed in such a way as to preferentially amplify RNA molecules of the selected snmRNA subtype. This can be performed whenever the snmRNA type of interest always exhibits a pair of sequence signatures at a short, fixed distance from the 5' and 3' termini. This is illustrated in the case of C/D snoRNAs that contain the box C and D motifs only a few nucleotides away from their 5' and 3' termini, respectively (23). Oligonucleotide-tagged RNA molecules are then submitted to RT-PCR using a pair of primers matching the tag as well as the adjacent terminal nucleotides of a typical C/D snoRNA specimen (**Fig. 2**). The 3'-terminal sequence of both primers, termed 2 and 3, matches the conserved box D and C motifs, respectively, thereby ensuring the preferential amplification of C/D snoRNAs.

1. DNA (or RNA) primers must be phosphorylated at their 5' end. Carry out the reaction in 15 μ L of 1X kinase buffer in the presence of 60 pmol of primer 1 and 10 U of T4 polynucleotide kinase through a 60-min incubation at 37°C. Dilute with 200 μ L of RNase-free water, extract by an equal volume of phenol-chloroform (1:1 [v/v]), and precipitate by ethanol using 1 μ L of glycogen (20 mg/mL) as a carrier. Wash the pellet with 70% ethanol, air-dry, and redissolve in 10 μ L of RNase-free water.
2. Ligate the extracted RNAs to phosphorylated primer 1 (*see Note 4*). Carry out the reaction in 20 μ L of 1X ligation buffer in the presence of 20 pmol of phosphorylated primer 1 and 20 U of T4 RNA ligase. Incubate overnight at 4°C. Then dilute with 200 μ L of RNase-free water and proceed as in **step 1**.



3. Reverse transcribe the ligation product (redissolved in RNase-free water) using primer 2, in a final reaction volume of 20 μL . Add 6 μL of RNase-free water to 5 μL of the solution of oligonucleotide-tagged RNA (obtained in **step 2**) and 1 μL of primer 2 solution (0.5 $\mu\text{g}/\mu\text{L}$). Heat the mixture at 70°C for 5 min and quickly chill on ice. Add immediately 4 μL of 5X first-strand buffer, 2 μL of 0.1 M DTT, and 1 μL of dNTP (10 mM each). Mix and incubate for 2 min at 42°C. Add 200 U of Superscript® II RNaseH-Reverse transcriptase (Life Technologies), and incubate for 2 h at 42°C.
4. Use the resulting cDNA product as template for PCR by *Taq* polymerase I with primers 2 and 3. Carry out the PCR reaction in a final volume of 100 μL in 1X *TaqI* buffer after the addition of 6 μL of MgCl_2 (25 mM), 2 μL of dNTPs mix (10 mM each), 2 μL of the reverse transcription reaction mix incubated in **step 3**, 1 μL of primer 2 and 3 solutions (0.5 $\mu\text{g}/\mu\text{L}$), and 2.5 U of *TaqI* DNA polymerase. Perform 30 cycles, with a 1 min-denaturation step at 90°C, a 1-min hybridization step at 55°C, and a 1-min elongation step at 72°C. Extract by an equal volume of phenol-chloroform (1:1 [v/v]) and precipitate by ethanol.
5. Digest the PCR product overnight at 37°C by *Bam*HI (the PCR amplification primers are designed to contain a *Bam*HI site), and clone in pKS vector by standard recombinant DNA methods (20). Individual clones can be manually sequenced (T7 sequenase version 2.0, U.S. Biochemical). About 70–80% of clones correspond to C/D snoRNA-like sequences (see **Note 5**).

3.3. Analysis of snmRNA Sequences

Individual cloned cDNAs may be significantly shorter than the corresponding snmRNAs determined by Northern blotting (see **Subheading 3.1.4.**). This is owing to the cloning strategy described in **Subheading 3.1.1.**, which may result in 5'-truncated cDNAs. For organisms with a completely sequenced genome, the missing 5' terminal sequence of the snmRNA can be inferred from the genomic sequence, taking into account the size determined by Northern blotting. All snmRNA sequences are then searched for the presence of known sequence or structure motifs typical of a particular snmRNA subtype, especially the two widespread C/D and H/ACA snoRNA families (see **Subheading**

Fig. 2. (see opposite page) Preferential amplification of specific RNA subset by RT-PCR procedure with specialized primers. C/D snoRNAs immunoprecipitated by the antifibrillar antibody have been oligonucleotide tagged at both ends. They are then amplified by RT-PCR using primers 2 and 3, which are designed to amplify only RNA molecules exhibiting the prevalent terminal structure of C/D snoRNAs, i.e., the RUGAUGA (box C) and CUGA (box D) positioned at 5 nt from their 5' end and 4 nt from their 3' end, respectively. Accordingly, both primers contain five and four degenerated positions to match the snoRNA nucleotides upstream from box C and downstream from box D, respectively.

3.3.1.). Motifs can be searched using the DNAMAN sequence analysis program. Folding of snmRNA structures can be predicted using the mfold version 3.0 by Zuker and Turner (24).

3.3.1. Search for snoRNA Sequences

3.3.1.1. C/D snoRNAs

1. Identify potential C/D snoRNA candidates on the basis of the box C (RUGAUGA) and box D (CUGA) signatures, located 2–8 nt from the 5' and 3' end, respectively. Take into account that a few box C variants are sometimes observed (see **Fig. 3** and **Note 6**).
2. Look for a potential complementarity between the nucleotides downstream from box D and upstream from box C (**Fig. 4**), which could provide for a typical 5'-3'-terminal stem (of at least 3 to 4 bp) of the RNA, thereby strongly supporting the C/D assignment (25).
3. If a 5'-3'-terminal stem is absent, look for an inverted repetition flanking the mature snoRNA-coding region in the genomic sequence, which may provide a functional substitute (see **Note 7**).
4. Confirm the assignment to the methylation guide type of C/D snoRNAs by looking for the presence of potential boxes D' (CUGA with at most one deviation) and C' (UGAUGA with at most two deviations) in that order, within the box C–box D interval (26). The bona fide box C' must be located more than 3 nt downstream from box D' and more than 10 nt upstream from box D (**Fig. 4**).
5. If the length of the candidate snmRNA sequence largely exceeds the usual size range of C/D snoRNAs (see **Note 8**), test the possibility that it corresponds to a C/D-H/ACA snoRNA chimera (11,27) by looking for H/ACA snoRNA structural signatures in its central region, as in **step 3** in **Subheading 3.3.1.2**.

3.3.1.2. H/ACA snoRNAs

Box H/ACA snoRNAs share a common secondary structure consisting of two large hairpin domains linked by a hinge containing the H motif and followed by a short tail including the ACA motif (28), as shown in **Fig. 5**.

1. Select candidates on the basis of an ACA trinucleotide located 3 nt from the 3' end (see **Note 9**).
2. Delineate all potential H box motifs (ANANNA, where N is any nucleotide) along each candidate RNA sequence and focus on those located at least 50 nt from the 5' or 3' end.
3. Using the Zuker program, fold separately each portion of the snoRNA sequence upstream and downstream from each H box selected in **step 2**.
4. For both the 5' and 3' domains, examine the five most stable structure predictions to identify any folding pattern resembling the typical stem–internal loop–stem organization (**Fig. 5**) of an H/ACA stem domain (28). Always use as foremost criterion the presence of a terminal helix of at least 5 bp positioned 1–3 nt

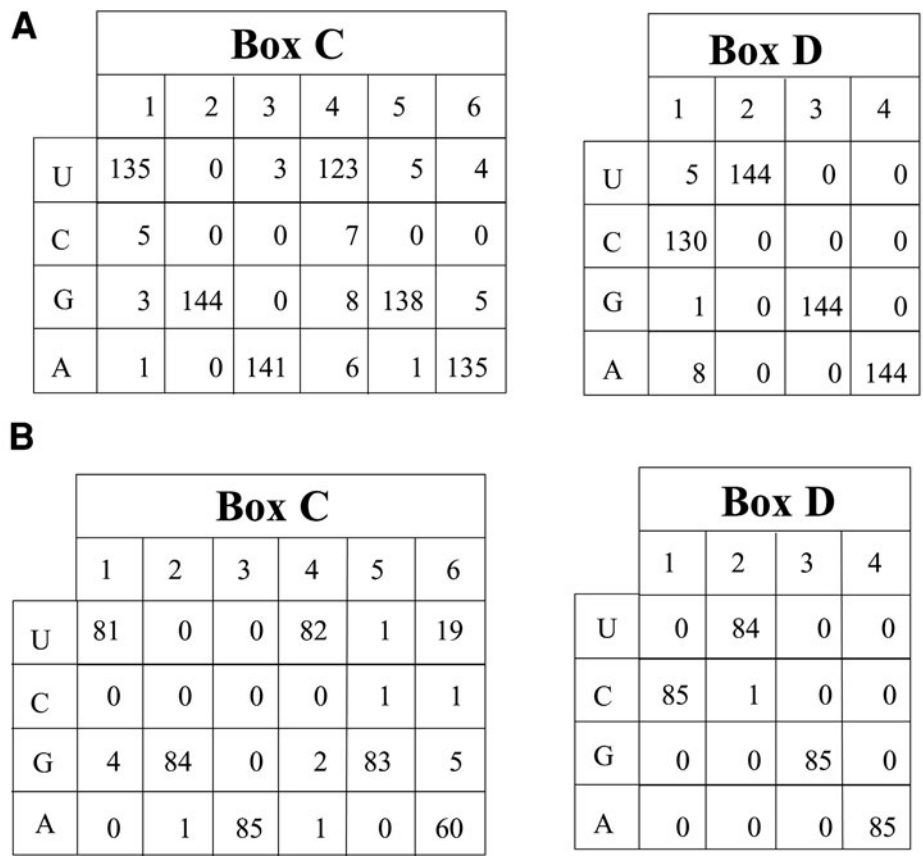


Fig. 3. Nucleotide conservation within C and D boxes. For each position in both boxes, the number of each of the four ribonucleotides has been totaled for all known C/D snoRNAs, in plant *Arobidopsis thaliana* (A) and mammals (B).

- upstream from the ACA (or H) box and enclosing a domain of at least 40 nt. Take into account that the single-strand hinge linking the two stem domains of an H/ACA snoRNA does not usually exceed 12 nt. Within each H/ACA stem domain, delineating the boundaries of the large internal loop may be problematic on a thermodynamic basis alone, owing to spurious complementarities between nucleotides in the loop. However, for H/ACA stem domains guiding pseudouridylation of an identified RNA target site (*see steps 1–3 in Subheading 3.3.3.*), the two bipartite antisense elements define unambiguously the loop boundaries on both strands.
5. In each potential H/ACA stem domain, try to identify a plausible apical stem (typically of at least 6 bp, possibly including a small internal loop) linking the two elements of the bipartite pseudouridylation guide RNA duplex (Fig. 5B). The

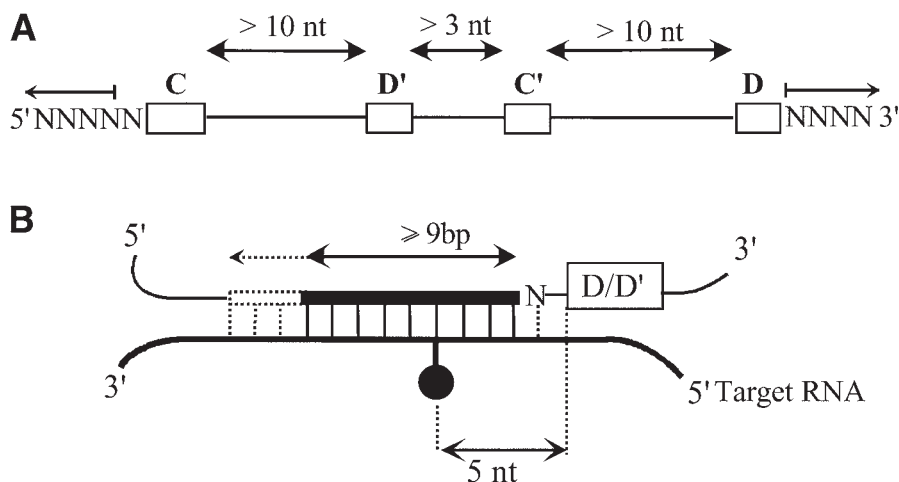


Fig. 4. C/D snoRNAs and their RNA target identification. (A) Typical structure of a methylation guide snoRNA (arrows in opposite orientation denote base complementarity); (B) canonical structure of guide duplex (with optional base pairings denoted by dots and position targeted for 2'-O-methylation by a solid circle).

3'-most nucleotide of this helix must be located between positions -14 and -17 relative to the downstream H or ACA box. In the absence of any properly positioned apical stem in at least one of the two folded large domains, discard this snmRNA as a presumptive H/ACA snoRNA of the pseudouridylation guide type (see **Note 10** and **Fig. 5**).

3.3.2. Identification of Sites Targeted by C/D snoRNAs

1. For each C/D candidate, identify all potential box D' motifs that are appropriately located (i.e., more than 10 nt downstream from box C and more than 20 nt upstream from box D) and followed by a potential box C' motif (see **step 4** in **Subheading 3.3.1.1**).
2. Search the sequence tracts immediately upstream from box D and presumptive box(es) D' for complementarities to rRNA or spliceosomal snRNAs of the organism of interest. Select each 9-nt-long tract extending upstream from position -2 relative to the D/D' box (**Fig. 4**). Using the DNAMAN program, search for the complementary sequence in your rRNA/snRNA sequence database for this organism, allowing any G.U pair (29). For every match, check whether the corresponding RNA duplex can be extended by a few additional base pairs in the direction opposite to the D/D' box. Select as likely antisense elements the sequence tracts exhibiting an rRNA or snRNA complementarity including at least 9 Watson-Crick base pairs and no more than 3 G.U. For each likely antisense ele-

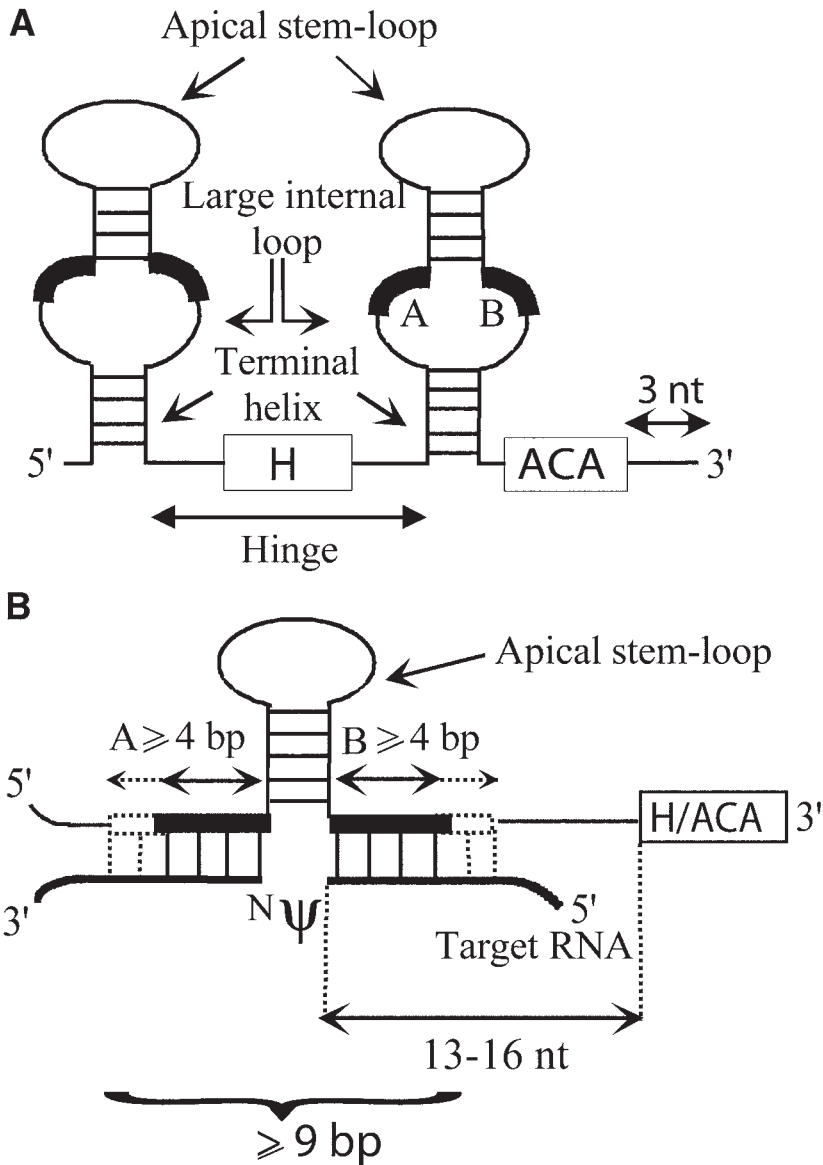


Fig. 5. H/ACA snoRNAs and their RNA target identification. **(A)** Typical structure of a pseudouridylation guide snoRNA; **(B)** canonical structure of pseudouridylation guide duplex (with optional base pairings denoted by dots and uridine targeted for pseudouridylation by ψ).

ment, the position in rRNA or snRNA that is targeted for 2'-*O*-methylation is paired to the fifth nucleotide upstream from box D or D' in the corresponding RNA duplex (**Fig. 4**).

3. Examine whether the target nucleotide corresponds to an rRNA or snRNA position previously found to be 2'-*O*-ribose methylated in a few eukaryal model organisms (**25,30–32**). The predicted methylation can be verified experimentally, by primer extension performed at low dNTP concentration (**33**). Remember that a subset of C/D snoRNA candidates, called orphan snoRNAs, remains without a predicted rRNA or snRNA target in various organisms (**10,16**).

3.3.3. Identification of Uridines Targeted by H/ACA snoRNAs

Nucleotides in the large internal loop of one or both large hairpin domains can form a bipartite guide duplex around the uridine targeted for pseudouridylation (**34**). While this uridine remains unpaired, as well as the adjacent downstream nucleotide (**Fig. 5B**), the combined length of the two flanking stems (each one of at least 4 bp) must be of at least 9 bp.

1. For each large stem domain of the H/ACA snoRNA, select the folding pattern(s) consistent with a pseudouridylation guide function, by reference to the most plausible apical helix (*see steps 4 and 5 in Subheading 3.3.1.2.*).
2. In the presumptive large internal loop, select the two 4-nt sequences abutting the apical stem, termed *tracts A* and *B* (**Fig. 5**), and test their potential to form a canonical guide duplex around any rRNA or snRNA uridine. Using the DNAMAN program, search your rRNA/snRNA database of the organism of interest for the presence of a matching 10-nt sequence of the form 5'(B*)UN(A*)3', where N stands for any nucleotide and (B*) and (A*) for 4-nt sequences complementary to tracts A and B, respectively. Tolerate G.U pair in this search, but discard matches including three or more G.U.
3. For the remaining matches, examine whether the RNA duplex can be extended by a few additional base pairs in both directions. Finally, consider only matches including at least 9 Watson-Crick base pairs.
4. Examine whether the predicted target uridine corresponds to a known rRNA or snRNA pseudouridine in eukaryal organisms examined so far (**35,36**). Predicted pseudouridines can be verified experimentally, by an RT approach following chemical modification by (*N*-cyclohexyl-*N'*-delta-(4-methyl-morpholinium)ethyl-carbodiimide p-tosylate) (CMC) and alkaline treatment of the RNA template (**36**). Remember that a subset of H/ACA snoRNA candidates, called orphan snoRNAs, remains without a predicted rRNA or snRNA target in various organisms (**10,16**).

4. Notes

1. After isolation of the RNA of choice by the TRIzol method (Gibco-BRL), we find it useful to perform an additional one or two phenol/chloroform extractions to remove residual amounts of proteins still attached to RNAs.

2. When running total RNA on denaturing polyacrylamide gels for size separation, try to minimize the running distance of RNA into the gel because otherwise you will have too much gel material from which the RNA has to be passively eluted: the more gel material you end up with, the larger the elution volume and hence the more difficult it will be to quantitatively precipitate size-selected RNA from that solution.
3. When employing the C-tailing method, sequence analysis of clones will be severely impeded when a primer downstream from the C-tail is used. According to our experience, sequencing through the C-tail to determine the snmRNA gene of interest gives very insufficient results. Hence, for sequence analysis a primer upstream from the snmRNA gene of interest is recommended.
4. Because the ligation reaction is not oriented in terms of 5' and 3' end reactivity, this step can lead to the production of various forms of primer-RNA concatemers, which may result in the cloning of inserts spanning two (or more) C/D snoRNA sequences.
5. C/D snoRNAs exhibiting a variant box C or D sequence may be underrepresented in the amplified products, particularly when the deviation occurs at the last and the first positions of the C and D boxes, respectively (*see Note 6*). This can also be the case for the few snoRNAs exhibiting an atypical length of sequence upstream from box C and downstream from box D, respectively. To ensure a better representation of such C/D snoRNAs in the library, modified versions of primers 2 and 3 can be designed and used in parallel in the RT-PCR step.
6. A substantial fraction of C/D methylation guide snoRNAs identified in plant *A. thaliana* and mammals (33 and 37%, respectively) exhibit a 1-nt variation over their box C motif (**Fig. 3**). Positions 2 and 3 following the initial purine of box C (i.e., G and A, respectively) are by far the most conserved while position 6 is the most variable. In mammals, the most frequent variant form of box C is RUGAUGU. In both *A. thaliana* and mammals, deviation from the consensus is much less frequent for box D than box C, and G and A at positions 3 and 4, respectively, remain strictly invariant (**Fig. 3**). The strongly conserved GA dinucleotides in each box form two tandem sheared G.A pairs essential for the recognition of the box C/D core motif by the 15.5-kDa (yeast snu13p) snoRNP protein (**37**).
7. Although the 5'-3' terminal stem is essential for snoRNA accumulation, several C/D snoRNAs lack a terminal helix (**25**). The accumulation of these is supported by external stem structures located within the flanking intronic sequences in the snoRNA precursor (**38**).
8. A few C/D snoRNAs that do not belong to the methylation guide type significantly depart from the consensus structure depicted in **Fig. 4**. These include U3, U8, and U13 snoRNAs, which all exhibit a much longer 5' extension upstream from box C, as well as an extended box C–box D interval. In C/D snoRNAs of the methylation guide type, the length of the box C–box D interval usually ranges from 50 to 100 nt. However, some substantially longer specimens can be found, which include C/D-H/ACA snoRNA chimeras guiding some snRNA methylations and pseudouridylations, together with a few C/D snoRNAs remaining without identified RNA target (**16**).

9. Two variant forms of the ACA motif, AUA and AGA, have been observed very infrequently.
10. Most H/ACA snoRNAs have sizes ranging from 120 to 140 nt. However, a substantial number of larger specimens have been detected, such as CD/H/ACA chimeras (*see step 5 in Subheading 3.3.1.1.*), H/ACA tandems (**39**), as well as some three-stem domain H/ACA snoRNAs (**17,18**). The enlarged size of a few additional specimens usually reflects the occasional presence of an additional hairpin within or between the two large stem domains (**Fig. 5**).

Acknowledgments

A.H. was supported by the German Human Genome Project through the BMBF (#01KW9966) and by an IZKF grant (Teilprojekt IKF3 G6, Münster). J.P.B. and J.C. were supported by laboratory funds from the Centre National de la Recherche Scientifique and Université Paul Sabatier, Toulouse, and by a grant from the Ministère de l'Éducation Nationale de la Recherche et de la Technologie (Programme de Recherche Fondamentale en Microbiologie et Maladies Infectieuses et Parasitaires, 2001–2002).

References

- 1 Couzin, J. (2002) Breakthrough of the year: small RNAs make big splash. *Science* **298**, 2296–2297.
- 2 Dennis, C. (2002) Small RNAs: the genome's guiding hand? *Nature* **420**, 732.
- 3 Dennis, C. (2002) The brave new world of RNA. *Nature* **418**, 122–124.
- 4 Huttenhofer, A., Brosius, J., and Bachellerie, J. P. (2002) RNomics: identification and function of small, non-messenger RNAs. *Curr. Opin. Chem. Biol.* **6**, 835–843.
- 5 Storz, G. (2002) An expanding universe of noncoding RNAs. *Science* **296**, 1260–1263.
- 6 Gottesman, S. (2002) Stealth regulation: biological circuits with small RNA switches. *Genes Dev.* **16**, 2829–2842.
- 7 Tuschl, T. (2002) Expanding small RNA interference. *Nat. Biotechnol.* **20**, 446–448.
- 8 Tuschl, T. (2003) Functional genomics: RNA sets the standard. *Nature* **421**, 220, 221.
- 9 Ambros, V. (2001) microRNAs: tiny regulators with great potential. *Cell* **107**, 823–826.
- 10 Bachellerie, J. P., Cavaille, J., and Huttenhofer, A. (2002) The expanding snoRNA world. *Biochimie* **84**, 775–790.
- 11 Kiss, T. (2002) Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell* **109**, 145–148.
- 12 Lau, N. C., Lim, L. P., Weinstein, E. G., and Bartel, D. P. (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294**, 858–862.
- 13 Lee, R. C. and Ambros, V. (2001) An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **294**, 862–864.

- 14 Rhoades, M. W., Reinhart, B. J., Lim, L. P., Burge, C. B., Bartel, B., and Bartel, D. P. (2002) Prediction of plant microRNA targets. *Cell* **110**, 513–520.
- 15 Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I., and Martienssen, R. A. (2002) Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**, 1833–1837.
- 16 Huttenhofer, A., Kieffmann, M., Meier-Ewert, S., O'Brien, J., Lehrach, H., Bachellerie, J. P., and Brosius, J. (2001) RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. *EMBP J.* **20**, 2943–2953.
- 17 Marker, C., Zemmann, A., Terhorst, T., Kieffmann, M., Kastenmayer, J. P., Green, P., Bachellerie, J. P., Brosius, J., and Huttenhofer, A. (2002) Experimental RNomics: identification of 140 candidates for small non-messenger RNAs in the plant *Arabidopsis thaliana*. *Curr. Biol.* **12**, 2002–2013.
- 18 Tang, T. H., Bachellerie, J. P., Rozhdestvensky, T., Bortolin, M. L., Huber, H., Drungowski, M., Elge, T., Brosius, J., and Huttenhofer, A. (2002) Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*. *Proc. Natl Acad. Sci. USA* **99**, 7536–7541.
- 19 Schmitt, A. O., Herwig, R., Meier-Ewert, S., and Lehrach, H. (1999) High density cDNA grids for hybridization fingerprinting experiments, in *PCR Applications: Protocols for Functional Genomics* (Innis, M. A., Gelfand, D. H., and Sninsky, J. J., eds.), Academic, San Diego, pp. 457–472.
- 20 Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning, A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- 21 Maier, E., Meier-Ewert, S., Ahmadi, A. R., Curtis, J., and Lehrach, H. (1994) Application of robotic technology to automated sequence fingerprint analysis by oligonucleotide hybridisation. *J. Biotechnol.* **35**, 191–203.
- 22 Kiss-Laszlo, Z., Henry, Y., Bachellerie, J. P., Caizergues-Ferrer, M., and Kiss, T. (1996) Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell* **85**, 1077–1088.
- 23 Cavaille, J., Vitali, P., Basyuk, E., Huttenhofer, A., and Bachellerie, J. P. (2001) A novel brain-specific box C/D small nucleolar RNA processed from tandemly repeated introns of a noncoding RNA gene in rats. *J. Biol. Chem.* **276**, 26,374–26,383.
- 24 Zuker, M. (1994) Prediction of RNA secondary structure by energy minimization. *Methods Mol. Biol.* **25**, 267–294.
- 25 Bachellerie, J. P. and Cavaille, J. (1997) Guiding ribose methylation of rRNA. *Trends Biochem. Sci.* **22**, 257–261.
- 26 Kiss-Laszlo, Z., Henry, Y., and Kiss, T. (1998) Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. *EMBO J.* **17**, 797–807.
- 27 Darzacq, X., Jady, B. E., Verheggen, C., Kiss, A. M., Bertrand, E., and Kiss, T. (2002) Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. *EMBO J.* **21**, 2746–2756.

- 28 Ganot, P., Caizergues-Ferrer, M., and Kiss, T. (1997) The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. *Genes Dev.* **11**, 941–956.
- 29 Cavaille, J. and Bachellerie, J. P. (1998) SnoRNA-guided ribose methylation of rRNA: structural features of the guide RNA duplex influencing the extent of the reaction. *Nucleic Acids Res.* **26**, 1576–1587.
- 30 Maden, B. E. (1990) The numerous modified nucleotides in eukaryotic ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol.* **39**, 241–303.
- 31 Ofengand, J. and Bakin, A. (1997) Mapping to nucleotide resolution of pseudouridine residues in large subunit ribosomal RNAs from representative eukaryotes, prokaryotes, archaebacteria, mitochondria and chloroplasts. *J. Mol. Biol.* **266**, 246–268.
32. Massenet, S., Mouglin, A., and Branlant, C. (1998) Posttranscriptional modifications in the U snRNAs, in *Modification and Editing of RNA: The Alteration of RNA Structure and Function* (Grosjean, H. and Benne, R. E., eds.), ASM Press, Washington, DC.
- 33 Maden, B. E., Corbett, M. E., Heeney, P. A., Pugh, K., and Ajuh, P. M. (1995) Classical and novel approaches to the detection and localization of the numerous modified nucleotides in eukaryotic ribosomal RNA. *Biochimie* **77**, 22–29.
- 34 Ganot, P., Bortolin, M. L., and Kiss, T. (1997) Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell* **89**, 799–809.
35. Ofengand, J. and Rudd, K. (2000) Bacterial, archaea, and organellar RNA pseudouridines and methylated nucleosides and their enzymes, in *Ribosome: Structure, Function, Antibiotics, and Cellular Interactions* (Garrett, R., Douthwaite, S., Liljas, A., Matheson, A., Moore, P. B., and Noller, H. E., eds.), ASM Press, Washington, DC.
36. Ofengand, J. and Fournier, M. J. (1998) The pseudouridine residues of rRNA: number, location, biosynthesis and function, in *Modification and Editing of RNA: The Alteration of RNA Structure and Function* (Grosjean, H. and Benne, R. E., eds.), ASM Press, Washington, DC.
- 37 Vidovic, I., Nottrott, S., Hartmuth, K., Luhrmann, R., and Ficner, R. (2000) Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment. *Mol. Cell* **6**, 1331–1342.
- 38 Darzacq, X. and Kiss, T. (2000) Processing of intron-encoded box C/D small nucleolar RNAs lacking a 5',3'-terminal stem structure. *Mol. Cell. Biol.* **20**, 4522–4531.
39. Kiss, A. M., Jady, B. E., Darzacq, X., Verheggen, C., Bertrand, E., and Kiss, T. (2002) A Cajal body-specific pseudouridylation guide RNA is composed of two box H/ACA snoRNA-like domains. *Nucleic Acids Res.* **30**, 4643–4649.