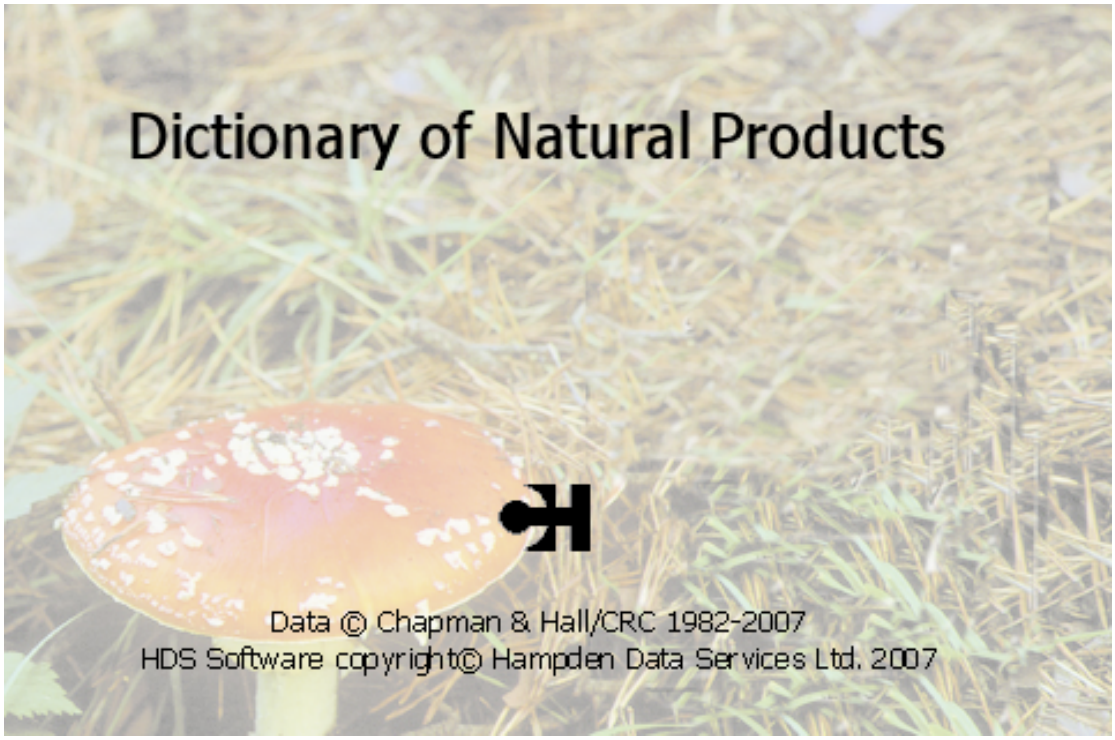

Dictionary of Natural Products
Data © Chapman & Hall/CRC 1982-2007
HDS Software copyright© Hampden Data Services Ltd. 2007

User Manual

Table Of Contents

Welcome 1
About the Database 1
What's New? 2
Minimum System Requirements 3
Technical Support 4
Caution 5
Antivirus Software and DNP 6
About 7
Editorial Team 8
Editorial 8
Contributors 8
International Advisory Board 8
Special Editors 8
Copyright Notices 10
Getting Started Guide 11
Installation 11
Working with multiple windows 13
The Search Form Window 14
The Hit List Window 16
The Entry Display Window 17
Text and Data Searching 18
Entering a Search Term 18
Available Fields 19
Selecting a Searchable Field 26
Using the Search Term Box 27
Using the Index Stem Box 28
Advanced Text and Data Searching 29
Truncation and Wildcards 29
Range Searching 30
Combining Search Terms (Boolean operators) 31
Linking Fields (Boolean operators) 32
Special Symbols 33
Searching the Database 34
Structure Searching 35
Manipulating Search Results 38
Examining Search Results 38
Viewing Entries 40
Combining Hits 44
Display and Print Options 45
Printing 48
Export Data 50
Edit Preferences 51
Toolbars and Menus 53
File Menu 53
Edit Menu 54
Display Menu 55
Search Menu 56
HitList Menu 57
View Menu 58
Window Menu 59
Type of Compound Codes 60
Overview 60
Aliphatic Natural Products 61
Alkaloids 64
Amino Acid and Peptide Natural Products 69

Benzofuranoids 70
Benzopyranoids 71
Carbohydrate Natural Products 72
Drugs 75
Flavonoids 81
Lignans 84
Miscellaneous Natural Products 85
Oxygen Heterocycles 86
Polycyclic Aromatic Natural Products 87
Polyketides 89
Polypyrroles 90
Steroid Natural Products 93
Tannins 94
Terpenoids 95
Abbreviations 103
General 103
Reference Tags 108
Amino Acids 110
α-Amino acids commonly found in peptides and proteins 110
Other Amino Acids 111
Frequently Asked Questions 113

Welcome

About the Database

The Dictionary of Natural Products on CD-ROM (DNP) is a chemical database containing over 199,000 compounds.

The information contained on the chemical database includes descriptive and numerical data on chemical, physical and biological properties of compounds; systematic and common names of compounds; literature references; structure diagrams and their associated connection tables.

The incorporation of data from the **Berdy Bioactive Natural Products** and **DEREP** databases extends the coverage of UV spectra and biological activity.

The CD-ROM contains powerful search software from Hampden Data Services to provide text and structure searching. Previous experience with online searching, or with the use of CD-ROM products is unnecessary, although a basic familiarity with Microsoft® Windows® is assumed.

Abbreviations are frequently used within the text.

Please refer to the database introduction for a complete description of data organisation and presentation in the Dictionary of Natural Products. The introduction can be found under the Chapman & Hall/CRC program group in the Start menu (You will need the Acrobat Reader installed on your computer to access the introduction).

This manual is also available in electronic format by pressing **F1** in the main program, or via the program group in the Windows Start menu.

What's New?

- **CAS Registry Numbers**

  Extra and related CAS numbers can include descriptive text to help identification, e.g. 123456-78-9 (hydrochloride). These tags will become more widespread as entries are updated.

  Extra CAS Registry Number(s)
  Compounds (main entry level, variant or derivative) may have more than one number associated with them. In such cases, the additional numbers representing the compound in question are termed 'Extra CAS Registry Numbers' and are held at compound level. Such numbers will typically refer to simple variations on the compound in question, e.g. stereoisomers, salts or duplicate number assignments.

  Related CAS Registry Number(s)
  Other registry numbers which cannot be assigned to a particular variant or derivative, but are none the less related to the entry compound are termed 'Related CAS Registry Numbers' and are held at entry level. These may refer to derivatives of the entry compound which are not present in the entry.

- **Data enhancements**

  Now over 199,000 compounds

Minimum System Requirements

Computer	PC with Pentium II processor or equivalent
Hard drive	40MB (standard installation) 550MB (hard-disk installation)
RAM Memory	32MB (or the minimum required by your operating system)
Display	800 x 600 (1024 x 768 or higher recommended)
Operating System	Microsoft® Windows® 2000 or XP

Technical Support

For support on installation or use of the program, please contact:

Technical Support Desk
Electronic Publishing Division
CRC Press LLC
6000 Broken Sound Parkway NW
Boca Raton FL 33487 USA
Tel: 888 318 2367
Tel: +1 561 361 6020 (Outside US)
Fax: +1 561 361 6075
E-mail: **e-products@crcpress.com**

If your query concerns your system setup rather than problems with searching the database, please be prepared to supply the following information to the Technical Support Desk staff when you call:

- description of the problem, including any error messages displayed
- make and model of PC, including size of RAM and hard disk
- make and model of CD drive
- version of operating system
- contents of SYSTEM.INI and WIN.INI files
- display driver details under Display Settings in the Control Panel

For comments on the data or suggestions for inclusion, please contact:

The Editors, The Dictionary of Natural Products
Chapman & Hall/CRC
24-25 Blades Court
Deodar Road
London SW15 2NU
United Kingdom
Fax: +44 (0)20 7017 6747
Email: **marion.schumacher@informa.com**

Caution

Treat all chemical substances as if they have dangerous properties. The publisher makes no representation, express or implied, with regard to the accuracy of the information contained in this Dictionary, and cannot accept any legal responsibility or liability for any errors or omissions that may be made.

The specific information in this publication on the hazardous and toxic properties of certain substances is included to alert the reader to possible dangers associated with the use of those compounds. The absence of such information should not however be taken as an indication of safety in use or misuse.

Antivirus Software and DNP

Please note that some antivirus software can cause structure searching to proceed very slowly. This seems to be especially so if the virus scanner is set to perform 'heuristic' scanning of all files. It may be possible to exclude the files in the installation directory from being scanned, but users are urged to exercise caution. Please consult your antivirus software for more details.

About

Published by Chapman & Hall/CRC
6000 Broken Sound Parkway NW
Boca Raton FL 33487 USA

A member of the Taylor & Francis Group. An Informa company.

CD ROM

ISBN 0-412-49150-8
ISSN 0966-2146

User Manual
May 2006

Editorial Team

Editorial

Executive Editor

J. Buckingham	Consultant

Project Editors

A. M. Hodgson	Consultant
S. P. Walford	Chapman & Hall/CRC

Editorial Advisor

L. Szabo	(alkaloids)	Semmelweis University, Budapest

Contributors

K. Baggaley
C. M. Cooper
M. W. Edwards
R. A. Hill
C. F. Pattenden
R. Purchase
P. H. Rhodes
A. D. Roberts
P. Thornton
T. J. Ward

International Advisory Board

J. D. Connolly	Glasgow University
C. Djerassi	Stanford University
D. J. Faulkner	University of California, San Diego
K. Mori	Science University of Tokyo
K. Nakanishi	Columbia University, New York
G. Ourisson	Universite Louis Pasteur, Strasbourg
R. A. Raphael*	Formerly of University of Cambridge
M. Shamma	Pennsylvania State University
Ch. Tamm	University of Basel

*Deceased

Special Editors

D. C. Ayres	(lignans)	Formerly of Queen Mary & Westfield College, London
B. W. Bycroft	(antibiotics and peptides)	University of Nottingham
P. M. Collins	(carbohydrates)	Birkbeck College, London
F. D. Gunstone	(lipids)	The Scottish Crop Research

		Institute
J. B. Harborne*	(flavonoids)	Reading University
E. Haslam	(tannins)	Sheffield University
R. A. Hill	(terpenoids and steroids)	Glasgow University
D. R. Kelly	(semiochemicals)	University of Wales, College of Cardiff
F. J. Leeper	(polypyrroles)	University of Cambridge
R. D. H. Murray	(coumarins)	Formerly of Glasgow University
A. D. Roberts	(antibiotics)	Consultant
I. W. Southon	(alkaloids)	Consultant

*Deceased

Copyright Notices

The Editors, The Combined Chemical Dictionary
Chapman & Hall/CRC
24-25 Blades Court
Deodar Road
London SW15 2NU
United Kingdom
Fax: +44 (0)20 7017 6747
Email: **marion.schumacher@informa.com**

Getting Started Guide

Installation

Insert the CD in the CD-ROM drive and click on the **Start** button and select **Run**.

In the command line type **D:\setup.exe** (where D is your CD-ROM drive letter) then click **OK**.

If you would like a shortcut added to your start menu (includes a shortcut to the help text, manual and introduction), check the appropriate box.

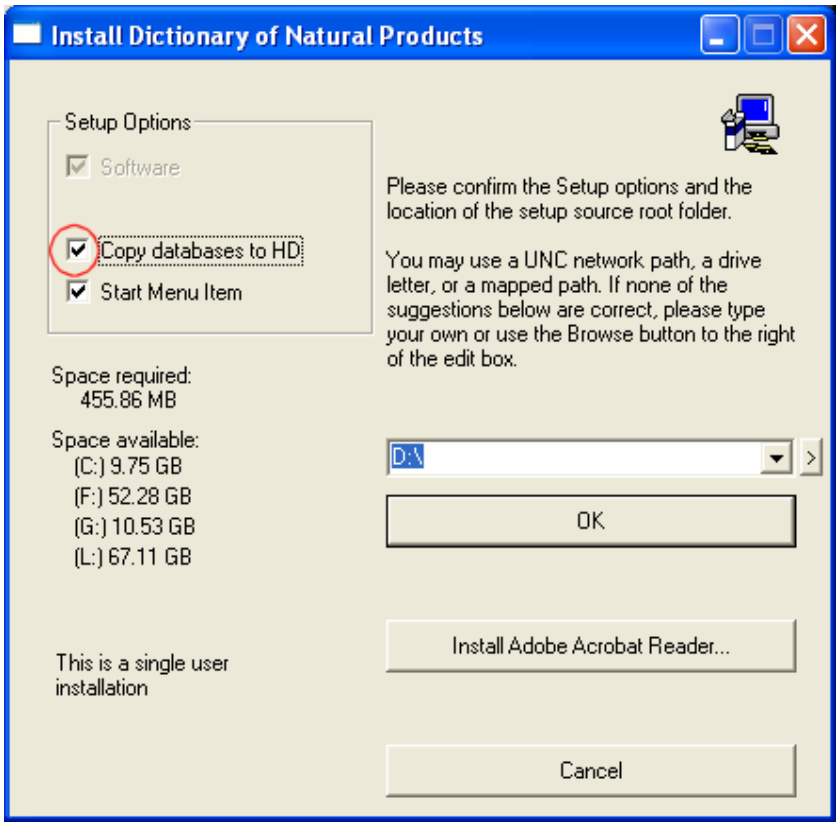
You have two Installation options:

- **Standard installation**

  The default installation copies the program files only to your computer's hard drive. Use this option if your computer has limited hard disk space. You will need to insert the CD every time you use the Dictionary of Natural Products.

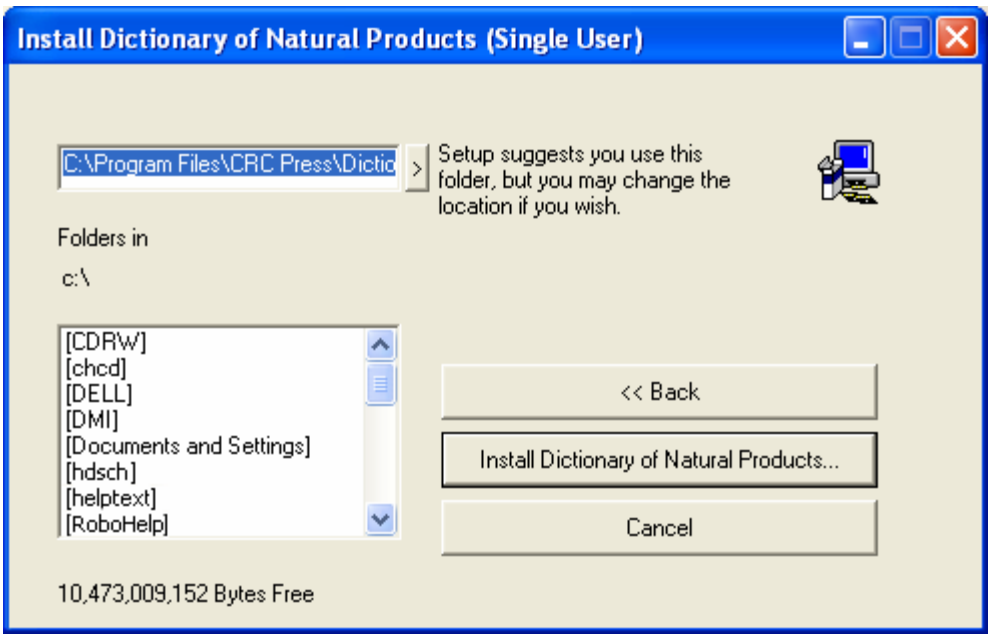
- **Hard disk installation**

  If you would like to copy the entire database to your hard drive please check the option **Copy database to HD.** Please note this will require approximately 550MB of hard disk space. Use this option if you have ample hard disk space (*Due to Windows requirements we recommend only using this option if you have at least 1GB free hard disk space, otherwise system performance may be affected*).



Check that the path to your CD-ROM drive is correct and then press **OK**.

You will be prompted to select a folder in which to install the software. If you wish the program to be installed to the default folder **C:\Program Files\CRC Press\Dictionaries**, just click **Install Dictionary of Natural Products**. Otherwise, type a new folder location and then click **Install Dictionary of Natural Products**.



For network installations please refer to the instructions provided with the network install disk.

Working with multiple windows

When you first start the Dictionary of Natural Products you are presented with the **Search Form** window. There are three main windows in the Dictionary of Natural Products:

The Search Form: Use this to prepare your text and structure queries

The Hit List: Shows the results of your search (this window will be displayed after you click **Search** on the **Search Form**)

The Entry Display: Displays the individual entries in the dictionary (this window will be displayed after you select an entry from the **Hit List**)

Windows can be arranged via the **Window** menu on the main toolbar. Click on the toolbar buttons to bring the desired window to the front:

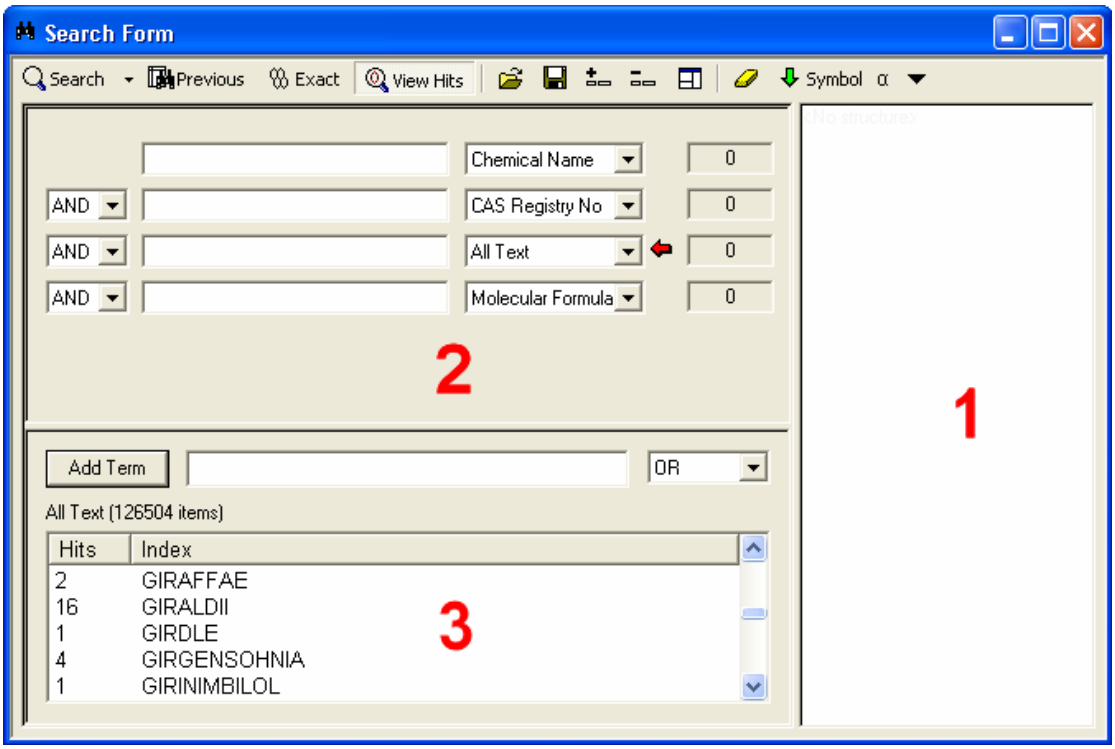
By default, the hit list and entry display appear side by side in the main program window and can be moved and resized by clicking and dragging the title bar or border of each window.

However, you can maximise the windows by clicking on the **Maximise button**  in the corner of each window. Clicking on the window toolbar buttons shown above will bring the desired window to the front. You can return to the windowed view by clicking the **Restore button** the corner of the window.


The Restore button (hilighted)

The Search Form Window

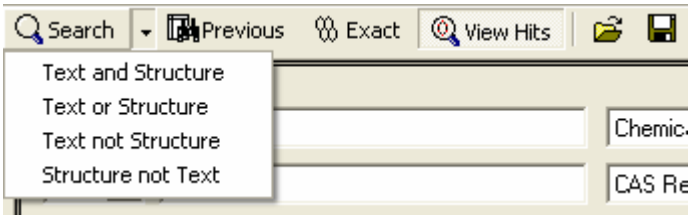
- The **Search Form** window is split into three panes. **1)** Structure search pane. **2)** Search terms pane (containing the **Search Term** boxes) **3)** Index pane (containing the **Index Stem** box). Click on the dividers between panes and drag to resize the panes.
- Four searchable fields are displayed by default in the search terms pane. You can change the field by clicking the field name and select the desired index from the 35 available **searchable fields** from the drop down list. Add or remove lines using the **toolbar icons**



- Text searching can be performed by either typing the search term directly into any of the search term boxes or by selecting the search term from the index.

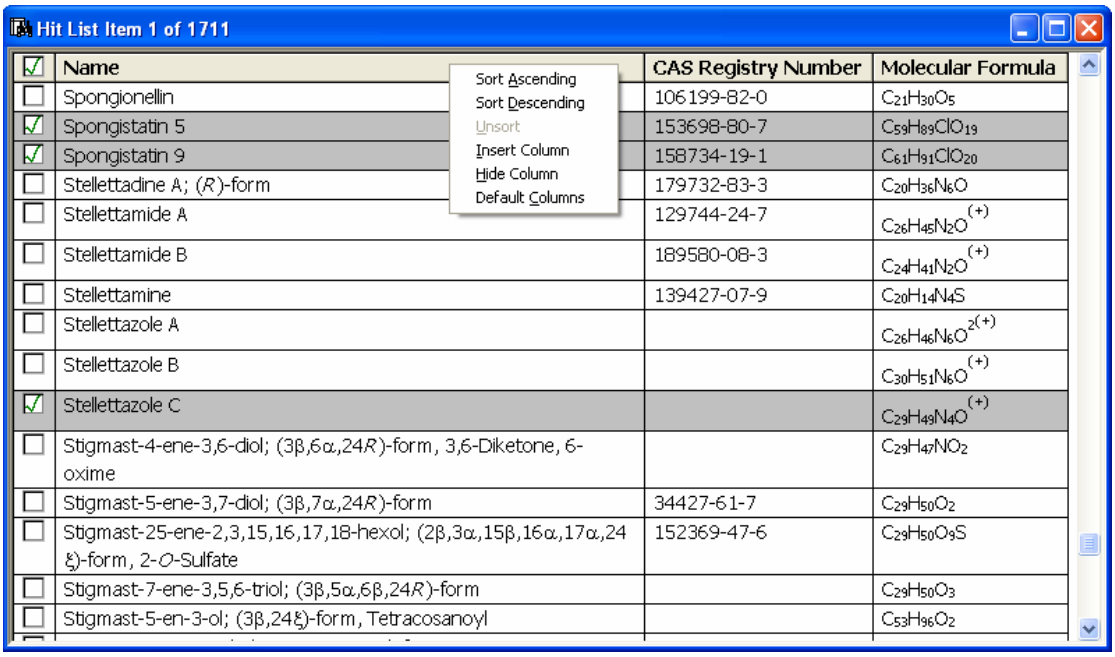
Type your search term into the Index Stem box to scroll down the index.

- Text searching supports **wildcards** for searching partial text strings, **range searching** in the numerical indexes and **Boolean operators** between fields and/or search terms.
- To draw a structure, first double-click on the structure window on the right to begin an editing session. To return the completed structure to the search screen press the green arrow.
- Click the arrow next to the structure icon if you wish to combine a text and structure search. A dropdown list is provided with four Boolean options (see below).
- To run the search, click the **Search** button.


Search
Previous
Exact
View Hits
Text and Structure
Text or Structure
Text not Structure
Structure not Text
Chemic
CAS Re

The Hit List Window

- Once the search is complete, the hit list window will present the search results in alphabetical order (based on chemical name).



- Many of the available fields can be added to the hit list window by right clicking on the column header and selecting **Insert Column**.
- Available options are **Sort Ascending**, **Sort Descending**, **Unsort**, **Insert Column**, **Hide Column** and **Default Columns**.
- In addition, columns can be rearranged by clicking and dragging the header to a new position.
- Mark an entry by clicking the box to the left of the chemical name. Marked entries appear highlighted in grey in the hit list.
- Selected (marked) entries can be printed separately.
- Click on a hit in the hit list to view the entry.
- New hit list formats can be saved by clicking **Save Settings** on the **Hit List** menu

The Entry Display Window

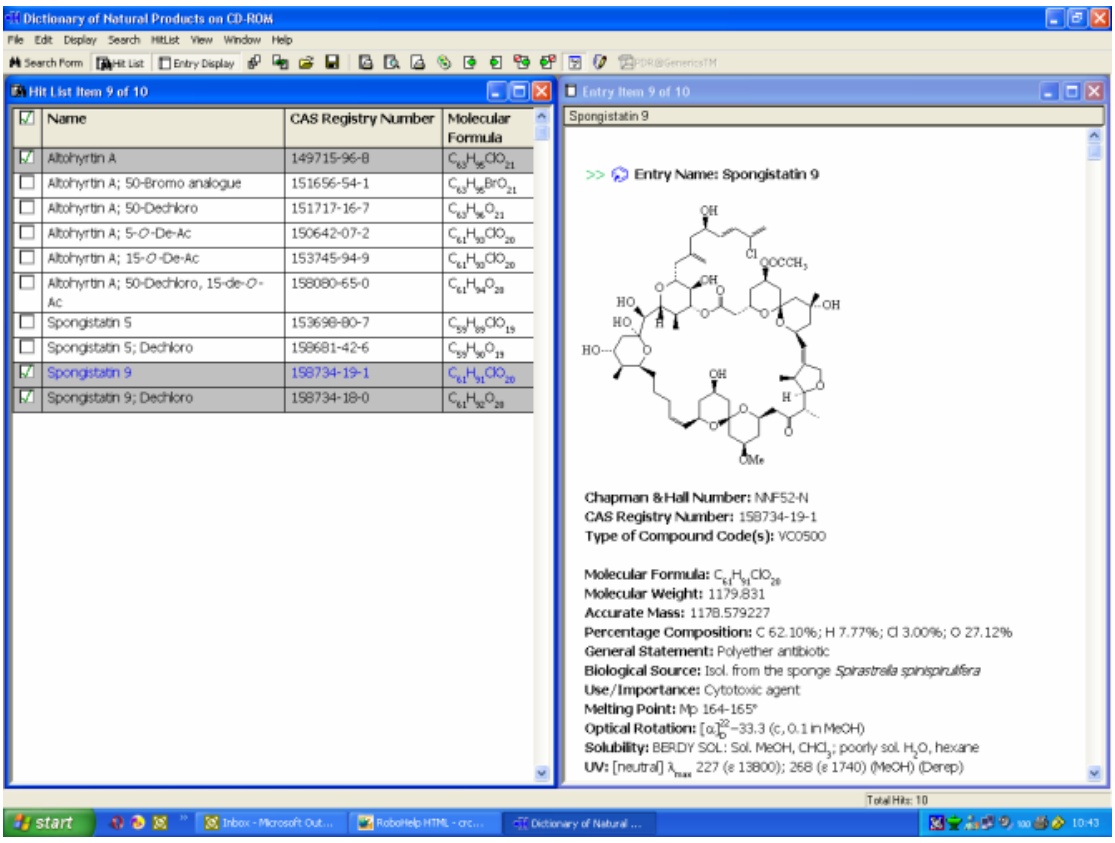
- By default, the entry display will appear in a window alongside the hit list.
- Entries are organised into a parent compound (which may be a parent acid, amine etc.), variants (*R*, *S*, *E*, *Z*-forms etc.) and derivatives (esters, ethers, salts etc.).

See the database introduction for more information

- A bibliography of the key references for each compound in the entry is presented at the end of the entry.
- The entry structure is displayed at the top of the page. These structures contain stereochemistry where appropriate. You can hide the entry structure by clicking on the **View Diagram** icon on the toolbar.
- Structures for each derivative can be viewed by clicking on the structure icon to the left of the derivative. A separate window will display the structure.

These structures do not contain any stereochemical information

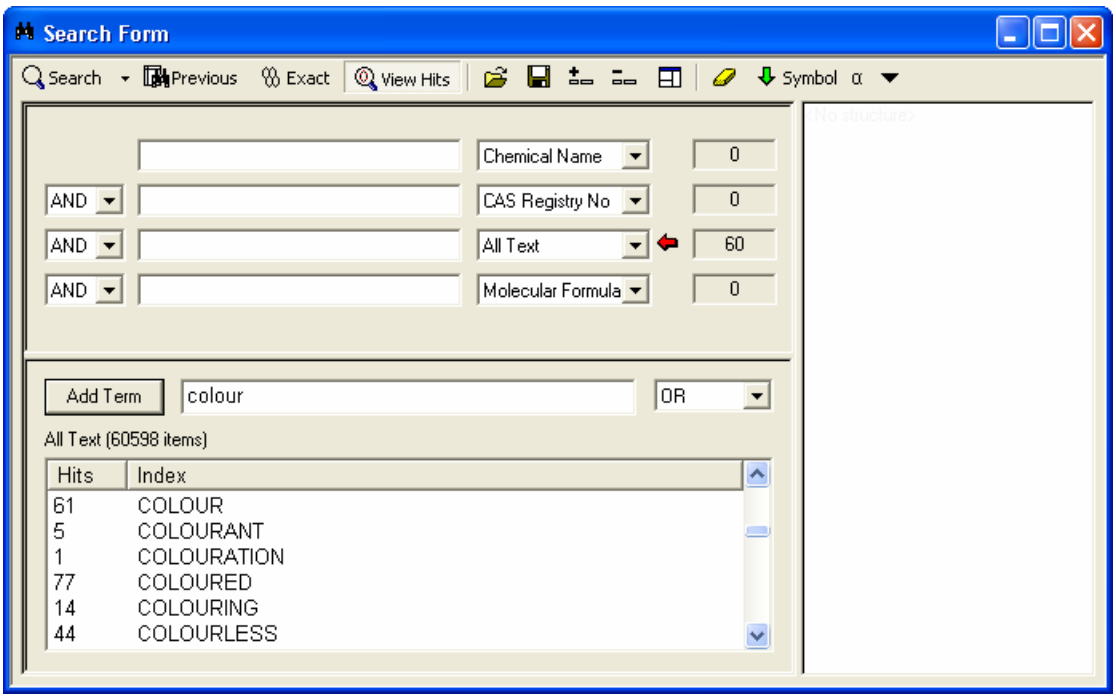
- Both the hit list and the entry can be printed via the respective icons on the toolbar.



Text and Data Searching

Entering a Search Term

The first screen you are presented with after starting the Dictionary of Natural Products is the **Search Form** window. From this window you can perform both text and structure searching of the entire database.



Available Fields

Accurate Mass

The accurate mass field contains all the accurate masses of compounds (quoted to 6 decimal places) reported in the entries. You can search for accurate mass ranges as well as for individual values.

All Entries

This field enables you to select all of the entries in the database. If you wish to retrieve all of the entries in **The Dictionary of Natural Products on CD-ROM**, type **Y** in the search box. You can use this field to carry out exclusion searches. To find all compounds that are not yellow, for example, enter the search terms as follows:-

All Entries Y

NOT All Text YELLOW

All Text

The **All Text** field contains all individual words and character strings from within a particular entry. These include physical description, synthesis, use of compound, biological source, therapeutic use, hazard and toxicity data, amino acid sequence as well as general/miscellaneous information.

Biological Source

This field contains concise information regarding a compound's natural source, e.g. 'Insect pheromone from the boll weevil, *Anthonomus grandis*', 'Metab. of *Streptomyces psammoticus*'.

This information is also searchable in the **All Text** field.

Biological Use/Importance

This field contains concise information about a compound's mechanism of action and therapeutic use, e.g. 'HIV reverse transcriptase inhibitor' or 'active against oomycete fungi'.

This information is also searchable in the **All Text** field.

Biosynthesis

This field contains a concise statement on the route by which a compound is biosynthesised in a producing organism, e.g. 'Major intermed. in the biosynth. of phaeomelanins'.

This information is also searchable in the **All Text** field.

Boiling Point

This field contains the boiling points reported in the entries in degrees Celsius. There is no need to put degree signs in your search term. You can search for a specific boiling point or for a boiling point **range**.

Boiling Point Pressure

This field contains the pressure at which the boiling point was measured if other than at ambient atmospheric pressure. You can search for a specific boiling point pressure or for a range of values.

CAS Registry Number

This field contains the registry numbers assigned to the compounds in the records by Chemical Abstract Services. Since registry numbers are provided primarily as an aid for further searching of the literature, multiple registry numbers assigned to different forms may be included in this field.

> **CAS Registry Number**
> Unless otherwise stated, this number applies to the compound (main entry level, variant or derivative) under which it is found.
>
> **Extra CAS Registry Number(s)**
> Compounds (main entry level, variant or derivative) may have more than one number associated with them. In such cases, the additional numbers representing the compound in question are termed 'Extra CAS Registry Numbers' and are held at compound level. Such numbers will typically refer to simple variations on the compound in question, e.g. stereoisomers, salts or duplicate number assignments.
>
> **Related CAS Registry Number(s)**
> Other registry numbers which cannot be assigned to a particular variant or derivative, but are none the less related to the entry compound are termed 'Related CAS Registry Numbers' and are held at entry level. These may refer to derivatives of the entry compound which are not present in the entry.
>
> Extra and related CAS numbers may contain descriptive to identify the compound, e.g. 123456-78-9 (hydrochloride).

When entering your search term in the **Search Term** box, the hyphens are not essential. e.g. 63968-64-9 and 63968649 are both valid for searching purposes.

C&H Number

This field contains the registry numbers assigned to the compounds in the records by Chapman & Hall/CRC. e.g. LFC01-I (also called Chapman & Hall Number)

Chemical Name

This field contains all chemical names and synonyms for each compound in the database, followed by the **naming system** (e.g. 9CI, BAN) where appropriate. It includes trade names, generic names for pharmaceuticals, CAS names, trivial names and semisystematic names. As there are many alternate spellings for chemicals it is best to select the term from the index. In cases where two or more compounds share the same name the duplicate name is followed by a double dagger symbol ‡.

8CI	Chemical Abstracts Service eighth collective index
9CI	Chemical Abstracts Service ninth collective index
BAN	British Approved Name
BSI	British Standards Institute
USAN	US Adopted Name
INN	International Nonproprietary Name (whether recommended or proposed)
ISO	International Standards Organisation
JAN	Japanese Accepted Name
JMAF	Japanese Ministry for Agriculture, Forestry and Fisheries
WSSA	Weed Science Society of America

Density

Use this field to search for relative density if known. Density values are quoted for liquids only. This field is range searchable.

Development Status

This field applies to drugs. It describes a compound's marketing status (where known) e.g. 'never marketed', 'reached phase III clinical trials (2002)'.

This information is also searchable in the **All Text** field.

The Development Status field is indexed by word and not by phrase, so search terms need to be combined using the Operator Toolbox. To search for compounds which have never been marketed, for example, enter the following search terms in the Development Status field:

never AND marketed

Dissociation Constant

This field contains values quoted for dissociation constants in the form of p*K*a values. You can search for a specific dissociation constant or for a range of values.

Hazard and Toxicity

This field contains data on reactive, physical and toxicological hazards, including LD_{50} values and flash point data. All the information in this field is displayed in red and is also searchable in the **All Text** field.

Hazard Flag

This field contains an indication of whether or not there is any hazard information related to the compound. **Y** indicates the presence of hazard information, and **N** indicates that there is no hazard information. Searches for specific hazards, e.g. carcinogenic, explosive, etc, should be carried out in the Hazard and Toxicity field.

Ion Charge

This field enables you to specify charged species in your search term.

Melting Point

This field contains the melting points reported in the entries in degrees Celsius. There is no need to put degree signs in your search term. You can search for a specific melting point or for a melting point **range**.

Metabolism

This field contains concise information on the route by which compounds are metabolised in humans or experimental animals, e.g. 'Human metabolite of morphine'.

This information is also searchable in the **All Text** field.

Molecular Formula

This field contains molecular formulae for most compounds in the database (following the Hill convention). Formulae are not generally given for characterisation derivatives or salts.
Note that this field is case sensitive, i.e. capitals and lower case must be used correctly.

The Molecular Formula index has been fragmented, i.e. a molecular formula is stored both in its entirety (e.g. C2H5NO) and as its components (e.g. C2, H5, N and O). Thus, to search for all compounds that contain $C_{30}H_{36}$ as part of their molecular formula, enter the following search terms in the **Molecular Formula** field:

C30 AND H36

Individual elements also support range searching. To search for more than one element in a range (e.g C14H18 - C14H20) perform separate searches for each element then combine the hits:

Search 1: C14
Search 2: H18 - H20
Search 3: Combine 1 AND 2

See the chapters **Range Searching** and **Combining Hits** for more information.

Molecular Weight

The Molecular Weight field contains the molecular weight calculated on the basis of molecular formula. The IUPAC published atomic weights of 1993 have been used as a standard.

You can search for molecular weight ranges, as well as individual values.

Optical Rotation

This field includes all the optical rotation [a] values reported in the entries, and is range searchable.

Partition Coeff. (Calc.)

This field contains calculated octanol-water partition coefficients for drugs on the database and is range searchable.

Percent Composition

Elemental analyses have been calculated for each compound based on molecular formula.

The Percent Composition field group contains data on the percentage composition for carbon, hydrogen and nitrogen to two decimal places. You must enter the search term exactly as it occurs in the database. Thus, to search for compounds containing 43.85% carbon, type the following into the Percent Composition search term box:

C=43.85

Or, to search for compounds containing between 18 and 18.5% nitrogen, type:

N=18.00 – N=18.50

References

The Reference field group is searched by selecting the appropriate reference field. The following fields are searchable:

AUTH	The surname of the first author
PTEE	The patenting company (for patent references)
RTAG	The reference tags
TITL	The abbreviated title of the journal
VOLN	The volume number of the journal
YEAR	The year of publication

You must enter the search term exactly as it occurs in the database; as such we recommend browsing the index to find the required search term. For example, to search for all entries containing 'Perkin 1' as the journal title type the following into the references index stem box:

TITL=J.C.S. Perkin 1

the index will scroll to the correct position as you type.

Refractive Index

This field includes all the refractive index (n_D) values (quoted to 4 decimal places) reported in the entries and is range searchable.

Rotation Conditions

This is a text field containing additional information relating to the optical rotation of a substance, e.g. the solvent.

This information is also searchable in the **All Text** field.

RTECS Accession No.

This field contains the RTECS® Accession Number(s) reported within an entry.

The RTECS® Accession Number refers to the NIOSH Registry of Toxic Effects of Chemical Substances and many entries within The Dictionary of Natural Products on CD-ROM contain one or more of these numbers. This data is displayed in red towards the end of an entry (or compound) in the form of a seven digit number preceded by a two letter code.

Source/Synthesis

This field contains information regarding the source or synthesis of a compound, e.g. 'Manuf. by redn. of 4-nitrobenzoic acid'.

Information regarding natural source can be found in the **Biological Source** field.

This information is also searchable in the **All Text** field.

Supplier

This field includes entries that contain compounds which are commercially available from Aldrich, Fluka, Rare Chemicals Library, Sigma and Supelco catalogues plus some other specialist catalogues. The supplier and catalogue numbers are listed.

For a list of contact web addresses for some of the main suppliers named in the Supplier field see below:

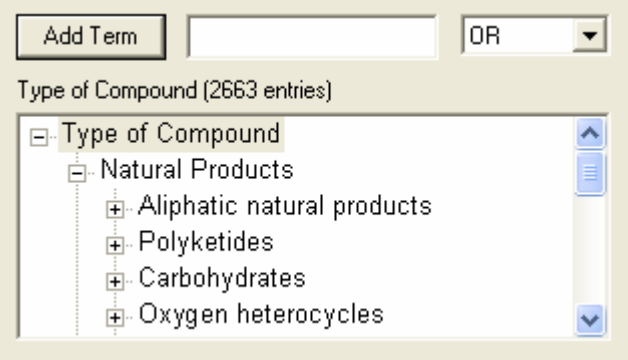
For Aldrich, Sigma, Riedel-de Haen, Supelco and Fluka see Sigma-Aldrich at **http://www.sigmaaldrich.com**
For Davos, see **http://www.davos.com**
For Hüls, see Degussa-Hüls AG at **http://www.huels.de**
For Gelest see **http://www.gelest.com**
For Acros see **http://www.acros.be**
For Amoco see BP at **http://www.bp.com**
For Oakwood see **http://www.oakwoodchemical.com**
For Lancaster see **http://www.lancastersynthesis.com**

Type of Compound
Use the Type of Compound field to search for specific categories of compounds, using class names for the compounds and the code letters assigned to certain compound classes in the Chapman & Hall/CRC Chemical Database:

Code Letter	Compound Class
V	Natural products
X	Drugs

Each subdivision of a class has been assigned a **Type of Compound** code. These codes appear towards the beginning of an entry (or compound) in the form of the two letters (which identify the compound class) and four numbers.

You must enter the search term exactly as it occurs in the database. Refer to the appendix for a complete list of Compound Codes. Alternatively, you may browse the list of compound codes and select the appropriate one from the index.



Type of Compound Words

This field contains all the text associated with the Type of Compound classifications.

Example: To search for all flavonoids on the database, enter the search term:

flavonoid*

Use/Importance

This field contains concise information about a compound's use (e.g. as a catalyst or useful synthetic intermediate) or significance, (e.g. a major industrial chemical).

Information regarding biological function can be found in the **Biological Use/Importance** field.

This information is also searchable in the **All Text** field.

UV Maxima

This field contains all the UV Maxima (λ_{max} values) reported in the entries and is range searchable.

The solvent text associated with the λ_{max} values can be searched in the UV Solvent field.

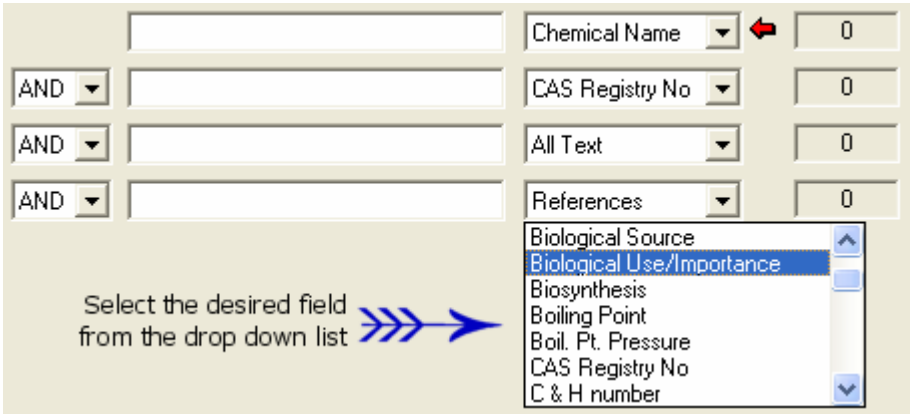
The data in this field may be followed by the term 'Berdy' or 'DEREP', indicating from which database the data originated. The absence of either of these terms implies that the data was abstracted directly from the primary literature.

UV Solvent

This is a text field containing the solvent information relating to the UV Maxima (λ_{max} values) of a substance.

Selecting a Searchable Field

You can search any of the 35 **available fields**. The default search screen list four fields (Chemical Name, CAS Registry No, Molecular Formula and All Text) but you can change the field using the **Field Selector**.

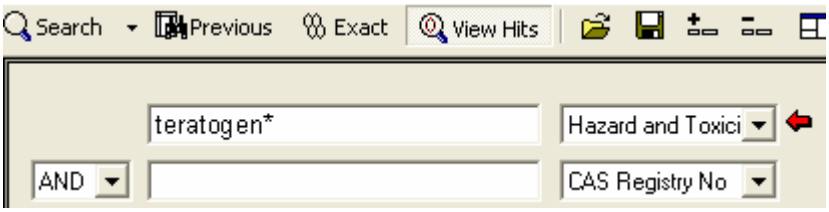


Once you have selected the required field the red arrow will point to the new field, indicating that this is now the active index. The contents of the active index are displayed in the index pane. You can now type your search term into the **Search Term box** or into the **Index Stem box** to browse the index.

Search term boxes can be added or removed using the two **toolbar icons** 

Using the Search Term Box

Search terms can be typed directly into the **Search Term** box or browsed in the index. Terms entered in the **Search Term** box can include **wildcard characters** and **ranges**. For example, to search for all compounds with teratogenic properties, type teratogen* in the **Hazard and Toxicity** field and click **Search.**

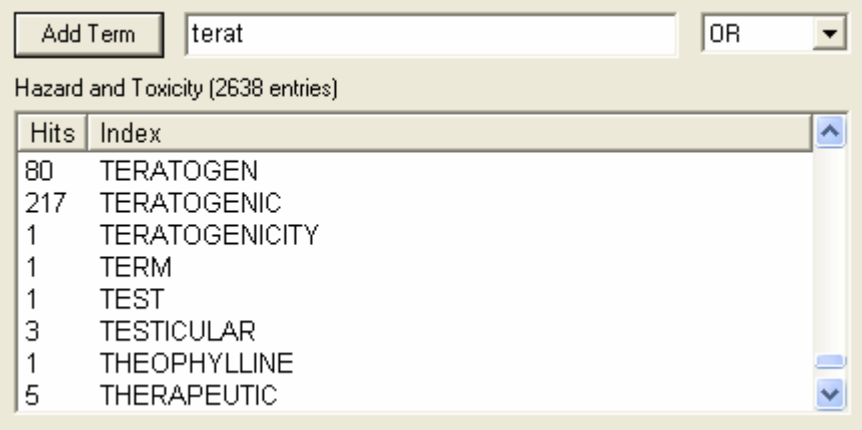


This will find all entries containing the string teratogen (e.g., teratogen, teratogenic, teratogenicity, teratogens) in the **Hazard and Toxicity field**.

Using the Index Stem Box

Alternatively, you can use the **Index Stem** box to browse the index:

1. Use the **Field Selector** to choose the appropriate index as described in the previous chapter. The active index is marked by a red arrow.
2. As you type the required search term in the **Index Stem** box, the index will scroll down to the correct point.
3. Double click on terms you wish to include in the search to add them to the **Search Term** box. (Alternatively, click the **Add Term** button)
4. If you wish to add more than one search term, select the appropriate **Boolean operator** from the drop down list and repeat steps 2 and 3. Selected numerical indexes also support the operators >, <, >=, <= and -.



5. Click **Search**

If the index pane is not visible on the search form, grab the divider at the bottom of the search form and drag upwards to reveal the index pane. Alternatively, click the **Reset Layout** *icon.*

Advanced Text and Data Searching

Truncation and Wildcards

Search terms in the **Search Term** box can include wildcard characters to represent any character or a string of characters.

Use * to represent a string of characters either within the term or at the end of the term (to truncate the term).

Use ? to represent a single character within a word.

You can use a combination of * and ? within a search term. For example, if you type **ETH?N*** in the Chemical Name field, you will retrieve entries which include terms starting with **ETHAN**, **ETHIN** and **ETHYN**.

Using a * at the end of the search term to truncate the term will save entering all the possible variations of the search term. For example, if you search for **hormon*** in the **All Text** field, you will retrieve all compounds containing the words **hormones, hormonal, hormone-dependent, hormone-releasing** etc.

The **Chemical Name** field can also be left truncated, allowing classes of compounds to be searched. For example, a search for ***toxin** will retrieve all compounds whose names end with toxin, such as **Amphitoxin, Brevetoxin B, Wikstrotoxin A** etc.

Range Searching

To search for a range of values, insert a hyphen between the two values in the **Search Term** box, remembering to insert a space between the values and the hyphen.

For example, to search the Melting Point field for a melting point within the range of 47 to 51 degrees Celsius, key in 47 - 51 or use the index to select the range.

You can use any of the following additional symbols:

>	greater than
>=	greater than or equal to
<	less than
<=	less than or equal to

When range searching, you will retrieve compounds where the recorded melting point falls within the search range, or where the recorded range overlaps with the search range. Thus a search for compounds with a melting point range of 47 to 51 degrees Celsius will retrieve compounds with a recorded melting point of, for example, 45-48 or 51-54 degrees Celsius.

Negative numbers are treated in the same manner, with the lower number first. To search for all compounds with a melting point between minus 10 and minus 5 degrees Celsius, enter the following search terms:

-10 - -5

Range searching operators can also be selected from the **Index Stem** drop down list.

Combining Search Terms (Boolean operators)

You can link the search terms using the **Boolean operators: AND**, **OR**, **NOT**. Numerical indexes also support the operators **<**, **>**, **<=**, **>=** and **-**.
Examples:

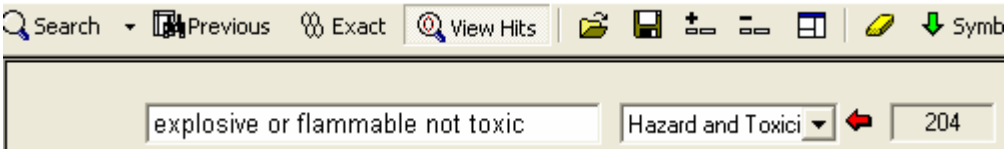
1. To find all compounds which are both antidepressants and analgesics, enter the following search terms in the **All Text** field:

analgesic* and antidepressant*

*You can also use **Type of Compound Codes** to search for a specific class of compounds or food source.*

2. To find all compounds which contain the words explosive or flammable but not toxic, enter the following search terms in the **Hazard and Toxicity** field:

explosive or flammable not toxic

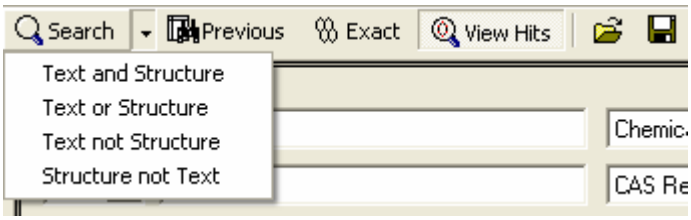


Linking Fields (Boolean operators)

Fields can be linked likewise using the drop down list to the left of the **Search Term** box.

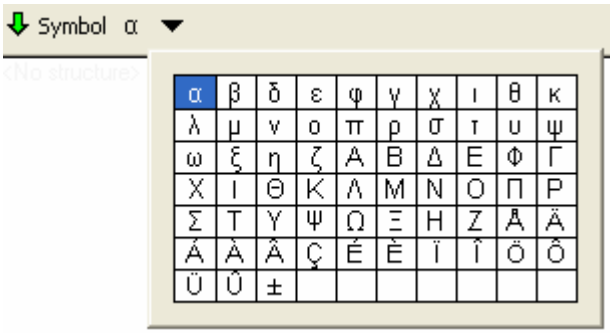
To find all terpenoid compounds that have been isolated from the genus *Taxus*, enter the search terms as follows:

Text and structure searches can also be combined. Click the arrow next to the search button if you wish to combine a text and structure search. A drop down list is provided with four Boolean options.



Special Symbols

The **Special Symbols** menu contains Greek characters, accents and other symbols you may need when using the **Search Term** box. The menu is located on the Search Form toolbar.



Click the black arrow to access the menu. Once the requires symbol has been selected, press the symbol button (green arrow) to enter the symbol into the **Search Term** box.

*Please note that while Greek characters and other special symbols are recognised in the Search Term box, they are ignored in the Index Stem box. So, to find α-Amylase in the Index, you need only type 'amylase' into the **Index Stem** box.*

Searching the Database

Once you have entered your text and/or structure search terms you can search the database.

Text Only Search

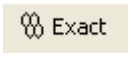
After you have entered the search terms, check all **Boolean operators between fields** are set as required (to change an operator, simply click the arrow next to the operator and select the correct one). Then click the **Search** button.

Structure Only Search

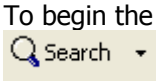
Draw your query and transfer it to the **Search Form**. (Comprehensive help on structure drawing can be accessed via the help menu in the structure drawing program).

Substructure Search: The default structure search finds any occurrence of the structure query in the database, whether a part of a molecule or a whole molecule.

Exact Structure Search: To search for the exact search query as drawn, click on the **Exact** button

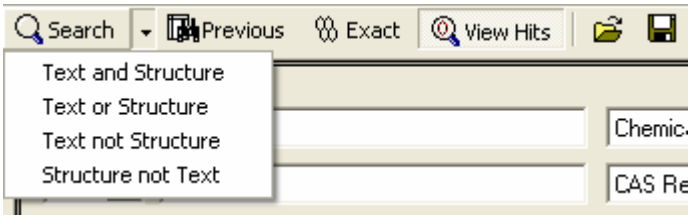


To begin the search click the **Search** button



Text and Structure Search

Prepare your text and structure queries as described above and click **Search** to perform a text and structure search. Other Boolean options for combining text and structure searches can be selected by clicking the arrow next to the search button.



Click on the appropriate Boolean option to run the search. The default action for clicking the search button is **Text AND Structure**.

The Search Progress indicator is located in the lower right hand corner of the screen. Click the red 'X' button to stop a search in progress.

Structure Searching

This topic provides a brief overview of the structure drawing process in the Dictionary of Natural Products. For comprehensive help on how to prepare a structure query, click on the help menu in the structure drawing program.

Important: *The structure drawing program provided with this software is not exclusive to Chapman & Hall/CRC dictionaries. As such, not all features are supported by the Dictionary of Natural Products. Please read the points below for tips on how to make the most from the structure drawing software.*

Double click on the search pane window on the right of the Search Form to open the structure drawing program.

The Structure Drawing Tool Palette

The Structure Drawing Tool Palette is located on the left side of the drawing screen. To change tool, move the cursor to the appropriate position on the palette and click once. With these tools you can draw bonds, rings (click on the ring tool and select the appropriate ring) and chains.

Atoms

Using the Pencil Tool you can draw atoms and bonds, add nodes or change existing nodes to a specific atom, shortcut, variable or G group.

The default atom is carbon and single bonds. Using the Pencil Tool and clicking draws C atoms until another atom is selected. Clicking and dragging draws a single bond until another bond type is selected.

Some common atoms are displayed in the Common Atoms Palette at the bottom of the screen. A single click on the symbol of an atom selects that atom for single use, a double-click selects it for multiple use.

Other atoms, shortcuts and variables can be selected by clicking on the Current Atom Box (located at the bottom left of the drawing screen). Either type in the symbol of the atom you want, or use the scroll bar to display the symbols and click on the atom that you want. Click on **Single** or **Multiple use** to make the selection temporary or permanent.

Atoms, Shortcuts and Variables can also be selected from the appropriate dialogues in the Draw Menu.

Modify an existing node by selecting the Atom, Shortcut or Variable that you want to use and then point to the existing node and click. Notice that the Pencil Cursor changes to display an "A" when pointing to an existing node.

Bonds

Use the Pencil Tool and click on an existing atom and then hold, drag and release the mouse button to join two atoms by a single bond.

Different Bond types can be selected by clicking on the Current Bond Box (located at the bottom of the drawing screen, it is on the left of the bonds displayed) or by using the Bond command in the Draw Menu. The Bond Selection dialogue box is displayed. Click on the

button alongside the bond you want to use followed by Single Use or Multiple Use to make the selection temporary or permanent.

Alternatively click on the appropriate icon in the Common Bonds Palette located at the bottom of the screen. The current bond value is used whenever you use the Pencil, Chain, or Ring tool.

Modify an existing bond by selecting the bond that you want to use from the bond selection dialogue box and then point to the existing bond and click. Notice that the Pencil Cursor changes to display a "--" when pointing to an existing bond. You can modify several bonds by selecting the bond(s) you want to change, using the Selection Tool, before picking the new bond type.

Shortcuts

Shortcuts are predefined groups of atoms, such as CO_2H and NH_2. Use shortcuts the same way as you would a single atom.

Select the Shortcut command to change the Current Atom to a Shortcut.

Click the radio button alongside the shortcut that you want to use, then click on Single Use to use that shortcut once or on Multiple Use to make that shortcut the default current atom. The selected shortcut will appear in the Current Atom box.

You can also select a shortcut by clicking the right mouse button on the current atom box to display the Atom Selection dialog box. Type in the appropriate shortcut symbol or use the scroll bar to display the shortcuts available and select the appropriate one.

Variables

Variables are used to represent generic atoms. An "X", for example, means any of the halogens (F, Cl, Br, I, or At) might be found at the position where the X is attached.

To change the Current Atom to a variable select the Variables command in the Draw menu to display the Variable Selection dialog box.

Click the radio button alongside the variable that you want to use, then click on Single Use to use that variable once or on Multiple Use to make that variable the default current atom. The selected variable will appear in the Current Atom box.

How to limit the number of hits from a substructure search

If you find a substructure search has produced too many hits, you can limit the number by restricting the number of connections on any one atom (node). You can specify the number of attachments to any atom from the QueryDef menu.

However, an easier way to do this is to explicitly draw the hydrogen atoms connected to the carbon or other element. This effectively caps off the atom, preventing any more connections being made.

Features not supported in the Dictionary of Natural Products

There are two main features of the structure drawing program not supported by Chapman & Hall/CRC dictionaries.

1. Stereochemistry: There is no need to specify stereochemistry in your structure query since the various stereoisomers are presented within the same entry. Any stereochemistry included in the structure query will be ignored.

2. Reactions: Reaction searching is not supported in Chapman & Hall/CRC dictionaries.

Structure Import/Export

The Import/Export facility within the structure drawing screen allows a **connection table** for a given compound on the database to be saved to disc or transferred to or from another drawing package.

A connection table is a 2-D representation of the structure and does not display any stereochemistry. All functional groups are drawn in full since the program uses these structures for the purpose of structure searching

Import

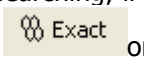
The structure formats available for importing structures are SMD (.SMD), DARC F1 (.DRC), MOLFILE and Alchemy (both .MOL) and the structures are stored with one structure per file.

To import a structure, click on the **File** menu within the structure drawing screen and select **Import**. The Import dialog box will appear on-screen allowing the choice of the appropriate structure file format. Select a format and a second dialog box will appear. Locate the desired structure file via selecting the appropriate drive letter and use of [..] tool. Once located, select the desired filename and click **Open**. As a result the structure will be imported and displayed on-screen as a connection table.

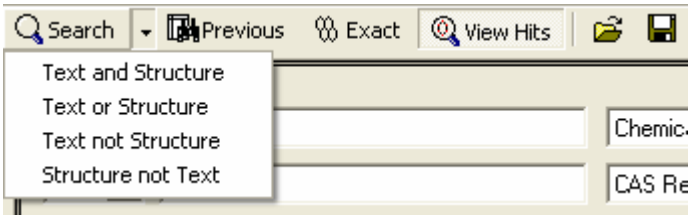
Export

The structure formats available for exporting a structure are SMD (.SMD), DARC F1 (.DRC), MOLFILE (.MOL) and SMILES (.SMI) and the structures are stored with one structure per file. As the structures are exported as queries, normalised and exact/normalised bonds become unspecified in the MOLFILE.

To export a structure, click on the **File** menu within the structure drawing screen and select **Export**. The Export dialog box will appear on-screen allowing the choice of the appropriate structure file format. Select a format and a second dialog box will appear. Locate the desired directory to where the diagram is to be exported via selecting the appropriate drive letter and use of [..] tool, provide a filename and click **Open**. As a result the structure will be exported as a connection table.

- Once you have prepared your structure query, click the green arrow to transfer your structure to the **Search Form**.
- The default search mode is substructure searching; if you wish to search for the exact structure, select the **Exact button** 

on the **Search Form**.
- By clicking on the arrow next to the **Search** button it is possible to combine text and structure searches. Select the appropriate Boolean option from the list. The default action of the **Search** button is **Text AND Structure**.

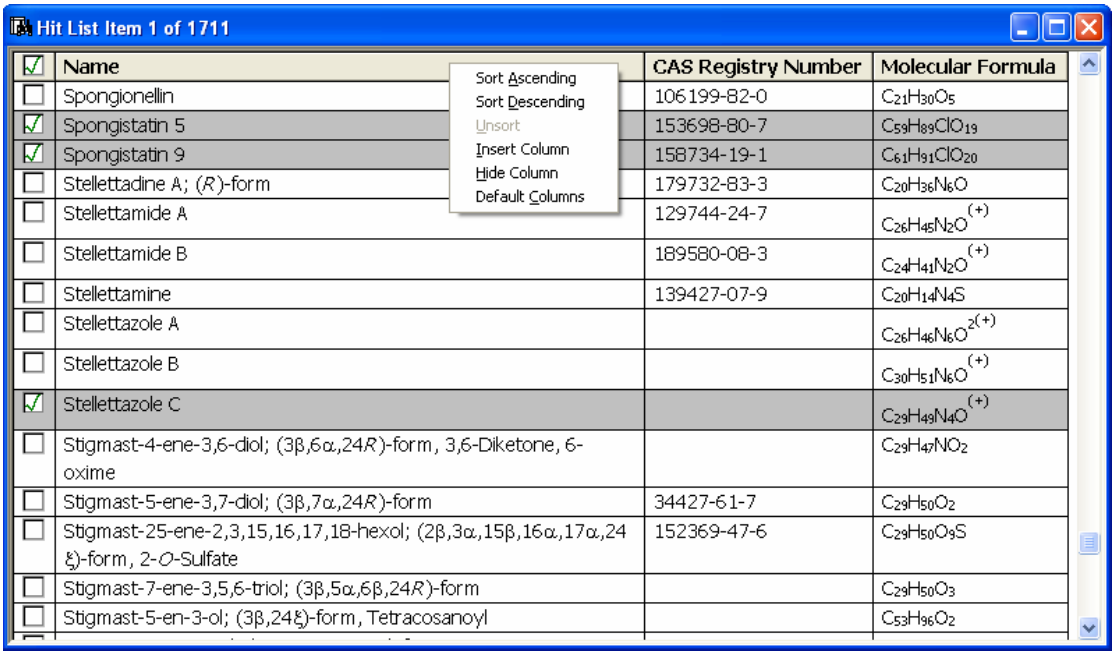


Trivial derivatives (e.g. salts, compounds containing common protecting/characterisation groups, etc) are not searchable and will not be displayed in the hit list.

Manipulating Search Results

Examining Search Results

After the search is complete you will be presented with the hit list containing your search results.
By default, the hit list contains four columns: **Mark/Unmark entry**, **Name**, **CAS Registry Number** and **Molecular Formula**.



- **Customising the Hit List**

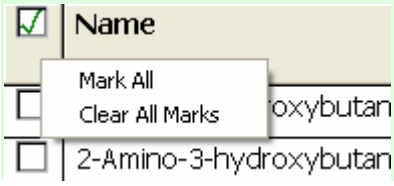
  Refer to the chapter **Display and Print Options** for Hit List customisation options.

- **Marking hits in the Hit List**

  Sometimes you may wish to print selected entries from the hit list without printing all the entries. To do this you can mark hits in the hit list by clicking on the **Mark/Unmark** checkbox to the left of the chemical name (or right click an entry and select mark/unmark from the drop-down menu).

  Marked hits are highlighted in grey with a green tick in the checkbox. To unmark a hit, click the checkbox again.

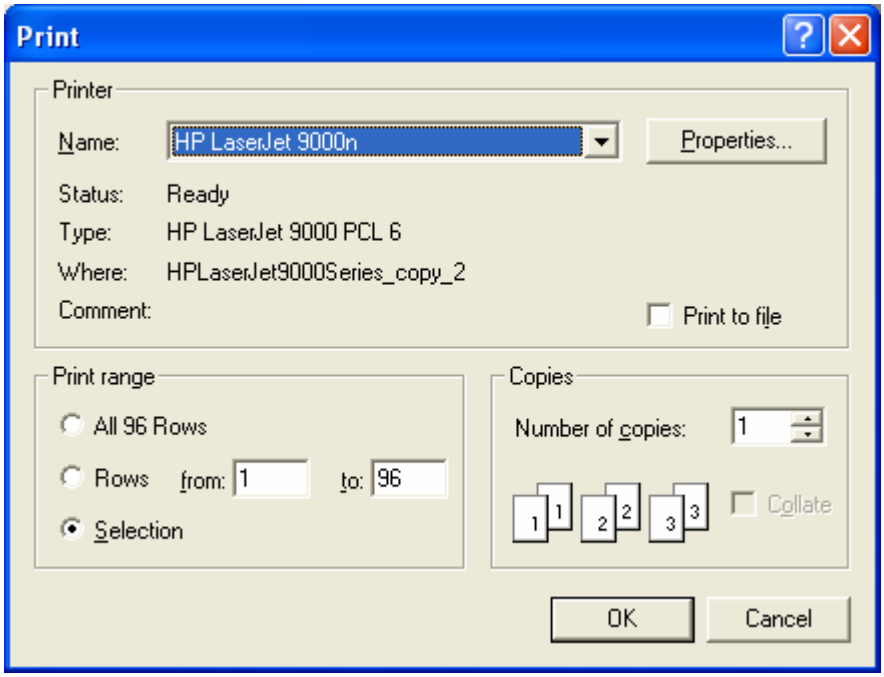
To unmark all entries in the hit list: right click on the column header and select ***Clear All Marks***



- **Printing the Hit List**

  To print the Hit List click the icon on the toolbar
  
  or select **Print-Hit List** from the **File menu**

  To print only the marked entries, click **Selection** in the print dialogue box.



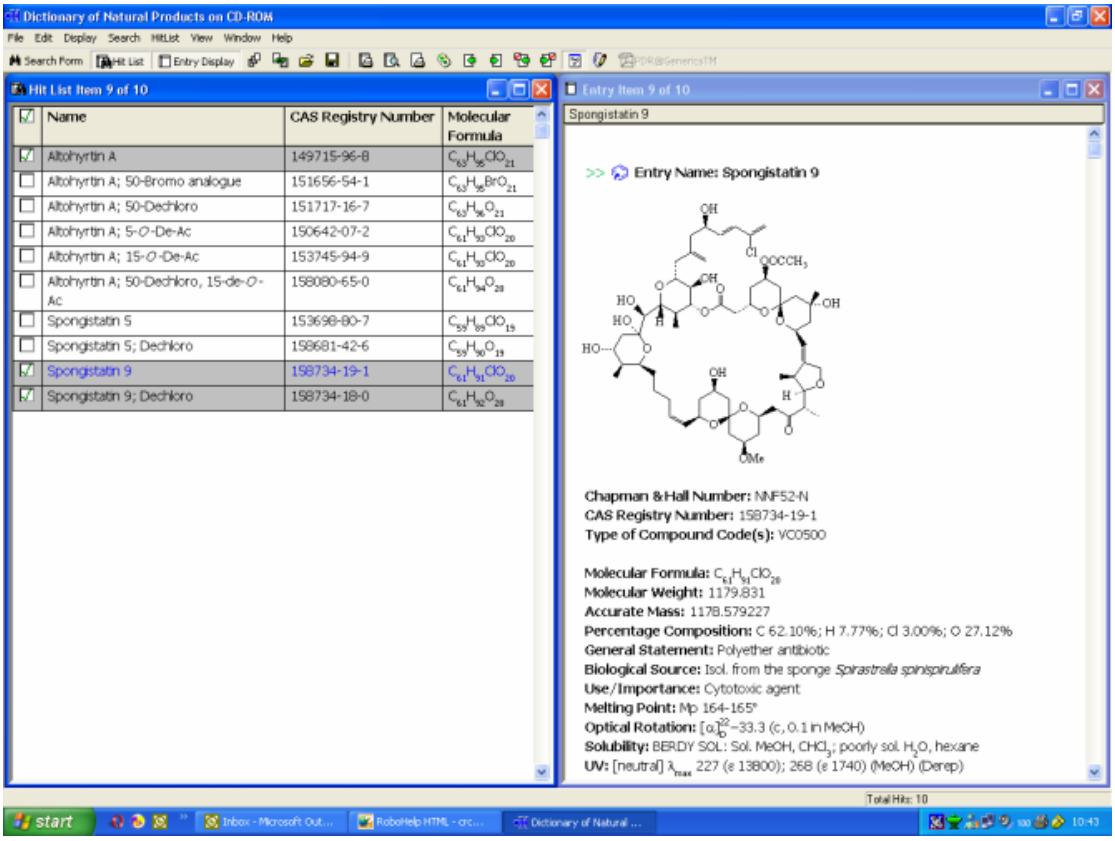
  It can sometimes be useful to see a print preview of the hit list so that changes can be made to column size/positions prior to printing. You can view a print preview by clicking the **Print Preview icon**  on the toolbar or via the **File** menu. See the chapter **Display and Print Options** for more information.

- **Viewing a hit**

  To display an entry simply click on the required hit. The entry appears in the **Entry Display** window alongside the hit list.

Viewing Entries

To display an entry simply click on the required hit. The entry appears in the **Entry Display** window alongside the hit list.



Entries are organised in the following order:

- **Entry Name and Synonyms**

  Duplicate Names

  In cases where two or more compounds share the same name the duplicate name is followed by the double dagger symbol ‡.

  Naming Systems

  Compound names are followed by the **naming system** where appropriate.

8CI	Chemical Abstracts Service eighth collective index
9CI	Chemical Abstracts Service ninth collective index
BAN	British Approved Name
BSI	British Standards Institute
USAN	US Adopted Name
INN	International Nonproprietary Name (whether recommended or proposed)
ISO	International Standards Organisation

JAN	Japanese Accepted Name
JMAF	Japanese Ministry for Agriculture, Forestry and Fisheries
WSSA	Weed Science Society of America

- **Entry Diagram**

  The main entry diagram appears towards the top of the entry. This diagram shows any relevant **stereochemistry** *(For natural products, the stereochemistry will usually refer to the naturally occurring form, although if the compound is a drug the stereochemistry will usually refer to the pharmacologically active isomer)* and numbering systems. The diagram can be hidden by clicking on the **View Diagram icon**  on the toolbar.

- **Variants/Derivatives**

  Variants
  Different forms of the entry compound. e.g. (*R*)-*form* or *Pyranose-form* etc.

  Derivative
  Common modifications to the entry compound. e.g. Me ester

  The entry compound is usually the most fundamental parent compound (e.g. carboxylic acid) regardless of whether or not it is a known compound.

- **Variant/Derivative structures**

  Structures for variants and important derivatives can displayed by clicking on the **Structure icon**  to the left of the variant/derivative. These structures are derived from a 2-D **connection table** and do not contain stereochemical information.

  A connection table is a 2-D representation of the structure and does not display any stereochemistry. All functional groups are drawn in full since the program uses these structures for the purpose of structure searching.

- **References**

  References for all known compounds in an entry are presented at the bottom of the entry.

  Each reference is followed by a reference tag indicating to which variant/derivative the reference belongs and the type of information contained in the article. Turn to the appendix for a list of abbreviations used in reference tags.

  Reference tag organisation:

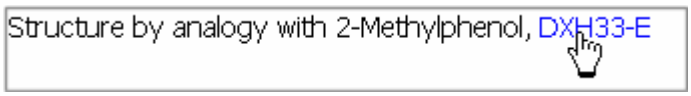
Properties at the beginning of the reference tag refer to the entry compound, whereas properties for derivatives follow the derivative descriptor.

e.g.
(*pmr, cmr, ms, R-form, Ac*) indicates that the article contains pmr, cmr, and ms data for the entry compound, as well as information on the R-form acetate derivative, whereas:

(*R-form, Ac, pmr, cmr, ms*) indicates that the article contains pmr, cmr and ms data for the R-form acetate derivative, but no information on the entry compound.

Cross References

Cross references to related compounds are indicated by a blue hyperlink consisting of the **Chapman & Hall number** of the target entry. Simply position your mouse pointer over this and the cursor will change to a hand shape. Click once and you will be taken to the cross referenced entry.



Display options

To **hide the diagram** click on the **View Diagram icon**  on the toolbar.

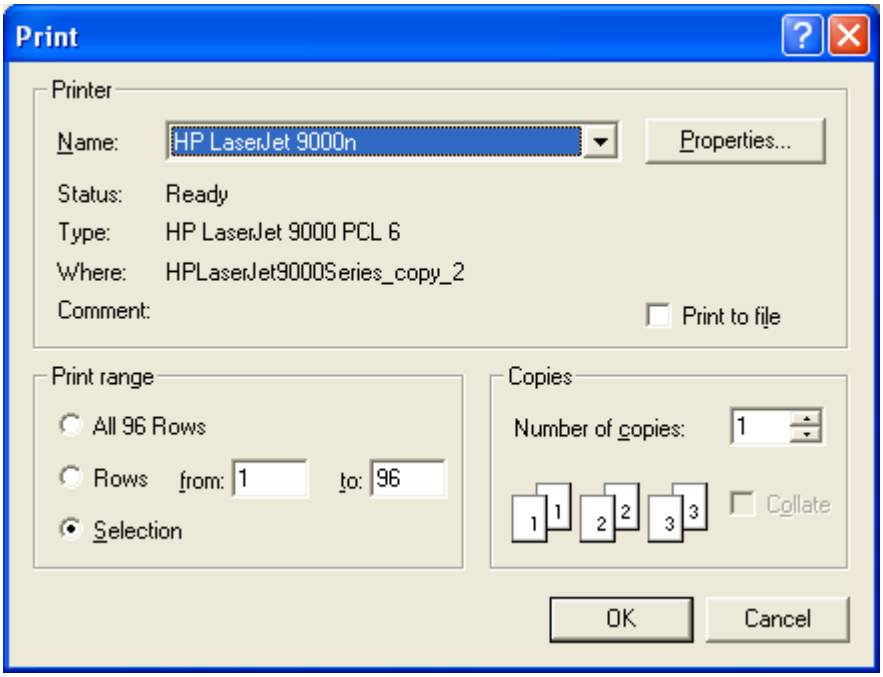
Toggle between **Full Fielded** (default) and **Standard** (compact) display from the **Options** toolbar menu.

Printing Entries

To print entries click the icon on the toolbar

or select **Print Entries** from the **file menu**

To print only the marked entries, click **Selection** from the print dialogue box



Refer to the chapter **Edit Preferences** for further customisation options.

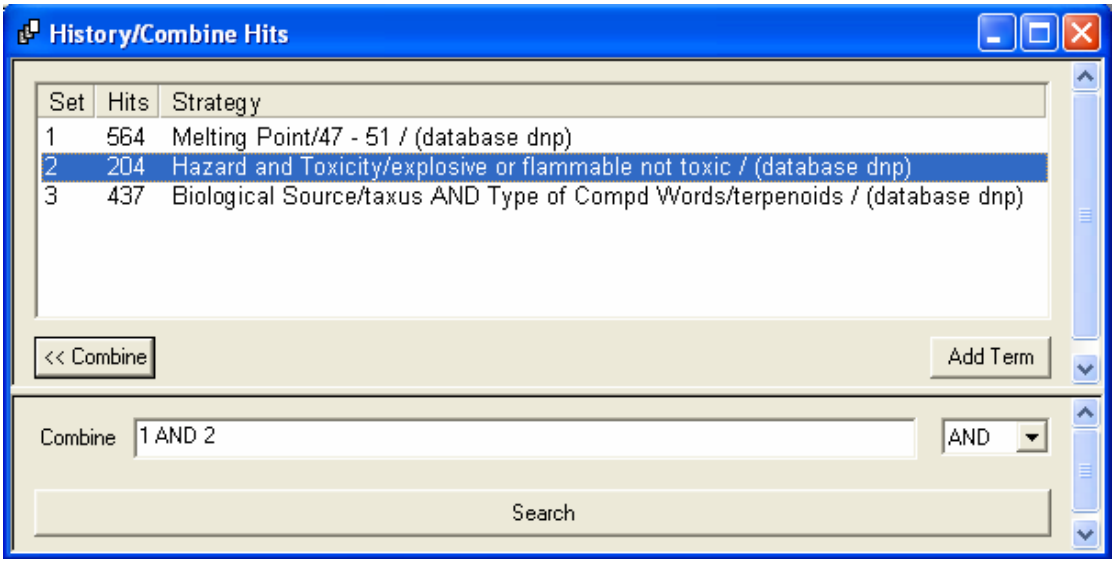
Please refer to the database introduction for a complete description of data organisation and presentation in the Dictionary of Natural Products. The introduction can be found under the Chapman & Hall/CRC program group in the Start menu (You will need the Acrobat Reader installed on your computer to access the introduction).

Combining Hits

Hit lists from the current search session may be combined using the Boolean operators **AND**, **OR** and **NOT**.

Click on the **Combine Hits icon** to display the **History/Combine Hits** dialogue box (alternatively, select **Combine Hits** from the **View** menu).

If, for example, you had carried out two searches and wanted to find all entries present in both searches 1 and 2 - you would double click on search 1, select the **AND** operator from the drop down list, double click search 2, then click on **Search** (alternatively, type the expression directly into the Combine Search Term box then click **Search**).



The resulting hit list is displayed in the **Hit list window** and saved as the next search number, which can be viewed by clicking on the **History List icon** or via the **View** menu.

Previous hit lists can be viewed in this way by simply selecting the desired hit list.

Display and Print Options

Customising the Hit List

These columns can be customised by right clicking on the column header. Available options are **Sort Ascending**, **Sort Descending**, **Unsort**, **Insert Column**, **Hide Column** and **Default Columns**.

Sort Ascending/Descending: Right click on the column header you wish to sort by and select these options to reorder the hit list.

*For example, to order the hit list from smallest molecular formula to largest, right click on the header **Molecular Formula** and select **Sort Ascending**.*

Unsort: Restores the original order

Insert/Hide Column: Use these options to add or remove columns. The new column is added to the right of the current column (change this option to left via the **Edit Preferences** dialogue box).

*For example, to include the optical rotation field, right click on the column header where you would like the new column added and select **Insert Column**. Scroll down the list until you find **Optical Rotation** and click **OK**.*

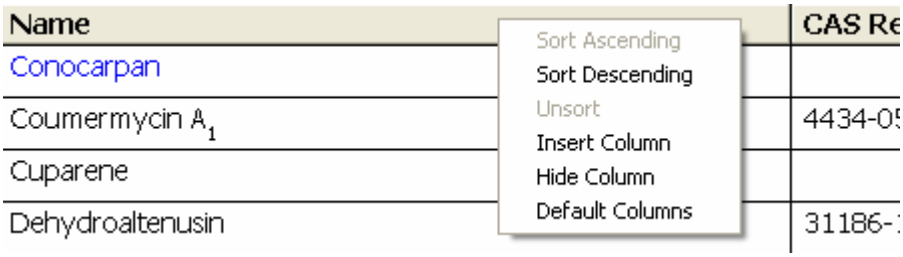
Default Columns: Restores the default columns

To move a column, click and drag a column header to the new position.

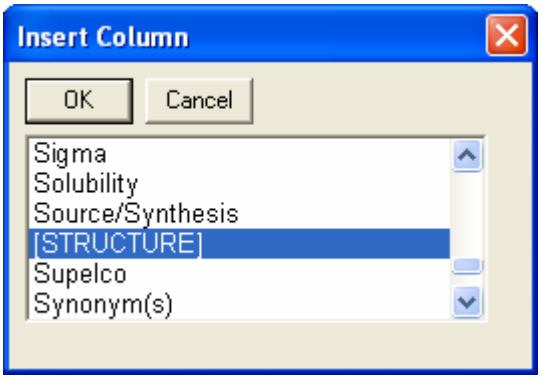
To save/load a customised layout, select **Save Settings** or **Open Settings** from the **HitList** toolbar menu.

Adding structures to the Hit List

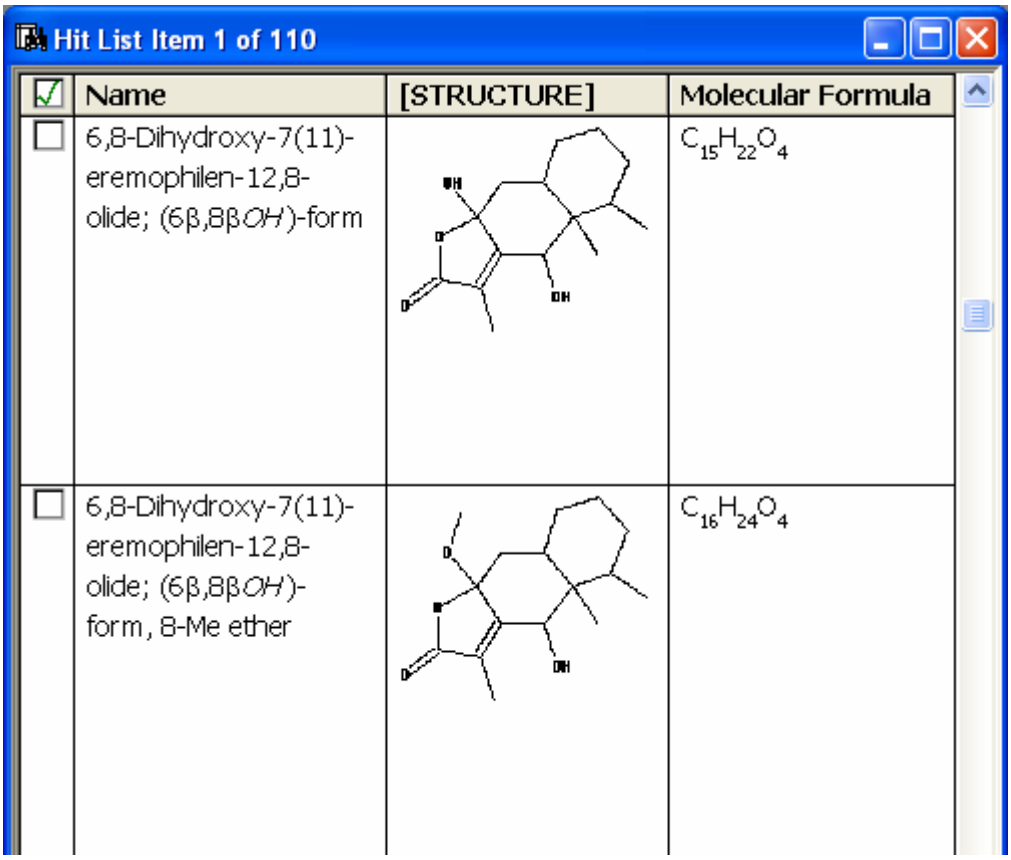
A useful feature of the Hit List view is the ability to add compound structures just as you would add an additional field. To add structures to the hit list, right-click on the column header and select **Insert Column**



Scroll down the list of available fields until you reach **[STRUCTURE]** then click **OK**.



• The structures will appear in a new column to the right of the current one. You can move columns by clicking and dragging the column header to a new position.



• Resize the columns by placing the mouse pointer on the dividers between columns until it changes to a double headed arrow ↔, then click and drag to resize.

Hit lists with several wide columns may be printed over more than one page.

Print Preview

To see how the hit list will appear once printed, select print preview from the File menu, or from the **Print Preview icon** on the toolbar

This can be particularly useful when printing hit lists with more than three columns, since these often need to be printed over more than one page. In such cases it is then possible to resize the hit list columns prior to printing.

When previewing a hit list containing marked entries, only the marked entries will be shown in the preview.

Customising the Entry Display

To **hide the diagram** click on the **View Diagram icon**  on the toolbar.

Toggle between **Full Fielded** (default) and **Standard** (compact) display from the **Display** toolbar menu.

Advanced display and print options can be set from the **Edit Preferences** dialogue box located in the **Edit** menu.

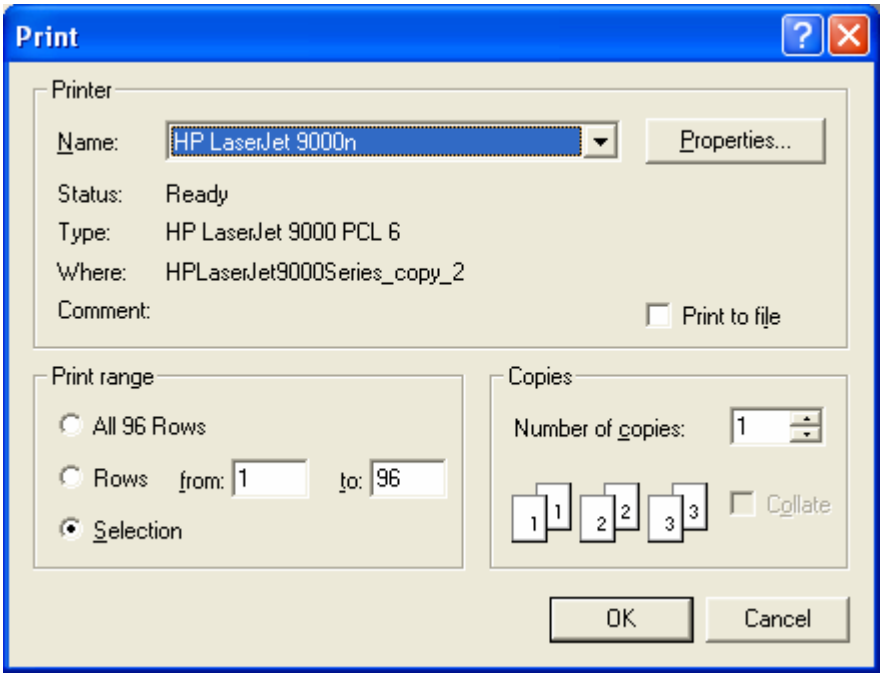
Printing

Printing the Hit List

To print the Hit List click the icon on the toolbar

or select **Print-Hit List...** from the **File menu**

To print only the marked rows, click **Selection** in the print dialogue box.



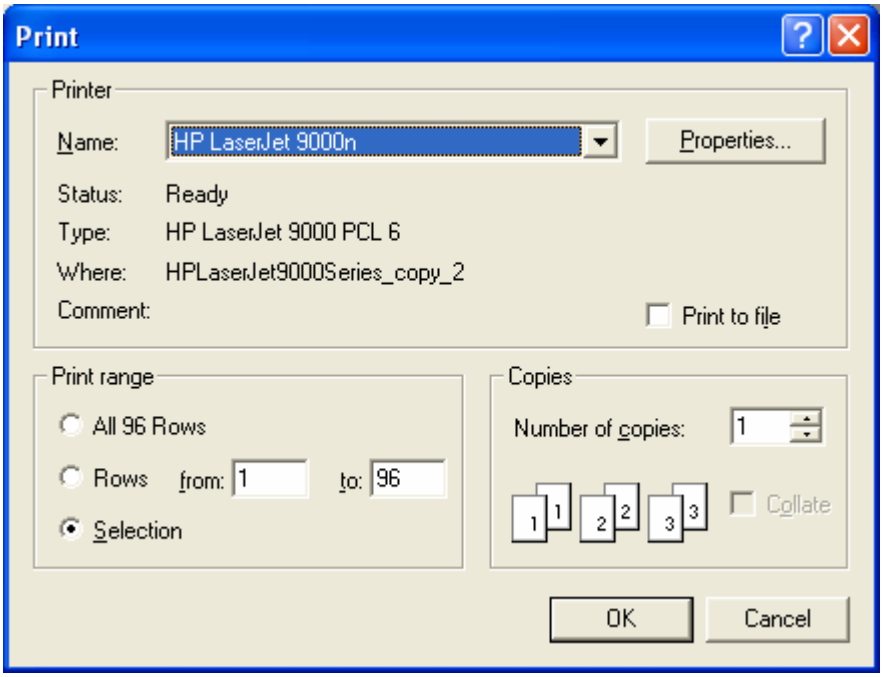
It can sometimes be useful to see a print preview of the hit list so that changes can be made to column size/positions prior to printing. You can view a print preview by clicking the **Print Preview icon**  on the toolbar, or via the **File** menu.

Printing Entries

To print entries click the icon on the toolbar

or select **Print - Entries...** from the **File menu**

To print only the marked entries, click **Selection** from the print dialogue box



When printing multiple entries, each entry begins on a new page. Select the option **Print Continuous** from the **Edit Preferences** dialogue to print without page breaks.

Printing the Search History

To print the history list select **Print - History...** from the **File menu**

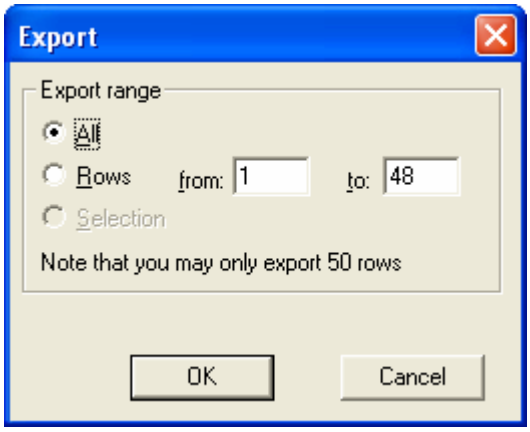
See the next chapter, 'Edit Preferences', for more print options.

Export Data

- **Exporting the Hit List**

  The contents of the Hit List can be saved in Microsoft® Excel format (.xls).

  With the Hit List displayed, select **Export Hit List** from the File menu.



  To export marked entries only, click on the button marked **Selection**.

  Up to 50 rows may be exported at any one time.

- **Exporting entries**

  Entries can also be saved in Microsoft® Excel format (.xls).

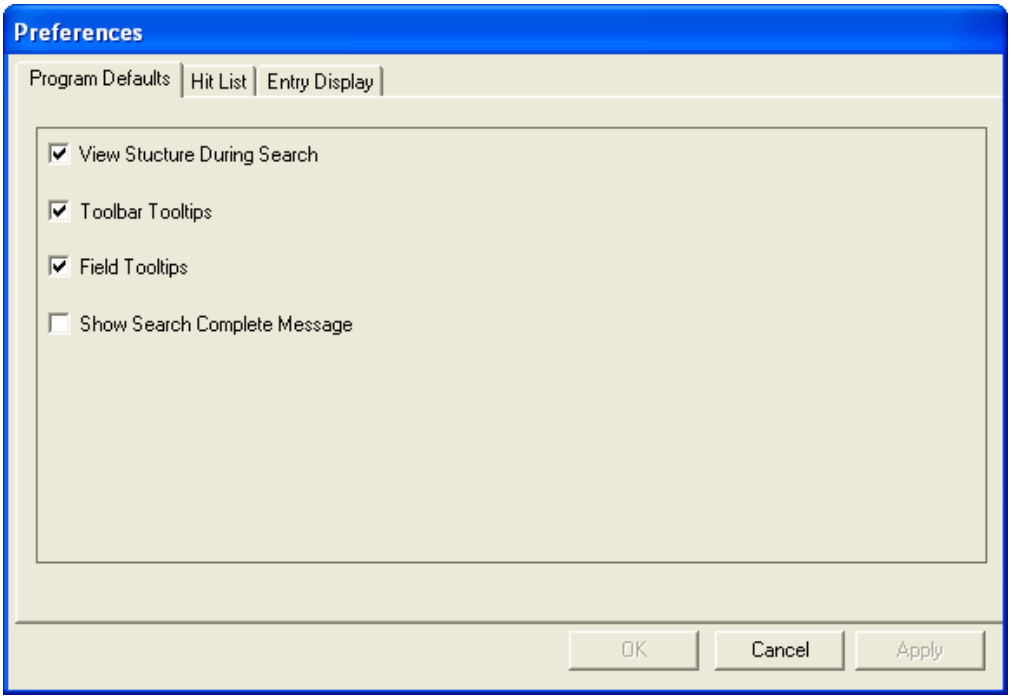
  With the Hit List displayed, select **Export Entries** from the File menu.

  Up to 50 entries may be exported at any one time.

Edit Preferences

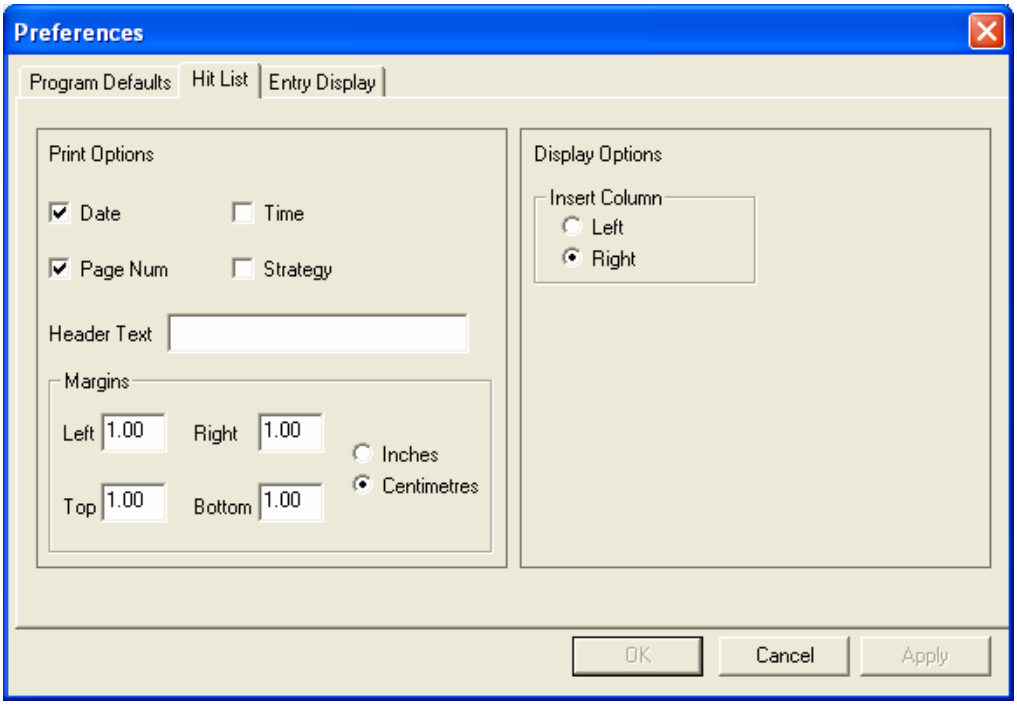
Many of the individual settings in the **Dictionary of Natural Products** can be accessed from the Preferences dialogue box, located in the **Edit** menu.

Program Defaults tab



Set default program options for the Dictionary of Natural Products.

Hit List tab

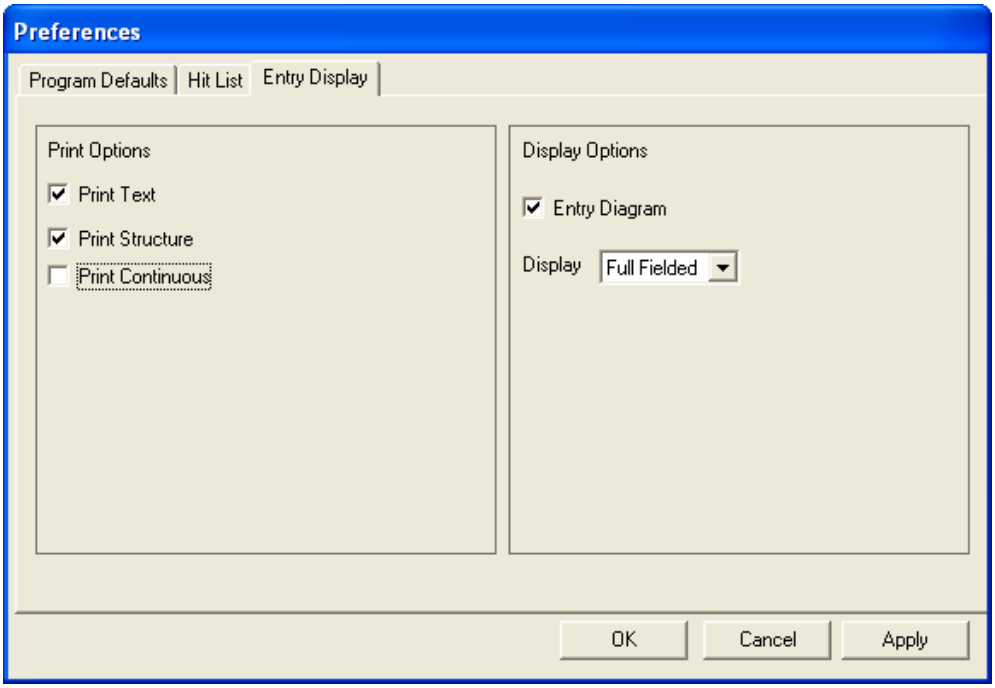


Use this dialogue box to set print and display defaults for the Hit List.

Print options: Choose header and footer text and margins.

Display Options: Choose whether new columns are added to the left or the right of the current column.

Entry Display tab



Use this dialogue box to set print and display defaults for the Entry Display.

Print Options: Choose whether text or structures are printed. Select the option **Print Continuous** to print without page breaks.

Display Options: Choose the format for displaying entries.

Toolbars and Menus

File Menu

- **New Strategy**

  Deletes the contents of the history window and Search Form

- **Open History…**

  Opens a previously saved history list

- **Save History**

  Saves the current history list

- **Save History As…**

  Saves the current history list

- **Print**

  Print History…

  Print the history list of past searches.

  Print Hit List…

  Print the current hit list. Equivalent to the toolbar icon 

  Print Entries…

  Print entries. Equivalent to the toolbar icon 

- **Print Preview (Hit List)…**

  It can sometimes be useful to see a print preview of the hit list so that changes can be made to column size/positions prior to printing. Alternatively, click the **Print Preview** icon on the toolbar 

- **Print Setup…**

  Adjust settings for your printer

- **Export Hit List**…

  Saves the current hit list in Microsoft® Excel (.xls) format

- **Export Entries**…

  Saves entries in Microsoft® Excel (.xls) format

- **Change Dictionary**…

  This option is for users who have more than one Chapman & Hall/CRC dictionary installed on their computer.

Edit Menu

- **Cut**

  Use to move text from one search term box to another. Highlight the text to want to move and select **Cut**. Place your cursor where you would like to paste the text and click **Paste**.

- **Copy Text**

  Copy text from the hit list or an entry

- **Copy Entry Diagram**

  Copy an entry diagram to be pasted into another application (such as Microsoft® Word).

- **Paste**

  Paste text into a search term box

- **Select All**

  Select the entire contents of an entry to be pasted into another application

- **Edit Preferences**
  Sets program defaults. See the chapter **Edit Preferences** for more details.

- **Structure Drawing**

  Access the structure drawing program

Display Menu

- **Next Hit**
  Display the next entry in the **Hit List**

- **Previous Hit**
  Display the previous entry in the **Hit List**.

- **Next Marked Hit**
  Display the next marked entry in the **Hit List**

- **Previous Marked Hit**
  Display the previous marked entry in the **Hit List**

- **Full Fielded**
  Change the **Entry Display** to Full Fielded view (this is the default setting)

- **Standard**
  Change the **Entry Display** to Standard (compact) display

- **Show Entry Diagram**
  Show/Hide the entry diagram from the **Entry Display**

Search Menu

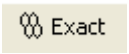
- **Search**

  Click to run a search. Equivalent to 

- **Search Previous Results**

  Sometimes a search will produce too many hits. Use this option to refine a hit list by searching within the previous results only. (Equivalent to the **Previous** icon on the **Search Form**).

- **Exact Search**

  Perform an exact structure search. Equivalent to 

- **View Structure Hits**

  Shows structure hits while the search is in progress.

  Enabling this option may cause structure searching to proceed more slowly.

- **Stop Structure Search**

  Stops a structure search. Equivalent to the **Stop** icon on the search progress indicator.

- **Load Form**

  Load a previously saved **Search Form** layout

- **Save Form**

  Save a **Search Form** layout.

- **Add Line**

  Add a line to the **Search Form**

- **Remove Line**

  Remove a line from the **Search Form**

- **Clear Form**

  Clear the contents of the **Search Form**

- **Reset Form Layout**

  Resets the dividers between panes in the **Search Form** to their default position. Equivalent to clicking on the Search Form.

HitList Menu

- **Open Settings**
  Open previously saved hit list settings

- **Save Settings**
  Save the current hit list layout

- **Default Columns**
  Restore the default columns on the hit list

View Menu

Show any of the following items via the View Menu:

- Toolbars
- Search Form
- Hit List
- History List
- Entry Display
- Combine Hits
- PDR Generics (Dictionary of Drugs only)

Window Menu

Use this menu to arrange the **Search Form, Hit List** and **Entry Display** windows.

- Cascade
- Tile Horizontally
- Tile Vertically
- Arrange Icons

Type of Compound Codes

Overview

Use the Type of Compound field to search for specific categories of compounds, using class names for the compounds and the code letters assigned to certain compound classes in the Chapman & Hall/CRC Chemical Database:

Code Letter	**Compound Class**
V	Natural products
X	Drugs

Each subdivision of a class has been assigned a **Type of Compound** code. These codes appear towards the beginning of an entry (or compound) in the form of the two letters (which identify the compound class) and four numbers.

You must enter the search term exactly as it occurs in the database.

Aliphatic Natural Products

The following is a list of all the **Type of Compound Codes** used for **Aliphatic Natural Products** on the database.

Code	***Classification***
VA0050	Saturated unbranched hydrocarbons
VA0100	Saturated unbranched alcohols
VA0150	Saturated unbranched acetates
VA0250	Saturated unbranched aldehydes and ketones
VA0300	Saturated unbranched carboxylic acids and lactones
VA0350	Saturated unbranched methyl esters
VA0370	Other saturated unbranched esters
VA0400	Unbranched alkenic hydrocarbons
VA0420	Unbranched alkenic alcohols
VA0450	Unbranched alkenic acetates
VA0500	Unbranched alkenic aldehydes and ketones
VA0600	Unbranched alkenic carboxylic acids and lactones
VA0650	Unbranched alkenic methyl esters
VA0700	Other unbranched alkenic esters
VA0800	Branched aliphatic hydrocarbons
VA0850	Branched aliphatic alcohols
VA0900	Branched aliphatic acetates
VA1050	Branched aliphatic aldehydes and ketones
VA1100	Branched aliphatic carboxylic acids
VA1150	Branched aliphatic methyl esters
VA1200	Other branched aliphatic esters
VA1250	Branched alkenic hydrocarbons
VA1300	Branched alkenic alcohols
VA1350	Branched alkenic acetates
VA1450	Branched alkenic aldehydes and ketones
VA1500	Branched alkenic carboxylic acids
VA1550	Branched alkenic methyl esters
VA1600	Other branched alkenic esters
VA1650	Acetylenic hydrocarbons
VA1700	Acetylenic alcohols
VA1750	Acetylenic acids and esters
VA1770	Allenes
VA1800	Miscellaneous acetylenes
VA1850	Monocarbocyclic alkanes
VA1900	Monocarbocyclic alkenes
VA1950	Monocarbocyclic alcohols
VA2000	Monocarbocyclic acetates
VA2100	Monocarbocyclic aldehydes and ketones
VA2150	Monocarbocyclic carboxylic acids and lactones
VA2200	Monocarbocyclic methyl esters
VA2250	Other monocarbocyclic esters

VA2300	Polycyclic alkanes
VA2320	Polycyclic alkenes
VA2350	Polycyclic alcohols
VA2400	Polycyclic acetates
VA2500	Polycyclic aldehydes and ketones
VA2550	Polycyclic carboxylic acids
VA2600	Polycyclic methyl esters
VA2620	Other polycyclic esters
VA2630	Cyclic polyenes
VA2650	Simple heteroalicyclics (one O)
VA2700	Simple heteroalicyclics (one N)
VA2750	Simple heteroalicyclics (one S)
VA2800	Simple heteroalicyclics (two O)
VA2850	Simple heteroalicyclics (two N)
VA2950	Simple heteroalicyclics (two S)
VA3000	Simple heteroalicyclics (one O, one N)
VA3050	Simple heteroalicyclics (one O, one S)
VA3100	Simple heteroalicyclics (one N, one S)
VA3150	Simple heteroalicyclics (miscellaneous heteroatoms)
VA3200	Bicycloheteroalicyclics (one O)
VA3250	Bicycloheteroalicyclics (one N)
VA3300	Bicycloheteroalicyclics (one S)
VA3350	Bicycloheteroalicyclics (two O)
VA3400	Bicycloheteroalicyclics (two N)
VA3450	Bicycloheteroalicyclics (two S)
VA3500	Bicycloheteroalicyclics (one O, one N)
VA3550	Bicycloheteroalicyclics (one O, one S)
VA3600	Bicycloheteroalicyclics (one N, one S)
VA3650	Bicycloheteroalicyclics (miscellaneous heteroatoms)
VA3700	Tricycloheteroalicyclics (one O)
VA3750	Tricycloheteroalicyclics (one N)
VA3800	Tricycloheteroalicyclics (one S)
VA3850	Tricycloheteroalicyclics (two O)
VA3900	Tricycloheteroalicyclics (two N)
VA3950	Tricycloheteroalicyclics (two S)
VA4000	Tricycloheteroalicyclics (one O, one N)
VA4050	Tricycloheteroalicyclics (one O, one S)
VA4100	Tricycloheteroalicyclics (one N, one S)
VA4150	Tricycloheteroalicyclics (miscellaneous heteroatoms)
VA4200	Polycycloheteroalicyclic compounds
VA6100	Prostaglandins, thromboxanes etc
VA6150	Oxylipins (including Eicosanoids)
VA6200	Simple thiocyanates and isothiocyanates
VA6300	Oxazolidinethiones
VA6400	2,2'-Bithiophenes
VA6500	Dithiophen-2,2'-yl ethynes
VA6600	Miscellaneous thiophenes
VA6650	Thiolactones
VA6700	Monoacylglycerols
VA6800	Diacylglycerols
VA6900	Triacylglycerols

VA7000	Phospholipids
VA7100	Glycolipids
VA7200	Sphingolipids
VA7300	Long-chain aromatic systems
VA7400	Disulfides, trisulfides
VA8000	Enediynes

Alkaloids

The following is a list of all the **Type of Compound Codes** used for **Alkaloid Natural Products** on the database.

Code	***Classification***
VX0100	Simple acyclic amine alkaloids with one N
VX0120	Simple acyclic amine alkaloids with two N
VX0150	Simple Guanidines
VX0200	Nitriles and isonitriles
VX0220	Simple isobutylamide alkaloids
VX0250	Simple amide alkaloids
VX0300	Simple pyrrolidine alkaloids
VX0340	Chromone alkaloids
VX0350	Flavonoid alkaloids
VX0360	Nicotine-like alkaloids
VX0380	Miscellaneous pyrrolidine alkaloids
VX0390	Tetramic acids
VX0400	Tropane alkaloids
VX0440	Simple pyrrolizidine alkaloids
VX0500	Pyrrolizidine alkaloids (macrocyclic lactones)
VX0520	Miscellaneous pyrrolizidine alkaloids
VX0620	Anabasine-like alkaloids
VX0660	Lobelia alkaloids
VX0680	Simple piperidine alkaloids
VX0690	Xestospongins
VX0700	Miscellaneous piperidine alkaloids
VX0760	Lythraceae alkaloids
VX0900	Quinolizidine alkaloids (two rings)
VX0920	Quinolizidine alkaloids (three rings)
VX0940	Quinolizidine alkaloids (four rings)
VX0960	Cylindricine alkaloids
VX0980	Miscellaneous quinolizidine alkaloids
VX1000	Azepine alkaloids
VX1020	Nicotinic acid derived alkaloids
VX1040	Pyridine alkaloids
VX1120	9B-Azaphenalene alkaloids
VX1140	Naphthalene-isoquinoline alkaloids
VX1160	Elaeocarpus alkaloids
VX1240	Galbulimima alkaloids
VX1260	Stemona alkaloids
VX1280	Lycopodium alkaloids
VX1300	Cytochalasan alkaloids
VX1360	Indolizidine alkaloids
VX1460	Simple anthranilic acid alkaloids
VX1480	Simple quinoline alkaloids
VX1520	Furanoquinoline alkaloids
VX1540	Pyranoquinoline alkaloids

VX1580 Miscellaneous quinoline alkaloids
VX1600 Quinazoline alkaloids
VX1620 Acridone alkaloids
VX1650 Acridone dimers
VX1690 Acridone-coumarin dimers
VX1700 Pyrido[2,3,4-*kl*]acridines
VX1720 1,4-Benzoxazin-3-one alkaloids
VX1760 Benzodiazepine alkaloids
VX1800 Cryptolepine type alkaloids
VX2000 Simple tyramine alkaloids
VX2005 α-Hydroxyphenethylamines
VX2010 Ephedra bases
VX2015 Miscellaneous phenethylamines
VX2020 Cinnamic acid amides
VX2100 Securinega alkaloids
VX2140 Betalain alkaloids
VX2200 Simple isoquinoline alkaloids
VX2250 Manzamine alkaloids
VX2310 Miscellaneous isoquinoline alkaloids
VX2320 Benzylisoquinoline alkaloids
VX2330 Pseudobenzylisoquinoline alkaloids
VX2340 Bisbenzylisoquinoline alkaloids (aryl links only)
VX2360 Bisbenzylisoquinoline alkaloids (one ether link)
VX2370 Bisbenzylisoquinolines (one ether/one aromatic link)
VX2380 Bisbenzylisoquinoline alkaloids (two ether links)
VX2390 Bisbenzylisoquinoline alkaloids (two ether/one aromatic link)
VX2400 Bisbenzylisoquinoline alkaloids (three ether links)
VX2430 Secobisbenzylisoquinoline alkaloids
VX2440 Cularine group alkaloids
VX2450 Secocularine alkaloids
VX2470 Quettamine alkaloids
VX2480 Dibenzopyrrocoline alkaloids
VX2500 Indenobenzazepine alkaloids
VX2520 Pavine alkaloids
VX2540 Isopavine alkaloids
VX2600 Proaporphine alkaloids
VX2620 Proaporphine-benzylisoquinoline alkaloid dimers
VX2640 Aporphine alkaloids
VX2700 Aporphine-benzylisoquinoline alkaloid dimers
VX2750 Oxoisoaporphine alkaloids
VX2780 Aristolochic acid alkaloids
VX2800 Aristolactam alkaloids
VX2820 Phenanthrene alkaloids
VX2840 Azafluoranthrene alkaloids
VX2900 Morphine alkaloids
VX2940 Erythrina alkaloids
VX2980 Dibenzazecine alkaloids
VX3000 Hasubanan alkaloids
VX3100 Protoberberine alkaloids
VX3130 Isoindolobenzazepine alkaloids
VX3140 Narceine alkaloids

VX3160 Protopine alkaloids
VX3180 Rhoeadine alkaloids
VX3200 Phthalideisoquinoline alkaloids
VX3220 Spirobenzylisoquinoline alkaloids
VX3240 Secoberberine alkaloids
VX3300 Benzo[*c*]phenanthridine alkaloids
VX3360 Phenethylisoquinoline alkaloids
VX3370 Homomorphinandienone alkaloids
VX3380 Homoaporphine alkaloids
VX3390 Homoproaporphine alkaloids
VX3400 Colchicine-like alkaloids
VX3410 Dibenzocycloheptylamine alkaloids
VX3420 Cephalotaxus alkaloids
VX3440 Homoerythrina alkaloids
VX3500 Amaryllidaceae alkaloids
VX3600 Mesembrenoid alkaloids
VX3690 Emetine group alkaloids
VX3700 Phenanthroindolizidine alkaloids
VX3760 Phenanthroquinolizidine alkaloids
VX4000 Simple indole alkaloids
VX4020 Simple bisindole alkaloids
VX4040 Simple tryptamine alkaloids
VX4080 Pyrrolnitrin-like alkaloids
VX4100 Physostigmine-like alkaloids
VX4110 Chaetocin-like alkaloids
VX4120 Evodia alkaloids
VX4140 Tryptamine alkaloid dimers
VX4160 Tryptamine alkaloid oligomers
VX4240 β-Carboline alkaloids
VX4300 Carbazole alkaloids
VX4350 Indolo[2,3-*a*]carbazole and related alkaloids
VX4400 Indolonaphthyridine alkaloids
VX4460 Ergot alkaloids
VX4620 Aristotelia alkaloids
VX4640 Monoterpenoid-derived indole alkaloid glycosides
VX4700 Camptothecine-like alkaloids
VX4740 Andranginine-type alkaloids
VX4780 Indoloquinolizine alkaloids
VX4800 Corynanthe alkaloids
VX4820 Corynanthe-tryptamine alkaloids
VX4840 Akagerine alkaloids
VX4860 Ajmalicine-like alkaloids
VX4900 Macroline alkaloids
VX4940 Oxindole alkaloids
VX4960 Gardneramine-oxindole alkaloids
VX4980 Pseudoindoxyl alkaloids
VX5000 Gelsemium alkaloids
VX5040 Yohimbinoid alkaloids
VX5100 Sarpagine alkaloids
VX5120 Ajmaline alkaloids
VX5140 Vobasine alkaloids

VX5180 Ervatamia alkaloids
VX5200 Akuammiline alkaloids
VX5220 Pleiocarpamine alkaloids
VX5240 Cinchona alkaloids
VX5260 Akuammicine alkaloids
VX5280 Strychnidine alkaloids and dimers
VX5320 Condylocarpan alkaloids
VX5360 Secodine alkaloids
VX5380 Bissecodine alkaloids
VX5400 Aspidosperma alkaloids
VX5500 Quebrachamine alkaloids
VX5540 Aspidofractine alkaloids
VX5560 Kopsane alkaloids
VX5580 Melodinus alkaloids
VX5700 Iboga alkaloids
VX5800 Pandoline alkaloids
VX5840 Pyridocarbazole alkaloids
VX5880 Uleine-dasycarpidan alkaloids
VX5900 Eburna alkaloids
VX5950 Hapalindoles
VX5980 Bisindole alkaloids
VX6000 Isoindoles
VX6070 Pyrrolo[4,3,2-*de*]quinoline alkaloids
VX6240 Monoterpenoid alkaloids
VX6260 Secologanin-derived monoterpenoid alkaloids
VX6300 Sesquiterpene alkaloids
VX6320 Macrocyclic sesquiterpene alkaloids
VX6340 Dendrobium alkaloids
VX6360 Nuphar alkaloids
VX6400 C_{19}-Diterpenoid alkaloids and 4-nor analogues
VX6420 C_{20} Diterpenoid alkaloids (Atisine type)
VX6460 Erythrophleum alkaloids
VX6480 Miscellaneous diterpene alkaloids
VX6490 Sesterterpene alkaloids
VX6500 Daphniphylline alkaloids
VX6640 Steroidal alkaloids (salamandra type)
VX6660 Steroidal alkaloids (jerveratrum type)
VX6680 Steroidal alkaloids (cerveratrum type)
VX6700 Steroidal alkaloids (conanine type)
VX6720 Steroidal alkaloids (spirosolane type)
VX6740 Steroidal alkaloids (solanidine type)
VX6760 Steroidal alkaloids (buxus type)
VX6780 Steroidal alkaloids (pregnane type)
VX6790 Miscellaneous steroidal alkaloids
VX6820 Azaanthracene alkaloids
VX6840 Azafluorene alkaloids
VX6900 Pyrazole alkaloids
VX6920 Imidazole alkaloids
VX6925 Cycloheptadiimidazoles
VX6930 Oxazole and benzoxazole alkaloids
VX6932 Isoazole alkaloids

VX6934 Spirobenzoxazoline alkaloids
VX6935 Simple thiazole and benzothiazole alkaloids
VX6936 Latrunculins
VX6937 Macrocyclic thiazole alkaloids
VX6940 Pyrazine and quinoxaline alkaloids
VX6950 Pyrrolo[1,2-*a*]pyrazines
VX6955 Morpholines
VX6960 Pyrimidines
VX6970 Ptilocaulins
VX6980 Triazaacenaphthylene alkaloids
VX6990 Tetrodotoxins
VX7000 Phenazine alkaloids
VX7005 Phenoxazines
VX7015 Pyrroloazepines
VX7010 Pyrrole alkaloids
VX7020 Putrescine alkaloids
VX7030 Acyclic spermine alkaloids
VX7040 Macrocyclic spermine alkaloids
VX7050 Acyclic spermidine alkaloids
VX7060 Homospermidine alkaloids
VX7070 Macrocyclic spermidine alkaloids
VX7100 Peptide alkaloids
VX7120 Amanita alkaloids
VX7200 Pyrrolo[2,3-*d*]pyrimidines
VX7300 Purines
VX7350 Pteridines and analogues
VX7400 Saxitoxins
VX7450 Miscellaneous metal complexes
VX7500 Miscellaneous monocyclic alkaloids
VX7540 Miscellaneous bicyclic alkaloids
VX7600 Miscellaneous polycyclic alkaloids
VX9000 Miscellaneous acyclic alkaloids
VX9100 Miscellaneous alkaloids with one ring
VX9200 Miscellaneous alkaloids with two rings
VX9300 Miscellaneous alkaloids with three rings
VX9400 Miscellaneous alkaloids with four rings
VX9999 Alkaloids of unknown or partially unknown structure

Amino Acid and Peptide Natural Products

The following is a list of all the **Type of Compound Codes** used for **Amino Acid and Peptide Natural Products** on the database.

Code	***Classification***
VV0050	Protein α-aminoacids
VV0100	Non-protein α-aminoacids
VV0120	β-Aminoacids
VV0130	Miscellaneous modified aminoacids
VV0140	Unsaturated aminoacids
VV0150	Diketopiperazines (dipeptide anhydrides)
VV0200	Dipeptides
VV0300	Tripeptides
VV0400	Oligopeptides (4-10 residues)
VV0450	Linear polypeptides
VV0500	Cyclic oligo- and polypeptides
VV0580	Aeruginosins
VV0600	Larger depsipeptides
VV0610	Small depsipeptides
VV0650	Bicyclic/polycyclic depspeptides
VV0660	Azinothricin-type depsipeptides (pyridazine containing)
VV0670	Actinomycin-type depsipeptides
VV0680	Thiopeptin/siomycin-type depsipeptides (thiazole containing)
VV0700	Penicillins
VV0800	Cephalosporins
VV0900	Carbapenems
VV0920	Monocyclic β-lactams (nocardicins and monobactams)
VV0950	Clavams
VV1000	Enzymes
VV2000	Other proteins
VV3000	Glycopeptides and glycoproteins
VV6000	Lipopeptides
VV9999	Peptides of unknown structure

Benzofuranoids

The following is a list of all the **Type of Compound Codes** used for **Benzofuranoid Natural Products** on the database.

Code	***Classification***
VH1000	Benzofurans
VH2000	Benzodifurans
VH3000	Isobenzofurans
VH3200	Angeolide group
VH3500	Flavagline compounds

Benzopyranoids

The following is a list of all the **Type of Compound Codes** used for **Benzopyranoid Natural Products** on the database.

Code	***Classification***
VI0030	1-Benzopyrans
VI0050	Furo-1-benzopyrans
VI0070	Pyrano-1-benzopyrans
VI0100	Non-oxygenated coumarins
VI0103	3-Oxygenated coumarins
VI0104	4-Oxygenated coumarins
VI0105	5-Oxygenated coumarins
VI0106	6-Oxygenated coumarins
VI0107	7-Oxygenated coumarins, unsubstituted
VI0116	7-Oxygenated coumarins, 6-substituted
VI0118	7-Oxygenated coumarins, 8-substituted
VI0120	7-Oxygenated coumarins with miscellaneous substituents
VI0200	8-Oxygenated coumarins
VI0304	3,4-Dioxygenated coumarins
VI0305	3,5-Dioxygenated coumarins
VI0306	3,6-Dioxygenated coumarins
VI0307	3,7-Dioxygenated coumarins
VI0308	3,8-Dioxygenated coumarins
VI0405	4,5-Dioxygenated coumarins
VI0406	4,6-Dioxygenated coumarins
VI0407	4,7-Dioxygenated coumarins
VI0408	4,8-Dioxygenated coumarins
VI0506	5,6-Dioxygenated coumarins
VI0507	5,7-Dioxygenated coumarins
VI0508	5,8-Dioxygenated coumarins
VI0607	6,7-Dioxygenated coumarins
VI0608	6,8-Dioxygenated coumarins
VI0708	7,8-Dioxygenated coumarins
VI3000	Trioxygenated coumarins
VI7000	Tetra- and penta-oxygenated coumarins
VI7500	Bis- and tris-coumarins
VI8000	Furanocoumarins
VI8500	Dihydrofuranocoumarins
VI9000	Pyranocoumarins
VI9500	Dihydropyranocoumarins
VI9600	2-Benzopyrans
VI9620	Azaphilone 2-benzopyrans
VI9650	Furo-2-benzopyrans
VI9670	Pyrano-2-benzopyrans
VI9700	Isocoumarins
VI9800	2-Benzothiopyrans
VI9999	Coumarins of unknown structure

Carbohydrate Natural Products

The following is a list of all the **Type of Compound Codes** used for **Carbohydrate Natural Products** on the database.

Code	***Classification***
VE0100	Tetroses
VE0200	*arabino*-Pentoses
VE0300	*lyxo*-Pentoses
VE0400	*ribo*-Pentoses
VE0500	*xylo*-Pentoses
VE0600	*allo*-Hexoses
VE0700	*altro*-Hexoses
VE0800	*galacto*-Hexoses
VE0900	*gluco*-Hexoses
VE1000	*gulo*-Hexoses
VE1100	*ido*-Hexoses
VE1200	*manno*-Hexoses
VE1300	*talo*-Hexoses
VE1400	Higher aldoses
VE1500	*erythro*-Pentuloses
VE1600	*threo*-Pentuloses
VE1700	*fructo*-Hexuloses
VE1800	*psico*-Hexuloses
VE1900	*sorbo*-Hexuloses
VE2000	*tagato*-Hexuloses
VE2100	Miscellaneous ketoses
VE2200	Higher ketoses
VE2300	1,2-Anhydrosugars
VE2400	1,3-Anhydrosugars
VE2500	1,4-Anhydrosugars
VE2600	1,5-Anhydrosugars
VE2700	1,6-Anhydrosugars
VE2800	2,3-Anhydrosugars
VE2900	2,5-Anhydrosugars
VE3000	2,6-Anhydrosugars
VE3100	3,4-Anhydrosugars
VE3200	3,6-Anhydrosugars
VE3300	4,6-Anhydrosugars
VE3400	5,6-Anhydrosugars
VE3500	Dianhydrosugars
VE3600	Other anhydrosugars
VE3700	Glycosyl halides and other 1-halogenosugars
VE3800	2-Halogenosugars
VE3900	3-Halogenosugars
VE4000	4-Halogenosugars
VE4100	5-Halogenosugars
VE4200	6-Halogenosugars
VE4300	Glycosylamines

VE4400 2-Amino-2-deoxysugars
VE4500 3-Amino-3-deoxysugars
VE4600 4-Amino-4-deoxysugars
VE4700 5-Amino-5-deoxysugars
VE4800 6-Amino-6-deoxysugars
VE4900 Aminodeoxyalditols
VE4950 Azidosugars
VE5000 Thioglycosides and 1-thiosugars
VE5100 Other thiosugars
VE5150 1-Deoxy sugars
VE5200 2-Deoxy sugars
VE5300 3-Deoxy sugars
VE5400 4-Deoxy sugars
VE5500 5-Deoxy sugars
VE5600 2,6-Dideoxy sugars
VE5700 3,6-Dideoxy sugars
VE5800 4,6-Dideoxy sugars
VE5850 Polydeoxy sugars
VE5900 6-Deoxyalloses
VE6000 6-Deoxyaltroses
VE6100 6-Deoxygalactoses
VE6200 6-Deoxyglucoses
VE6300 6-Deoxyguloses
VE6400 6-Deoxyidoses
VE6500 6-Deoxymannoses
VE6600 6-Deoxytaloses
VE6650 Other deoxy sugars
VE6700 Unsaturated sugars; 1-enes
VE6800 Unsaturated sugars; 2-enes
VE6900 Unsaturated sugars; 3-enes
VE7000 Unsaturated sugars; 4-enes
VE7100 Unsaturated sugars; 5-enes
VE7150 Other unsaturated sugars
VE7200 Branched chain sugars
VE7300 Dicarbonyl sugars; glycos-2-uloses
VE7400 Dicarbonyl sugars; glycos-3-uloses
VE7500 Dicarbonyl sugars; glycos-4-uloses
VE7600 Dicarbonyl sugars; glycos-5-uloses
VE7700 Diuloses
VE7800 Dialdoses
VE7900 Aldonic acids
VE8000 Glycuronic acids
VE8100 Aldaric acids
VE8200 Ketoacid sugars
VE8300 Higher sugar acids
VE8400 Other sugar acids
VE8600 Tetritols
VE8700 Pentitols
VE8800 Hexitols
VE8900 Higher alditols
VE9000 Cyclitols

VE9100 Other polyols
VE9150 Sugar phosphates
VE9200 Disaccharides
VE9300 Oligosaccharides
VE9400 Polysaccharides
VE9500 Miscellaneous carbohydrate antibiotics
VE9600 Glycosinolates
VE9700 Cyanogenic glycosides
VE9900 Nucleosides
VE9999 Carbohydrates of unknown or partially unknown structure

Drugs

The following is a list of all the **Type of Compound Codes** used for **Drugs** on the database.

Code	*Compound Class*
XA	**Drugs**
XA0020	Abortifacients
XA0050	Acaricides
XA0070	Adenosine receptor agonists
XA0080	Adenosine receptor antagonists
XA0090	Adrenergic - neurone blocking drugs
XA0101	α-Adrenoceptor agonists
XA0111	β-Adrenoceptor agonists
XA0150	α-Adrenoceptor antagonists
XA0200	β-Adrenoceptor antagonists
XA0370	Aldehyde dehydrogenase inhibitors
XA0380	Aldose reductase inhibitors
XA0400	Aldosterone antagonists
XA0450	Aminoglycoside antibiotics
XA0500	Anabolic agents
XA0550	Anaesthetics, general
XA0600	Anaesthetics, local
XA0650	Analgesics
XA0660	Analgesics - NSAID
XA0670	Analgesics - opioid
XA0700	Androgenic agents
XA0750	Angiographic media
XA0800	Angiotensin-converting enzyme inhibitors
XA0810	Angiotensin receptor agonists
XA0820	Angiotensin receptor antagonists
XA0850	Anorectic agents
XA0860	Antacids
XA0900	Anthelmintics
XA0920	Antiageing drugs
XA0950	Antiallergic agents
XA0960	Antiamoebic agents
XA0970	Antianaemic agents
XA1000	Antiandrogens
XA1020	Antianginal agents
XA1030	Antianxiety agents
XA1050	Antiarrhythmic agents
XA1100	Antiasthmatic agents
XA1150	Antibacterial agents
XA1200	Antibiotics
XA1300	Anticholinesterases
XA1340	Anticoagulant antagonists
XA1350	Anticoagulants
XA1360	Anticoccidial agents

XA1370 Anticolitis agents
XA1400 Anticonvulsants
XA1450 Antidepressants
XA1460 Antidiarrhoeal agents
XA1470 Antidiuretic agents
XA1500 Antiemetics
XA1525 Antiestrogens
XA1530 Antifibrinogens
XA1535 Antifibrinolytic agents
XA1550 Antifungal agents
XA1625 Anti-HIV agents
XA1640 Antihyperglycaemic agents
XA1645 Antihyperlipidaemic agents
XA1650 Antihypertensive agents
XA1660 Antihypoglycaemic agents
XA1670 Antihypotensive agents
XA1700 Antiinflammatory agents
XA1730 Antileishmanial agents
XA1750 Antileprotic agents
XA1800 Antimalarials
XA1850 Antimanic agents
XA1920 Antimigraine agents
XA1930 Antimitotic agents
XA1950 Antineoplastic agents
XA1980 Antioxidants and free radical scavengers
XA2050 Antiparkinsonian agents
XA2070 Antiperspirants
XA2100 Antiprotozoals
XA2130 Antipsychotics
XA2150 Antipyretics
XA2180 Antisense oligonucleotides
XA2200 Antiseptics
XA2220 Antischistosomes
XA2230 Antisickling compounds
XA2250 Antispasmodics
XA2270 Antisympathetic agents
XA2280 Antithrombins
XA2290 Antithrombotic agents
XA2300 Antithyroid agents
XA2320 Antitrichomonal agents
XA2330 Antitrypanosomal agents
XA2340 Antitubercular agents
XA2350 Antitussives
XA2370 Antiulcerogenic agents
XA2400 Antiviral agents
XA2450 Aromatase inhibitors
XA2500 Astringents
XA2520 ATPase inhibitors
XA2530 Atrial natriuretic peptide receptor agonists
XA2540 Atrial natriuretic peptide receptor antagonists
XA2545 Benzodiazepine binding-site agonists

XA2547 Benzodiazepine binding-site antagonists
XA2548 Benzodiazepine binding-site inverse agonists
XA2550 Bombesin receptor agonists
XA2560 Bombesin receptor antagonists
XA2575 Bradykinin receptor agonists
XA2576 Bradykinin receptor antagonists
XA2600 Bronchodilators
XA2605 Calcitonin-gene related peptide agonists
XA2608 Calcitonin-gene related peptide antagonists
XA2610 Calcium channel activators
XA2620 Calcium channel blockers
XA2625 Calcium metabolism modifying agents
XA2630 Cannabinoid receptor agonists
XA2635 Cannabinoid receptor antagonists
XA2650 Carbonic anhydrase inhibitors
XA2680 Carboxypeptidase inhibitors
XA2700 Cardiac depressants
XA2750 Cardiac glycosides
XA2760 Cardiac stimulants
XA2800 Carminatives
XA2850 Central depressants
XA2900 Central stimulants
XA2950 Cephalosporins
XA3050 Chelating agents
XA3060 Chloride channel activators
XA3070 Chloride channel blockers
XA3080 Cholecystokinin receptor agonists
XA3090 Cholecystokinin receptor antagonists
XA3100 Choleretic agents
XA3110 Cholinesterase reactivators
XA3120 Cholinoceptor - muscarinic agonists
XA3130 Cholinoceptor - muscarinic antagonists
XA3140 Cholinoceptor - nicotinic agonists
XA3150 Cholinoceptor - nicotinic antagonists
XA3250 Contraceptives
XA3300 Contraceptives, oral
XA3400 Contrast media
XA3550 Corticosteroids
XA3600 Corticotrophin analogues
XA3605 Counter irritants
XA3610 Cyclooxygenase inhibitors
XA3620 Dehydropeptidase inhibitors
XA3640 Dental caries prophylactics
XA3650 Dermatological agents
XA3670 Diagnostic agents
XA3680 Diamine oxidase inhibitors
XA3750 Digestive agents
XA3780 Dihydrofolate reductase inhibitors
XA3850 Diuretics
XA3870 DNA intercalating agents
XA3880 Dopa decarboxylase inhibitors

XA3950 Dopamine β-hydroxylase inhibitors
XA3960 Dopamine receptor agonists
XA3970 Dopamine receptor antagonists
XA3972 Dopamine reuptake inhibitors
XA4050 Emetics
XA4080 Endothelin receptor agonists
XA4085 Endothelin receptor antagonists
XA4095 Enzyme inhibitors
XA4100 Enzymes
XA4110 Enzyme substrates
XA4130 Estrogens
XA4140 Excipients
XA4150 Expectorants
XA4155 Fertility agents
XA4157 Fibrinolytic agents
XA4160 Formyl receptor agonists
XA4162 Formyl receptor antagonists
XA4175 GABA receptor agonists
XA4180 GABA receptor antagonists
XA4182 GABA reuptake inhibitors
XA4190 Gallstone dispersing agents
XA4200 Ganglion blocking agents
XA4205 Gastric motility agents
XA4210 Gastric proton pump inhibitors
XA4220 Gastric secretion inhibitors
XA4280 Glucocorticoid antagonists
XA4300 Glucocorticoids
XA4305 Glutamate receptor agonists
XA4306 Glutamate receptor antagonists
XA4309 Glycine receptor agonists
XA4310 Glycine receptor antagonists
XA4315 G-protein activators
XA4320 G-protein antagonists
XA4500 Haemostatic agents
XA4507 Hepatoprotective agents
XA4508 Histamine receptor agonists
XA4510 Histamine H_1-receptor antagonists
XA4515 Histamine H_2-receptor antagonists
XA4520 Histamine H_3-receptor antagonists
XA4530 Histidine decarboxylase inhibitors
XA4550 HMG CoA reductase inhibitors
XA4630 5-Hydroxytryptamine receptor agonists
XA4635 5-Hydroxytryptamine receptor antagonists
XA4637 5-Hydroxytryptamine reuptake inhibitors
XA4650 Hypnotics
XA4660 Hypotensive agents
XA4670 Hypothalamic hormones
XA4750 Immunomodulators
XA4780 Immunostimulants
XA4800 Immunosuppressants
XA4830 Inotropic agents

XA4850 Insecticides
XA4920 Keratolytic agents
XA4925 β-Lactamase inhibitors
XA4930 Laxatives
XA4933 Leukotriene receptor agonists
XA4935 Leukotriene receptor antagonists
XA4940 LHRH receptor agonists
XA4945 LHRH receptor antagonists
XA4960 Lipoxygenase inhibitors
XA4970 Luteolytic agents
XA4972 Macrolide antibiotics
XA4976 Melatonin receptor agonists
XA4977 Melatonin receptor antagonists
XA4980 Metal-poisoning antidotes
XA5050 Mineralocorticoids
XA5055 Mineral supplements
XA5060 Miotic agents
XA5070 Molluscicides
XA5100 Monoamine oxidase inhibitors
XA5150 Mucolytic agents
XA5200 Muscle relaxants - skeletal
XA5210 Muscle relaxants - smooth
XA5220 Mydriatic agents
XA5420 Nasal decongestants
XA5500 Neuromuscular blocking agents - competitive
XA5510 Neuromuscular blocking agents - depolarising
XA5600 Neuropeptide Y receptor agonists
XA5610 Neuropeptide Y receptor antagonists
XA5611 Neuroprotective agents
XA5612 Neurotensin receptor agonists
XA5613 Neurotensin receptor antagonists
XA5614 Neurotransmitter release modulating agents
XA5615 Neutral endopeptidase inhibitors
XA5620 Nitrergic stimulants
XA5630 Nitric oxide synthetase inhibitors
XA5640 Nootropic agents
XA5642 Noradrenaline reuptake inhibitors
XA5650 Nutritional agents
XA5720 Oligonucleotides
XA5730 Opioid receptor agonists
XA5740 Opioid receptor antagonists
XA5770 Ovulation-inducing agents
XA5800 Oxytocic agents
XA5850 Penicillins
XA5940 Phosphodiesterase inhibitors
XA5960 Phospholipase inhibitors
XA5962 Photosensitisers
XA5965 Pituitary hormones
XA5967 Platelet-activating factor receptor agonists
XA5969 Platelet-activating factor receptor antagonists
XA5970 Platelet aggregation inhibiting agents

XA5971 Platelet aggregation inducers
XA5975 Potassium channel activators
XA5980 Potassium channel blockers
XA5990 Progestogen receptor antagonists
XA6000 Progestogens
XA6020 Prolactin release inhibitors
XA6050 Prostanoid receptor agonists
XA6052 Prostanoid receptor antagonists
XA6055 Proteinase inhibitors - aminoproteinase
XA6057 Proteinase inhibitors - metalloproteinase
XA6061 Proteinase inhibitors - serineproteinase
XA6063 Proteinase inhibitors - thiolproteinase
XA6065 Protein kinase inhibitors
XA6070 Psychotropic agents
XA6170 Purine receptor agonists
XA6175 Purine receptor antagonists
XA6200 Radiopaque agents
XA6220 Radiopharmaceutical agents
XA6230 Radioprotective agents
XA6235 Radiosensitisers
XA6240 5α-Reductase inhibitors
XA6242 Renin inhibitors
XA6245 Respiratory stimulants
XA6250 Reverse transcriptase inhibitors
XA6300 Sedatives
XA6320 Sensory irritants
XA6375 Sodium channel activators
XA6380 Sodium channel blockers
XA6390 Somatostatin receptor agonists
XA6395 Somatostatin receptor antagonists
XA6410 Sulfonamides
XA6420 Sunscreen agents
XA6430 Sympathomimetic agents
XA6450 Tachykinin receptor agonists
XA6455 Tachykinin receptor antagonists
XA6500 Tetracyclines
XA6540 Thromboxane synthetase inhibitors
XA6550 Thyroid hormones
XA6570 Toxins - general
XA6575 Toxins - neurotoxins
XA6600 Tranquillisers
XA6801 Uptake inhibitors
XA6820 Urease inhibitors
XA6850 Uricosuric agents
XA6950 Vasoconstrictors
XA7000 Vasodilators
XA7170 Vasopressin receptor agonists
XA7175 Vasopressin receptor antagonists
XA7200 Vitamins

Flavonoids

The following is a list of all the **Type of Compound Codes** used for **Flavonoid Natural Products** on the database.

Code	***Classification***
VK0010	Anthocyanidins and anthocyanins; one O substituent
VK0020	Anthocyanidins and anthocyanins; two O substituents
VK0030	Anthocyanidins and anthocyanins; three O substituents
VK0040	Anthocyanidins and anthocyanins; four O substituents
VK0050	Anthocyanidins and anthocyanins; five O substituents
VK0060	Anthocyanidins and anthocyanins; six O substituents
VK0070	Anthocyanidins and anthocyanins; seven O substituents
VK0080	Anthocyanidins and anthocyanins; eight O substituents
VK0090	Anthocyanidins and anthocyanins; nine O substituents
VK0095	Pyranoanthocyanidins
VK1000	Flavans
VK1100	Flavan-3-ols
VK1200	Leucoanthocyanidins
VK1250	Flavan-4-ols
VK1300	Peltogynoid flavonoids
VK1500	Proanthocyanidin flavonoids
VK2000	Biflavonoids and polyflavonoids
VK3000	Isoflavones; no O substituent
VK3010	Isoflavones; one O substituent
VK3020	Isoflavones; two O substituents
VK3030	Isoflavones; three O substituents
VK3040	Isoflavones; four O substituents
VK3050	Isoflavones; five O substituents
VK3060	Isoflavones; six O substituents
VK3070	Isoflavones; seven O substituents
VK3080	Isoflavones; eight O substituents
VK3090	Isoflavones; nine O substituents
VK3100	Isoflavanones
VK3200	Simple rotenoid flavonoids
VK3250	12*a*-Hydroxyrotenoid flavonoids
VK3300	Dehydrorotenoid flavonoids
VK3400	Simple pterocarpan flavonoids
VK3450	6*a*-Hydroxypterocarpan flavonoids
VK3500	Pterocarpene flavonoids
VK3550	Pterocarpanone and pterocarpenequinone flavonoids
VK3600	Isoflavans
VK3650	Isoflavanquinones
VK3670	Isoflavanols
VK3680	Isoflav-2-enes
VK3700	Isoflav-3-enes
VK3720	3-Arylcoumarin flavonoids
VK3750	Coumestan flavonoids

VK3770 Coumaranochromene flavonoids
VK3800 α-Methyldeoxybenzoin flavonoids
VK3820 2-Arylbenzofuran flavonoids
VK3850 Biisoflavans
VK4000 Neoflavonoids
VK5000 Flavones
VK5010 Flavones; one O substituent
VK5020 Flavones; two O substituents
VK5030 Flavones; three O substituents
VK5040 Flavones; four O substituents
VK5050 Flavones; five O substituents
VK5060 Flavones; six O substituents
VK5070 Flavones; seven O substituents
VK5080 Flavones; eight O substituents
VK5090 Flavones; nine O substituents
VK5210 Flavonols; one O substituent
VK5220 Flavonols; two O substituents
VK5230 Flavonols; three O substituents
VK5240 Flavonols; four O substituents
VK5250 Flavonols; five O substituents
VK5260 Flavonols; six O substituents
VK5270 Flavonols; seven O substituents
VK5280 Flavonols; eight O substituents
VK5290 Flavonols; nine O substituents
VK6010 Chalcone flavonoids; one O substituent
VK6020 Chalcone flavonoids; two O substituents
VK6030 Chalcone flavonoids; three O substituents
VK6040 Chalcone flavonoids; four O substituents
VK6050 Chalcone flavonoids; five O substituents
VK6060 Chalcone flavonoids; six O substituents
VK6070 Chalcone flavonoids; seven O substituents
VK6080 Chalcone flavonoids; eight O substituents
VK6090 Chalcone flavonoids; nine O substituents
VK6100 Aurone flavonoids
VK6200 Dihydrochalcone flavonoids
VK6300 Flavanones; no O substituents
VK6310 Flavanones; one O substituent
VK6320 Flavanones; two O substituents
VK6330 Flavanones; three O substituents
VK6340 Flavanones; four O substituents
VK6350 Flavanones; five O substituents
VK6360 Flavanones; six O substituents
VK6370 Flavanones; seven O substituents
VK6380 Flavanones; eight O substituents
VK6390 Flavanones; nine O substituents
VK6410 Dihydroflavonols; one O substituent
VK6420 Dihydroflavonols; two O substituents
VK6430 Dihydroflavonols; three O substituents
VK6440 Dihydroflavonols; four O substituents
VK6450 Dihydroflavonols; five O substituents
VK6460 Dihydroflavonols; six O substituents

VK6470 Dihydroflavonols; seven O substituents
VK6480 Dihydroflavonols; eight O substituents
VK6490 Dihydroflavonols; nine O substituents
VK6500 Furanoflavonoids
VK6600 Diarylpropane flavonoids
VK6700 Cinnamylphenol flavonoids
VK6800 Homoisoflavonoids
VK8300 Cyclised *C*-isopentenylated flavonoids
VK9999 Flavonoids of unknown or partially unknown structure

Lignans

The following is a list of all the **Type of Compound Codes** used for **Lignan Natural Products** on the database.

Code	***Classification***
VO0050	Simple dibenzylbutane lignans
VO0100	Side-chain oxygenated dibenzylbutane lignans
VO0150	Saturated dibenzylbutyrolactone lignans
VO0200	Unsaturated dibenzylbutyrolactone lignans
VO0250	9,9'-Epoxytetrahydrofuranoid lignans
VO0280	7,8'-Epoxytetrahydrofuranoid lignans
VO0300	7,9'-Epoxytetrahydrofuranoid lignans
VO0350	7,7'-Epoxytetrahydrofuranoid lignans
VO0370	Furanoid lignans
VO0400	Simple furofuranoid lignans
VO0450	Side-chain oxygenated furofuranoid lignans
VO0470	Miscellaneous furofuranoid lignans
VO0500	Simple aryltetralin lignans
VO0550	Side-chain oxygenated aryltetralin lignans
VO0600	Aryltetralin lactone lignans
VO0650	Apolignans
VO0670	7',7'-Cyclolignans (cyclobutanes)
VO0700	Naphthalenoid lignans
VO0750	Dibenzocyclooctadiene lignans
VO0800	Norlignans
VO1500	Neolignans
VO1600	Flavonolignans

Miscellaneous Natural Products

The **Type of Compound Code** used for **Miscellaneous Natural Products** on the database is as follows:

Code	***Classification***
VZ9999	Natural products of unknown structure

Oxygen Heterocycles

The following is a list of all the **Type of Compound Codes** used for **Oxygen Heterocycle Natural Products** on the database.

Code	***Classification***
VF1000	β-Lactones
VF2000	Furans
VF3000	Butanolides
VF4000	Pyrans
VF5000	Pentanolides
VF5100	Avenaciolide group
VF5200	Glauconic acid group
VF6000	2-Pyrones
VF7000	4-Pyrones
VF8000	Spiroketals

Polycyclic Aromatic Natural Products

The following is a list of all the **Type of Compound Codes** used for **Polycyclic Aromatic Natural Products** on the database.

Code	***Classification***
VQ2000	Naphthalenes
VQ2100	Furonaphthalenes
VQ2200	Pyranonaphthalenes
VQ2400	Dinaphthyl ethers
VQ2500	Binaphthyls
VQ2600	Perylenes
VQ2700	Duclauxin group
VQ3000	Naphthoquinones with no O substituents
VQ3010	Naphthoquinones with one O substituent
VQ3020	Naphthoquinones with two O substituents
VQ3030	Naphthoquinones with three O substituents
VQ3040	Naphthoquinones with four O substituents
VQ3050	Naphthoquinones with five O substituents
VQ3060	Naphthoquinones with six O substituents
VQ3090	Benzochromanquinones
VQ3100	Benzoisochromanquinones
VQ3300	Indenes
VQ3400	Indan-1-spirocyclohexanes
VQ3450	Anthracenes
VQ3500	Furoanthracenes
VQ3700	Pyranoanthracenes
VQ4000	9,10-Anthraquinones with no O substituents
VQ4010	9,10-Anthraquinones with one O substituent
VQ4020	9,10-Anthraquinones with two O substituents
VQ4030	9,10-Anthraquinones with three O substituents
VQ4040	9,10-Anthraquinones with four O substituents
VQ4050	9,10-Anthraquinones with five O substituents
VQ4060	9,10-Anthraquinones with six O substituents
VQ4070	9,10-Anthraquinones with seven O substituents
VQ4080	9,10-Anthraquinones with eight O substituents
VQ4100	1,2- and 1,4-Anthraquinones
VQ4200	Pyrano[*b*]anthraquinones
VQ4300	Anthracyclinones
VQ4350	Benzoanthracyclinones
VQ4400	Benz[*b*]anthraquinones
VQ4800	Phenanthrenes
VQ5000	9,10-Phenanthraquinones with no O substituent
VQ5010	9,10-Phenanthraquinones with one O substituent
VQ5020	9,10-Phenanthraquinones with two O substituents
VQ5030	9,10-Phenanthraquinones with three O substituents
VQ5040	9,10-Phenanthraquinones with four O substituents
VQ5050	9,10-Phenanthraquinones with five O substituents

VQ5060 9,10-Phenanthraquinones with six O substituents
VQ5070 9,10-Phenanthraquinones with seven O substituents
VQ5080 9,10-Phenanthraquinones with eight O substituents
VQ5100 1,4-Phenanthraquinones
VQ6000 Extended quinones
VQ7000 Ansaquinones
VQ7500 Phenalenes
VQ7600 Acenaphthalenes
VQ7700 Fluorenes
VQ9000 Miscellaneous polycyclic aromatics

Polyketides

The following is a list of all the **Type of Compound Codes** used for **Polyketide Natural Products** on the database.

Code	***Classification***
VC0050	Linear polyketides
VC0070	Marine halogenated acetogenins
VC0080	Annonaceae acetogenins
VC0100	Macrolide polyketides
VC0150	Lactone polyketides
VC0170	Miscellaneous small polyketides
VC0180	Malyngamides and related amides
VC0200	Ansamycins and related polyketides
VC0250	Miscellaneous macrolide polyketides
VC0300	Polyenes
VC0400	Linear tetracyclines
VC0450	Angucyclines
VC0500	Polyether antibiotics
VC0550	Spirolides and pinnatoxins
VC0600	Aflatoxins and related substances
VC9999	Polyketides of unknown structure

Polypyrroles

The following is a list of all the **Type of Compound Codes** used for **Polypyrrole Natural Products** on the database.

Code	***Classification***
VY0900	Tetrapyrrole precursors
VY0905	Porphyrins and porphyrinogens
VY0910	Haems and metal-free haems
VY0915	Bile pigments (bilins)
VY0920	Chlorophylls and derivatives
VY0925	Bacteriochlorophylls and derivatives
VY0930	Vitamin B12 precursors
VY0935	Vitamin B12 variants
VY0940	Geoporphyrins
VY0945	Miscellaneous polypyrroles

Simple Aromatic Natural Products

The following is a list of all the **Type of Compound Codes** used for **Simple Aromatic Natural Products** on the database.

Code	***Classification***
VG0005	Simple benzene derivatives
VG0010	Simple phenols
VG0020	Simple benzyl alcohols
VG0030	Simple benzaldehydes
VG0035	Simple aryl ketones
VG0040	Simple benzoic acids and esters
VG0050	Phenylacetic acid derivatives
VG0060	Simple phenylpropanoids
VG0070	Miscellaneous aryl derivatives
VG0300	Benzoquinones with no O substituents
VG0310	Benzoquinones with one O substituent
VG0320	Benzoquinones with two O substituents
VG0330	Benzoquinones with three O substituents
VG0340	Benzoquinones with four O substituents
VG370	Prenylated benzoquinones
VG0380	Furanobenzoquinones
VG0390	Polyphenylbenzoquinones
VG0400	Hydroquinones
VG0450	Diphenylmethanes
VG0460	Acylphloroglucinols
VG0500	Benzophenones with no O substituent
VG0501	Benzophenones with one O substituent
VG0502	Benzophenones with two O substituents
VG0503	Benzophenones with three O substituents
VG0504	Benzophenones with four O substituents
VG0505	Benzophenones with five O substituents
VG0506	Benzophenones with six O substituents
VG0507	Benzophenones with seven O substituents
VG0508	Benzophenones with eight O substituents
VG0520	Dibenzofurans
VG0530	Griseofulvins
VG0535	Dibenzo[*b*,*d*]pyrans
VG0540	Dibenzo[*b*,*e*]pyrans
VG0550	Xanthones with no O substituent
VG0551	Xanthones with one O substituent
VG0552	Xanthones with two O substituents
VG0553	Xanthones with three O substituents
VG0554	Xanthones with four O substituents
VG0555	Xanthones with five O substituents
VG0556	Xanthones with six O substituents
VG0557	Xanthones with seven O substituents
VG0558	Xanthones with eight O substituents
VG0600	Unchlorinated depsidones
VG0610	Chlorinated depsidones
VG0620	Dimeric unchlorinated depsides

VG0630 Dimeric chlorinated depsides
VG0640 Trimeric unchlorinated depsides
VG0650 Trimeric chlorinated depsides
VG0660 Tetrameric depsides
VG1000 Diphenyl ethers
VG2000 Biphenyls
VG3000 Dibenzyls
VG4000 Stilbenes
VG5000 Stilbene polymers
VG7000 Diarylalkyls
VG7500 Terphenyls
VG7600 Pulvinone group
VG9800 Tropolone derivatives
VG9999 Simple aromatic natural products of unknown structure

Steroid Natural Products

The following is a list of all the **Type of Compound Codes** used for **Steroid Natural Products** on the database.

Code	***Classification***
VT0100	Estrane steroids (aromatic ring A)
VT0150	Nonaromatic 19-norandrostane steroids (estranes)
VT0250	Androstane steroids
VT0300	Androstane-17-carboxylic acid steroids (etianic acids)
VT0400	19-Norpregnane steroids
VT0450	Pregnane steroids
VT0500	Pregn-20-yne steroids (17-ethynylandrostanes)
VT0550	23,24-Dinorcholan-22-oic acid steroids (pregnane-20-carboxylic acids)
VT0650	24-Norcholan-23-oic acid steroids
VT0750	Cardanolide steroids
VT0800	Cholan-24-oic acid steroids
VT0850	Other cholane steroids
VT0900	Bufanolide steroids
VT0950	Homocholane steroids (26,27-dinorcholestanes)
VT1000	27-Norcholestane steroids
VT1050	Neutral cholestane steroids
VT1100	Cholestanoic acid steroids
VT1150	Ecdysteroids
VT1200	Spirostane steroids
VT1250	Furostane steroids
VT1300	Ergostane steroids (excluding withanolides and brassinolides)
VT1400	Withanolide and brassinolide steroids
VT1550	Stigmastane steroids
VT1700	Gorgostane and other cyclopropacholestane steroids
VT2850	Vitamin D_3 (cholecalciferol) metabolites and analogues
VT2900	Vitamin D_2 (ergocalciferol) metabolites and analogues
VT9999	Steroids of unknown structure

Tannins

The following is a list of all the **Type of Compound Codes** used for **Tannin Natural Products** on the database.

Code	***Classification***
VM6000	Simple gallate ester tannins
VM6100	Hexahydroxydiphenoyl ester tannins
VM6200	Dehydrohexahydroxydiphenoyl ester tannins
VM6300	Elaeocarpusinoyl ester tannins
VM6400	Dehydrochebuloyl ester tannins
VM6500	Chebuloyl ester tannins
VM6600	Brevifoloyl ester tannins
VM6700	Dehydrodigalloyl ester tannins
VM6800	Valoneoyl ester tannins
VM6820	Valoneoyl monolactone tannins
VM6850	Valoneoyl dilactone tannins
VM6900	Sanguisorbyl ester tannins
VM7000	Flavogallonoyl ester tannins
VM7050	Flavogallonoyl dilactone tannins
VM7100	Tetrahydroxybenzofuran dicarboxylate tannins
VM7200	Macaranoyl ester tannins
VM7300	Tergalloyl ester tannins
VM7350	Tergalloyl monolactone tannins
VM7400	Trilloyl ester tannins
VM7500	Euphorbinoyl ester tannins
VM7550	Dehydroeuphorbinoyl ester tannins
VM7600	Gallagyl ester tannins
VM7605	Flavotannins
VM7650	Terchebuloyl ester tannins
VM7700	Mallatoyl ester tannins
VM7750	Woodfordinoyl ester tannins
VM7800	Flavonotannins

Terpenoids

The following is a list of all the **Type of Compound Codes** used for **Terpenoid Natural Products** on the database.

Code	***Classification***
VS0050	Hemiterpenoids
VS0100	Acyclic monoterpenoids
VS0150	Irregular acyclic monoterpenoids
VS0200	Halogenated dimethyloctane monoterpenoids
VS0220	Ochtodane monoterpenoids
VS0240	1-Ethyl-1,3-dimethylcyclohexane monoterpenoids
VS0260	1-Ethyl-2,4-dimethylcyclohexane monoterpenoids
VS0300	Cyclopropane monoterpenoids
VS0350	Cyclobutane monoterpenoids
VS0400	Iridoid monoterpenoids
VS0420	10-Alkyliridoid monoterpenoids
VS0440	Secoiridoid monoterpenoids
VS0450	Other cyclopentane monoterpenoids
VS0500	*p*-Menthane monoterpenoids
VS0520	*m*-Menthane monoterpenoids
VS0540	*o*-Menthane monoterpenoids
VS0600	Other cyclohexane monoterpenoids
VS0700	Cycloheptane monoterpenoids
VS0800	Camphane monoterpenoids
VS0850	Fenchane monoterpenoids
VS0900	Pinane monoterpenoids
VS0950	Carane monoterpenoids
VS1000	Thujane monoterpenoids
VS1050	Miscellaneous bicyclic monoterpenoids
VS1200	Tricyclic monoterpenoids
VS1300	Simple farnesane sesquiterpenoids
VS1310	Homofarnesane sesquiterpenoids
VS1320	Furanoid farnesane sesquiterpenoids
VS1400	Irregular acyclic sesquiterpenoids
VS1420	Miscellaneous cyclobutane sesquiterpenoids
VS1430	Cyclopentane sesquiterpenoids
VS1450	Cyclofarnesane sesquiterpenoids
VS1460	Rearranged cyclofarnesane sesquiterpenoids
VS1470	Humbertiane sesquiterpenoids
VS1500	Bisabolane sesquiterpenoids
VS1520	Miscellaneous cyclohexane sesquiterpenoids
VS1530	Cycloheptane sesquiterpenoids
VS1540	Cyclooctane sesquiterpenoids
VS1550	Cyclobisabolane sesquiterpenoids
VS1600	Elemane sesquiterpenoids
VS1650	Simple germacrane sesquiterpenoids
VS1660	12,6-Germacranolide sesquiterpenoids

VS1670 12,8-Germacranolides and furanogermacrane sesquiterpenoids
VS1680 Nor- and homogermacrane sesquiterpenoids
VS1690 Secogermacrane sesquiterpenoids
VS1700 Cyclogermacrane sesquiterpenoids
VS1710 Lepidozanes and bicyclogermacrane sesquiterpenoids
VS1720 Humulane sesquiterpenoids
VS1730 Caryophyllane sesquiterpenoids
VS1735 Caryolane sesquiterpenoids
VS1740 Bicyclohumulane sesquiterpenoids
VS1745 Dunniane sesquiterpenoids
VS1750 Cuparane sesquiterpenoids
VS1760 Cyclolaurane sesquiterpenoids
VS1800 Herbertane sesquiterpenoids
VS1850 Laurane sesquiterpenoids
VS1900 Trichothecane sesquiterpenoids
VS1950 Simple eudesmane sesquiterpenoids
VS1970 12,6-Eudesmanolide sesquiterpenoids
VS1975 12,8-Eudesmanolides and furanoeudesmane sesquiterpenoids
VS1980 Agarofuran eudesmane sesquiterpenoids
VS1990 Secoeudesmane sesquiterpenoids
VS2000 Noreudesmane sesquiterpenoids
VS2010 Emmotin sesquiterpenoids
VS2020 Oppositane sesquiterpenoids
VS2040 Farfugin sesquiterpenoids
VS2050 Cycloeudesmane sesquiterpenoids
VS2060 Gorgonane sesquiterpenoids
VS2100 Simple eremophilane sesquiterpenoids
VS2110 Eremophilanolide and furanoeremophilane sesquiterpenoids
VS2120 Seco- and abeoeremophilane sesquiterpenoids
VS2130 Noreremophilane sesquiterpenoids
VS2140 Chiloscyphane sesquiterpenoids
VS2150 Aristolane sesquiterpenoids
VS2160 Nardosinane sesquiterpenoids
VS2170 Brasilane sesquiterpenoids
VS2180 Cacalol sesquiterpenoids
VS2200 Valerane sesquiterpenoids
VS2220 Miscellaneous rearranged eudesmane sesquiterpenoids
VS2225 Iphionane sesquiterpenoids
VS2250 Cadinane sesquiterpenoids
VS2260 Nor- and secocadinane sesquiterpenoids
VS2270 Alliacane sesquiterpenoids
VS2280 Oplopane sesquiterpenoids
VS2290 Mutisianthol sesquiterpenoids
VS2300 Drimane sesquiterpenoids
VS2310 Coloratane sesquiterpenoids
VS2320 Nor- and secodrimane sesquiterpenoids
VS2380 Xanthane sesquiterpenoids
VS2390 Carabrane sesquiterpenoids
VS2400 Simple guaiane sesquiterpenoids
VS2410 12,6-Guaianolide sesquiterpenoids
VS2420 12,8-Guaianolide sesquiterpenoids

VS2430 Dimeric guiaiane sesquiterpenoids
VS2440 Seco-, cyclo-, abeo- and norguaiane sesquiterpenoids
VS2450 Pseudoguaiane sesquiterpenoids
VS2470 Seco-, cyclo-, abeo- and norpseudoguaiane sesquiterpenoids
VS2500 Aromadendrane sesquiterpenoids
VS2520 Cycloaromadendrane sesquiterpenoids
VS2540 Secoaromadendrane sesquiterpenoids
VS2600 Cubebane sesquiterpenoids
VS2620 Ivaxillarane sesquiterpenoids
VS2650 Patchoulane sesquiterpenoids
VS2660 Rearranged patchoulane sesquiterpenoids
VS2710 Valerenane sesquiterpenoids
VS2750 Africanane sesquiterpenoids
VS2760 Lippifoliane sesquiterpenoids
VS2780 Himachalane sesquiterpenoids
VS2790 Allohimachalane sesquiterpenoids
VS2800 Longipinane sesquiterpenoids
VS2850 Longifolane sesquiterpenoids
VS2900 Longibornane sesquiterpenoids
VS3000 Pinguisane sesquiterpenoids
VS3050 Thapsane sesquiterpenoids
VS3080 Fukinane sesquiterpenoids
VS3100 Picrotoxane sesquiterpenoids
VS3180 Daucane sesquiterpenoids
VS3190 Isodaucane sesquiterpenoids
VS3200 Perforane sesquiterpenoids
VS3350 Pacifigorgiane sesquiterpenoids
VS3380 Asteriscane sesquiterpenoids
VS3400 Illudane sesquiterpenoids
VS3420 Protoilludane sesquiterpenoids
VS3430 Sterpurane sesquiterpenoids
VS3440 Illudalane sesquiterpenoids
VS3470 Isolactarane sesquiterpenoids
VS3475 Merulane sesquiterpenoids
VS3480 Lactarane sesquiterpenoids
VS3490 Tremulane sesquiterpenoids
VS3500 Marasmane sesquiterpenoids
VS3550 Furodysin sesquiterpenoids
VS3560 Furodysinin sesquiterpenoids
VS3600 Botrydial sesquiterpenoids
VS3700 Spirovetivane sesquiterpenoids
VS3750 Acorane sesquiterpenoids
VS3800 Chamigrane sesquiterpenoids
VS3810 Secochamigrane sesquiterpenoids
VS3850 Miscellaneous spirosesquiterpenoids
VS3900 Cedrane sesquiterpenoids
VS3920 Isocedrane sesquiterpenoids
VS3950 Zizaane sesquiterpenoids
VS3960 Prezizaane sesquiterpenoids
VS4000 Clovane sesquiterpenoids
VS4200 Precapnellane sesquiterpenoids

VS4250 Capnellane sesquiterpenoids
VS4260 Rearranged capnellane sesquiterpenoids
VS4300 Hirsutane sesquiterpenoids
VS4310 Rearranged hirsutane sesquiterpenoids
VS4400 Pentalenane sesquiterpenoids
VS4450 Silphinane sesquiterpenoids
VS4460 Silphiperfoliane sesquiterpenoids
VS4470 Presilphiperfoliane sesquiterpenoids
VS4500 Isocomane sesquiterpenoids
VS4630 Panasinsane sesquiterpenoids
VS4700 Modhephane sesquiterpenoids
VS4750 Quadrane sesquiterpenoids
VS4770 Campherenane sesquiterpenoids
VS4780 α-Santalane sesquiterpenoids
VS4790 β-Santalane sesquiterpenoids
VS4800 Sativane sesquiterpenoids
VS4820 Copacamphane sesquiterpenoids
VS4850 Sinularane sesquiterpenoids
VS4960 Copaane sesquiterpenoids
VS5000 Ishwarane sesquiterpenoids
VS5020 Rotundane sesquiterpenoids
VS5040 Thujopsane sesquiterpenoids
VS5050 Bourbonane sesquiterpenoids
VS5070 Gymnomitrane sesquiterpenoids
VS5090 Miscellaneous monocyclic sesquiterpenoids
VS5100 Miscellaneous bicyclic sesquiterpenoids
VS5200 Miscellaneous tricyclic sesquiterpenoids
VS5320 Tetracyclic sesquiterpenoids
VS5350 Phytane diterpenoids
VS5360 Irregular linear diterpenoids
VS5370 Linear homo- and nor- diterpenoids
VS5380 Prenylbisabolane diterpenoids
VS5390 10,15-Cyclophytane diterpenoids
VS5400 Labdane diterpenoids
VS5430 Secolabdane diterpenoids
VS5450 Norlabdane diterpenoids
VS5460 Halimane diterpenoids
VS5470 Rearranged labdane diterpenoids
VS5480 Colensane diterpenoids
VS5500 Clerodane diterpenoids
VS5530 Nor-, seco- and abeoclerodane diterpenoids
VS5550 Abietane diterpenoids
VS5560 Furanoabietane diterpenoids
VS5565 Icetexane diterpenoids
VS5570 Secoabietanes and secofriedoabietane diterpenoids
VS5580 Nor- and homoabietane diterpenoids
VS5590 Abeoabietane diterpenoids
VS5600 Dimeric abietane diterpenoids
VS5620 13,16-Cycloabietane diterpenoids
VS5630 17(15$\rightarrow$16)-Abeoabietane diterpenoids
VS5650 Totarane diterpenoids

VS5660 Nagilactone diterpenoids
VS5700 Pimarane diterpenoids
VS5710 Rosane diterpenoids
VS5720 Erythroxylane diterpenoids
VS5730 Paraguerane diterpenoids
VS5735 Isoparguerane diterpenoids
VS5740 Devadarane diterpenoids
VS5750 Isopimarane diterpenoids
VS5760 Rearranged pimarane and isopimarane diterpenoids
VS5770 Norpimarane and norisopimarane diterpenoids
VS5800 Cassane and vouacapane diterpenoids
VS5850 Cleistanthane and isocleistanthane diterpenoids
VS5950 Isocopalane and spongiane diterpenoids
VS5960 Seco-, nor and abeospongiane diterpenoids
VS5980 Podocarpane diterpenoids
VS6000 Kaurane diterpenoids
VS6010 Phyllocladane diterpenoids
VS6020 Norkaurane diterpenoids
VS6030 Secokaurane diterpenoids
VS6040 Rearranged kaurane diterpenoids
VS6050 Beyerane diterpenoids
VS6060 Nor- and secobeyerane diterpenoids
VS6080 Villanovane diterpenoids
VS6090 Thyrsiflorane diterpenoids
VS6100 Atisane diterpenoids
VS6150 Trachylobane diterpenoids
VS6160 Helifulvane diterpenoids
VS6180 Aphidicolane diterpenoids
VS6200 Gibberellins
VS6210 Rearranged gibberellins
VS6220 Leucothol diterpenoids
VS6300 Grayanotoxane diterpenoids
VS6400 Cembrane diterpenoids
VS6410 Norcembrane diterpenoids
VS6420 Rearranged cembrane diterpenoids
VS6440 Eunicellane diterpenoids
VS6450 Asbestinane diterpenoids
VS6460 Sphaerane diterpenoids
VS6470 Briarane diterpenoids
VS6500 Dolabellane diterpenoids
VS6510 Modified dolabellane diterpenoids
VS6540 Dolastane diterpenoids
VS6550 Modified dolastane diterpenoids
VS6560 Cyathane diterpenoids
VS6570 Sphaeroane diterpenoids
VS6580 Verrucosane diterpenoids
VS6590 Modified verrucosane diterpenoids
VS6600 Casbane diterpenoids
VS6610 Jatrophane diterpenoids
VS6615 Euperfoliane diterpenoids
VS6620 9,13-Cyclojatrophane diterpenoids

VS6621 Presegetane diterpenoids
VS6622 Segetane diterpenoids
VS6624 Pepluane diterpenoids
VS6625 Paraliane diterpenoids
VS6650 Lathyrane diterpenoids
VS6660 Rhamnofolane diterpenoids
VS6680 Daphnane diterpenoids
VS6700 Tigliane diterpenoids
VS6710 Ingenane diterpenoids
VS6720 Jatropholane and secojatropholane diterpenoids
VS6750 Fusicoccane diterpenoids
VS6770 Valparane diterpenoids
VS6780 Mulinane diterpenoids
VS6800 Spatane diterpenoids
VS6810 Seco- and abeospatane diterpenoids
VS6880 Verticillane diterpenoids
VS6900 Taxane diterpenoids
VS6950 11(15→1)-Abeotaxane diterpenoids
VS7000 Trinervitane diterpenoids
VS7010 Kempane diterpenoids
VS7020 Amphilectane diterpenoids
VS7030 Cycloamphilectane diterpenoids
VS7040 Adociane diterpenoids
VS7100 Xenicane diterpenoids
VS7110 Nor-, seco- and cycloxenicane diterpenoids
VS7150 Xeniaphyllane diterpenoids
VS7160 Viscidane diterpenoids
VS7180 Eremane diterpenoids
VS7190 Prenyleudesmane diterpenoids
VS7200 Prenylgermacrane diterpenoids
VS7210 Prenylbicyclogermacrane diterpenoids
VS7220 Lobane diterpenoids
VS7230 Pachydictyane diterpenoids
VS7240 Cneorubine diterpenoids
VS7250 Serrulatane and biflorane diterpenoids
VS7260 Decipiane diterpenoids
VS7270 Sacculatane diterpenoids
VS7280 Obtusane diterpenoids
VS7290 Irieol diterpenoids
VS7300 Sphenolobane diterpenoids
VS7310 Miscellaneous monocyclic diterpenoids
VS7320 Miscellaneous bicyclic diterpenoids
VS7330 Miscellaneous tricyclic diterpenoids
VS7340 Miscellaneous tetracyclic diterpenoids
VS7400 Acyclic sesterterpenoids
VS7410 Noracyclic sesterterpenoids
VS7420 Cyclohexane sesterterpenoids
VS7440 Cericerane sesterterpenoids
VS7460 Bicyclic sesterterpenoids
VS7500 Cheilanthane sesterterpenoids
VS7520 Ophiobolane sesterterpenoids

VS7540 Scalarane sesterterpenoids
VS7550 Methyl- and dimethylscalarane sesterterpenoids
VS7580 Miscellaneous sesterterpenoids
VS7600 Linear triterpenoids
VS7620 Botryococcene triterpenoids
VS7700 Protostane and fusidane triterpenoids
VS7750 Lanostane triterpenoids
VS7800 Cycloartane triterpenoids
VS7900 Cucurbitane triterpenoids
VS7950 Dammarane triterpenoids
VS8000 Tirucallane/euphane triterpenoids
VS8050 Apotirucallane triterpenoids
VS8100 Intact tetranortriterpenoids
VS8120 Ring cleaved tetranortriterpenoids
VS8130 Rearranged tetranortriterpenoids
VS8200 Quassinoid nortriterpenoids
VS8205 C_{25} Quassinoid triterpenoids
VS8230 Baccharane triterpenoids
VS8250 Lupane triterpenoids
VS8260 3(2→1)-Abeolupane triterpenoids
VS8270 Nor-, friedo- and secolupane triterpenoids
VS8300 Oleanane triterpenoids
VS8310 Nor-, seco- and abeooleanane triterpenoids
VS8350 Taraxerane triterpenoids
VS8360 Nor-, seco- and cyclotaraxerane triterpenoids
VS8400 Multiflorane triterpenoids
VS8450 Glutinane triterpenoids
VS8500 Friedelane triterpenoids
VS8510 Nor- and secofriedelane triterpenoids
VS8520 Pachysanane triterpenoids
VS8550 Taraxastane triterpenoids
VS8650 Ursane triterpenoids
VS8660 Nor-, seco- and abeoursane triterpenoids
VS8700 Bauerane triterpenoids
VS8720 Hopane triterpenoids
VS8730 Nor-, seco- and abeohopane triterpenoids
VS8760 Moretane triterpenoids
VS8770 Neohopane triterpenoids
VS8800 Fernane triterpenoids
VS8850 Adianane triterpenoids
VS8870 Filicane triterpenoids
VS8880 Arborinane triterpenoids
VS8900 Stictane triterpenoids
VS8950 Gammacerane triterpenoids
VS9000 Serratane triterpenoids
VS9050 Onocerane triterpenoids
VS9080 Polypodane triterpenoids
VS9100 Malabaricane and isomalabaricane triterpenoids
VS9300 Miscellaneous triterpenoids
VS9350 Iridal group norterpenoids
VS9400 Tetraterpenoids

VS9450 Megastigmane norterpenoids
VS9700 Apocarotenoids
VS9750 Norterpenoid tobacco constituents
VS9800 Polyterpenoids
VS9900 Meroterpenoids
VS9910 Hop meroterpenoids
VS9999 Terpenoids of unknown structure

Abbreviations

General

The following is a selection of the most common Database Abbreviations.

Abbreviation	Name
[α]	specific rotation
acac	acetylacetonato
Ac	acetyl
ACGIH	American Conference of Governmental Industrial Hygienists
Ac_2O	acetic anhydride
AcOH	acetic acid
ADI	Acceptable Daily Intake
alk.	alkaline
amorph.	amorphous
ANSI	American National Standards Institute
anhyd.	anhydrous
approx.	approximately
aq.	aqueous
asym.	asymmetrical, unsymmetrical
B	base
BAN	British Approved Name
biol.	biological
bipy	2,2'-bipyridine
BOC	tert-Butyloxycarbonyl
Bp	boiling point
br	broad
BSI	British Standards Institution
Bu	butyl
bwd	bird (wild)
Bz	benzyl
c.	concentration
ca.	(circa) about
CAS	Chemical Abstracts Service
ccp	cubic close packed
cdt	1,5,9-cyclododecatriene
C_6H_6	benzene
C_5Me_5	pentamethylcyclopentadienyl
CNS	central nervous system
cod	1,5-cyclooctadiene
col.	colour, coloration
comly.	commercially
compd(s)	compounds(s)
conc.	concentrated
const.	constant
constit.	constituent
coord	coordinate(d), coordination
cot	1,3,5,7-cyclooctatetraene

Cp	cyclopentadienyl
C_5Ph_5	pentaphenylcyclopentadienyl
cryst.	crystal(s)
cv	cultivar
CVD	chemical vapour deposition
Cy	cyclohexyl
d	density
dba	dibenzylideneacetone
dck	duck
dec.	decomposes, decomposition
degradn.	degradation
depe	1,2-bis(diethylphosphino)ethane
descr.	described
diars	diarsine (generalised ligand)
dil.	dilute, dilution
dimorph.	dimorphic
diphos	diphosphine (generalised ligand)
diss.	dissolves, dissolved
dissoc.	dissociates
dist.	distil, distillation
DMA	dimethylacetamide
DMF	dimethylformamide
dmpe	1,2-bis(dimethylphosphino)ethane
dmpm	bis(dimethylphosphino)methane
DMSO	dimethyl sulfoxide
dppe	1,2-bis(diphenylphosphino)ethane
dppm	bis(diphenylphosphino)methane
dppp	1,3-bis(diphenylphosphino)propane
EDTA	ethylenediaminetetracetate(4-)
ee	enantiomeric excess
Eg	band gap (electron volts)
en	ethylenediamine
equilib.	equilibrium
esp.	especially
Et	ethyl
EtOAc	ethyl acetate
EtOH	ethanol
EtOH aq.	aqueous ethanol
evapn.	evaporation
exp.	exposure
exp.	experimental
fac	facial
Fc	ferrocenyl
fl. p.	flash point
fluor.	fluoresces, fluorescence
formn.	formation
Fp	freezing point
g	gram(s)
$\Delta G0_f$	standard free energy of formation
Glc	β-D-glucopyranosyl
gpg	guinea pig

ham	hamster
$\Delta H0_f$	standard enthalpy of formation
hcp	hexagonal close packed
hydrol.	hydrolyses, hydrolysed, hydrolysis
ihl	inhalation
im	imidazolato
ims	intramuscular
INN	International Non-proprietary Name
inorg.	inorganic
insol.	insoluble
intermed.	intermediate
ipr	intraperitoneal
ISO	International Standards Organisation
ivg	intravaginal
ivn	intravenous
JAN	Japanese Accepted Name
JMAF	Japanese Ministry for Agriculture, Forestry and Fisheries
K	temperature (Kelvin)
L	generalised ligand
LC	lethal concentration
LD	Lethal dose; LD_{50} : a dose which is lethal to 50% of the animals tested
M	relative molecular mass (formula weight)
M	metal
m	medium
mcd	magnetic circular dichroism
Me	methyl
MEL	maximum exposure limit
MeOH	methanol
mer	meridional
mes	mesityl (1,3,5-trimethylphenyl)
Me_2CO	acetone
misc.	miscible
misc.	miscellaneous
mixt.	mixture
mky	monkey
MOCVD	metal-organic chemical vapour deposition
mod.	moderately
Mp	melting point
mus	mouse
n	index of refraction (e.g. *n*D20 for 20° and sodium light)
nbd	norbornadiene
nqr	nuclear quadrupole resonance spectrum
obt.	obtained
oc	open cup
oep	octaethylporphyrinato
OES	occupational exposure standard
O_h	octahedral
op	optical purity
org.	organic
orl	oral
ox	oxalato

Ph	phenyl (C_6H_5)
pH	Measure of soln. acidity where pH = $\log_{10}$ (1/[H+]) where [H+] is the hydrogen ion concentration
phen	1,10-phenanthroline
phys.	physical
pK	Measure of dissoc. const. (K) where pK= Log_{10} (1 /K)
pm	picometres
PMDET	pentamethyldiethylenetriamine
polarog.	polarography
polym.	polymerised, polymerisation
ppm	parts per million
Pr	propyl
prob.	probably
purifn.	purification
Py	pyridine
pz	pyrazolato
R	generalised alkyl group
rbt	rabbit
ref.	reference
rel.	relative(ly)
r.t.	room temperature
s	strong
S0	standard entropy
scu	subcutaneous
skn	skin
sl.	slightly
sol.	soluble
soln(s)	solution(s)
solv(s)	solvent(s)
soly.	solubility
sp.	species (singular)
spar.	sparingly
spp.	species (plural)
ssp.	subspecies
subl.	sublimation, sublimes
tbp	trigonal bipyramidal
T_d	tetrahedral
Tf	triflate
THF	tetrahydrofuran
tht	tetrahydrothiophene
TLV	Threshold Limit Value
TMED	tetramethylethylenediamine
tpp	tetraphenylporphyrinato
triphos	triphosphine (generalised ligand)
Ts	tosyl
μeff	effective magnetic moment (in Bohr magnetons μB)
unsatd.	unsaturated
USAN	United States Adopted Name
uv	ultraviolet spectrum
v.	very
var.	variety

vis.	visible
vol.	volume
w	weak
WSSA	Weed Science Society of America
X	generalised anion, usually halide
Z	Benzyloxycarbonyl

Reference Tags

The following is a selection of the most common Reference Tags used.

Abbreviation	Name
abs config	absolute configuration
anal	analysis
bibl	bibliography
biodistribn	biodistribution
biosynth	biosynthesis
cd	circular dichroism
chromatog	chromatography
cmr	^{13}C nuclear magnetic resonance spectrum
config	configuration
conformn	conformation
cryst struct	X-ray crystal structure determination
deriv(s)	derivative(s)
detn	determination, detection
dsc	differential scanning calorimetry
dta	differential thermal analysis
ed	electron diffraction
electrochem	electrochemistry, cyclic voltammetry
em	electron microscopy
epr	electron paramagnetic (spin) resonance spectrum
esca	electron spectroscopy for chemical analysis
exafs	extended X-ray diffraction fine structure
fab-ms	fast atom bombardment mass spectroscopy
glc	gas-liquid chromatography
haz	hazard
hist	historically significant publication
hplc	high performance liquid chromatography
ir	infrared spectrum
isol	isolation
isom	isomerism
manuf	manufacture
metab	metabolism
mineral	mineralogy
ms	mass spectrum
nmr	nuclear magnetic resonance spectrum
occur	occurrence
ord	optical rotatory dispersion
pe	photoelectron spectroscopy
pharmacol	pharmacology
photol	photolysis
pmr	proton (^{1}H) nuclear magnetic resonance spectrum
polarog	polarography
powder struct	X-ray powder structure determination
props	properties (chemical or physical)
Raman	Raman spectrum
resoln	resolution

rev	review
sepn	separation
soly	solubility
spectra	
struct	structure
synonyms	
synth	synthesis
tautom	tautomerism
tga	thermogravimetric analysis
theory	MO calculations etc.
tlc	thin layer chromatography
tox	toxicity
trans	transition(s)
use(s)	
uv	ultraviolet spectrum
uv-vis	ultraviolet visible spectrum

Amino Acids

α-Amino acids commonly found in peptides and proteins

Abbreviation		Name
Ala	A	alanine
Arg	R	arginine
Asn	N	asparagine
Asp	D	aspartic acid
Cys	C	cysteine
Glu	E	glutamic acid
Gln	Q	Glutamine
Gly	G	glycine
His	H	histidine
Ile	I	isoleucine
Leu	L	leucine
Lys	K	lysine
Met	M	methionine
Phe	F	phenylalanine
Pro	P	proline
Ser	S	serine
Thr	T	threonine
Trp	W	tryptophan
Tyr	Y	tyrosine
Val	V	valine

Other Amino Acids

Abbreviation	Name
βAad	3-aminoadipic acid
Aad	2-aminoadipic acid
A2bu	2,4-diaminobutyric acid
Abu	2-aminobutanoic acid
εAhx	6-aminohexanoic acid
Ahx	2-aminohexanoic acid (norleucine)
Aib	α–aminoisobutyric acid
2-MeAla	2-methylalanine
bAla	b-alanine
Ape	2-aminopentanoic acid (norvaline)
A2pm	2,6-diaminopimelic acid
Apm	2-aminopimelic acid
A2pr	2,3-diaminopropionic acid
Asp(NH_2)	asparagine
Asx	asparagine or aspartic acid
Avl	2-aminopentanoic acid (norvaline)
Cit	citrulline
Cya	cysteic acid
Dab	2,4-diaminobutyric acid
Dpm	2,6-diaminopimelic acid
Dpr	2,3-diaminopropionic acid
Gla	4-carboxyglutamic acid
Glp	5-oxoproline (pyroglutamic acid)
pGlu	5-oxoproline (pyroglutamic acid)
<Glu	5-oxoproline (pyroglutamic acid)
Glu(NH_2)	glutamine
Glx	glutamine or glutamic acid
Hcy	homocysteine
Hse	homoserine
Hsl	homoserine lactone
Hyl	5-hydroxylysine
5Hyl	5-hydroxylysine
Hyp	4-hydroxyproline
4Hyp	4-hydroxyproline
aIle	alloisoleucine
alloIle	alloisoleucine
Iva	isovaline
Met(O)	methionine *S*-oxide
MetO	methionine *S*-oxide
$MetO_2$	methionine *S.S*-dioxide
Mur	muramic acid
Neu	neuraminic acid
Neu5Ac	*N*-acetylneuraminic acid
Nle	norleucine
Nva	norvaline
Orn	ornithine

5-oxo-Pro	5-oxoproline (pyroglutamic acid)
Sar	sarcosine
Ser(P)	phosphoserine
alloThr	allothreonine
aThr	allothreonine
Thx	thyroxine
Tyr(I_2)	3,5-diiodotyrosine
Tyr(SO_3H)	*O*4-sulfotyrosine
Xaa	unspecified amino acid

Frequently Asked Questions

Q. Whenever I open the electronic version of this help file, my computer tries dialling an internet connection.

A. This help text contains links to web sites which can cause your modem to dial if it is set to dial-up automatically. To resolve, set your modem to use manual dialling. (From this help text, click: Tools > Internet Connection, click on the 'Connections' tab and select 'Never dial a connection').

Q. Can I install this copy of DNP over a previous version?

A. We strongly recommend deleting your old copy first. This software contains software updates as well as updates to the data, so deleting previous versions first will avoid potential conflicts.

Q. How do I uninstall DNP?

A. There is no specific uninstall program for DNP. To remove the software from your computer simply delete the installation directory (c:\Program Files\CRC Press\Dictionaries if you accepted the default directory during installation) via Windows Explorer. Take care when deleting directories from your computer. If you accidentally delete the wrong files you can restore them from the recycle bin.

Q. Structure searching takes longer than expected.

A. Substructure searching for small fragments may take longer than expected. However, if you are finding structure searching is generally slow it may be due to your virus scanner. This seems to be especially so if the virus scanner is set to perform 'heuristic' scanning of all files. It may be possible to exclude the files in the installation directory from being scanned, but users are urged to exercise caution. Please consult your antivirus software for more details.

Q. Structures appear on the hit list but they don't print.

A. This is a printer driver problem affecting Windows XP and some HP PCL6 printer drivers. This problem can be circumvented by using PostScript printer drivers instead of the usual PCL6 drivers. Please consult your printer's documentation for more information.

Q. I can only see the entry and not the hit list (or vice versa).

A. You have maximised the windows. Either use the window toolbar buttons to switch between windows or restore the windows using the restore button.

Dictionary of Natural Products on CD-ROM

This introduction screen gives access to (*a*) a general introduction to the scope and content of DNP on CD-ROM, followed by (*b*) an extensive review of the different types of natural product and the way in which they are organised and categorised in DNP.

You may access the section of your choice by clicking on the appropriate line below, or you may scroll through the text forwards or backwards from any point.

Introduction to the DNP database *page* 3

Data presentation and organisation **3**
- Derivatives and variants 3
- Chemical names and synonyms 4
- CAS Registry Numbers 6
- Diagrams 7
- Stereochemical conventions 7
- Molecular formula and molecular weight 8
- Source 9
- Importance/use 9
- Type of Compound 9
- Physical Data 9
- Hazard and toxicity information 10
- Bibliographic References 11
- Journal abbreviations 12
- Entry under review 12

Description of Natural Product Structures 13

Aliphatic natural products **15**
- Semiochemicals 15
- Lipids 22

Polyketides **29**

Carbohydrates **35**

Oxygen heterocycles **44**

Simple aromatic natural products **45**

Benzofuranoids **48**

Benzopyranoids **49**

Flavonoids *page* **51**

Tannins **60**

Lignans **64**

Polycyclic aromatic natural products **68**

Terpenoids **72**
Monoterpenoids 73
Sesquiterpenoids 77
Diterpenoids 101
Sesterterpenoids 118
Triterpenoids 121
Tetraterpenoids 131
Miscellaneous terpenoids 133
Meroterpenoids 133

Steroids **135**
The sterols 140

Aminoacids and peptides **148**
Aminoacids 148
Peptides 150
β-Lactams 151
Glycopeptides 153

Alkaloids **154**
Alkaloids derived from ornithine 154
Alkaloids derived from lysine 156
Alkaloids derived from nicotinic acid 158
Alkaloids of polyketide origin 159
Alkaloids derived from anthranilic acid 160
Alkaloids derived wholly or in part from phenylalanine or tyrosine 163
Isoquinoline alkaloids 163
Alkaloids derived from tryptophan 181
Monoterpenoid indole alkaloids 184
Terpenoid alkaloids 196
Steroidal alkaloids 200
Imidazole alkaloids 202
Oxazole alkaloids 202
Thiazole alkaloids 203
Pyrazine and quinoxaline alkaloids 203
Pyrrole alkaloids 203
Putrescine alkaloids 203
Spermine and spermidine alkaloids 203
Peptide alkaloids 204
Purines 205
Pteridines and analogues 205

Polypyrroles **206**

Introduction to the DNP database

The Chapman & Hall/CRC Chemical Database is a structured database holding information on chemical substances. It includes descriptive and numerical data on chemical, physical and biological properties of compounds; systematic and common names of compounds; literature references; structure diagrams and their associated connection tables. The ***Dictionary of Natural Products on CD-ROM*** is a subset of this database and includes all compounds contained in the *Dictionary of Natural Products* (Main Work and Supplements).

The *Dictionary of Natural Products* (DNP) is the only comprehensive and fully-edited database on natural products. It arose as a daughter product of the well-known *Dictionary of Organic Compounds* (DOC) which, since its inception in the 1930s has, through successive editions, always been a leading source of natural product information.

In the early 1980s, following the publication of the Fifth Edition of DOC, the first to be founded on database methods, the Editors and contributors for the various classes of natural products embarked on a programme of enlargement, rationalisation and classification of the natural product entries, while at the same time keeping the coverage up-to-date. In 1992 the results of this major project, which had grown to match DOC in size, were separately published in both book (7 volumes) and CD-ROM format, leaving DOC with coverage of only the most widely distributed and/or practically important natural products. DNP compilation has since continued unabated by a combination of an exhaustive survey of current literature and of historical sources such as reviews to pick up minor natural products and items of data previously overlooked.

The compilation of DNP is undertaken by a team of academics and freelancers who work closely with the in-house editorial staff at Chapman & Hall. Each contributor specialises in a particular natural product class (e.g. alkaloids) and is able to reorganise and classify the data in the light of new research so as to present it in the most consistent and logical manner possible. Thus the compilation team is able to reconcile errors and inconsistencies.

The resulting CD-ROM version, which is re-issued every six months, represents an extremely well organised dictionary documenting virtually every known natural product.

A valuable feature of the design is that closely related natural products (e.g. where one is a glycoside or simple ester of another) are organised into the same entry, thus simplifying and bringing out the underlying structural and biosynthetic relationships of the compounds. Structure diagrams are drawn and numbered in the most consistent way according to best stereochemical and biogenetic relationships. In addition, every natural product is indexed by structural/biogenetic type under one of more than 1000 headings, allowing the rapid location of all compounds in the category, even where they have undergone biogenetic modification and no longer share exactly the same skeleton.

There is extensive (but not complete) coverage of natural products of unknown structure, and the coverage of these is currently being enhanced by various retrospective searches.

Data presentation and organisation

Derivatives and variants

In the database, closely related compounds are grouped together to form an ***entry***. Stereoisomers and derivatives of a parent compound are all listed under one entry. The compounds in the ***Dictionary of Natural Products*** are grouped

together into approximately 40,000 entries. The structure of an entry is shown below.

Entry (parent compound)
Derivatives
Variants (stereoisomers or other closely-related compounds)
Derivatives of the variant

A simple entry covers one compound, with no derivatives or variants. A composite entry will start with the entry compound, then may have:

one or more derivatives at entry level
one or more variants of the entry
one or more derivatives of the variant.

Variants may include stereoisomers, e.g. (*R*)-form, *endo*-form; members of a series of natural products with closely related structures such as antibiotic complexes.

For example, Trienomycins are often treated as variants although their structures may be more varied.

Derivatives may include hydrates, complexes, salts, classical organic derivatives, substitution products and oxidation products etc. Derivatives may exist on more than one functional group of an entry compound.

The following techniques are among those used to bring together related substances in the same entry:

(*a*) **Glycosides** are given as derivatives of the parent aglycone, except for those glycosides which have an extensive literature in their own right (e.g., Digoxin)

(*b*) **Acyl derivatives** are extremely common and are listed under the parent compound, again unless it has as extensive literature of its own

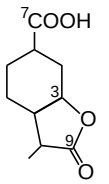
(*c*) **N-Alkyl and O-Alkyl derivatives** such as methyl ethers of phenols are similarly given under the parent compound.

Data Types

The format of a typical entry is given in Fig. 1, and shows the individual types of data that may be present in an entry.

Chemical names and synonyms

All the names discussed below can be searched using the Chemical Name field. Compounds have been named so as to facilitate access to their factual data by keeping the nomenclature as simple as possible, whilst still adhering to good practice as determined by IUPAC (the International Union of Pure and Applied Chemistry). A great deal of care has been taken to achieve this aim as nearly as possible. Some intentional departures from IUPAC terminological principles are occasionally made to clarify the nomenclature of natural products. For example, compounds containing both lactone and –COOH groups are often named using two principal functional groups:



p-Menthan-9,3-olid-7-oic acid

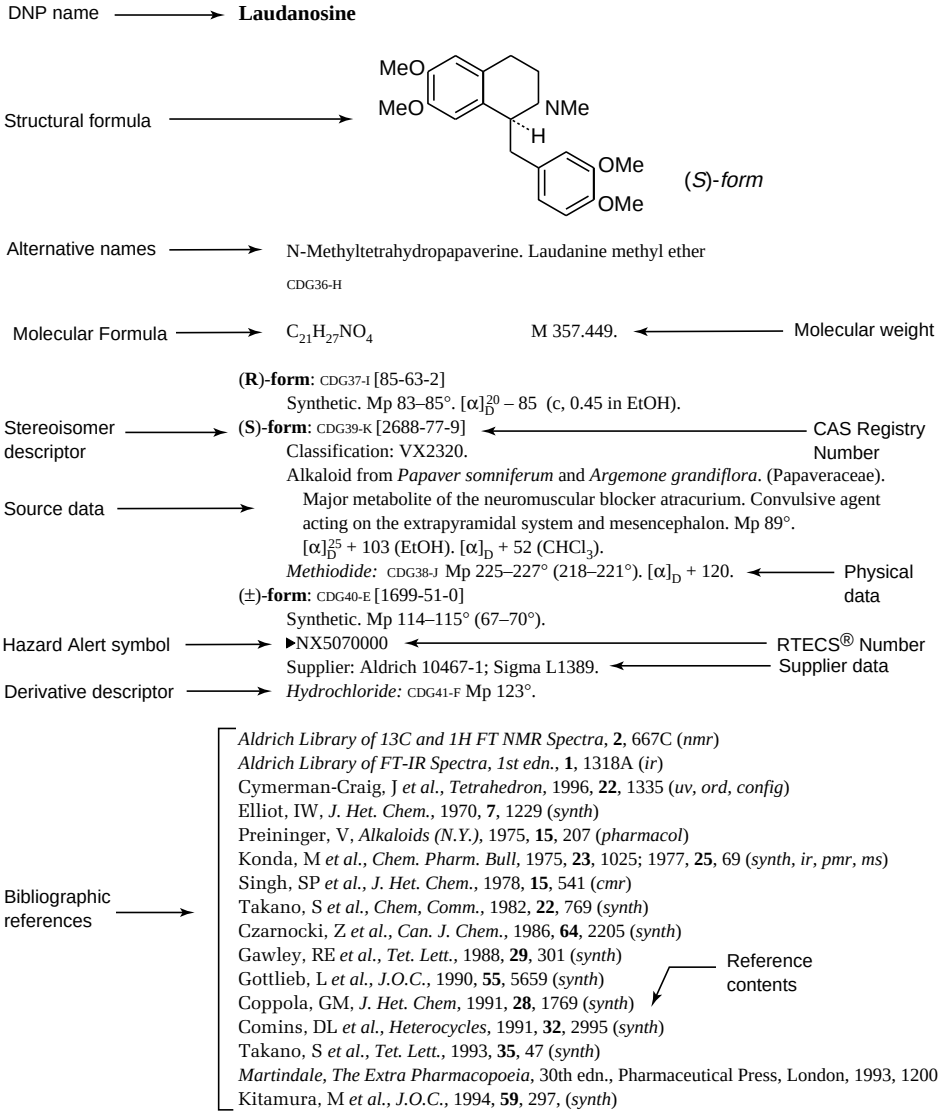
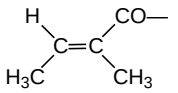


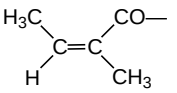
Fig. 1. Sample entry from database

(*a*) There are many instances in the primary literature of compounds being named in ways which are gross violations of good IUPAC practice, e.g., where the substituents are ordered non-alphabetically. These have been corrected.

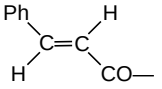
(*b*) The number of trivial names used for acylating substituents has been kept to a minimum but the following are used throughout.



Tigloyl = *E*-2-methyl-2-butenoyl



Angeloyl = *Z*-2-methyl-2-butenoyl



Cinnamoyl = 3-phenyl-2-propenoyl

Many other trivial appellations have from time to time appeared in the literature for other acyl groups (e.g., Senecioyl = 3-methyl-2-butenoyl,

Feruloyl = 3-(4-hydroxy- 3-methoxyphenyl)-2-propenoyl or 4-hydroxy-3-methoxycinnamoyl) but the systematic forms are usually employed except in a few cases where the shortened form is used to abbreviate a very long and unwieldy derivative descriptor as much as possible (e.g., for some of the complex flavonoid glycosides).

(*c*) The term **prenyl** for the common 3-methyl-2-butenyl substituent, $(H_3C)_2C{=}CHCH_2{-}$, is used throughout.

(*d*) Names which are known to be duplicated within the chemical literature (not necessarily within DNP), are marked with the sign†.

CAS Registry Numbers

CAS Registry Numbers are identifying numbers allocated to each distinctly definable chemical substance indexed by the Chemical Abstracts Service since 1965 (plus retrospective allocation of numbers by CAS to compounds from the sixth and seventh collective index periods). The numbers have no chemical significance but they provide a label for each substance independent of any system of nomenclature.

In DNP, much effort has been expended to ensure that accurate CAS numbers are given for as many substances as possible.

If a CAS number is not given for a particular compound, it may be (*a*) because CAS have not allocated one, (*b*) very occasionally, because an editorial decision cannot be made as to the correct number to cite, or (*c*) because the substance was added to the DNP database at a late stage in the compilation process, in which case the number will probably be added to the database soon.

At the foot of the DNP entry, immediately before the references, may be shown additional registry numbers. These are numbers which have been recognised by the DNP editors or contributors as belonging to the entry concerned but which cannot be unequivocally assigned to any of the compounds covered by the entry. Their main use will be in helping those who need to carry out additional searches, especially online searches in the CAS or other databases, and who will be able to obtain additional hits using these numbers. Clearly, discretion is needed in their use for this purpose.

Additional registry numbers may arise for a variety of reasons:

(*a*) A number may refer to stereoisomers or other variants of the main entry compound or its derivatives, which may or may not be mentioned in the entry but for which no physical properties or other useful information is available.

For example, the DNP entry for Carlic acid [56083-49-9] states that it has so far been obtained in solution as a mixture of (*E*) and (*Z*)-forms. The additional registry numbers given are those of the (*E*) and (*Z*) isomers [67381-73-1] and [67381-74-2].

(*b*) A CAS number may refer to a mixture, in which case it is added to the DNP entry referring to the most significant component. It may refer to a hydrate, salt, complex, etc. which is not described in detail in the DNP entry.

(*c*) Replaced numbers, duplicate numbers and other numbers arising from CAS indexing procedure or, occasionally, from errors or inconsistencies by CAS, are also reported. For example, the DNP entry *scyllo*-Inositol [488-59-5] contains an additional registry number for D-*scyllo*-Inositol [41546-32-1]. Since *scyllo*-Inositol is a meso-compound, the number is erroneous. More generally, CAS frequently replace a given number with one that more accurately represents what they now know about a substance, and the replaced number remains on their files and is given in DNP as an additional number.

(*d*) In the case of compounds with more than one stereogenic centre, additional registry numbers frequently refer to levels of stereochemical

description which cannot be assigned to a particular stereoisomer described in the entry.

For example, the CHCD entry for 2-Amino-3-hydroxy-3-phenylpropanoic acid (*β*-Hydroxyphenylalanine, 9CI) has a general CAS number [1078-17-7] and CAS numbers for all four optically active diastereoisomers [7352-06-9, 32946-42-2, 109120-55-0, 6524-48-4] as well as the two possible racemates [2584-74-9] [2584-75-0]. However, among the additional registry numbers quoted are the following:

[7687-36-7] – number for *erythro*-*β*-Hydroxyphenylalanine
[50897-27-3] – number for *β*-Hydroxy-L-phenylalanine
[68296-26-4] – number for *β*-Hydroxy-D-phenylalanine
[39687-93-9] – general number for the methyl ester, hydrochloride which cannot be placed under any of the individual stereoisomers of this compound described in the entry.

(*e*) Numbers may refer to derivatives similar to those described in the DNP entry for which no data is available, or which have not yet been added to the entry.

(*f*) Some DNP entries refer to families of compounds, such as the entry for Calcitonin where only the porcine and human variants are described in detail. The additional registry numbers given in this entry are those of a number of other species variants which appear to have been identified according to CAS but for which no attempt has been made to collate full data for DNP.

Diagrams

In each entry display there is a single diagram which applies to the parent entry. Separate diagrams are not given for variants or derivatives.

Every attempt has been made to present the structures of chemical substances as accurately as possible according to current best practice and IUPAC recommendations. In drawing the formulae, as much consistency as possible between closely related structures has been aimed at. Thus, for example, sugars have been standardised as Haworth formulae and, wherever possible in complex structures, the rings are oriented in the standard Haworth manner so that structural comparisons can quickly be made. In formulae the pseudoatom abbreviations Me, Et and Ac for methyl, ethyl and acetyl respectively, are used only when attached to a heteroatom. Ph is used throughout whether attached to carbon or to a heteroatom. Other pseudoatom abbreviations such as Pr^{i} for isopropyl and Bz for benzoyl are not used in DNP.

Care must be taken with the numbering of natural products, as problems may arise due to differences in systematic and non-systematic schemes. Biogenetic numbering schemes which are generally favoured in DNP may not always be contiguous, e.g., where one or more carbon atoms have been lost during biogenesis.

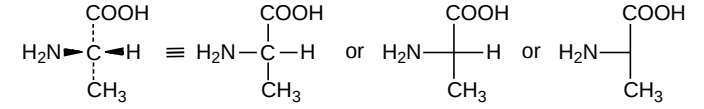
Structures for derivatives can be viewed in **Structure Search**, but remember that these structures are generated from connection tables and may not always be oriented consistently.

Stereochemical conventions

Where the absolute configuration of a compound is known or can be inferred from the published literature without undue difficulty, this is indicated. Where only one stereoisomer is referred to in the text, the structural diagram indicates that stereoisomer. Wherever possible, stereostructures are described using the Cahn-Ingold-Prelog sequence-rule *(R,S)* and *(E,Z)* conventions but, in cases where these are cumbersome or inapplicable, alternatives such as the

α,β-system are used instead. Alternative designations are frequently presented in such cases.

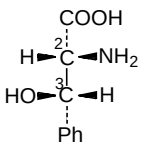
The structure diagrams for compounds containing one or two chiral centres are given in DNP as Fischer-*type* diagrams showing the stereochemistry unequivocally. True Fischer diagrams in which the configuration is implied by the North-South-East-West positions of the substituents are widespread in the literature; they are quite unambiguous but need to be used with caution by the inexperienced. They cannot be reoriented without the risk of introducing errors.



Where only the relative configuration of a compound containing more than one chiral centre is known, the symbols (*R**) and (*S**) are used, the lowest-numbered chiral centre being arbitrarily assigned the symbol (*R**).

For racemic modifications of compounds containing more than one chiral centre the symbols (*RS*) and (*SR*) are used, with the lowest-numbered chiral centre being arbitrarily assigned the symbol (*RS*). The racemate of a compound containing one chiral centre only is described in DNP as (±)-.

In comparing CAS descriptors with those given in DNP, it is important to remember that the order of presentation of the chirality labels in CAS is itself based on the sequence rule priority and not on any numbering scheme, for example the CAS descriptor for the structure illustrated is [*S*-(*R**,*S**)].



The relative stereochemical label (*R**,*S**) is first applied with the *R** *applying to the chiral centre of higher priority* (C-3). The absolute stereochemical descriptor (*S*)- is then applied changing *R** to *S* for the chiral centre of higher priority and *S** to *R* for the chiral centre of lower priority (C-2). For further details, see the current CAS Index Guide.

For simplicity, the enantiomers of bridged-ring compounds, such as camphor, are described simply as (+)- and (–)-. Although camphor has two chiral centres, steric restraints mean that only one pair of enantiomers can be prepared.

For further information on the (*R,S*)-system, see Cahn, R,S *et al*, *J. Chem. Soc*., 1951, 612; *Experientia*, 1956, **12**, 81; *Angew. Chem. Int. Ed. Engl*., 1966, **5**, 383.

Where appropriate, alternative stereochemical descriptors may be given using the D, L or α,β-systems. For a fuller description of these systems, consult *The Organic Chemist's Desk Reference* (Chapman & Hall, 1995).

Molecular formula and molecular weight

The elements in the molecular formula are given according to the Hill convention (C, H, then other elements in alphabetical order). The molecular weights given are formula weights (or more strictly, molar masses in daltons) and are rounded to one place of decimals. In the case of some high molecular mass substances such as proteins the value quoted may be that taken from an original literature source and may be an aggregate molar mass.

Molecular formulae are included in DNP for all derivatives which are natural products and so are readily searchable, whether they are documented as derivatives or have their own individual entry. Molecular formulae are not in

general given for salts, hydrates or complexes (e.g. picrates) nor for most "characterisation" derivatives such as acetates and methyl ethers of complex natural products.

Where a derivative appears to have characterised only as a salt, the properties of the salt may be given under the heading for the derivative. In such cases the data is clearly labelled, e.g., Mp 179° (as hydrochloride).

Source

The taxonomic names for organisms given throughout are in general those given in the primary literature. Standardisation of minor orthographical variations has been carried out. Data in this field may be searched under **Source/Synthesis** or **All Text**. Standards used are: Brummitt, R.K. (1992) *Vascular Plant Families and Genera*, Royal Botanic Gardens, Kew; Willis, J.C. (1973) *A Dictionary of the Flowering Plants*, Cambridge University Press, Cambridge; Gozmany, L. (1990) *Seven Language Thesaurus of European Animals*, Chapman & Hall London; Chemical Abstracts Service.

Importance/use

Care has been taken to make the information given on the importance and uses of chemical substances as accurate as possible. Data in this field may be searched under **Use/Importance** or **All Text**.

Type of Compound

All natural products are classified under one of more than 1050 headings according to structural type, e.g., daucane sesquiterpenoid, pyrrolizidine alkaloid, withanolide. Each structural type is assigned as a type of compound code, e.g., VG0300, VX0150. Type of compound words and type of compound codes may both be searched in Menu and Command search.

The full type of compound code index is given in Table 3, page 128 of the printed User Manual, and in the Description of Natural Product Structures that follows, each descriptive paragraph is followed by its Type of Compound code(s).

Physical Data

Appearance

Natural products are considered to be colourless unless otherwise stated. Where the compound contains a chromophore which would be expected to lead to a visible colour, but no colour is mentioned in the literature, the DNP entry will mention this fact if it has been noticed by the contributor.

An indication of crystal form and of recrystallisation solvent is often given but these are imprecise items of data; most organic compounds can be crystallised from several solvent systems and the crystal form often varies. In the case of the small number of compounds where crystal behaviour has been intensively studied (e.g. pharmaceuticals), it is found that polymorphism is a very common phenomenon and there is no reason to believe that it is not widespread among organic compounds generally.

Melting points and boiling points

The policy followed in the case of conflicting data is as follows:

(*a*) Where the literature melting points are closely similar, only one figure (the highest or most probable) is quoted.

(*b*) Where two or more melting points are recorded and differ by several degrees (the most likely explanation being that one sample was impure), the lower figure is given in parentheses, thus: 139° (134–135°).

(*c*) Where quoted figures differ widely and some other explanation such as polymorphism or incorrect identity seems to be the most likely explanation, both figures are quoted without parentheses, thus Mp 142°, Mp 205–206°.

(*d*) Known cases of polymorphism or double melting point are noted.

Boiling point determination is less precise than that of melting points and conflicting boiling point data is not usually reported except when there appears to be a serious discrepancy between the different authors.

Optical rotations

These are given whenever possible, and normally refer to what the DNP contributor believes to be the best-characterised sample of highest chemical and optical purity. Where available an indication of the optical purity (op) or enantiomeric excess (ee) of the sample measured now follows the specific rotation value.

Specific rotations are dimensionless numbers and the degree sign which was formerly universal in the literature has been discontinued.

Densities and refractive indexes

Densities and refractive indexes are now of less importance for the identification of liquids than has been the case in the past, but are quoted for common or industrially important substances (e.g. monoterpenoids), or where no boiling point can be found in the literature.

Densities and refractive indexes are not quoted where the determination appears to refer to an undefined mixture of stereoisomers.

Solubilities

Solubilities are given only where the solubility is unusual. Typical organic compounds are soluble in the usual organic solvents such as ether and chloroform, and virtually insoluble in water. The presence of polar groups (OH, NH_2 and especially COOH, SO_3H, NR^+) increases water solubility.

pK_a values

pK_a values are given for both acids and bases. The pK_b of a base can be obtained by subtracting its p*K*a from 14.17 (at 20°) or from 14.00 (at 25°).

Spectroscopic data

Spectroscopic data such as uv wavelengths and extinction coefficients are given only where the spectrum is a main point of interest, or where the compound is unstable and has been identified only by spectroscopic data.

In many other cases, spectroscopic data can be rapidly located through the references quoted.

Hazard and toxicity information

General

Toxicity and hazard information is highlighted by the sign ▶, and has been selected to assist in risk assessments for experimental, manufacturing and manipulative procedures with chemicals.

The field of safety testing is a complex, difficult and rapidly expanding one, and while as much care as possible has been taken to ensure the accuracy of reported data, the *Dictionary* must not be considered a comprehensive source on hazard data. The function of the reported hazard data is to alert the user to possible hazards associated with the use of a particular compound, but the absence of such data cannot be taken as an indication of safety in use, and the Publishers cannot be held responsible for any inaccuracies in the reported information, neither does the omission of hazard data in DNP imply an absence of this data from the literature. Widely recognised hazards are included however, and where possible key toxicity reviews are identified in the references. Further advice on the storage, handling and disposal of chemicals is given in *The Organic Chemist's Desk Reference*.

Finally, it should be emphasised that any chemical has the potential for harm if it is carelessly used. For many newly isolated materials, hazardous properties may not be apparent or may have been cited in the literature. In addition, the toxicity of some very reactive chemicals may not have been evaluated for ethical reasons, and these substances in particular should be handled with caution.

*RTECS® Accession Numbers**

Many entries in DNP contain one or more RTECS® Accession Numbers. Possession of these numbers allows users to locate toxicity information on relevant substances from the NIOSH *Registry of Toxic Effects of Chemical Substances*, which is a compendium of toxicity data extracted from the scientific literature.

For each Accession Number, the RTECS® database provides the following data when available: substance prime name and synonyms; date when the substance record was last updated; CAS Registry Number; molecular weight and formula; reproductive, tumorigenic and toxic dose data; and citations to aquatic toxicity ratings, IARC reviews, ACGIH Threshold Limit Values, toxicological reviews, existing Federal standards, the NIOSH criteria document program for recommended standards, the NIOSH current intelligence program, the NCI Carcinogenesis Testing Program, and the EPA Toxic Substances Control Act inventory. Each data line and citation is referenced to the source from which the information was extracted.

Bibliographic References

The selection of references is made with the aim of facilitating entry into the literature for the user who wishes to locate more detailed information about a particular compound. Thus, in general, recent references are preferred to older ones, particularly for chiral compounds where optical purity and absolute configuration may have been determined relatively recently. The number of references quoted cannot therefore be taken as an indication of the relative importance of a compound, and the references quoted for important substances may not be the most significant historically.

References are given in date order except for references to spectroscopic library collections, which sort at the top of the list, and those to hazard/toxicity sources which sort at the bottom.

The content of most references is indicated by means of suffixes, known as reference tags. A list of the most common ones is given in Table 4, p. 145 of the printed User Manual. For references describing a minor natural product which has been included in DNP as a derivative of a parent compound, the reference tag may be the identifying name of the natural product, e.g. (Laciniatoside II).

Some reference suffixes are now given in **boldface** type, where the editors consider the reference to be particularly important, for example the best synthesis giving full experimental details and often claiming a higher yield than previously reported methods.

In some entries, minor items of information, particularly the physical properties of derivatives, may arise from references not cited in the entry.

Journal abbreviations

In general these are uniform with the *Chemical Abstracts Service Source Index* (CASSI) listing except for a short list of very common journals:

DNP ABBREVIATION	***CASSI***
Acta Cryst. (and sections thereof)	Acta Crystallogr. (and sections thereof)
Annalen	Justus Liebigs Ann. Chem.
Chem. Comm.	J. Chem. Soc., Chem. Commun.
J.A.C.S.	J. Am. Chem. Soc.
J.C.S. (and various subsections thereof)	J. Chem. Soc. (and various subsections thereof)
J. Het. Chem.	J. Heterocycl. Chem.
J.O.C.	J. Org. Chem.
Tet. Lett.	Tetrahedron Lett

Entry under review

The database is continually updated. When an entry is undergoing revision at the time of a CD-ROM release (for example by the addition of further derivatives or references), this is indicated by a message at the head of the entry.

Description of Natural Product Structures

This **Description of Natural Product Structures** is adapted from the printed version of ***Dictionary of Natural Products***, and revised for ***Dictionary of Natural Products on CD-ROM***

The purpose of this general introduction and review is to facilitate access to the DNP Type of Compound Index which in turn leads on to the individual DNP entries. The order of main sections is the same as in the Type of Compound Index, and within the main sections the order of description of types of compound in general parallels the order in which they appear in the Type of Compound Index (except in the case of aliphatic natural products). Throughout this Description, the names of natural products which are not specially illustrated here but which are documented in DNP entres are given in **boldface type**. (The names used in this Description may not necessarily be the Dictionary entry names: use the Compound Name Index to locate substances if necessary.)

The various classes of natural product are described in respect of: (*a*) structural characteristics; (*b*) nomenclature, including Chemical Abstracts nomenclature; (*c*) biogenesis; (*d*) general biochemical significance and (*e*) any other information about the class which is of general importance. For detailed information about individual natural products it is necessary to locate the compound within its entry, which will in turn facilitate access to the primary literature.

Contents

Aliphatic natural products *page* **15**
Semiochemicals 15
Lipids 22

Polyketides **29**

Carbohydrates **35**

Oxygen heterocycles **44**

Simple aromatic natural products **45**

Benzofuranoids **48**

Benzopyranoids **49**

Flavonoids **51**

Tannins **60**

Lignans **64**

Polycyclic aromatic natural products **68**

Terpenoids *page* **72**
Monoterpenoids 73
Sesquiterpenoids 77
Diterpenoids 101
Sesterterpenoids 118
Triterpenoids 121
Tetraterpenoids 131
Miscellaneous terpenoids 133
Meroterpenoids 133

Steroids **135**
The sterols 140

Aminoacids and peptides **148**
Aminoacids 148
Peptides 150
β-Lactams 151
Glycopeptides 153

Alkaloids **154**
Alkaloids derived from ornithine 154
Alkaloids derived from lysine 156
Alkaloids derived from nicotinic acid 158
Alkaloids of polyketide origin 159
Alkaloids derived from anthranilic acid 160
Alkaloids derived wholly or in part from phenylalanine or tyrosine 163
Isoquinoline alkaloids 163
Alkaloids derived from tryptophan 181
Monoterpenoid indole alkaloids 184
Terpenoid alkaloids 196
Steroidal alkaloids 200
Imidazole alkaloids 202
Oxazole alkaloids 202
Thiazole alkaloids 203
Pyrazine and quinoxaline alkaloids 203
Pyrrole alkaloids 203
Putrescine alkaloids 203
Spermine and spermidine alkaloids 203
Peptide alkaloids 204
Purines 205
Pteridines and analogues 205

Polypyrroles **206**

Aliphatic natural products (VA)

A wide variety of small aliphatic and alicyclic compounds occur in nature. Because they are diverse, no attempt will be made here to give a general account; for information on specific aliphatics, their individual entries should be consulted. Accounts are given below, however, of the *semiochemicals*, which are structurally diverse but which are defined functionally and include many of the aliphatic compounds of greatest biochemical importance and current research interest, and of the *lipids*, which are structurally more or less well defined.

In the Type of Compound Index, the aliphatic compounds included in the Dictionary are simply classified by functional group and ring/chain structure.

A wide range of aliphatic compounds is documented as secondary products from natural starting materials, e.g. in cooked foods, but these are outside the scope of the Dictionary.

Semiochemicals

Semiochemicals are defined as chemicals that mediate communication between individual organisms. The word is derived from the Greek word *semeion* which means mark, sign or signal. Although some semiochemicals are released purposefully (sex pheromones, the scent of flowers), others are released as a consequence of normal metabolism, but nevertheless still convey information. An example of the latter is the attraction of the tsetse fly to carbon dioxide in the breath of cattle. It is of no consequence whatsoever to the fly that the 'signal' is intentional or otherwise.

Most workers exclude the consumption of food (but not its detection) and the use of defence chemicals but again not their subsequent detection (cf. the stink of the skunk) as semiochemical interactions. The study of semiochemical interactions is termed chemical ecology. An alternative nomenclature emphasises the transfer of information, and uses the term *infochemical*, instead of semiochemical.

When semiochemicals act between members of the same species they are known as *pheromones* from the Greek *pherein* to carry and *horman* to excite or stimulate. They are the external counterparts of hormones which act as messengers between organs in the body. Pheromones have been divided into two categories based on temporal criteria. Releaser pheromones elicit a response which is immediate and usually behavioural, whereas primer pheromones cause longer term physiological changes.

Pheromonal systems are usually the most highly developed semiochemical systems because the species directly benefits from any improvement. Highly developed in this context means that release of the pheromone is efficient and timely and that the receiver has a sensitive and selective detection system. Moreover because most pheromones are involved in reproductive functions (mate attraction, courtship and copulation), increased efficacy is immediately apparent in higher fecundity.

Semiochemicals that act between members of different species are called allelochemics and these are further divided into *allomones* which cause an effect favouring the emitter (such as the stink of the skunk) and *kairomones* which favour the receiver (e.g. the odour of a prey species that attracts its predator). Semiochemicals which favour both the receiver and the emitter are known as *synomones*. A good example of this is the scent of a flower which attracts bees to feed on nectar and pollinate them.

The most widely known semiochemicals are the volatile sexual attraction

pheromones of insects but volatility, although common, is not a prerequisite. Semiochemicals may also be transferred by touch (e.g. the aphrodisiac polypeptides of the golden hamster) or in solution (fish maturation factors). They range in structure from carbon dioxide, ethylene (an attractant for bark beetles), aliphatics, mono and polycyclic hetero and carbocycles through steroids to polypeptides. All classes of organisms including yeast, corals, crustacea, newts, fungi, plants, insects, spiders, mites, fish and mammals employ some form of semiochemical. However only limited information is available for amphibians, reptiles and birds.

Insect pheromones

There has been much interest in the use of insect pheromones for the monitoring and control of insect pests in the field, gardens, stored products, food processing factories and in commercial kitchens.

Semiochemicals of social insects, ants, termites, locusts, heteroptera and the coniferphagus choristoneura have been reviewed (see bibliography); other representative classes are reviewed below.

The structures of more lepidopteran pheromones are known than all other pheromones together and moreover they are chemically fairly homogenous. Typically they consist of an unbranched carbon chain with an even number of carbon atoms, terminated by an oxygen containing group (acetate, alcohol or aldehyde) with 0–4 alkenic bonds located predominantly towards the hydrocarbon terminus of the chain.

The 1992 compilation of attractants for lepidoptera and other species by Arn *et al.* includes 2292 semiochemicals from 1068 species, but only 264 unique chemical structures. The most common attractants are *Z*9–14Ac (168 cases), *Z*11–14Ac (157), *Z*11–16Ac (133), *E*11–14Ac (124) and *Z*7–12Ac (111). Taken together these five chemicals are used by 30% of all species; however 80% of species use at least two components. Approximately 40% of the structures have a terminal acetate group and the remaining 60% are evenly divided between aldehydes, alcohols and hydrocarbons with a few 2-ketogroups. Recently a nitrate terminal group was reported and presumably others remain to be discovered. The commonest chain lengths in decreasing order of abundance are C14, C12, C16 and C18. Approximately 40% are monoenes and 40% dienes, with the double bonds predominantly located at the (ω-3) and (ω-5) positions.

The biosynthesis of lepidopteran pheromones is controlled by a neuropeptide (PBAN, pheromone biosynthesis activating neuropeptide). The pathway commences with saturated fatty acids which undergo desaturation and successive losses of terminal acetate units, followed by modification of the terminal group. Δ^9 and Δ^{11} desaturases for dodecanoic and tetradecanoic acids (e.g. **Bombykol**) have been identified.

Methyl substituted straight-chain pheromones are fairly common in nature. These range from simple systems with a sole methyl group remote from any functionality or at the *antiso* position (ω-2) through *n*, (*n* + 6) and the more common *n*, (*n* + 4) dimethyl compounds.

It is notable that *n*, (*n* + 4) polymethyl compounds have predominantly the *R*-configuration. This results from incorporation of propionate into the normal fatty acid biosynthetic pathway, in which the 2-pro-*R* proton is removed from propionate. An interesting example which has been investigated in considerable detail is **4*R*,8*R*-Dimethyldecanal**, the aggregation pheromone of flour beetles.

2,6-Dichlorophenol is the female-produced pheromone for a number of ticks. Several other species use **Phenol** and ***p*-Cresol**. However these are frequently insufficient to elicit mating and a contact pheromone, **Cholesteryl oleate**, has been identified from the American dog tick.

The aggregation pheromone of the acarid mite *Lardoglyphus konoi* is **Lardolure.**

Fruit flies (Diptera) and some species of wasp (Hymenoptera) use simple spiroketals as sex pheromones. **1,7-Dioxaspiro[5.5]undecane** is the major female produced sex attractant of the olive fly *Bactrocera oleae* (formerly *Dacus oleae*). Two other minor components, the 3-hydroxy and 4-hydroxy derivatives were also identified.

Simple spiroketal ring systems have also been identified as the structural motif of a wide range of other pheromones, e.g. rectal gland secretions of male Asian fruit flies, **Chalcogran**, the aggregation pheromone of the bark beetle *Pityogenes chalcographus*, aggression inhibiting pheromones of the wasp *Paravespula vulgaris*, sexual attractants of several species of *Andrena* bees and cephalic secretions of cleptoparasitic bees.

Attack of a tree by a bark beetle is orchestrated by a complex score of semiochemical notes. Initial attack by pioneers (females in monogamous species) occurs in response to tree volatiles such as **Myrcene**. Upon landing gallery excavation commences, followed by mating and release of pheromones by the males. The synergistic blend of tree volatiles and the pheromones produced by both sexes initiates mass attack, which overwhelms the tree's defences. Meanwhile fungal and yeast spores are introduced into the tree adventitiously or from special chambers (mycangi) on the shoulders of the beetles. These micro-organisms proliferate and block the sap channels which prevents the galleries being flooded with resin and may produce the beetles' aggregation pheromone. In some species such as the ambrosia beetle the larvae feed on the fungus. Finally when the tree is fully colonised deterrent pheromones are produced by the beetle, or by the fungi.

Bridged spiroketals are predominantly found as pheromones of bark beetles and usually have a 6,8-dioxabicyclo[3.2.1]octane skeleton. The first structure to be elucidated was that of **Brevicomin**, the female-produced aggregation pheromone of the western pine beetle, *Dendroctonus brevicomis*. Different beetle species produce different isomeric Brevicomin compositions. Male western pine beetles produce **Frontalin** which is the last component to be added to the synergistic blend of *exo*-Brevicomin and Myrcene (from the host tree) which initiates mass attack of the host tree by the beetles.

Some other notable examples are **Multistriatin**, the pheromone of the European elm bark beetle, *Scolytus multistriatis*, responsible for 'Dutch Elm' disease, the male-produced pheromones of the swift moth, *Hepialus hecta* which are one carbon higher homologues of the bark beetle pheromones described above, and **Lineatin** which has the same carbon skeleton as **Grandisol**.

Rodents

The behaviour and development of the house mouse (*Mus musculus*) is determined by a complex system of pheromonal effects mostly mediated by urine. The best characterised of these are **3,4-Dehydro-*exo*-brevicomin** and **2-*sec*-Butyl-2-thiazoline** which are testosterone-dependant aggression promotors, isolated from the urine of adult male mice. One of the most interesting developmental factors is the acceleration of puberty by volatiles from adult urine. Recent evidence suggest that 3,4-Dehydro-*endo*-brevicomin is bound to a lipocalin (a small peptide) in mouse urine. When the urine is deposited it releases the volatile ligand which promotes investigation by other mice. When immature mice pick up the lipocalin it accelerates development into puberty.

Dimethyl disulfide has been identified in male rabbit pellets and rat preputial tissue. It is released in the vaginal secretions of the female hamster and acts as a potent attractant.

Fish

The sexual development of many species of fish is determined by the presence of steroids in the water. **17α, 20β-Dihydroxypregn-4-en-3-one** is detected by the medial olfactory tracts of male goldfish and causes gonadotrophin secretion. The *in vitro* biosynthesis of several steroids has been demonstrated in the ovaries of the female during the final stages of oocyte maturation. Male Atlantic salmon are similarly affected by the corresponding 20-sulfate and **Testosterone**.

Herbivores

Hoofed mammals have many specialised glands for scent marking. The tail gland in the red deer (*Cervus elaphus*) produces **Phenol, *m*-Cresol, Benzoic acid, 4-Ethylphenol, Dimethyl sulfone, *o*-Cresol, 3-Phenylpropanoic acid** and **Phenylacetic acid**. Many of these are common to other herbivore and carnivore species.

Herbivores avoid areas which carnivores have marked. Treatment of conifer saplings with red fox urine, a synthetic mixture of its volatile constituents or **Isopentenyl methyl sulfide** suppressed feeding on the saplings by the snowshoe hare (*Lepus americanus*); which is common prey for the red fox in winter. Similarly trees in orchards were protected from feeding by voles and gophers by these materials and by **2,5-Dihydro-2,4,5-trimethylthiazole**, a constituent of red fox faeces. **(*Z*)-7-Dodecenyl acetate** which is best known as a lepidopteran sexual attractant in at least 111 species, is also produced by female elephants at the time of ovulation.

Carnivores

The Carnivores (dogs, cats etc.) show a wide range of social behaviour ranging from solitary individualism to hunting in packs. Territory marketing with urine, faeces and various glands is common and social interactions initiated by sniffing are common even amongst habitually solitary species.

The characteristic ‘skunky odour’ of fox urine is due to **Methyl 2-phenylethyl sulfide** and Isopentenyl methyl sulfide, which is also a component of wolf urine and comprises more than 50% of the volatiles from female coyote urine. Bis(3-isopentenyl) sulfide and **Isopentanethiol** are also found in coyote urine and reach a maximum at oestrus. The latter is also a component of the defensive ‘scent’ of the striped skunk, together with a series of crotyl thiols. **2,2-Dimethylthietane** (Mustelan) was isolated from the anal glands of the mink and the polecat which also contains **3,3-Dimethyl-1,2-dithiolane** and Bis(3-isopentenyl) sulfide. **(2*S*)-2-Propylthietane** is the major malodorous substance from the anal gland of the stoat (*Mustela erminea*).

The anal sacs of the red fox secrete a mobile strongly smelling liquid containing a complex mixture of short chain carboxylic acids, **Trimethylamine** and diamines such as **Putrescine** and **Cadaverine**. The volatile fatty acids are produced by anaerobic bacteria.

The anal sacs of the dog (*Canis familiaris*) and the coyote (*Canis latrans*) also contain short chain carboxylic acids and Trimethylamine. Carboxylic acids and medium chain esters, plus **Indole** and **Hexanol** were found in the anal gland secretion of the aardwolf (*Proteles cristatus*), but no aliphatic amines were detected. Indole is also present in the anal sac secretion of the ferret (*Mustelo furo*), but the components which elicit the greatest attraction for conspecifics are **2-Propylthietane**, **3-Propyl-1,2-dithiolane** and ***cis*** and ***trans*-2,3-Dimethylthietane**.

The supracaudal gland is present in most Canidae and is most highly developed in the less social species. Sniffing the gland is common in social interactions. The secretion of the gland has a floral odour and it is often described as the 'violet gland'. **Dihydroactinidiolide** and other volatile terpenes have been detected in the distillate from the secretion of the red fox *Vulpes vulpes* but very little is known about the other constituents.

Humans

The apocrine glands, which are mostly found in the armpit and on the head, are believed to be the source of human pheromones. **5α-Androst-16-en-3-one** (the principal component of boar taint) and other steroids are secreted by males from the armpits and modified by the microflora on the skin. Armpit extracts have been shown to modify the menstrual cycle of women with irregular periods and the curious phenomenon of menstrual synchronisation which occurs between cohabiting women is believed to be mediated pheromonally. The odour of steroids and the similar odour of musk has always played a significant role in human culture, such as the base note of perfumes, the burning of incense and the highly prized truffle.

Yeast

Saccharomyces cerevisiae (Bakers' yeast) has two haploid mating types designated as a and α which release the pheromones a-factor and α-factor (respectively). These cause arrest of the cell cycle in the G1 phase and development of pear shaped extrusions of the cell called shmoos. Agglutination, cell fusion and nuclear fusion gives a zygote that produces a/α diploid buds. Cells are able to discriminate amongst potential partners and select those with the highest pheromone production. This has been termed courtship.

The genes MFα1 and MFα2 encode two α-factor precursors, consisting of 165 and 120 amino acids respectively. These are translocated into the classical secretory pathway where they are glycosylated on asparagine and proteolysed to give mature α-factor which is a 13 aminoacid peptide (WHWLQLKPGQPMY) and released.

MFA1 and MFA2 are the two genes which encode 36 and 38 amino acid peptide precursors of two forms of a-factor. They are synthesised and undergo proteolysis in the cytoplasm to give 15-mers. These have a C-terminal -CAAX motif, where C is **Cysteine**, A is an aliphatic aminoacid and X is variable. The Cysteine undergoes *S*-farnesylation, the terminal tripeptide is cleaved, and finally methylation of the terminal cysteine gives the a-factors (YIIKG(V or L)FWDFA(S-Far)C-OMe). These both contain 12 aminoacid residues but differ at residue 6.

The genes STE2 and STE3 encode the receptors for α and a-factor respectively. The proteins are predicted to have seven potential transmembrane domains, analogous to the β-adrenoceptor and rhodopsin receptors which interact with mammalian G-proteins. Therefore G-proteins may also participate in the mating signal transduction pathway in yeast.

Five unique pheromones have been identified from seven mating types of the ciliated protozoa *Euplotes raikovi*. These consist of 38–40 aminoacid residue polypeptides with an amino terminal aspartic acid residue and six conserved half cystines, but otherwise they have little homology although E*r*-1189 and E*r*-2 are bound equally well by some monoclonal antibodies. Different mating types producing the same pheromone are able to mate and one type which secretes E*r*-20 is unable to mate with any of the others.

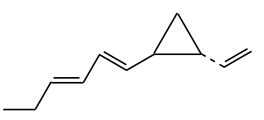
Conjugative plasmid transfer by *Enterococcus faecalis* is induced by a pheromone released by the prospective recipient.

Albone, E.S. (1984) in *Mammalian Semiochemistry: The Investigation of Chemical Signals Between Mammals*, Wiley, Chichester, pp. 1–360.
Aldrich, J.R. (1988) *Ann. Rev. Entomol.*, **33**, 211.
Ali, M.F. *et al*, (1990) *Biol. Rev. Cambridge Phil. Soc.*, **65**, 227.
Anderson, D. *et al*, (1990) *J. Mol. Biol.*, **216**, 1.
Arn, H., Toth, M. and Priesner, E. (1986) *List of Sex Pheromones of Lepidoptera and Related attractants*, OILB-SROP/IOBC-WRPS, Paris.
Arn, H., Toth, M. and Priesner, E. (1992) *List of Sex Pheromones of Lepidoptera and Related attractants*, 2nd edn, OILB-SROP/IOBC-WRPS, Paris, pp. 1–179.
Ashby, M.N. *et al*, (1992) *Prod. Natl. Acad. Sci. USA*, **89**, 4613.
Attygalle, A.B. *et al*, (1984) *Chem. Soc. Rev.*, **13**, 245.
Baker, R. *et al*, (1984) *Nat. Prod. Rep.*, **1**, 299.
Bedoukian, P. (1994) *Perfum. Flavor*, **19**, 1 (*rev*).
Bell, W.J. and Carde, R.T. eds. (1984) *The Chemical Ecology of Insects*, Chapman & Hall, New York.
Boivin, T.L.B. (1987) *Tetrahedron*, **43**, 3309.
Brand, J.M. *et al*, (1979) *Prog. Chem. Org. Nat. Prod.*, **37**, 1.
Byers, J.A. (1991) *Biol. Rev. Cambridge Phil. Soc.*, **66**, 347.
Calkin, R.R. *et al*. (1994) *Perfumery Practice and Principles*, Wiley.
Claus, R. *et al.* (1979) *Experientia*, **35**, 1674.
Claus, R. *et al.* (1981) *Experientia*, **37**, 178.
Clissold, D. *et al.* (1994) *Nat. Prod. Rep.*, **11**, 621.
Dicke, M. *et al.* (1988) *Functional Ecology*, **2**, 131.
Eriotou-Bargiota *et al.* (1992) *Biochemistry*, **31**, 551.
Evans, D.A. *et al.* (1973) *Chem. Soc. Rev.*, **2**, 75.
Featherstone, C. (1990) *Trends Biochem. Sci.*, **15**, 169.
Fields, S. (1990) *Trends Biochem. Sci.*, **15**, 270.
Fiori, P.L. *et al.* (1990) *J. Protozool.*, **37**, 187.
Francke, W. (1984) in *Advances in Invertebrate Reproduction*, **3** (eds. E. Engels *et al.*), Elsevier Science Publishers B.V., Amsterdam, pp. 493.
Freestone, N.P. (1988) *Performance Chemicals*, August, 26–29.
Galli, D. *et al.* (1991) *J. Bacteriol.*, **173**, 3029.
Gehrung, S. *et al.* (1990) *J. Cell Biol.*, **111**, 1451.
Hamilton, J.G.C. (1992) *Parasitol. Today*, **8**, 130.
Harborne, J.B. (1993) *Nat. Prod. Rep.*, **10**, 327.
He, B. *et al.* (1991) *Prod. Natl. Acad. Sci. USA*, **88**, 11373.
Hirsch, J.P. *et al.* (1992) *Bioessays*, **14**, 367.
Hrycyna, C.A. *et al.* (1992) *J. Biol. Chem.*, **267**, 10457.
Hummer, H.E. and Miller, T.A. (eds.) (1984) *Techniques in Pheromone Research*, Springer, Berlin.
Jackson, C.L. *et al.* (1991) *Cell*, **67**, 389.
Jones, O.T. (1985) in *Insecticides* (D.H. Hutson and T.R. Roberts, eds.), Wiley, Chichester, pp. 311–373.
Karlson, P. *et al.* (1959) *Nature*, **183**, 155.
Kelly, D.R. (1990) *Chem. Brit.*, 124.
Kelly, D.R. (1993) in *The Chemistry of Natural Products*, 2nd edn (ed. R.H. Thomson), Blackie, Glasgow, pp. 382.
Kelly, D.R. (1996) *Chem. Biol.*, **3**, 595–602.
Kluge, A.F. (1986) *Heterocycles*, **24**, 1699.
Ko, H.A. *et al.* (1990) *J. Biol. Chem.*, **265**, 21652.
Kurjan, J. (1992) *Annu. Rev. Physiol.*, **54**, 639.
Lake, M. (1989) *Scents and Sensuality. The Essence of Excitement*, John Murray, London.
Law, J.H. *et al.* (1971) *Ann. Rev. Biochem.*, **40**, 533.
Lin, G. (1993) *Pure Appl. Chem.*, **65**, 1233 (rev, synth, insect pheromones).
Lipke, P.N. *et al.* (1992) *Microbiol. Rev.*, **56**, 180.
Luporini, P. *et al.* (1992) *Dev. Genetics*, **13**, 9.
Lutz, R.J. *et al.* (1992) *Prod. Natl. Acad. Sci. USA*, **89**, 3000.
Marcus, S. *et al.* (1990) *Biochem. Biophys. Res. Commun.*, **172**, 1310.
Marr, R.S. *et al.* (1990) *J. Biol. Chem.*, **265**, 20057.

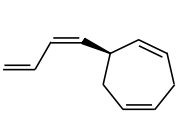
Mayer, M.S. and McLaughlin, J.R. (1991) *Handbook of Insect Pheromones and Sex Attractants*, CRC Press, Boca Raton.
Mori, K. (1992) *Total Synth. Nat. Prod.*, **9**, 1
Mori, K. (1997) *Chem. Commun.*, 1153 (*rev, pheromones*)
Millar, J. and Haynes, K., Eds (1997) *Methods in Chemical Ecology*, Chapman & Hall, N.Y., 1997
Morse, D. *et al.* (1986) *J. Chem. Ecol.*, **12**, 335.
Mundy, B.P. *et al.* (1977) *Heterocycles*, **6**, 51.
Ortenzi, C. *et al.* (1990) *J. Cell Biol.*, **111**, 607.
Perron, F. *et al.* (1989) *Chem. Rev.*, **89**, 1617.
Pontius, L.T. *et al.* (1992) *J. Bacteriol.*, **174**, 1821, 3152.
Poucher, W.A. (1991) *Poucher's Perfumes, Cosmetics and Soaps*, Volume 1, The Raw Materials of Perfumery, 9th edn (A.J. Jouhar ed.), Chapman & Hall, London, pp. 236–237.
Prestwich, G.D. (1983) *Sci. Am.*, **249**, 68.
Prestwich, G.D. (1984) *Ann. Rev. Entomol.*, **29**, 201.
Proteau, G. *et al.* (1990) *Biochem. Biophys. Res. Commun.*, **170**, 182.
Rad, M.R. *et al.* (1992) *Mol. Gen. Genet.*, **236**, 145.
Raffioni, S. *et al.* (1992) *Prod. Natl. Acad. Sci. USA*, **89**, 2071.
Rose, M.D. (1991) *Ann. Rev. Microbiol.*, **45**, 539.
Schafer, W.R. *et al.* (1990) *Science*, **249**, 1133.
Schneider, D. (1992) *Naturwissenschaften*, **79**, 241 (*rev, pheromones*)
Silk, P.J. *et al.* (1988) *Ann. Rev. Entomol.*, **33**, 83.
Silverstein, R.M. (1981) *Science*, **213**, 1326.
Singer, A.G. (1991) *J. Steroid Biochem. Mol. Biol.*, **39**, 627.
Solomon, F. (1991) *Ann. Rev. Cell Biol.*, **7**, 633.
Stoddart, D.M. (1976) *Mammalian Odours and Pheromones*. The Institute of Biology's Studies in Biology no. 73, Edward Arnold, London.
Stoddart, D.M. (1988) Human odour culture: a zoological perspective, in *Perfumery: The psychology and biology of fragrance* (eds S. Van Toller and G.H. Dodd), Chapman & Hall, London, pp. 1–18.
Tesseire, P.J., Ed. (1995) *Chemistry of Fragrant Substances*, Allured Carol Stream, Il., USA.
Trotter, K.M. *et al.* (1990) *Plasmid*, **24**, 57.
Tumlinson, J.H. (1989) *Pure Appl. Chem.*, **61**, 559.

Algae

The marine brown algae have evolved a unique pheromone system, the known members of which are based on a range of acyclic or aliheterocyclic alkenes. The majority of these are C_{11}, examples being the Dictyopterenes, exemplified by Dictyopterene A and Desmarestene. Some *Fucus* and *Sargassum* species employ the smaller molecule **1,3,5-Octatriene** (various stereoisomers). These molecules modulate the navigation of the male gametophytes in the marine environment over the very short ranges (~1 mm) required for fertilisation. They are biosynthesised from highly unsaturated fatty acids by pathways not yet understood; the precursor unsaturated acids such as eicosahexaenoic, are of a type not found in terrestrial plants.



Dictyopterene A



Desmarestene

There are some known examples of low molecular-weight compounds fulfilling a similar function in freshwater algae, but in these cases the pheromones are terpenoid, as exemplified by **Sirenin**. In general, knowledge of pheromone systems in lower plants is still too fragmentary to generalise that compounds of these series represent all of the possibilities. This point is underlined by the

recent characterisation of **Lurlene**, of a totally different structural type, as the sex pheromone of the terrestrial green flagellate alga *Chlamydomonas allensworthii*.

Jaenicke, L. *et al*, *Angew. Chem. Internat. Ed. Engl.*, 1982, **21**, 643–653.
Wirth, D. *et al*, (1992), *Helv. Chim. Acta*, **75**, 734–744.

Lipids

Lipids have been defined in different ways at different times and there is still no agreed definition of the term. Recent proposals are based mainly on chemical structure and, in turn, on the underlying biosynthetic pathways; two recent definitions read: 'lipids are fatty acids and their derivatives, and substances related biosynthetically or functionally to these compounds' and 'lipids are compounds based on fatty acids or closely-related compounds such as the fatty alcohols and the sphingosine bases'. This definition includes all the major groups of materials generally recognised as lipids: it incorporates sterol esters but not the free sterols.

Fatty acids (VA0300, VA0600, VA1100, VA1500)

Many fatty acids are still known by their trivial names (e.g. **Palmitic, Linoleic**). These were often related to the original source of the acid and were given before the structure could be adequately defined.

Systematic names indicate the chain-length and the nature, position, and configuration of unsaturated centres as in the following examples:

Trivial	Systematic	Abbreviation
Palmitic	hexadecanoic	16 : 0
Oleic	*cis*-9-octadecenoic	18 : 1 (9*Z*)
Arachidonic	all-*cis*-5,8,11,14-eicosatetraenoic	20 : 4 (*n*-3)

The systematic names are often replaced by abbreviations of the form A: B(C). A indicates the number of carbon atoms in the molecule, B represents the number of unsaturated centres which are usually *cis*-(*Z*-) alkenic, and C indicates the position and configuration of the unsaturation. Organic chemists number the chain from the carboxyl group (COOH = 1) hence 9Z for oleic acid but there are times when it is more appropriate to count from the methyl end and to use symbols such as ω3 or *n*-3 to indicate the position of the unsaturated centre closest to the CH_3 group. In this case it is assumed that all unsaturation is methylene – interrupted and has *cis*-(*Z*-) configuration.

The number of natural fatty acids which has been reported may exceed 1000 though only 20–50 are of common concern. From a survey of all these structures it is possible to make a number of general statements. These are essentially true, particularly in respect of the more common and important acids, but there are significant exceptions to each statement. These statements were first based on chemical structure but it is clear that they also reflect underlying biosynthetic pathways.

(i) Natural fatty acids – both saturated and unsaturated – are straight-chain compounds with an even number of carbon atoms. This holds for the great majority of structures and for the more common acids. Chain lengths range from two to more than eighty carbon atoms but are most commonly between C_{12} and C_{22}. Despite the validity of this generalisation acids with an odd number of carbon atoms (e.g. **Nonadecanoic**) occur as do those with branched structures

(e.g. **Isopalmitic, Anteisononadecanoic**) or with carbocyclic units (e.g. **Sterculic**).

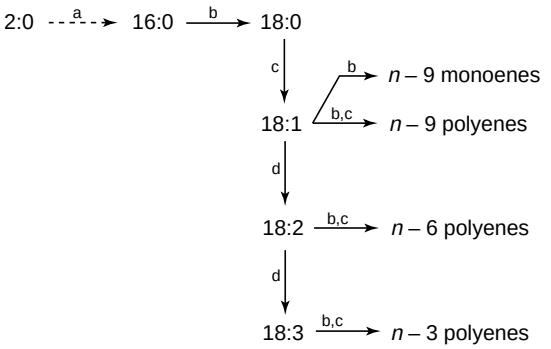
(ii) Acids with one unsaturated centre are usually alkenic compounds with *cis*- (*Z*-) configuration and with the double bond in one of a limited number of preferred positions. This is most commonly Δ^9 (e.g. **Oleic**) or *n*-9 (e.g. **Erucic**) but the double bond can occur in other positions (e.g. **Petroselinic**) and monoacetylenic acids are also known (e.g. **Tariric**).

(iii) Polyunsaturated acids are mainly poly-alkenic (*cis-/Z*-configuration) with a methylene-interrupted arrangement of double bonds, i.e. double bonds are separated from each other by one CH_2 group. The pattern of 1,4-unsaturation is characteristic of fatty acids and differs from that in isoprenoids which is usually 1,3-(conjugated) or 1,5-. Polyunsaturated fatty acids occur in biochemically related families and the two most important are the *n*-6 family based on **Linoleic acid** and the *n*-3 family based on **α-Linolenic acid** (see discussion on biosynthesis below). Some acids have conjugated unsaturation which is both *cis* and *trans* (e.g. **Eleostearic, Calendic, Parinaric**), some have mixed ene/yne unsaturation both conjugated (e.g. **Isanic**) and non-conjugated (e.g. **Crepenynic**), and there is a small group of acids in which unsaturation is not entirely methylene-interrupted (e.g. **Columbinic**).

(iv) Fatty acids rarely have functionality apart from the carboxyl group and the various types of unsaturation discussed above but acids are known with fluoro, hydroxy, keto, and epoxy groups. Two important examples are **Ricinoleic** (12-hydroxyoleic) and **Vernolic** (*cis*-12,13-epoxyoleic).

These generalisations have a biosynthetic basis and even some of the exceptions can be accommodated in general biosynthetic schemes with only minor modification.

The major biosynthetic pathways leading to fatty acids are summarised in Figure 1. In the *de novo* pathway leading to saturated fatty acids, acetate (the primer) condenses with malonate (the extender) to produce a C_4 oxo acid which is reduced in three steps to butanoate. This cycle of condensation and reduction continues until, most commonly, palmitate has been obtained, though in lauric oils rich in 12 : 0 the process stops mainly at the C_{12} level. The malonate is itself derived from acetate by carboxylation in the presence of a biotin enzyme and the carbon dioxide lost during condensation is that derived during carboxylation so that the carbon atoms in butanoate and in the longer chain acids are entirely acetate-derived.



a = *de novo* synthesis of saturated acids (mainly palmitic), b = chain-elongation, c = Δ^9-desaturation, d = desaturation of Δ^9-monoene (plants only), e = desaturation between an existing double bond and the carboxyl group (occurs rarely in plants but commonly in animals).

Figure 1. Fatty acid biosynthesis.

If the acetate is replaced by a different primer then other fatty acids are produced. This can be propionate (major product: heptadecanoate), 2-methyl-propionate (*iso* acids), or 2-methylbutanoate (*anteiso* acids).

The chain-elongation process is similar in outline to the *de novo* process but differs in some significant details. It operates with both saturated and unsaturated acids and occurs with either acetate or malonate. Erucic acid is made from Oleic acid by two chain-elongation steps:

$$18:1\ (9Z) \rightarrow 20:1\ (11Z) \rightarrow 22:1\ (13Z)$$

Both Oleic acid and Erucic acid are *n*-9 monoenes. This chain-elongation process is also important in the biosynthesis of polyenes (see below).

The most common route to monoene acids involves Δ^9 desaturation. This oxygen-requiring process occurs in plants, animals and microorganisms and furnishes acids with a *cis*-double bond between carbon atoms 9 and 10, e.g. 9-hexadecenoic, 9-octadecenoic (Oleic).

Further desaturation of Oleic acid to the 9,12-diene (Linoleic) and 9,12,15-triene (α-Linolenic) occurs only in plants. The additional double bonds assume a methylene interrupted pattern and lie between the existing double bond and the methyl group. Animals requiring these acids for the production of *n*-6 and *n*-3 polyene acids must obtain them through their dietary intake.

The bioconversion of 22 : 5 to 22 : 6 (and presumably of 22 : 4 to 22 : 5) may occur by a more complicated pathway than that suggested in Figure 2 involving the sequence 22 : 5 → 24 : 5 (elongation) → 24 : 6 (desaturation) → 22 : 6 (chain shortening)

Desaturation between an existing double bond and the carboxyl group occurs only rarely in plants (e.g. γ-Linoleic acid) but readily in animals. The additional double bonds have *cis*- configuration and are introduced in a methylene-interrupted pattern.

Some of these acids are required for the maintenance of health and are designated essential fatty acids (efa). It is possible to produce efa deficiency in

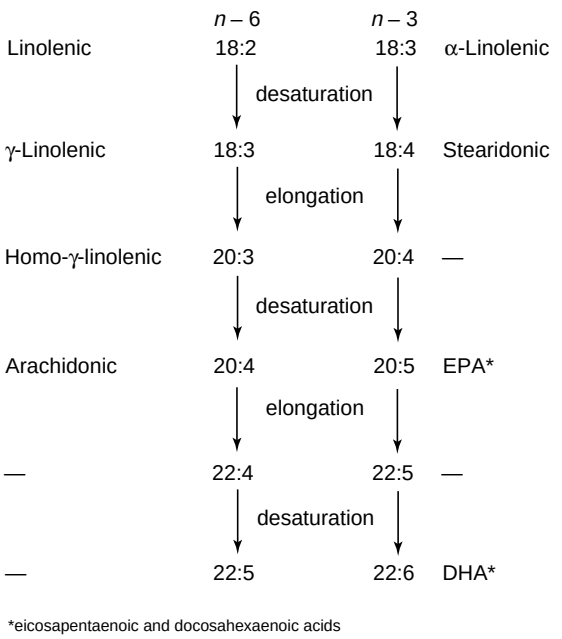


Figure 2. The *n*-6 and *n*-3 families of polynsaturated acids based on Linoleic and α-Linolenic acids.

experimental animals but this state is rarely observed in humans who generally ingest adequate quantities of Linoleic and α-Linolenic acid. Efa deficiency is most likely to be observed under conditions where the normal enzymic processes – especially desaturation – are impaired.

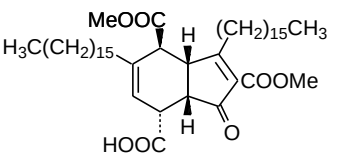
A wide variety of unusual fatty acids and phospholipids are found in sponges, and these arise by totally different biosynthetic pathways.

Oxylipins (VA6150)

Three C_{20} acids – 20 : 3 (*n*-6), 20 : 4 (*n*-6), and 20 : 5 (*n*-3) – are precursors of the PG1, PG2 and PG3 series of prostaglandins and of many other C_{20} metabolites. These are known collectively as eicosanoids and are products of the eicosanoid cascade.

The term oxylipin has been coined relatively recently to describe the class of natural product, of which prostaglandins are the most widespread, that are produced from C_{20} and in some cases C_{18} fatty acid precursors in at least one stage of mono- or dioxygenase-dependent oxidation. Since it is now known that C_{20} precursors are not universal, the term oxylipin is to be preferred to the previous term eicosanoid.

The widest variety of structural types is found in marine organisms where ring formation may produce three- (e.g. **Constanolactones**), five (e.g. **Ecklonialactone A**) or six- (e.g. Manzamenone A) membered rings.



Manzamenone A

Prostaglandins, prostacyclins and thromboxanes (VA6100)

The eicosanoid or arachidonic acid cascade produces **Prostaglandins**, **Prostacyclin** and **Thromboxanes** that mediate a wide range of physiological responses. They have short half lives and thus have limited clinical application, however synthetic analogues are being used as drugs.

In mammals, the arachidonate-derived prostaglandins play an important role in maintaining homeostasis.

Prostaglandins $F_{2\alpha}$ and E_2 are unexpectedly also encountered in marine algae and invertebrates. It appears that at least in corals these arise via lipoxygenase metabolism rather than the mammalian prostaglandin H synthase (cyclooxygenase) pathway.

Baxter, A.D. *et al.* (1986) *Chem. Ind. (London)*, 510 (*synth*).
Bentley, P.H. (1973) *Chem. Soc. Rev.*, **2**, 29 (*synth*).
Collins, P.W. *et al.* (1993) *Chem. Rev.*, **93**, 1533 (*synth*).
Djerassi, C. *et al.*, *Acc. Chem. Res.*, 1991, **24**, 64 (sponges)
Gerwick, W.H. (1993), *Chem. Rev.*, **93**, 1807–1823 (marine oxylipins)
Hart, T.W. (1988) *Nat. Prod. Rep.*, **5**, 1.
Lai, S.M.F. *et al.* (1984) *Nat. Prod. Rep.*, **1**, 409.
Lands, W.E.M. (1991) *Annu. Rev. Nutr.*, **11**, 41 (*biosynth*).
Newton, R.F. *et al.* (1984) *Synthesis*, 449 (*synth*).
Smith, W.L. (1992) *Am. J. Physiol.*, **263**, F181 (*biosynth, action*).

Glycerides (VA6700, VA6800, VA6900)

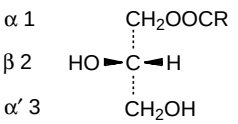
Fatty acids occur naturally as esters of **Glycerol** or of some other hydroxy compound or as amides of long-chain amines such as **Sphingenine**. The less common long-chain alcohols occur as esters or as ethers. Triacylglycerols are major storage lipids whilst phospholipids (see below) are important membrane constituents.

Acylglycerols are esters of glycerol and fatty acids. Partial glycerides are important intermediates in metabolism and triacylglycerols are the major constituents of natural fats and oils. In DNP glycerides are named as glycerol triesters, e.g. entry name = **Glycerol tri-9-octadecenoate**.

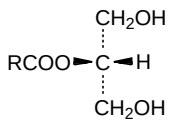
In order to designate the stereochemistry of glycerol-containing components, the carbon atoms of glycerol are numbered stereospecifically. When the glycerol molecule is drawn in a Fischer projection with the secondary hydroxyl group to the left of the central prochiral carbon atom, then the carbons are numbered 1, 2 and 3 from top to bottom. Molecules which are stereospecifically numbered in this fashion have the prefix '*sn*' immediately preceding the term 'glycerol' in the name of the compound to distinguish them from compounds which are numbered in a conventional fashion. The prefix '*rac*' in front of the full name shows that the compound is an equal mixture of both antipodes and '*x*' is used if the configuration is unknown or unspecified.

Any glycerolipid will be chiral when the substituents at the *sn*-1 and *sn*-3 positions are different. If both substituents are long-chain acyl groups then the optical rotation will be extremely small.

Monoacylglycerols (monoglycerides) (VA6700) are fatty acid monoesters of glycerol and exist in two isomeric forms:

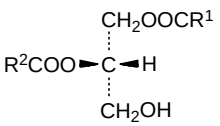


1-Monoacyl-*sn*-glycerol
(α-isomer)

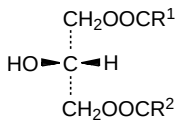


2-Monoacyl-*sn*-glycerol
(β-isomer)

Diacylglycerols (diglycerides) (VA6800) are fatty acid diesters of glycerol and also occur in two isomeric forms:

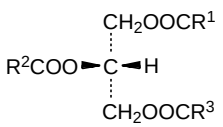


1,2-Diacyl-*sn*-glycerol
(α,β-isomer)



1,3-Diacyl-*sn*-glycerol
(α,α′-isomer)

Triacylglycerols (triglycerides) (VA6900) are fatty acid triesters of glycerol. The fatty acids may be all different, two different, or all alike.



Triacyl-*sn*-glycerol

These materials occur as mixtures with various acyl chains which may show some fatty acid specificity. As a consequence particular fatty acids may be concentrated in or excluded from particular positions in the glycerol ester. To produce seed oils with more than 67% of a particular acid it may be necessary to modify the acylating enzymes by genetic manipulation.

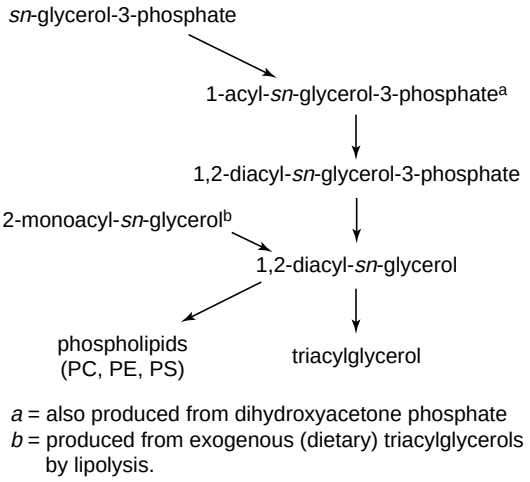


Figure 3. Biosynthesis of the major lipid classes.

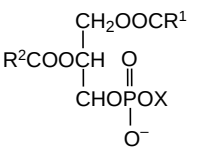
In plants glycerolipids are produced by wholly endogenous pathways but in animals there are additional routes by which dietary lipids are modified. The lipid composition of animals is influenced by dietary intake but is not completely defined by it.

1,2-Diacyl-*sn*-glycerols (Figure 3) are key intermediates in the biosynthesis of both triacylglycerols and phospholipids and are produced mainly from *sn*-glycerol-3-phosphate (a product of carbohydrate metabolism) by acylation of both free hydroxyl groups in separate stages followed by dephosphorylation. Further acylation of the *sn*-3 hydroxyl group gives triacylglycerols.

Phospholipids and sphingolipids (VA7000, VA7200)

Phospholipids and sphingolipids are constituents of cell membranes and they play an essential role in the synthesis of plasma lipoproteins and in the transduction of messages from cell surfaces to second messengers that control cellular processes. **Phosphatidylcholine** (Lecithin) is the most abundant phospholipid.

Sphingenine – the most common of the long-chain bases – is produced from Palmitic acid (as its CoA derivative) and **Serine** as shown in Figure 4. Such compounds are then acylated at the NH_2 group to give ceramides and further modified at the primary hydroxyl group to give sphingolipids (Figure 5).



X	Name of class	Abbreviation
H	phosphatidic acid	PA
$CH_2CH_2N^{\oplus}H_3$	phosphatidylethanolamine	PE
$CH_2CH_2N^{\oplus}Me_3$	phosphatidylcholine	PC
$CH_2CH(N^{\oplus}H_3)COOH$	phosphatidylserine	PS
$CH_2CH(OH)CH_2OH$	phosphatidylglycerol	PG
$C_6H_{11}O_6$	phosphatidylinositol	PI

Figure 4. Structures of the major phospholipids.

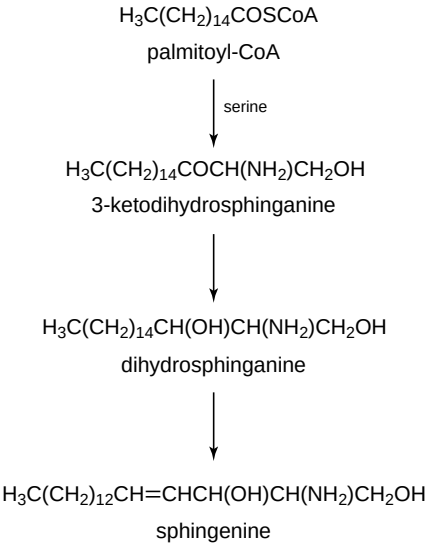


Figure 5. Biosynthesis of Sphingenine.

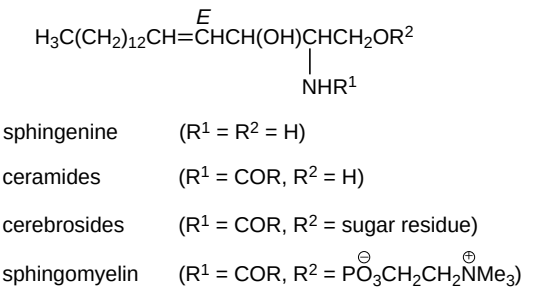


Figure 6. Structures of sphingolipids.

Alternatively the free hydroxyl group is converted to an appropriate phosphate ester to produce a phospholipid. Dietary triacylglycerols can be hydrolysed to 2-monoacyl-*sn*-glycerols and then reacylated to diacylglycerols and triacylglycerols.

Fattorusso, E. *et al.* (1997) *Prog. Chem. Org. Nat. Prod.*, **72**, 215 (*rev, glycolipids*)
Gunstone, F.D. *et al.* (1994) *The Lipid Handbook*, 2nd edn, Chapman & Hall, London.
Gunstone, F.D. (1996) *Fatty Acid and Lipid Chemistry*, Blackie, London.
Jie, M.S.F.L.K. *et al.* (1997) *Nat. Prod. Rep.*, **14**, 163 (*rev*).

Polyketides (VC)

Fungi have the ability to produce a very wide range of structural types of metabolite which are derived from a poly-β-ketomethylene chain. This chain is formed by condensation of an acetyl unit (or other acyl unit) with malonyl or methylmalonyl units, with concomitant decarboxylation as in fatty acid biosynthesis but without the reduction of the intermediate β-dicarbonyl system. The resulting polyketide chain can take part in internal aldol-type condensations to give aromatic systems characterised by an alternating oxygenation pattern. Alternatively reduction or partial reduction of the carbonyls during biosynthesis can give rise to nonaromatic metabolites. One method of classifying polyketides is by the number of acetate (or propionate) units in a metabolite; however, this has the disadvantage of separating structurally similar types. The vast array of polyketides is treated in DNP according to a mixture of structural, biosynthetic and functional criteria. The advantage of this approach is that related compounds are listed together. Aromatic polyketides are listed under the appropriate aromatic grouping.

Herbert, R.B. (1989) *The Biosynthesis of Secondary Metabolites*, 2nd edn, Chapman & Hall, London.

O'Hagan, D. (1991) *The Polyketide Metabolites*, Ellis Horwood, New York.

O'Hagan, D. (1992) *Nat. Prod. Rep.*, **9**, 447.

O'Hagan, D. (1995) *Nat. Prod. Rep.*, **9**, 447.

Simpson, T.J. (1991) *Nat. Prod. Rep.*, **12**, 1.

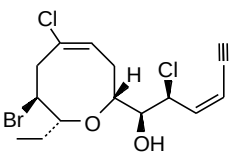
Turner, W.B. *et al.* (1983) *Fungal Metabolites II*, Academic Press, London.

Linear polyketides (VC0050)

This section contains a small number of polyketides that do not contain carbocyclic or macrolide ring systems.

Marine halogenated acetogenins (VC0070)

Marine metabolites include a series of halogenated polyketides particularly from red algae (*Laurencia* spp.). The metabolites contain, along with bromine and chlorine substituents, oxygen heterocycles, acetylenes and allenes. A typical example is Bermudenynol.

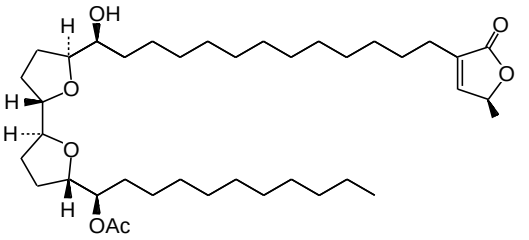


Bermudenynol

Faulkner, D.J. (1996) *Nat. Prod. Rep.*, **13**, 75.

Annonaceae acetogenins (VC0080)

The Annonaceae are a large family of tropical and subtropical trees. Several species contain compounds of apparent polyketide origin typified by the first example of this class, Uvaricin. They contain from 35 to 38 carbons, one, two or less commonly three tetrahydrofuran rings, a γ-lactone and various other oxygen functions and are characterised by a three carbon unit joined onto a long aliphatic chain. The determination of the stereochemistry of this group is often very difficult since they are generally waxy, amorphous compounds unsuitable for X-ray analysis.

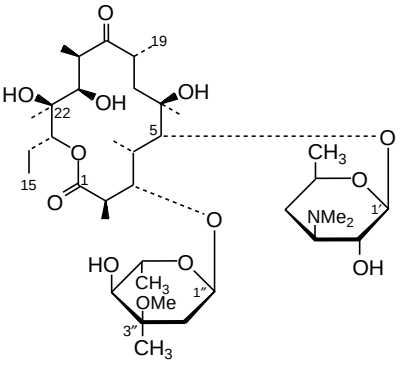


Uvaricin

Cavé, A. *et al*. (1993) *Recent Adv. Phytochem*, **27**, 167
Cavé, A. (1997) *Progr. Chem. Org. Nat. Prod*, **70**, 81
Figadere, B. *et al*. (1996) *Stand. Nat. Prod. Chem.*, **18**, 193 (*rev, synth*).
Gu, Z.M. *et al*. (1995) *Recent Adv. Phytochem.* **29**, 249
Gu, Z.M. *et al.* (1997) *J. Nat. Prod*., **60**, 242 (*chromatog, ms*)
Rupprecht, J.K. *et al.* (1990) *J. Nat. Prod.*, **53**, 237.
Zafra-Polo, M.C. *et al*. (1996) *Phytochemistry*, **42**, 253.
Zeng, L. *et al.* (1996) *Nat. Prod. Rep.*, **13**, 275.

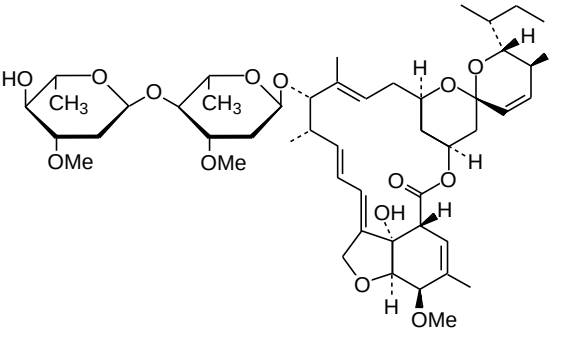
Macrolides and lactone polyketides (VC0100, VC0150)

Macrolide antibiotics are metabolites of *Streptomyces* and *Micromonospora* spp. Many antibiotics classified as macrolides have been reported for which full structures are not described. Structurally, macrolides are a class of complex glycosidic lactones; the aglycone is normally a 12–16 membered macrocyclic ring and one to three neutral or aminosugar residues are linked to the aglycone via ether linkages. Many of the aglycones have also been isolated from the fermentation broths, often from mutant strains, but these are usually devoid of biological activity. Erythromycin is a typical macrolide antibiotic.



Erythromycin

The related milbemycins and avermectins are a group of 16-membered macrocyclic lactones that possess an oxygen heterocyclic ring fused to the lactones.



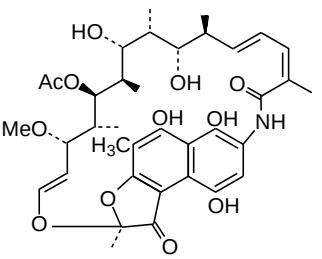
Avermectin A_{1a}

Blizzard, T. *et al.* (1990) *Recent Progr. Chem. Synth. Antibiot.*, 65 (*synth*).
Davies, H.G. *et al.* (1986) *Nat. Prod. Rep.*, **3**, 87 (*Avermectins, Milbemycins*).
Davies, H.G. *et al.* (1991) *Chem. Soc. Rev.*, **20**, 211; 271.
Fukagawa, Y. *et al.* (1988) *Life Sci. Rep.*, **6**, 267.
Kornis, G.I. *et al.* (1991) *ACS Sympos. Ser.* 443 (*Avermectins, Milbemycins*).
Nakata, M. *et al.* (1993) in *Studies in Natural Product Chemistry*, Vol. 12, (ed. Atta-ur-Rahman), Amsterdam, Elsevier, pp. 35 (*synth*).
Neuzil, J. *et al.* (1986) *Folia Microbiol. (Prague*), **31**, 402 (*biosynth*).
O'Hagan, D. (1989) *Nat. Prod. Rep.* **6**, 205 (*biosynth*).
Omura, S. (1984) *Macrolide Antibiotics, Chemistry, Biology and Practice*, Academic Press, London (*general*).
Omura, S. (1986) in *Biotechnology*, Vol. 4, (ed. H. Page), VCH, Weinheim, pp. 359 (*general*).
Paterson, I. *et al.* (1985) *Tetrahedron*, **41**, 3569 (*synth*).
Seno, E.T. *et al.* (1986) in *The Bacteria*, Vol IX, (eds S.W. Queener *et al.*), Academic Press, Orlando, pp. 231 (*biosynth*).
Tatsuta, K. (1990) *Recent Prog. Chem. Synth. Antibiot.*, 1 (*synth*).

Ansamycins and related polyketides (VC0200)

Ansamycins are benzenoid or naphthalenoid aromatic compounds in which non-adjacent positions are bridged by an aliphatic chain to form a cyclic structure. One of the aliphatic-aromatic junctions is always an amide bond. They are produced by *Streptomyces*, *Nocardia* and *Micromonospora* spp. and have also been isolated from plant sources; although for the latter, the involvement of microorganisms has not been ruled out.

The natural ansamycins may be subdivided according to the nature of the aromatic moiety and the length of the aliphatic chain. The major group contains a naphthalenoid moiety and a 17 carbon aliphatic chain. Rifamycin is a typical member of this group. The differences in structure are not merely of chemical interest but indicate a profound difference in biological activity. Members of this group show selective antibacterial activity and inhibit RNA polymerase. The benzenoid ansamycins with a 15-C chain include the **Ansamitocins** and the related **Maytansine**; these compounds show pronounced antitumour activity.



Rifamycin

Antosz, F.J. (1978) in *Kirk-Othmer Encyclopedia of Chemical Technology* (eds M. Grayson *et al.*) Wiley, NY, **2**, 852 (*isol*).
Crandall, L.W. *et al.* (1986) in *The Bacteria* Vol IX (ed. S.W. Queener) Academic Press, Orlando, pp. 360 (*isol*).
Isobe, M. (1990) *Recent Prog. Chem. Synth. Antibiot.*, 103 (*synth*).
Lancini, G. (1983) in *Biotechnology*, Vol 2, (ed. G. Lancini) VCH, Weinheim, Ger., pp. 231 (*biosynth*).
Lancini, G. (1986) in *Biotechnology*, Vol 4 (ed. H. Page) VCH, Weinheim, Ger., 431.
O'Hagan, D. (1989) *Nat. Prod. Rep.*, **6**, 205.
Reider, P.J. *et al.* (1984) in *The Alkaloids* (ed. R. Brossi *et al.*) Academic Press, **23**, 71.
Rickards, R.W. *et al.* (1991) *Stud. Nat. Prod. Chem.*, **9**, 431.
Smith, C.R. *et al.* (1984) in *Alkaloids, Chemical and Biological Perspectives* (ed. S.W. Pelletier) Wiley, New York. **2**, 149.
Traxler, P. *et al.* (1982) *J. Antibiot.*, **35**, 1361 (*biosynth*).

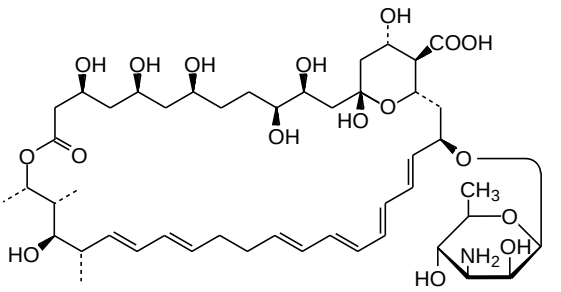
Polyenes (VC0300)

The group of antibiotics known collectively as polyenes is characterised by a large lactone ring (20−44 membered) containing a series of conjugated double bonds. This leads to the sub-division of the group into trienes, tetraenes etc. The macrolide ring is often linked by a hydroxyl group to an aminosugar unit and may have an aliphatic side chain possibly terminating with an aromatic residue.

Streptomyces are the usual producing organisms, and to date over 200 polyenes have been claimed. However, only some of these have established structures. One reason for the paucity of structural information is that they are often mixtures of closely related compounds. The advent of HPLC has enabled better separation to be obtained and has indicated that many polyenes previously considered to be defined were in fact mixtures of the same components but in different proportions.

The macrolide ring is probably derived from acetate and propionate, otherwise little is known about their detailed mechanism of biosynthesis.

Nystatin is a typical polyene antibiotic showing antifungal activity.



Nystatin A_1

Beau, J.M. (1990) *Recent Prog. Chem. Synth. Antibiot.*, 135 (*synth*).

Bolard, J. (1986) *Biochim. Biophys. Acta.*, **864**, 257 (*props*).

Crandall, L.W. *et al.* (1986) in *The Bacteria*, Vol IX (ed. S.W. Queener) Academic Press, Orlando.

Omura, S. (1984) *Macrolide Antibiotics, Chemistry, Biology and Practice*, Academic Press, London.

Rinehart, K.L. (1983) *Biotechnology*, **1**, 581 (*ms*).

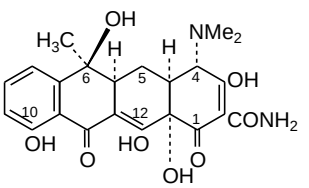
Rychnovsky, S.D. (1990) *Acta Pharm. Nord.*, **2**, 155.

Simpson, T.J. (1985) *Nat. Prod. Rep.*, **2**, 321 (*biosynth*).

Thomas, A.H. (1986) *J. Antimicrob. Chemother.*, **17**, 269 (*action*).

Linear tetracyclines (VC0400)

The tetracyclines, which contain a polyhydronaphthacene nucleus, form a small but very important group of antibiotics. Many of the *Streptomyces* metabolites have been used clinically since their discovery in the late 1940s. They are active against gram-positive and gram-negative bacteria, spirochaetes, mycoplasmas and rickettsiae. In addition they display significant amoebicidal activity and have efficacy in some diseases caused by large viruses. They have veterinary applications in promoting growth and feed efficiency. They are second to the β-lactam group in terms of clinical use and exhibit low toxicity and good oral absorption. Their mode of action is by the inhibition of protein biosynthesis.



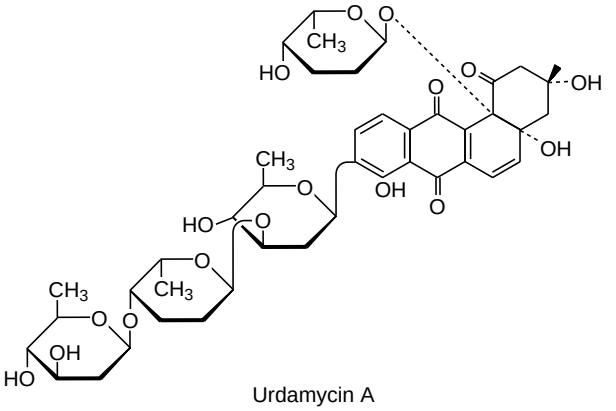
Tetracycline

The biochemistry of tetracycline production has been extensively studied using mutant strains and cell-free systems to identify a variety of intermediates. Biosynthetically tetracycline antibiotics are derived from oligoketides.

Aszalos, A. (1985) *Chromatographia*, **20**, 313 (*hplc*).
Hlavka, J.J. *et al.* (1985) *The Tetracyclines, Handbook of Experimental Pharmacology*, Springer, Heidelberg, pp. 78.
Krohn, K. *et al.* (1989) *Prog. Chem. Org. Nat. Prod.*, **55**, 37.
Mitscher, L.A. (1978) *The Chemistry of the Tetracycline Antibiotics*, Marcel Dekker, New York.
Mooibroek, S. *et al.* (1987) *Can. J. Chem.*, **65**, 357 (*cmr, bibl*).
Turner, W.B. *et al.* (1983) *Fungal Metabolites II*, Academic Press, London.

Angucyclines (VC0450)

The angucycline antibiotics are related to the tetracyclines but they have an angular arrangement of the tetracyclic ring system as in Urdamycin A. Angucyclinones are defined as natural products with a benz[*a*]anthracene nucleus but no hydrolysable sugar moieties whereas the term angucycline includes those with hydrolysable sugars.



Urdamycin A

Rohr, J. *et al.* (1992) *Nat. Prod. Rep.*, **9**, 103.

Polyether antibiotics (VC0500)

The majority of polyethers are characterised by a linear series of tetrahydrofuran and tetrahydropyran residues, frequently linked by spiroketal systems. These compounds always terminate with a carboxylic acid residue or a simple ester function thereof. Some polyethers also carry a sugar unit linked to a hydroxy group on one of the tetrahydropyran rings. The most common sugar residue is 4-*O*-methylamicetose.

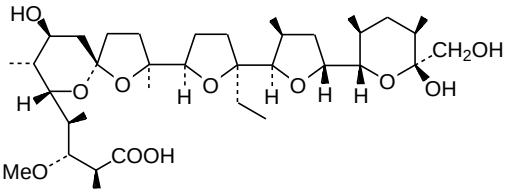
More than 1000 polyether antibiotics have been isolated so far, mostly as metabolites of *Streptomyces* spp., although some *Streptoverticillium, Actinomadura, Nocardia* and *Dactylosporangium* spp. are also reported to produce them. Polyethers are generally produced as a series of closely related compounds e.g. the major component may possess methyl substituents on each of the cyclic ether units, but in addition small amounts of ethyl homologues may also be present.

Chemical subdivision is based on the number of spiroketal functionalities, and the presence or absence of a sugar residue.

Polyethers possess the ability to bind and transport certain ions, and each antibiotic has its own ion specificity. For this reason they are important biochemical tools in studying the role of cations in biological systems. The antibiotics show a wide range of activities, being active against gram-positive organisms and mycobacteria, fungi and yeasts, but because of their toxicity,

these properties have found little application. Their uses to date are mainly as feed additives.

Biosynthetically, the polyethers are polyketide in origin. The major building blocks are acetate, propanoate, and butyrate. There is evidence to suggest the intermediacy of an epoxide in the formation of the tetrahydrofuran and tetrahydropyran systems. Monensin A is a typical polyether antibiotic.



Monensin A

Berdy, J. (1986) in *Biotechnology*, Vol 4, (ed. H. Page) VCH, Weinheim, Ger., pp. 494.

Crandall, L.W. *et al.* (1986) in *The Bacteria*, Vol IX (ed. S.W. Queener *et al.*) Academic Press, Orlando, pp. 385.

Denyer, S.P. *et al.* (1983) *Antibiotics*, Society for Applied Bacteriology, Washington, 77.

Dutton, C.J. *et al.* (1995) *Nat. Prod. Rep.*, **12**, 165.

Robinson, J.A. (1991) *Prog. Chem. Org. Nat. Prod.*, **58**, 1.

Siegel, M.M. *et al.* (1987) *Biomed. Environ. Mass Spectrom.*, **14**, 29 (*ms*).

Westley, J.W. (ed.) (1982) *Polyether Antibiotics*, Marcel Dekker, NY, (2 vols).

Westley, J.W. (1986) *J. Nat. Prod.*, **49**, 35 (*biosynth*).

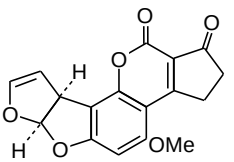
Wieranga, W. (1981) in *Total Synthesis of Natural Products* (ed. J. Ap'Simon) Wiley, New York.

Yonemitsu, O. *et al.* (1990) *Recent Prog. Chem. Synth. Antibiot.*, 447 (*synth*).

Aflatoxins and related substances (VC0600)

Structurally, aflatoxins consist of a hydrogenated difurano-moiety fused to a substituted coumarin. The naturally occurring aflatoxins are acutely toxic and extremely carcinogenic compounds produced by *Aspergillus* spp. Metabolism of these compounds by microbial and animal species or chemical transformation leads to a number of equally potent aflatoxin derivatives. Toxic effects centre primarily on the liver.

The formation of the principal toxin, Aflatoxin B_1, has been studied in considerable detail. The results are consistent with a pathway from a single decaketide chain via a series of intermediates e.g. **Averufin** and **Sterigmatocystin**. The other aflatoxins are formed from Aflatoxin B_1.



Aflatoxin B_1

Heathcote, J.G. *et al.* (1978) *Aflatoxins, Chemical and Biological Aspects*, Elsevier, Amsterdam.

Lacey, J. (1987) *Trichothecenes and other Mycotoxins*, Wiley, New York.

Simpson, T.J. (1984) *Nat. Prod. Rep.*, **1**, 287 (*biosynth*).

Steyn, P.S. *et al.* (1980) in *Biosynthesis of Mycotoxins*, Academic Press, New York (*biosynth*).

Steyn, P.S. *et al.* (eds) (1986) *Mycotoxins and Phycotoxins*, Elsevier, Amsterdam.

Carbohydrates (VE)

This is an abbreviated account dealing only with aspects of carbohydrate chemistry relevant to natural products. For a fuller coverage including synthetic carbohydrates, see the companion disc *Dictionary of Carbohydrates on CD-ROM.*

Carbohydrates comprise a family of polyhydroxy aldehydes, ketones and acids, together with linear and cyclic polyols. They are diverse because they exist as a wide range of stereoisomers.

These compounds are the most abundant and widespread organic substances in nature and are essential constituents of all living matter. They are the most important (in terms of volume and availability) of the non-nitrogenous

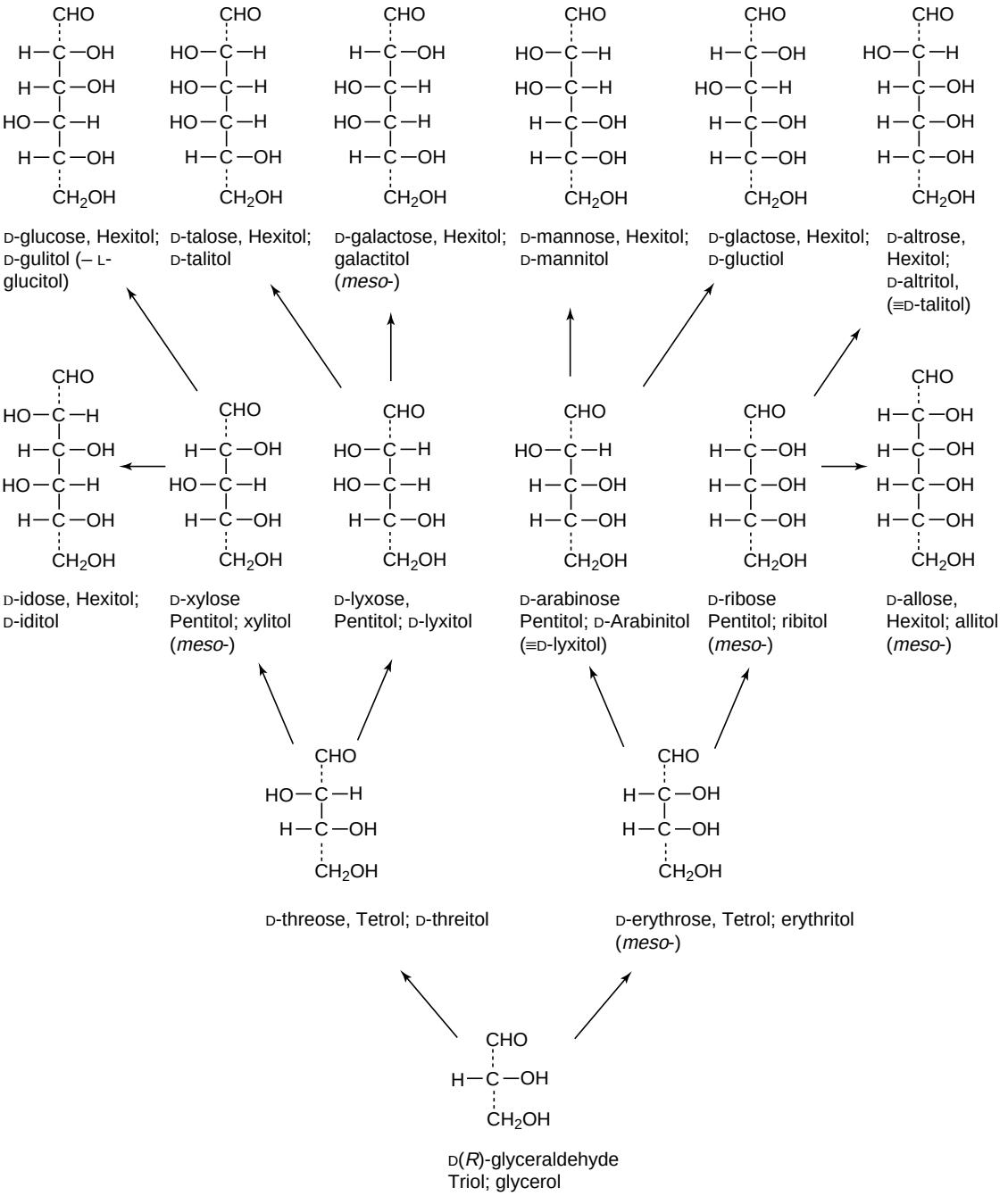


Figure 7. The fundamental aldoses, their corresponding alditols and their notional relationships to glyceraldehyde, as illustrated for sugars of the D-series.

constituents of the chiral pool and are extremely important in chiral synthesis. Of the 36 possible stereoisomeric pentoses, pentuloses, hexoses and hexuloses only D-glucose, D-fructose, D-galactose, D-mannose and L-arabinose occur naturally in the free state, but only the first two are found in significant amounts.

Photosynthesis is the means by which plants produce sugars from carbon dioxide and water. In brief, it occurs by carbon dioxide being transferred to D-*erythro*-pentulose-1,5-diphosphate to give, via an unstable β-keto-6-carbon acid, two molecules of D-glyceric acid-3-phosphate, from which hexoses, for example, D-fructose 1,6-diphosphate and D-glucose 1-phosphate can be formed. Animals on the other hand use the reverse of the glycolysis metabolic pathway to produce glucose from proteins and fats utilising phosphoenolpyruvate as an intermediate. Most of the routes used by nature to interconvert sugars occur by way of enzymic reactions on nucleoside diphosphate sugars, particularly **Uridine diphosphate glucose** (UDPG) which gives D-galactose on epimerization at C-4, D-glucuronic acid by oxidation at C-6 and D-xylose by decarboxylation of this acid. Deoxygenation at C-6 and configuration changes at C-4 and C-5 give L-rhamnose and by similar means the commonly occurring D-sugars may be transformed into members of the L-series.

Fundamental aldoses and ketoses (VE0100–VE2200)

Figures 7, 8 and 9 illustrate the derivation of the (natural and non-natural) fundamental monosaccharides by the notional chain-lengthening process from their C_3 parents giving rise to sugars of the D- or L-series (illustrated here for D-sugars).

In DNP, the cyclic forms of the sugars are normally illustrated as Haworth formulae as shown in Figure 8. Wherever possible, the sugar components of complex molecules such as antibiotics are shown in the Dictionary in the standard Haworth orientation so that rapid configurational comparison with

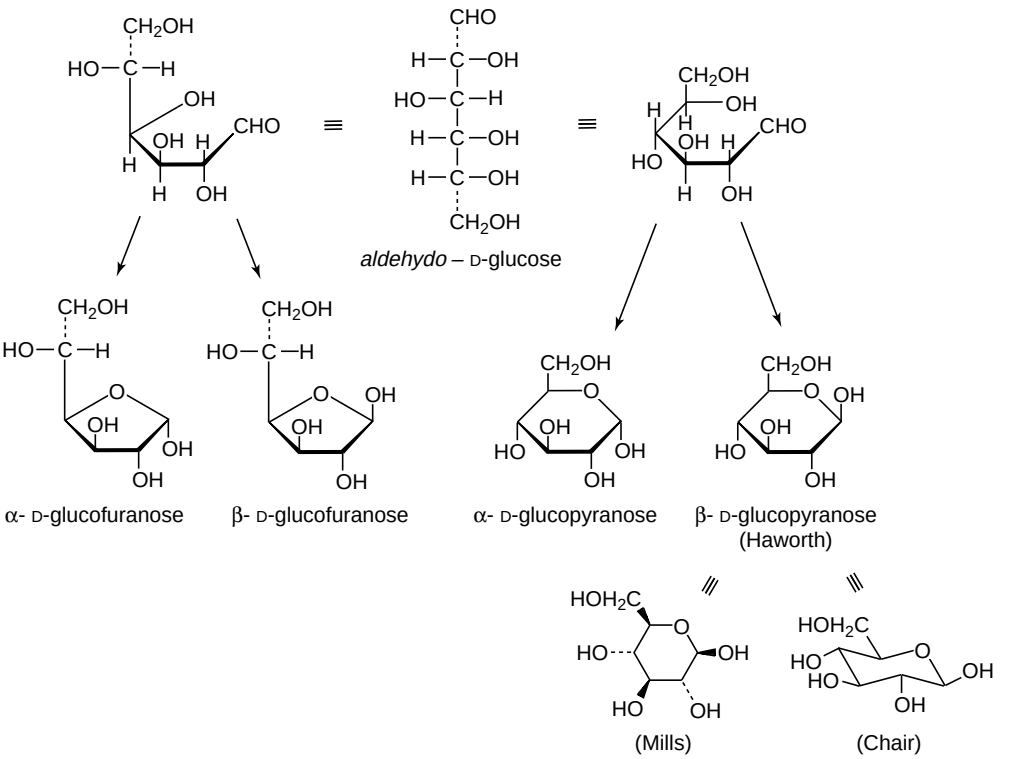


Figure 8. Anomerism and ring formation in a simple aldose as exemplified by D-glucose.

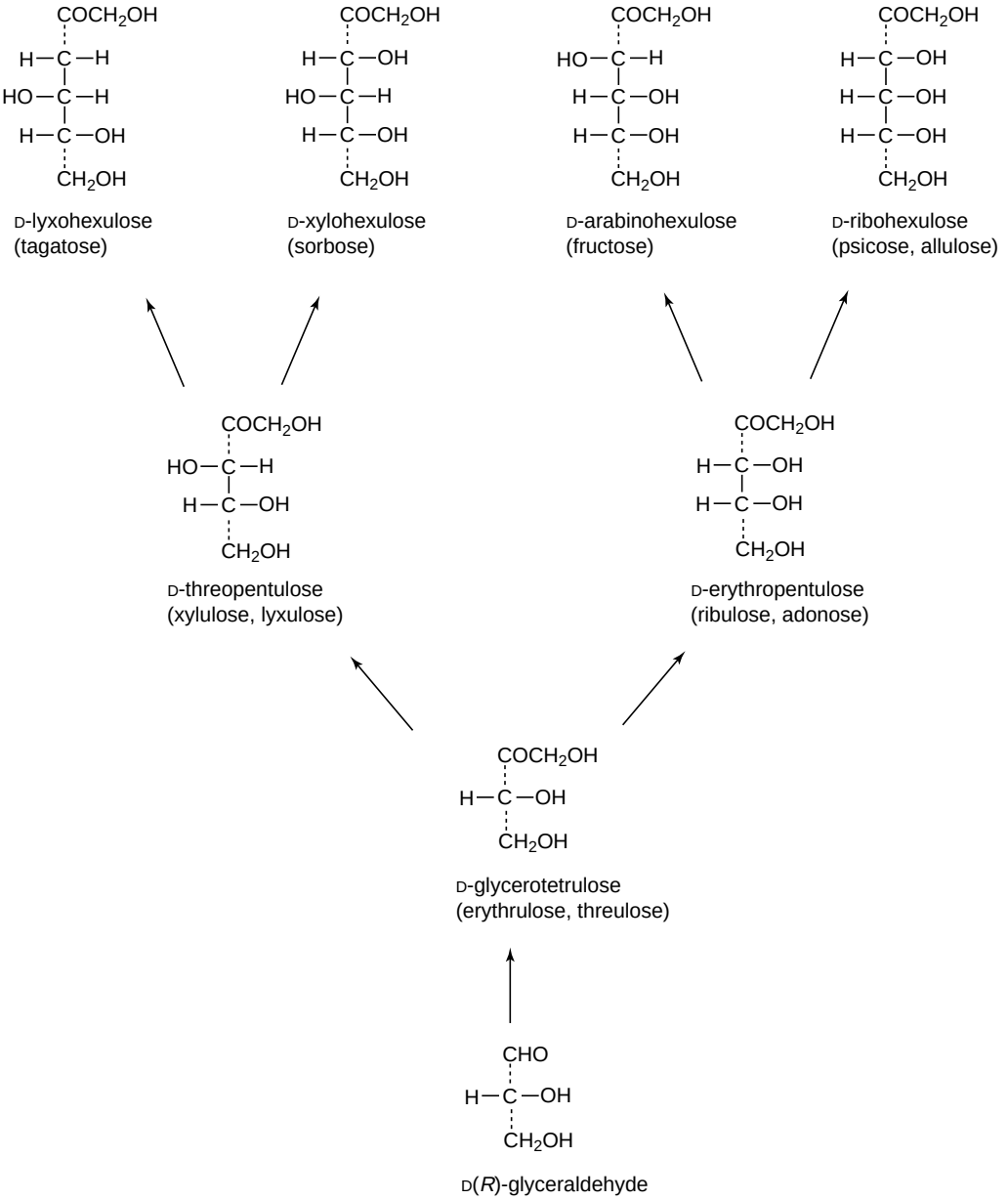
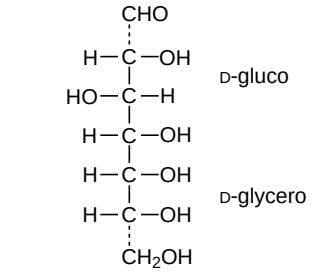


Figure 9. The fundamental ketoses and their notional relationships to glyceraldehyde, illustrated for sugars of the D-series.

other related structures can be made. Alternative representations for β-D-glucopyranose are shown in Figure 8 according to the planar ring (Mills formula) and chair conventions which are often encountered. The latter should be used with caution since it has implications of conformational preference which may not correspond to reality in all cases.

In the Type of Compound index the simple sugars documented in DNP are classified into their various stereoisomeric subgroups.

Higher sugars having more than six carbon atoms, some of which occur naturally, are named using two prefixes, one defining the relative configuration at the last four carbon atoms in the chain (C-2 to C-5 in a hexose), and the other, which appears first in the name, defining the configuration at the remaining chiral centre(s).



D-*glycero*-D-*gluco*-heptose

Modified aldoses and ketoses (VE2600–VE8400)

The number of natural sugars increases when modified forms of these 36 fundamental sugars are considered. Thus in addition to the common occurrence of combined forms of the five sugars mentioned above, D-allose, D-talose, D-arabinose, D-ribose, D-xylose, L-lyxose, D-psicose, L-sorbose and D-tagatose are also found as their derivatives in varying quantities. It is rather surprising that the vast number of naturally occurring carbohydrate compounds are derived from so few sugars in this pool. The shortfall is made up by the occurrence of so called modified sugars such as deoxy-, amino-, thio-, branched chain, and higher sugars in addition to various alditols, cyclitols and sugar acids.

Bacteria contain several sugars that are unique to their constitution. **Muramic acid**, glycosidically linked to **N-Acetylglucosamine**, is the disaccharide repeating unit that forms the peptidoglycan of gram-negative bacterial cell walls. Several rare deoxy-sugars such as **Paratose** and **Tyvelose** are components of the 'antigenic' outer cell wall and these inner and outer regions are linked through a unique ketodeoxyoctulosonic acid (Kdo). In gram-positive bacteria, teichoic acids, which are large polymers of the phosphates of D-ribitol or glycerol, form up to 50% of the cell wall.

Bacterially-produced antibiotics are a rich source of rarer sugars. For example, the **Neomycins, Kanamycins** and **Paromomycins** contain a variety of amino-sugars. **Streptomycin** also contains a branched-chain sugar and the anthracycline, **Daunomycin**, possesses a dideoxyamino-sugar. **Nojirimycin**, the α-glucosidase inhibitor, is 5-amino-5-deoxyglucose and the enediyne antitumor antibiotic, **Calicheamicin**, has 4,6-dideoxy-4-thio-*ribo*-hexose as part of its carbohydrate structure.

Replacement of a hydroxyl hydrogen, other than at the anomeric position by an alkyl or aryl group, gives an ether named as an *O*-substituted sugar (e.g. 3-*O*-methylglucose). This, however, is not the case when the anomeric hydroxyl is involved. The product is then named as a glycoside when the aglycone is obtained from a relatively simple alcohol (e.g. methyl β-D-glucopyranoside not 1-*O*-methyl-β-D-glucose). For more complicated aglycones the prefixes glycosyloxy or *O*-glycosyl can be used. For example 3α- (β-D-glucopyranosyloxy)-5-β-pregnan-20-ol can also be named as 3-*O*-(-β-D-glucopyranosyl)-5β-pregnane-3α, 20-diol. The latter system is often used to name oligosaccharides (see later). However, when the anomeric hydroxyl is esterified the product can be referred to as a 1-*O*-substituted sugar or alternatively, the glycosyl prefix can be used (e.g. 1-*O*-acetyl-α-D-glucopyranose or α-D-glucopyranosyl acetate); for esters of phosphoric acid 1-phospho-α-D-glucopyranose or α-D-glucopyranosyl phosphate may be used, but D-glucose 1-phosphate is in common use.

N-Glycosides can be conveniently named glycosylamines. Thus, for example, N^4-(2-acetamido-2-deoxy-β-D-glucopyranosyl)asparagine can also be called 2-acetamido-N^1-(β-aspartyl)-2-deoxy-β-D-glucopyranosylamine.

In many natural products hydroxy groups other than at the anomeric centre of the sugar are replaced by a thiol-group, an amino-group or a hydrogen atom. Compounds arising from these changes are named respectively as thio-sugars (e.g. 3-thioglucose), aminodeoxysugars (e.g. 3-amino-3-deoxyglucose) and deoxy sugars but in this case the configuration of the remaining asymmetric carbons must be described with a prefix (e.g. 3-deoxy-*ribo*-hexose not 3-deoxyglucose).

Branched-chain sugars (VE7200)

Carbon chain branching in sugars can arise biogenetically in two ways; either *C*-bonded hydrogen atoms are replaced as in HO-H→ HO-R in which case the products are *C*-substituted derivatives of the normal straight-chain compounds and are classified as members of the 'dehydro' group, or else hydroxyl functions are replaced as in HO-H→ R-H. In the naming of the latter class the 'deoxy' prefix is included to denote the absence of the hydroxyl substituent at the branching carbon atom, and members can be described as belonging to the 'deoxy' group of branched chain sugars. (e.g. 3-*C*-methyl-D-glucose and 3-deoxy-3-*C*-methyl-D-glucose are the respective names of compounds obtained by replacing in glucose either the hydrogen at *C*-3 or the hydroxyl at *C*-3 by methyl).

Carbohydrate acids VE7900, VE8000, VE8100, VE8200)

The following four types of carbohydrate acids occur in nature for which named examples are given for compounds derived from glucose: aldonic acids (VE7900) (D-**Gluconic acid**) which are formed when the aldehydic function in an aldose is oxidized; aldaric acids (VE8100) (**D-Glucaric acid**) which are dicarboxylic acids formed by oxidation of the aldehydic groups and hydroxymethyl groups in aldoses; uronic acids (VE8000) (**D-Glucuronic acid**) and ketoaldonic acids (VE8200) (**D-*arabino*-Hex-2-ulosonic acid**) which are formed by oxidation of the hydroxymethyl groups in aldoses and ketoses respectively.

Glycopyranosides (e.g. methyl) and esters (e.g. benzyl) of the last two acids are named in the following way: benzyl(methyl-α-D-glucopyranosid)uronate for the former and benzyl(methyl α-D-*arabino*-hex-2-ulopyranosid)onate for the latter.

Alditols (VE8600–VE8900)

The polyols obtained by reduction of the aldehyde function of an alditol (or the keto function of a ketol) are known as alditols. An example is **Mannitol**. They are named by a straightforward extension of the rules used for aldoses. The alditol corresponding to a chiral sugar may be *meso*-, e.g. **Galactitol**.

Cyclitols (VE9000)

The polyhydroxycycloalkanes, known as cyclitols, are a group of natural products closely related to the carbohydrates proper, of which the most important are the inositols (1,2,3,4,5,6-cyclohexanehexols). Trivial names are often used but systematic rules have been introduced to assign configurations at each enumerated ring carbon atom and this requires the application of a recommended numbering convention. Further information on the various descriptions of stereochemistry for these compounds can be obtained by the inspection of the individual Dictionary entries. It should be noted that some *meso*-isomers in the series can have optically active derivatives.

***myo*-Inositol** is the most abundant cyclitol, occurring both free and as its derivatives. Its hexaphosphate is phytic acid and occurs in large amounts in grain. *myo*-Inositol and its derivatives are universally present in cells and the 1,4,5-trisphosphate plays a vital role as a secondary messenger, which mediates mobilization of intercellular calcium ions.

Anderson, L. (1972) in *The Carbohydrates*, (ed. W. Pigman and D. Horton) Academic Press, **IA**, 519.

Angyal, S.J. and Anderson, L. (1959) *Adv. Carbohydr. Chem.*, **14**, 135.

Posternak, Th. (1965) *The Cyclitols*, Holden-Day, San Francisco.

Reitz, A.B. (1991) *Inositol Phosphates and Derivatives*, ACS Symposium Series, Washington, DC.

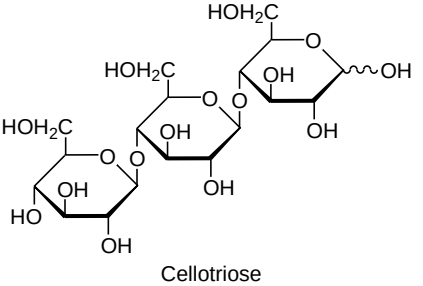
Disaccharides (VE9200)

These are formed by sugars combining with each other by way of glycosidic links. If the anomeric position of a sugar is attached to the anomeric oxygen of another, then non-reducing disaccharides such as the glucosylglucoside, **Trehalose**, or the fructosyl-glucoside, **Sucrose**, are formed. The latter compound is widespread, occurring in most plants. It constitutes a significant part of man's diet in Europe and the USA, being produced in pure form on a larger scale than any other monomeric organic compound. If the glycosidic bond in a disaccharide is to a non-anomeric hydroxyl of the second sugar, then so-called reducing disaccharides are formed, **Maltose** and **Lactose** being typical examples.

Oligosaccharides and polysaccharides (VE9300, VE9400)

These are obtained by repetition of the glycosylation process on reducing disaccharides. The former are most often found as glycosides in, for example, plants, antibiotics and some glycoproteins. Polysaccharides are the most abundant form of carbohydrates. **Cellulose**, for example, is an industrial raw material with world consumption approaching 10^9 tons p.a. It is the principal constituent of plant cell walls providing their structural strength. **Starch** and **Glycogen** are found preponderantly in plants and animals respectively where they serve as energy reserves. Whereas glucose is the building unit for the previous three polymers, **Chitin**, which is found in the shells of arthropods, is a polymer of 2-acetamido-2-deoxyglucose.

Oligosaccharides with a free hemiacetal group are named as glycosylglycosylglycoses as illustrated by Cellotriose, for example, which is β-D-glucopyranosyl-(1 → 4)-β-D-glucopyranosyl-(1 → 4)-D-glucose (*Chemical Abstracts* inserts *O*-locants as in *O*-β-D-glucopyranosyl-(1 → 4)-*O*-β-D-glucopyranosyl-D-glucose). Branched oligosaccharides use square brackets in the name to designate branching, e.g. α-D-glucopyranosyl-(1 → 4)-[α-D-glucopyranosyl (1 → 6)]-D-glucopyranose. Polysaccharides use an extended form of this nomenclature.



Cellotriose

Glycosidic bonds in naturally occurring oligosaccharides and glycosides are formed in natural glycosylations which take place primarily by way of the nucleoside diphosphate sugars as follows:

$$\text{nucleoside DP sugar} + \text{ROH} \rightarrow \text{sugar OR} + \text{nucleoside DPH}$$

Disaccharides or their phosphates are produced when ROH is a sugar or a sugar phosphate. Polysaccharide biosynthesis is basically similar but requires an oligomer primer as an acceptor; glycogen synthesis follows the course:

$$\text{primer} + \text{UDPG} \rightarrow \text{G}\alpha\text{-primer} + \text{UDP}$$

there being one enzyme present which catalyses the formation of 1,4-bonds and another responsible for glycosylations at position 6. The biosynthesis of cellulose and other polysaccharides is basically similar, UDP being the nucleoside diphosphate used predominantly. However, starch synthesis depends rather on adenosine diphosphate.

Carbohydrates are important components of a variety of significant biopolymers, the most obvious being DNA and RNA which contain **2-Deoxy-D-ribose** and **D-Ribose** respectively. Glycoproteins and glycolipids (described elsewhere) are another important group, although the carbohydrates are usually present as oligomers, comprising sugars drawn from the following: 2-acetamido-2-deoxy-D-glucose, 2-acetamido-2-deoxy-D-galactose, D-mannose, D-galactose, sialic acid and L-fucose, which are glycosidically attached to the protein. They are widely distributed in all living organisms, occurring bound to cell membranes and in a free soluble form. Their role has only recently been appreciated as it becomes evident that they serve as recognition sites for a variety of intra- and intermolecular communication events. They function as specific binding sites for enzymes, hormones, soluble toxins, bacteria, and viruses. They are also implicated in cell-cell recognition and interaction, such as cell adhesion which is particularly important in inflammatory diseases. They play a role in the development of normal cells and metastasis in cancer cells. The structural basis for blood ABO(H) and Lewis group antigenicities resides in the oligosaccharide portions of blood cell glycolipids and the associated secreted glycoproteins.

Glycosaminoglycans (included in VE9400)

Mucopolysaccharides are another group of high molecular weight sugars usually formed by polymerization of a disaccharide. Dermatan sulfate, for example, contains L-iduronic acid and D-galactosamine 4-sulfate as the repeating unit. Others in this group are **Chondroitin**, **Keratan sulfate**, **Hyaluronic acid** and **Heparin** which are found in body fluids and associated with connective tissue. Mucopolysaccharidosis comprises several rare and fatal metabolic diseases, among them being Hurler's syndrome and Hunter's disease, in which some of these compounds accumulate to abnormal levels in the tissues of affected individuals.

Plant glycosides

Glycosides of many different aglycones are ubiquitous in the plant kingdom. Many are formed from phenols, polyphenols, steroidal and terpenoidal alcohols by glycosidic attachment to sugars. In a majority of cases D-glucose is present but L-rhamnose, D- and L- fucose and L-arabinose occur quite frequently. Of the pentoses, L-arabinose is more common than D-xylose and the sugars often occur as oligosaccharides. For example, **Digitonin** and **Digitoxin** from *Digitalis* contain respectively a branched heteropentasaccharide, comprising two glucose

and two galactose molecules and xylose, and a linear homotrisaccharide of 2,6-dideoxy-*β*-D-*ribo*-hexose.

Tannins (see separate section below) are a special case of plant glycoside.

In general, the plant glycosides are so numerous that their sugar components are not reported as such in the carbohydrate section of the Type of Compound Index, except for a few special classes described below.

Casu, B. (1985) *Adv. Carbohydr. Chem. Biochem.*, **43**, 51.
Collins, P.M. (1997) *Dictionary of Carbohydrates*, Chapman & Hall.
Courtois, J.E. and Percheron, F. (1970) in *The Carbohydrates*, (eds W. Pigman and D. Horton) Academic Press, **II**, 213.
Hassid, W.Z. (1970) in *The Carbohydrates*, (eds W. Pigman and D. Horton) Academic Press, **IIA**, 301.
Hughes, R.C. (1983) *Glycoproteins*, Chapman & Hall, London.
Jeanloz, R.W. (1970) in *The Carbohydrates*, (eds W. Pigman and D. Horton) Academic Press, **IIB**, 590.
Lee, M.D. *et al.* (1991) *Acc. Chem. Res.*, **24**, 235.
Lemieux, R.U. (1978) *Chem. Soc. Revs.*, **7**, 423.
Lindberg, B. (1990) *Adv. Carbohydr. Chem. Biochem.*, **48**, 279.
Mallams, A.K. (1988) in *Carbohydrate Chemistry* (ed. J.F. Kennedy) Clarendon Press, Oxford, pp. 73.
Montreuil, J. (1980) *Adv. Carbohydr. Chem. Biochem.*, **37**, 157.
Rademacher, T.W. *et al.* (1988) *Ann. Rev. Biochem.*, **57**, 785.
Schmidt, O. Th. (1956) *Fortschr. Chem. Org. Naturst.*, **13**, 70.
Sharon, N. and Lis, H. (1993) *Sci. Am.*, 74.
Umezawa, S. (1974) *Adv. Carbohydr. Chem. Biochem.*, **30**, 111.
IUPAC/IUBMB Joint Commission on biochemical nomenclature of Carbohydrates, *Pure Appl. Chem.*, 1996, **68**, 1919.

Aminoglycoside antibiotics

Aminoglycosides constitute a large and diverse group of metabolites produced by both bacteria and *Streptomyces*. The group covers those antibiotics containing a highly functionalised cyclohexane aglycone, i.e. cyclitol, glycosidically linked to amino or neutral sugar residues.

The streptamine cyclitol sub-group are *Streptomyces* metabolites which have broad spectrum activity. They find clinical use both as topical and systemic agents but exhibit varying degrees of oto- and nephrotoxicity thus limiting their application. In addition many are rendered ineffective by resistant strains carrying aminoglycoside inactivating enzymes.

The group of aminoglycosides containing inositol and/or monoamino-cyclitols, which are produced by *Streptomyces* and *Nocardia* spp. represent a much smaller subdivision.

In the Type of Compound Index, the aminoglycoside antibiotics are listed under one or more structural categories (e.g. cyclitol), but a general bibliography is given here.

Gambardella, P. *et al.* (1985) *J. Chromatogr.*, **348**, 229 (*hplc*).
Gero, S.D. *et al.* (1984) *Stud. Org. Chem. (Amsterdam)*, **17**, 79 (*synth*).
Grisebach, H. (1978) *Adv. Carbohydr. Chem. Biochem.*, **35**, 122 (*biosynth*).
Inchauspe, G. *et al.* (1985) *J. Antibiot.*, **38**, 1526 (*hplc*).
Okachi, R. *et al.* (1984) *Drugs Pharm. Sci.*, **22**, 329 (*biosynth*).
Rinehart, K.L. (ed.) (1980) *Aminocyclitol Antibiotics*, Amer. Chem. Soc.
Schubert, J. *et al.* (1986) *Justus Liebigs Ann. Chem.*, 2009 (*synth*).
Umezawa, H. *et al.* (1982) *Handbook of Experimental Pharmacology*, Springer, Berlin, 62 (*biosynth*).
Umezawa, S. (1986) in *Biotechnology*, Vol 4 (ed. H. Page), VCH, Weinheim, Ger., pp. 309.

Whelton, A. *et al.* (1982) *The Aminoglycosides*, Marcel Dekker, New York.
Williams, N.R. (ed.) (1986) *Carbohydr. Chem.*, Royal Society of Chemistry, London, **18**, 176.

Nucleosides (VE9900)

These are glycosides of purines, pyrimidines and other heterocyclic bases. The well-known quartet of **Adenosine, Guanosine, Cytosine** and **Thymidine** are fundamental biomolecules essential to life through their participation in the structure of DNA and RNA. A small number of 'hypermodified' nucleosides such as **Wybutosine** occur in bacterial nucleic acids.

A more prolific source of different nucleoside structures is the nucleoside antibiotics which are analogues of the essential purine and pyrimidine nucleosides. They consist of a sugar linked to a base either via a ring nitrogen or through a ring C atom (the latter are designated *C*-nucleosides). Structurally they are rather diverse but a subclassification is given by Isono (*loc. cit.*). Although most of the compounds are *Streptomyces* metabolites, fungal and bacterial products have also been identified.

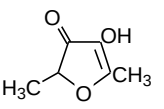
Many nucleosides show interesting antiviral and antitumour properties, and find wide application as biochemical tools. The commercial importance of synthetic and semisynthetic compounds has led to a concentration of research effort on synthetic methods and several good reviews on this subject are available.

A few nucleotides (nucleoside phosphate conjugates) are also found, e.g. **Agrocin 84**.

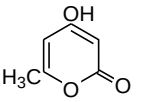
Brown, E.G. (1991) *Methods Plant Biochem.*, **5**, 53.
Buchanan, J.G. (1982) *Top. Antibiotic Chem.*, **6**, 229.
Buchanan, J.G. (1983) *Prog. Chem. Org. Nat. Prod.*, **44**, 243 (*C-nucleosides*).
De las Heras, F.G. *et al.* (1990) *Recent Prog. Chem. Synth. Antibiot.*, 321.
Eckardt, K. (1983) *J. Nat. Prod.*, **46**, 544.
Garner, P.P. (1988) *Stud. Nat. Prod. Chem.*, **1**, 397 (*synth*).
Grisebach, H. (1978) *Adv. Carbohydr. Chem. Biochem.*, **35**, 122 (*biosynth*).
Hobbs, J.B. (1993) in *The Chemistry of Natural Products*, 2nd edn (ed. R.H. Thomson), Blackie, Glasgow, pp. 259.
Isono, K. (1988) *J. Antibiot.*, **41**, 1711 (*biosynth, struct*).
McCloskey, J.A. (1990) *Methods Enzymol.*, **193**, 771 (*anal, ms*).
Secrist, J.A. *et al.* (eds) (1989) *Nucleosides Nucleotides*, **8**, parts 5 and 6 (*rev*).
Suhadolnik, R.J. (1979) *Nucleosides as Biological Probes*, Wiley, New York.
Suhadolnik, R.J. (1981) *Antibiotics (N.Y.)*, **4**, 353 (*biosynth*).
Townsend, L.B. (1988) *Chem. of Nucleosides and Nucleotides Vol. 1*, Plenum Press, New York.
Williams, N.R. (ed.) (1986) *Carbohydrate Chemistry*, Royal Society of Chemistry, London, **18**, 176, 190.

Oxygen heterocycles (VF)

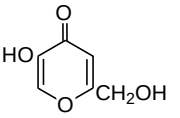
Many simple natural products contain basic oxygen heterocycles – for example the furan derivative, Furaneol, the pyran-2-one derivative, Triacetic acid lactone and the 4-pyrone, Kojic acid. Although most of these simple oxygen heterocyclic compounds can be seen to be derived from polyketides or carbohydrates, some have unknown biosynthetic origins. The oxygen heterocycles are listed under the headings: β-Lactones (VF1000), Furans (VF2000), Butanolides (VF3000), Pyrans (VF4000), Pentanolides (VF5000), 2-Pyrones (VF6000) and 4-Pyrones (VF7000). Natural products that contain these substructures in terpenoid, steroid or alkaloid skeletons are not listed here.



Furaneol



Triacetic acid lactone



Kojic acid

Davies-Coleman, M.T. *et al.* (1989) *Prog. Chem. Org. Nat. Prod.*, **55**, 1.

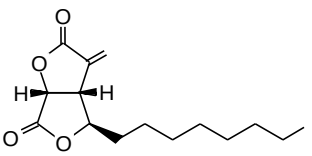
Dickinson, J.M. (1993) *Nat. Prod. Rep.*, **10**, 71.

Ley, S.V. (1991) in *Heterocycles in Bioorganic Chemistry*, (eds J. Bergman *et al.*), RSC, London.

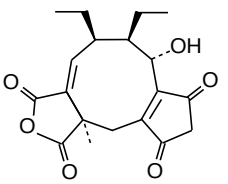
Turner, W.B. *et al.* (1983) *Fungal Metabolites II*, Academic Press, London.

Avenaciolide and glauconic acid groups (VF5100, VF5200)

Separate listings are given for the bislactones related to Avenaciolide and the dimeric anhydrides related to Glauconic acid.



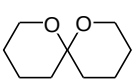
Avenaciolide



Glauconic acid

Spiroketals (VF8000)

There are a number of biologically active spiroketals, clearly of acetate origin, exemplified by 1,7-Dioxaspiro[5.5]undecane, a sex pheromone of the olive fly (see also Aliphatic Natural Products above).



1,7-Dioxaspiro[5.5]undecane

Simple aromatic natural products (VG)

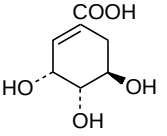
Simple benzene derivatives

These may be of terpenoid, polyketide or shikimate origin. Those of terpenoid origin, such as the aromatic *p*-menthanes are listed in the terpenoid section. Since there is a large number of benzenoid compounds they have been subdivided into simple benzenes (VG0005), simple phenols (VG0010), simple benzyl alcohols (VG0020), simple benzaldehydes (VG0030), simple aryl ketones (VG0035), simple benzoic acids (VG0040), phenylacetic acid derivatives (VG0050) and simple phenylpropanoids (VG0060). Benzoquinones are listed according to their number of oxygen substituents. (VG0300–VG0330) with a separate code for prenylated representatives (VG0370).

Fungi are a prolific source of simple benzoquinones which in the main arise by the polyketide route.

Dewick, P.M. (1995) *Nat. Prod. Rep.*, **12**, 101, 579.
Gill, M. (1993) in *The Chemistry of Natural Products*, 2nd edn (ed. R.H. Thomson), Blackie, Glasgow, pp. 60.
Herbert, R.B. (1989) *The Biosynthesis of Secondary Metabolites*, 2nd edn, Chapman & Hall, London.
Simpson, T.J. (1984) in *The Chemistry of Natural Products* (ed. R.H. Thomson), Blackie, Glasgow, pp. 107.
Simpson, T.J. (1991) *Nat. Prod. Rep.*, **8**, 573 (*biosynth*).
Thomson, R.H. (1971) *Naturally Occurring Quinones*, 2nd edn, Academic Press, London.
Thomson, R.H. (1987) *Naturally Occurring Quinones III*, Chapman & Hall, London.
Turner, W.B. *et al.* (1983) *Fungal Metabolites II*, Academic Press, London.
Tyman, J.H.P. (1979) *Chem. Soc. Rev.*, **7**, 499 (*long chain phenols*).

Shikimic acid is derived from glucose in plants *via* the shikimate pathway. Shikimic acid is the biogenetic precursor of the aromatic amino acids, **Phenylalanine, Tyrosine** and **Tryptophan**. As the shikimate pathway is found in plants but not in animals there is a great deal of interest in targeting shikimate pathway enzymes for control of plant growth, particularly after the success of Glyphosate as a herbicide.



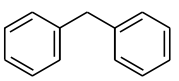
Shikimic acid

The shikimic acid pathway feeds many biosynthetic routes including those involving *p*-aminobenzoic acid, anthranilic acid, cinnamic acid and other phenylpropanoids and hence to many other classes of natural products including the flavonoids and lignans.

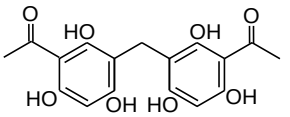
Bentley, R. (1990) *Crit. Rev. Biochem. Mol. Biol.*, **25**, 307.
Campbell, M.M. *et al.* (1993) *Synthesis*, 165 (*synth*).
Conn, E.E. *et al.* (1986) *Recent Adv. Phytochem.*, **20**, 1.
Dewick, P.M. (1992) *Nat. Prod. Rep.*, **9**, 153 (*biosynth*).
Floss, H.G. (1979) *Recent Adv. Phytochem.*, **12**, 59.
Haslam, E. (1993) *Shikimic Acid*, Wiley, Chichester.

Diphenylmethanes, acylphloroglucinols and benzophenones (VG0450, VG0460, VG0500–VG0506)

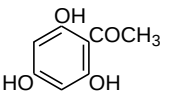
A small number of simple diphenylmethanes occur naturally but there is a growing number of acylphloroglucinols being identified. Acylphloroglucinol derivatives may have more than one diphenylmethane linkage and various alkyl substituents. They are formed by coupling of aromatic units. Other couplings can lead to benzophenone derivatives. The benzophenones, like the benzoquinones, are subdivided in the Type of Compound index according to the number of oxygen substituents.



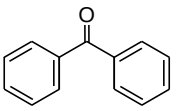
Diphenylmethane
1,1′-Methylenebisbenzene, 9CI



Typical acylphloroglucinol derivative



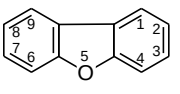
Acetylphoroglucinol



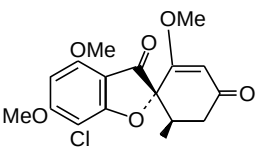
Benzophenone
Diphenylmethanone, 9CI

Dibenzofurans, griseofulvins, dibenzopyrans and xanthones (VG0520, VG0530, VG0535, VG0550–VG0556)

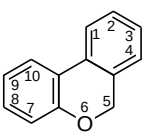
Biogenetically this group of compounds may arise by coupling of aromatic rings as for **Usnic acid** or by ring cleavage of polycyclic aromatic compounds. The xanthones are again listed in subsections according to their number of oxygen substituents.



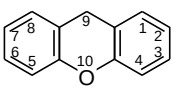
Dibenzofuran, 9CI



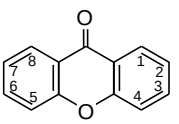
Griseofulvin



Dibenzo[*b,d*]pyran



Dibenzo[*b,e*]pyran
9*H*-Xanthene, 9CI

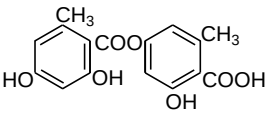


Xanthone
9*H*-Xanthen-9-one, 9CI

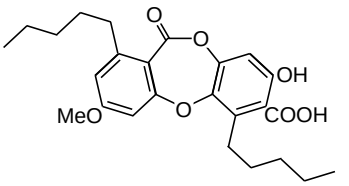
Afzal, M. *et al.* (1980) *Heterocycles*, **14**, 1173.
Bennett, G.J. *et al.* (1989) *Phytochemistry*, **28**, 967.
Sargent, M.V. (1984) *Prog. Chem. Org. Nat. Prod.*, **45**, 103.
Sultanbawa, M.U.S. *et al.* (1980) *Tetrahedron*, **36**, 1465.
Turner, W.V. (1971) *Fungal Metabolites*, Academic Press, London.

Depsides and depsidones; other lichen substances (VG0600–VG0660)

Depsides are esters of polyketide aromatic acids with polyketide phenols such as Lecanoric acid. Depsidones have an additional ether linkage to form a seven membered ring as in Colensoic acid. Depsides and depsidones are predominantly found in lichens, and often carry one or more chlorine substituents; halogenated depsides and depsidones are indexed separately.



Lecanoric acid
(Depside)

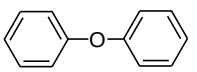


Colensoic acid
(Depsidone)

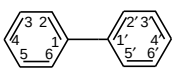
Sargent, M.V. (1984) *Prog. Chem. Org. Nat. Comp.*, **45**, 103.

Diphenyl ethers, biphenyls, dibenzyls and stilbenes (VG1000, VG2000, VG3000, VG4000–VG5000)

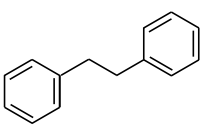
Diphenyl ethers and biphenyls probably arise by radical coupling mechanisms whereas dibenzyls and stilbene derivatives may be derived from a mixed shikimate-polyketide pathway. A large number of stilbene derivatives have been isolated from *Morus* species.



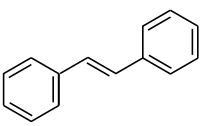
Diphenyl ether



Biphenyl



Dibenzyl

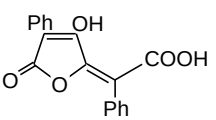


Stilbene

Gorham, J. (1995) *Biochemistry of the Stilbenoids*, Chapman & Hall.
Nomora, T. (1988) *Prog. Chem. Org. Nat. Prod.*, **53**, 87.

Diarylalkyls, terphenyls and the pulvinone group (VG7000, VG7500, VG7600)

Diarylalkyls having more than four carbons separating the aromatic rings may be of mixed biogenetic origin. The terphenyls and the pulvinone group are strictly neolignans (see below) as they arise from two molecules of a phenylpropanoid.

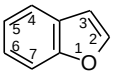


Pulvinic acid

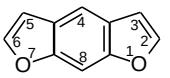
Gill, M. *et al.* (1987) *Prog. Chem. Org. Nat. Prod.*, **51**, 1.
Turner, W.B. *et al.* (1983) *Fungal Metabolites II*, Academic Press, London.

Benzofuranoids (VH)

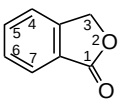
Simple benzofurans (VH1000), benzodifurans (VH2000) and isobenzofurans (VH3000) (including phthalides) are listed here. Dimeric phthalides are Diels–Alder adducts belonging to the Angeolide Group (VH3200). 2-Phenylbenzofurans are probably derived biogenetically from stilbenes.



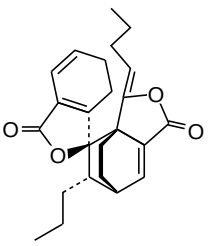
Benzofuran
Benzo[*b*]furan, 9CI



Benzodifuran
Benzo[1,2-*b* : 5,4-*b*′]furan, 9CI



Isobenzofuran
1(3*H*)-Isobenzofuranone, 9CI



Angeolide

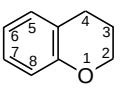
Dean, F.M. (1963) *Naturally Occurring Oxygen Ring Compounds*, Butterworths, London.

Livingstone, R. (1973) in *Rodd's Chemistry of Carbon Compounds*, Elsevier, Amsterdam, Vol IVA, Suppl., 1984.

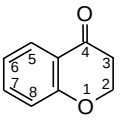
Benzopyranoids (VI)

1-Benzopyrans (VI0030)

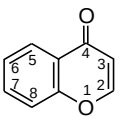
Historically 1-benzopyrans have been known as chromans, chromanones, chromones, and 2- and 3-chromenes but in DNP the simpler members are named systematically. A large number of natural products, of both polyketide and shikimate derivation, occur naturally. The coumarins are the largest class of 1-benzopyran derivatives. They are found mainly in higher plants. Most natural coumarins are oxygenated at C-7; **Umbelliferone** (7-hydroxycoumarin) being regarded as the structural and biogenetic parent of the more highly oxygenated coumarins. Prenylation at carbon and oxygen is common in a large number of coumarins. The prenyl groups found in coumarins exhibit the greatest number of biogenetic modifications including cyclisation to dihydropyrans, pyrans, dihydrofurans and furans. In the Type of Compound Index the very numerous coumarins are subdivided into classes of manageable size according to their oxygen substitution pattern (VI0100–VI7500), with separate sections for natural products having additional rings; furo-1-benzopyrans (VI0050) and pyrano-1-benzopyrans (VI0070).



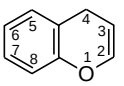
Chroman
3,4-Dihydro-
2*H*-1-benzopyran,
9CI



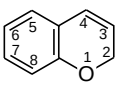
Chromanone
2,3-Dihydro-
4*H*-1-benzopyran-
4-one, 9CI



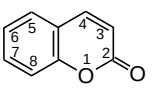
Chromone
4*H*-1-Benzopyran-4-one,
9CI



2-Chromene
β-Chromene
4*H*-1-Benzopyran, 9CI



3-Chromene
α-Chromene
2*H*-1-Benzopyran, 9CI



Coumarin
2*H*-1-Benzopyran-2-one,
9CI

Ellis, G.P. (ed.) (1977) *Chromenes, Chromones and Chromanones*, Wiley, New York.

Gill, M. (1993) in *The Chemistry of Natural Products*, 2nd edn (ed. R.H. Thomson), Blackie, Glasgow, pp. 65.

Livingstone, R. (1977) in *Rodd's Chemistry of Carbon Compounds*, Vol. IVE, Suppl. 1990.

Murray, R.D.H. *et al.* (1982) *The Natural Coumarins*, Wiley, Chichester.

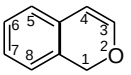
Murray, R.D.H. (1991) *Prog. Chem. Org. Nat. Prod.*, **58**, 83.

Murray, R.D.H. (1997) *Prog. Chem. Org. Nat. Prod.*, **72**, 1.

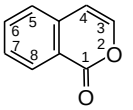
2-Benzopyrans (VI9600)

Compared to the 1-benzopyrans, the 2-benzopyrans are less common. They are normally of polyketide origin. The isochromene nucleus is found in fungal

metabolites such as **Citrinin**. Isocoumarins (VI9700) are the largest class of 2-benzopyran derivatives.



Isochromene
Isobenzopyran
1*H*-2-Benzopyran, 9CI



Isocoumarin
1*H*-2-Benzopyran-1-one, 9CI

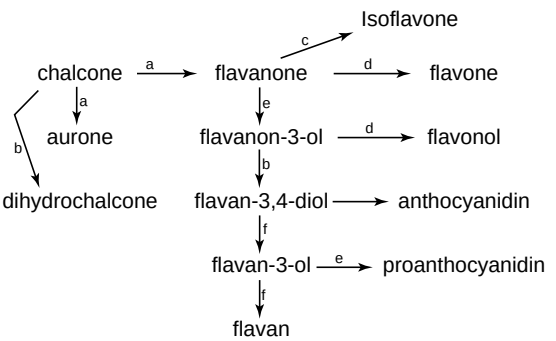
Hill, R.A. (1986) *Prog. Chem. Org. Nat. Prod.*, **49**, 1.
Livingstone, R. (1977) in *Rodd's Chemistry of Carbon Compounds*, Vol. IVE, Suppl. 1990.
Murray, R.D.H. (1995) *Nat. Prod. Rep.*, **12**, 477.
Turner, W.B. *et al.* (1983) *Fungal Metabolites II*, Academic Press, London.

Flavonoids (VK)

The flavonoids are a large group of natural products which are widespread in higher plants but also found in some lower plants including algae. The anthocyanidins are responsible for flower colour in the majority of angiosperms, but colourless flavonoids are also widespread and abundant. A variety of biological functions is fulfilled by various members of the series, but many metabolic and extracellular roles doubtless remain to be discovered.

Flavonoids fall into two major categories according to whether the central heterocyclic ring is unsaturated or not. When unsaturation is present, as in anthocyanins, flavones and flavonols, the molecule is planar (occasionally distorted, e.g. by the substitution of the 2′-hydroxyl group in a 3-*O*-methyl flavonol). Saturated flavonoids (flavanones, flavans) have one or more chiral centres. Optical activity may also be present in flavonoids due to the presence of glycosidic substituents.

Flavonoids can be classified according to their biosynthetic origin. Some flavonoid types are both intermediates in biosynthesis as well as end-products, which can accumulate in plant tissues. These include chalcones (the first formed C_{15} structure derived from malonyl coenzyme A and *p*-coumaryl coenzyme A), flavanones, flavanon-3-ols and flavan-3,4-diols. Other classes are only known as end-products of biosynthesis, e.g. anthocyanins, flavones and flavonols. Two further classes of flavonoid are those in which the 2-phenyl sidechain of flavonoid isomerises to the 3-position (giving rise to isoflavones and related isoflavonoids) and then to the 4-position (giving rise to the neoflavonoids).



Biosynthetic relationship of flavonoids
a = cyclisation, b = bioreduction, c = aryl migration,
d = dehydrogenation, e = hydroxylation,
f = dehydroxylation

Flavonoids may also be classified according to molecular size. While the majority of flavonoids are monomeric, an increasing number of dimeric and oligomeric structures are being described. Most biflavonoids are based on carbon-carbon linking of two similar flavone units, but mixed dimers (e.g. flavonylflavanones) are known. The highest molecular weight flavonoids are the oligomeric and polymeric proanthocyanidins, derived biosynthetically from flavan-3-ols.

Most flavonoids occur naturally associated with sugars in conjugated form and within any one class may be characterised as monoglycosidic, diglycosidic, etc. Glycosidic complexity is considerable. There are, for example, over 2,000 glycosides of the flavones and flavonols that have been isolated to date. (There is a considerable number of glycosides isolated in the course of earlier work which have only been partially characterised structurally and which may or may not be identical with fully characterised glycosides isolated later.) Mono-, di- and trisaccharides may be linked through a phenolic hydroxyl; and one or more

such OH groups may carry a sugar substitution. Acylated *O*-glycosides are known, where aromatic or aliphatic acids are linked through the 6-hydroxyl of a glucose moiety. A special group of mainly flavone-based *C*-glycosides occurs in plants. Sulfated conjugates are common in the flavone and flavonol series, where the sulfation may be on a phenolic hydroxyl and/or on an aliphatic hydroxyl of a glycoside moiety.

A fairly considerable number of *C*-glycosylated flavonoids occur naturally. These are readily distinguished from *O*-glycosyl derivatives by their resistance to acid hydrolysis. They commonly have one or two sugar residues directly linked by a carbon-carbon bond at the C-1 of the sugar to the 6- or 8-position of the flavone nucleus. Thus, the flavone **Apigenin** can occur with a glucose at C-6 and C-8 (**Isovitexin**) or at C-8 (**Vitexin**) or at both C-6 and C-8 (**Vicenin 2**). Other apigenin *C*-glycosides are known where the carbon linked sugar is arabinose, galactose or xylose or two of these monosaccharides.

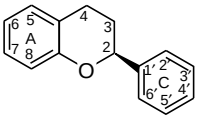
C-Glycosides of flavones commonly occur both as such and with further sugars *O*-glycosidically linked. These other glycosides readily lose their *O*-linked sugar(s) on acid hydrolysis. Such *O*-glycosidic residues may be attached either to a hydroxyl of the *C*-sugar or directly to one of the free phenolic groups. Acylated *C*-glycosides have been described, e.g. the 2″-*p*-coumarate of Vitexin. Many flavone *C*-glycosides are known and they are widely distributed throughout the plant kingdom. By contrast, *C*-glycosides of other classes of flavonoid (e.g. flavonols, flavanones, isoflavones) are of rare occurrence.

Agrawal, P.K. (ed.) (1989) *Carbon-13 NMR of Flavonoids*, Elsevier, Amsterdam.
Barron, D. *et al.* (1996) *Phytochemistry*, **43**, 921 (prenylated flavonoids).
Dean, M. (1963) *Naturally Occurring Oxygen Ring Compounds*, Butterworths, London.
Donnelly, D.M.X. *et al.* (1995) *Nat. Prod. Rep.*, **12**, 321 (isoflavonoids, neoflavonoids).
Ferreira, D. *et al*. (1996) *Nat. Prod. Rep.*, **13**, 411 (proanthocyanins).
Gabor, M. (1986) *The Pharmacology of Benzopyrone Derivatives*, Akademiai Kiado, Budapest.
Geissman, T.A. (ed.) (1962) *The Chemistry of Flavonoid Compounds*, Pergamon Press, Oxford.
Harborne, J.B. (1967) *Comparative Biochemistry of the Flavonoids*, Academic Press, London.
Harborne, J.B. Mabry, T.J. and Mabry, H. (eds) (1975) *The Flavonoids*, Chapman & Hall, London.
Harborne, J.B. and Mabry, T.J. (eds) (1982) *The Flavonoids: Advances In Research*, Chapman & Hall, London.
Harborne, J.B. (ed.) (1988) *The Flavonoids: Advances in Research Since 1980*, Chapman & Hall, London.
Harborne, J.B. (ed.) (1989) *Methods in Plant Biochemistry, Volume 1. Plant Phenolics*, Academic Press, London.
Harborne, J.B. (ed.) (1994) *The Flavonoids: Advances in Research Since 1986*, Chapman & Hall, London.
Harborne, J.B. *et al*. (1995) *Nat. Prod. Rep.*, **12**, 639 (anthocyanins).

In general there are two parallel systems of nomenclature, one based on trivial names such as flavan and chalcone as the parent structure and the other based on systematic chemical names, such as 3,4-Dihydro-2*H*-1-benzopyran (CA) for flavan. The latter becomes cumbersome and easy to get wrong in cases of polysubstitution. There are also two systems of ordering the substituents around the flavan nucleus: one in which the A-and B-ring substituents precede C-ring substituents (e.g. 3,5,7, 3′,4′-pentahydroxyflavone); and one in which the substituents are ordered numerically (e.g. 3,3′, 4′, 5,7-pentahydroxyflavone). There are additionally two conventions for drawing flavonoid formulae, with the heterocyclic oxygen at the top and with the heterocyclic oxygen at the bottom.

In this Dictionary, the semisystematic flavan-type nomenclature is given precedence, substituents are ordered numerically and structures are drawn with

the oxygen heterocyclic atom at the bottom. In the Type of Compound Index, most of the main classes of flavonoid are subdivided according to the number of oxygen substituents.



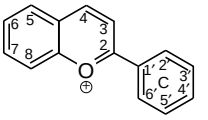
Flavan
3,4-Dihydro-2-phenyl-2*H*-1-benzopyran
(*S*)-form shown

Further information about the nomenclature and numbering of each subclass of flavonoid is given below.

A wealth of trivial names have been used for flavonoids. Some names indicate the class of compound. For example, the ending ‘inidin’ denotes an anthocyanidin (e.g. **Pelargonidin**) and the ending ‘etin’ a flavonol (e.g. **Quercetin**). Likewise glycosides of Quercetin have related names such as **Quercitrin** (the 3-rhamnoside), **Isoquercitrin** (the 3-glucoside) and **Quercimeritrin** (the 7-glucoside). However, there is little consistency in such use and many names have been derived from the generic or specific name of the plant source (e.g. **Tricin** from *Triticum*, **Corniculatusin** from *Lotus corniculatus*). A key to the trivial names most widely used for flavones and flavonols may be found in Harborne (1988). There are a considerable number of duplications of trivial names both between different flavonoids and between flavonoids and other classes of natural product, e.g. terpenoids, alkaloids.

Anthocyanidins (VK0010–VK0070)

Anthocyanidins are intensely coloured plant pigments found throughout vascular plants (they are replaced by purple betalain (alkaloidal) pigments in one order of higher plants, the Centrospermae or Caryophyllales). The flavylium chromophore in e.g. **Cyanidin** is cationic, being associated *in vivo* with organic acid anions. The sugar-free anthocyanidin aglycones are relatively few and vary according to the number and position of hydroxy and methoxy substituents. Structural complexity is associated with the sugar substituents that are present in the water-soluble anthocyanins. The anthocyanins range from simple structures such as cyanidin 3-glucoside (**Chrysanthemin**) to **Ternatin A1**, a delphinidin derivative which is substituted by seven glucose, four *p*-coumaric acid and one malonic acid moiety. Some third of all the known anthocyanins have malonic acid (or other aliphatic dicarboxylic acid) residues linked through sugar and are zwitterionic in their properties.

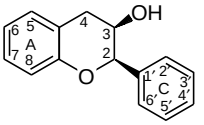


Flavylium (2-phenylbenzopyrylium)

Flavans, Flavanols and Leucoanthocyanidins (VK1000, VK1100, VK1200)

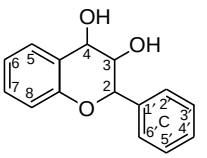
Flavans are formed by reduction of flavanones with flavan-3-ols as intermediates. This is apparent from the facts that they may co-occur with the related flavanone and that they have the same 2*S* configuration. There are a small number of natural flavans, most of which are lipid soluble and occur notably as leaf surface constituents. **4′,7-Dihydroxy-8-methylflavan**, for example, is a phytoalexin formed in the daffodil following fungal inoculation.

The flavan-3-ols (or catechins) make up by far the largest class of monomeric flavans. Two substances with the 3,3′,4′, 5,7-pentahydroxy substitution pattern, namely **Catechin** and **Epicatechin**, are extremely widespread. Most flavan-3-ols, such as Catechin, are of the 2*R*, 3*S* configuration. Those with the 2*R*, 3*R* configuration are prefixed with 'epi', e.g. Epicatechin. Those with a 2*S* configuration are distinguished by the enantio (*ent*-) prefix.



Flavan-3-ol
3,4-Dihydro-2-phenyl-2*H*-1-benzopyran-3-ol
(2*R*,3*R*)-form shown

The term leucoanthocyanidin is used to designate all monomeric flavanoids which produce coloured anthocyanidins by cleavage of a C-O bond on heating with mineral acid. In addition to flavans and flavan-3-ols, there occur flavan-3,4-diols and also a fourth but small class of flavans, the flavan-4-ols. Flavan-3,4-diols are of biosynthetic importance, since they have recently been recognised as the immediate precursors of the anthocyanins. Most naturally occurring 3,4-diols have been obtained by extracting the heartwood of legume trees.



Flavan-3,4-diol
3,4-Dihydro-2-phenyl-2*H*-1-benzopyran-3,4-diol

Proanthocyanidins (VK1500)

Proanthocyanidin is the preferred name for condensed tannins (or flavolans), a series of flavan-3-ol oligomers which are usually based on a C-C link from the 8-position of one flavan unit to the 4-position of a second unit. As with the monomeric leucoanthocyanidins, they produce coloured anthocyanidins on heating with mineral acid, but they have the additional property of binding to protein. The best known proanthocyanidins are procyanidins, based on catechin and/or epicatechin units, and oligomers up to the hexamer have now been found in plants.

The interflavonoid linkage in proanthocyanidins is indicated in the same way as for polysaccharides, the bond and its direction being contained in parentheses (4→). The configuration of the interflavonoid bond at C-4 is indicated by the IUPAC $\alpha\beta$ nomenclature within the above parentheses. Thus two common procyanidin dimers are described as **Epicatechin-(4β→8)-catechin** and ***ent*-Epicatechin-(4α→8)-epicatechin** respectively. A considerable number of doubly linked proanthocyanidins are known, where there is a second linkage through C-2 to O-7. The naming of such compounds can be accommodated in the same general way, e.g. one such compound is **Epicatechin-(2β→7,4α→8)-epicatechin**. Many oligomeric proanthocyanidins with molecular sizes greater than the hexamer, have been isolated from plants but their stereochemistries have yet to be determined.

Biflavonoids and polyflavonoids (VK2000)

The structural variety present in biflavonoids is best illustrated with reference to dimers of Apigenin (4′,5,7-trihydroxyflavone). **Amentoflavone** is the dimer in

which two apigenin units are linked by a carbon-carbon bond from the 8-position of one unit to the 3‴ of the other. A range of *O*-methyl ethers of this basic structure occur naturally. Biapigenins with other C-C linkages have been discovered, where the linkage is 3′-3‴, 3-8″, 3-3‴, 6-8″, 8-8″, 6-6″, or 6-5‴. Linkage through a C-O-C bond may also occur, as in **Hinokiflavone**, where the two apigenin units are linked at the 6 and 4‴ positions.

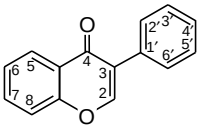
Mixed biflavonoids are also possible, e.g. flavone-flavanone dimers, as well as compounds based on two flavanone units (e.g. **Rhusflavanone**). The first triflavonoid has been reported recently, based on three units of Luteolin (3′,4′,5,7-tetrahydroxyflavone). Biflavonoids have a distinctive distribution pattern. There are major occurrences in gymnosperms, mosses and ferns and a more limited presence in some 15 angiosperm families.

Isoflavonoids (VK3000–VK3100)

Isoflavonoids are based on the 3-phenylchroman skeleton that is biogenetically derived by an aryl migration from a flavanone precursor. They have a very limited distribution in the plant kingdom and are almost entirely restricted to the subfamily Papilionoideae of the Leguminosae. They are found very occasionally in about 18 other angiosperm families and there are isolated occurrences in mosses and gymnosperms. Another striking feature about the isoflavonoids is their major presence in lipophilic plant extracts in the free state and the relative rarity of glycosidic derivatives.

Some isoflavonoid isolations reported from microorganisms are almost certainly spurious, and associated with contamination from the culture medium.

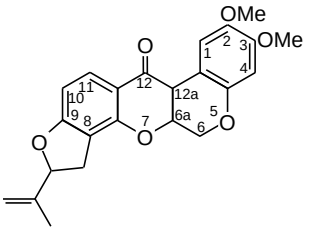
The largest class of isoflavonoids are the isoflavones (VK3000–VK3070). There are simple structures such as **Genistein** (4′,5,7-trihydroxyisoflavone) but also a wealth of prenylated derivatives. The prenyl sidechains may ring-close on adjacent hydroxyl groups, giving rise to tetracyclic and pentacyclic compounds. The related isoflavanones (VK3100), in which the 2,3-bond is reduced, are much rarer than the isoflavones.



Isoflavone
3-Phenyl-4*H*-1-benzopyran-4-one

Rotenoid flavonoids (VK3200–VK3300)

Rotenoids are a class of isoflavonoid characterised by the presence of an extra carbon atom in an additional heterocyclic ring. This system is derived by oxidative cyclisation of a 2′-methoxyisoflavone. Rotenoids characteristically possess insecticidal and piscicidal activity, as shown by Rotenone, one of the parent structures. Besides rotenoids proper, there are a small number of 12*a*-hydroxyrotenoid (VK3250) and dehydrorotenoid (VK3300) flavonoids, in which there is a double bond introduced at the 6*a*–12*a* position.

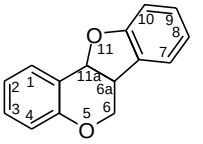


Rotenone
1,2,12,12*a*-Tetrahydro-8,9-dimethoxy-2-(1-methylethenyl)[1]benzopyrano[3,4-*b*]furo[2,3-*h*][1]benzopyran-6(6*aH*)-one, 9CI

The numbering system most used by natural products scientists for Rotenone is shown but other schemes have been used and it must be noted that the CA scheme differs. Various numbering schemes have also been used for the cyclised prenyl side-chain in Rotenone and similar compounds.

Pterocarpans (VK3400–VK3550)

Pterocarpans contain a tetracyclic ring system derived from the basic isoflavone skeleton by an ether linkage between the 4- and 2′-positions. The systematic numbering is distinctive for this particular carbon skeleton. The majority of natural pterocarpans have been obtained from phytoalexin studies, so that in general they possess antifungal activity. They are conveniently subdivided into simple pterocarpan flavonoids, 6*a*-hydroxypterocarpan flavonoids and pterocarpene flavonoids, in which unsaturation is introduced at the 6*a*–11*a* position. The best known structure is **Pisatin**, a 6*a*-hydroxypterocarpan which is the phytoalexin of the pea plant. Many isoprenylated pterocarpans have been described and these substances constitute the second largest group of isoflavonoids after the isoflavones. The commonly used numbering system corresponds with the CA scheme.

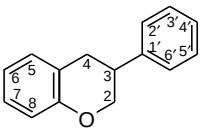


Pterocarpan
6*a*,11*a*-Dihydro-6*H*-benzofuro[3,2-*c*][1]benzopyran, 9CI

Although pterocarpans have two chiral centres, only *R*,*R* and *S*,*S* configurations are sterically possible. Most pterocarpan phytoalexins that have been isolated are laevorotatory and have the 6*aR*, 11*aR* absolute configuration; a few are dextrorotatory and can be assigned to the 6*aS*, 11*aS* series.

Isoflavans (VK3600–VK3700)

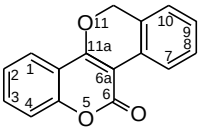
Isoflavans are another class of isoflavonoid which have been mainly isolated as phytoalexins after fungal inoculation of plant tissues. They are also metabolites of dietary isoflavones. **Equol** (4′,7-dihydroxyisoflavan) which has been isolated from the urine of mammals, has estrogenic activity. The numbering system of isoflavans is the same as that of the isoflavones.



Isoflavan
3,4-Dihydro-3-phenyl-2*H*-1-benzopyran

Coumestan flavonoids (VK3750)

One final group of isoflavonoids, numerically important in terms of numbers of structures, are the coumestans. Like the isoflavans and many isoflavones, they exhibit weak estrogenic activity in mammals. The simplest structure is **Coumestrol** (7,9-dihydroxycoumestan) but a variety of prenylated derivatives have also been characterised. The numbering system used is the same as in the pterocarpan series and coincides with the CA systematic numbering.



Coumestan
6*H*-1-Benzofuro[3,2-*c*][1]benzopyran-6-one,9CI

Neoflavonoids (VK4000)

This term refers to a small group of C_{15} naturally occurring substances structurally and biogenetically related to the flavonoids and isoflavonoids. They have a limited distribution, occurring with isoflavonoids in the subfamily Papilionoideae of the Leguminosae. Other families where they have been encountered are the Guttiferae, Rubiaceae, Passifloraceae, Compositae and Polypodiaceae.

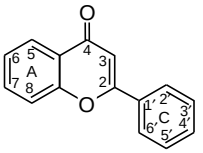
There are three main subdivisions of structures: the 4-arylcoumarins, the dalbergiones and the dalbergiquinols. Representative structures, all isolated from *Dalbergia* species, are the ring-closed **Dalbergin** and the two related ring-opened compounds, **4-Methoxydalbergione** and **Obtusaquinol**. Prenylated derivatives of the 4-arylcoumarins have been characterised in the Guttiferae.

Flavones and Flavonols (VK5000–VK5280)

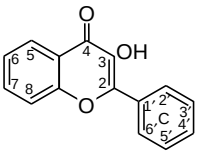
Flavones are a class of polyhydroxyflavonoid based on the structure of **Flavone** (2-phenyl-4*H*-1-benzopyran-4-one or phenylchromone) which itself occurs naturally as a farina on *Primula* plants. Flavonols are flavones with a 3-hydroxy substituent and they share the same nomenclature. It is convenient to separate these two classes, mainly because so many structures are known; some 1000 aglycones and over 2,000 glycosides. They differ in their spectroscopic and chromatographic properties and can readily be distinguished by these means. They are biosynthetically distinct, flavones being formed by oxidation of flavanones, flavonols by oxidation of dihydroflavonols. There are also differences in the way they occur naturally; *C*-glycosides are common in the flavone series but rare among flavonols.

In the DNP Type of Compound Index they are subdivided according to the number of O substituents (including methylenedioxy groups): *C*-methylation and *C*-prenylation is relatively common.

Free lipophilic flavones and flavonols occur at the upper surface of leaves in the wax or in bud exudates. There are also many *O*-glycosides, which are found within the leaf in the cell vacuole and in other parts of the plant. There are at least 200 different glycosides of **Quercetin** and 250 of the related flavonol, **Kaempferol**. (The principal derivatives of such widespread parent flavonoids have their own entries in DNP and it is important to use the indexes to locate a particular glycoside which may be documented in one of these subsidiary entries).



Flavone
2-Phenyl-4*H*-1-benzopyran-4-one

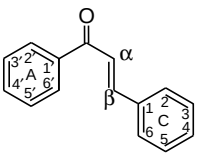


Flavonol
3-Hydroxy-2-phenyl-4*H*-1-benzopyran-4-one

In DNP individual flavonols are named both as derivatives of an n-hydroxyflavonol and as derivatives of an (n + 1) hydroxyflavone, allowing their rapid location through the indexes whichever name is employed. The flavonoid alkaloids e.g. **Ficine**, are described under the alkaloids (VX 0350)

Chalcones and dihydrochalcones (VK6010–VK6080, VK6200)

Chalcones are open-chain C_6-C_3-C_6 compounds, the first intermediates of flavonoid biosynthesis. They occur sporadically in plants as yellow pigments, some 200 structures being known. The numbering system of chalcone substituents differs from that in ringclosed flavonoids.



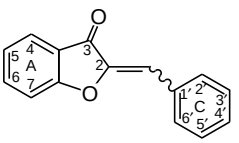
Chalcone
1,3-Diphenyl-2-propen-1-one, 9CI

Note that the numbering of the A ring is the same in both systems of nomenclature, but the C ring is unprimed in the semitrivial chalcone system and carries a double prime if systematic numbering is used (the α- and β-positions becoming 2 and 3 respectively). The majority of chalcones have hydroxy/methoxy substituents at the 2′,4,4′,6′-positions, and a significant number of prenylated derivatives are known.

In dihydrochalcones, the double bond in the α-β-position is reduced and the compounds are colourless. The numbering system is the same as in the chalcone series. They are less common than chalcones and occur variously in higher plants, ferns and liverworts.

Aurone flavonoids (VK6100)

Aurones are a small group of yellow pigments, based on the 2-benzylidenecoumaranone nucleus. These are formed by oxidation of chalcones and may co-occur with the related chalcone precursors. The numbering system differs from that in the chalcone series, so that the most common hydroxylation pattern, that of the pigment **Aureusidin**, is 3′,4,4′,6-tetrahydroxyaurone. Note the potential occurrence of geometrical isomers.

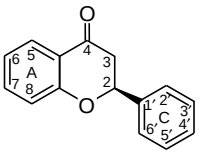


Aurone
2-(Phenylmethylene)-3(2*H*)-benzofuranone, 9CI

The auronols (2-hydroxy-2-benzylcoumaranones) are a closely related series of colourless compounds, with only a few members so far described.

Flavanones (VK6300–VK6380)

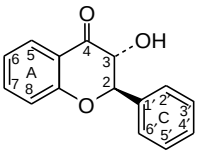
Flavanones are 2,3-dihydro-2-phenyl-4*H*-1-benzopyran-4-ones. The simplest known natural flavanone is the 7-hydroxy derivative, while the commonest is 4′,5,7-trihydroxyflavanone (**Naringenin**). Flavanones are isomeric with chalcones and arise biosynthetically from them by a reaction catalysed by an isomerase. They have a centre of chirality at C-2 and usually occur in optically active form with the 2*S*-configuration. They commonly occur as glycosides. A variety of more complex derivatives with methyl and/or prenyl substituents has been described. Flavanones have a wide occurrence in plants.



Flavanone
2,3-Dihydro-2-phenyl-4*H*-1-benzopyran-4-one, 9CI
(*S*)-form shown

Dihydroflavonols (VK6410–VK6470)

Dihydroflavonols can be described as 3-hydroxyflavanones or as flavanon-3-ols. They are formed biosynthetically by oxidation at C-3 of flavanones, without inversion at C-2, and are the immediate precursors by a further oxidation of the flavonols. Dihydroflavonols have two chiral centres at C-2 and C-3; most naturally occurring compounds possess the (2*R*,3*R*) stereochemistry. Dihydroflavonols such as **Dihydroquercetin** have a wide occurrence in nature being present in the free state in woody plant tissues. They also occur in glycosidic combination in other plant parts.



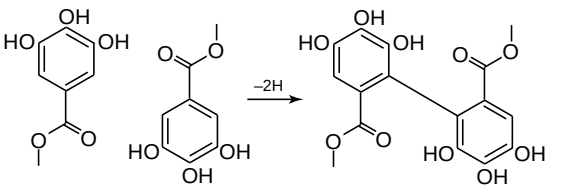
Dihydroflavonol
2,3-Dihydro-3-hydroxy-2-phenyl-4*H*-1-benzopyran-4-one, 9CI
(2*R*,3*R*)-form shown

Tannins (VM)

Plant polyphenols (vegetable tannins) are secondary metabolites widely distributed in the plant kingdom. They are based upon two broad structural themes:

(*a*) Condensed proanthocyanidins in which the fundamental structural unit is the phenolic flavan-3-ol (catechin) nucleus.
(*b*) Galloyl and hexahydroxydiphenoyl esters and their derivatives.

These metabolites are almost invariably found as multiple esters of 3,4,5-trihydroxybenzoic (gallic) acid with D-glucose and a great many can be envisaged as derived from the key biosynthetic intermediate *β*-1,2,3,4,6-pentagalloyl-D-glucose. Derivatives of hexahydroxydiphenic acid are assumed to be formed by oxidative coupling of vicinal galloyl ester groups in a galloyl-D-glucose ester.



Gallic acid is most frequently encountered in plants in the form of esters. These may be classified into several broad categories:

(*a*) Simple gallate ester tannins. (VM6000).
(*b*) Depside metabolites (gallotannins).
(*c*) Hexahydroxydiphenoyl and Dehydrohexahydroxydiphenoyl ester tannins (ellagitannins) based upon:
  (i) $^{4}C_{1}$ conformation of D-Glucose.
  (ii) $^{1}C_{4}$ conformation of D-Glucose.
  (iii) ‘open-chain’ derivatives of D-Glucose.
(*d*) Dimers and higher oligomers formed by oxidative coupling of monomers, principally those of class (iii) above.

Four distinctive and principal pathways are presumed to lead from *β*-1,2,3,4,6-pentagalloyl-D-glucose and to give, by appropriate chemical embellishment the various classes of metabolites (Figure 10).

Ellagitannin metabolites fall into two broad categories – monomeric species formed by intramolecular C-C oxidative coupling and oligomeric species formed by intermolecular C-O coupling (Figure 11). Numerous intramolecular C-C linked ester groups have been located in the monomers and similarly various intermolecular C-O linking ester groups have been defined in the formation of the oligomeric structures. The principal members of these two classes of ester group are shown in Figures 12–15.

The nomenclature and numbering of the more complex types of tannin is difficult. CA names them as complex carbohydrate esters or (in the more complicated cases, e.g. **Vescalagin** as stereoparents, with closely related natural products being named as derivatives, e.g. **Castavaloninic acid** = 25-*O*-(6-carboxy-2,3,4-trihydroxyphenyl)vescalagin, 9CI. In such cases the numbering is arbitrary. In DNP, limited numbering of structures is shown only when it is strictly necessary.

The various subclasses of tannin are separately listed in the Type of Compound Index according to the esterifying acid group, beginning with simple gallate esters and proceeding to the more complex types.

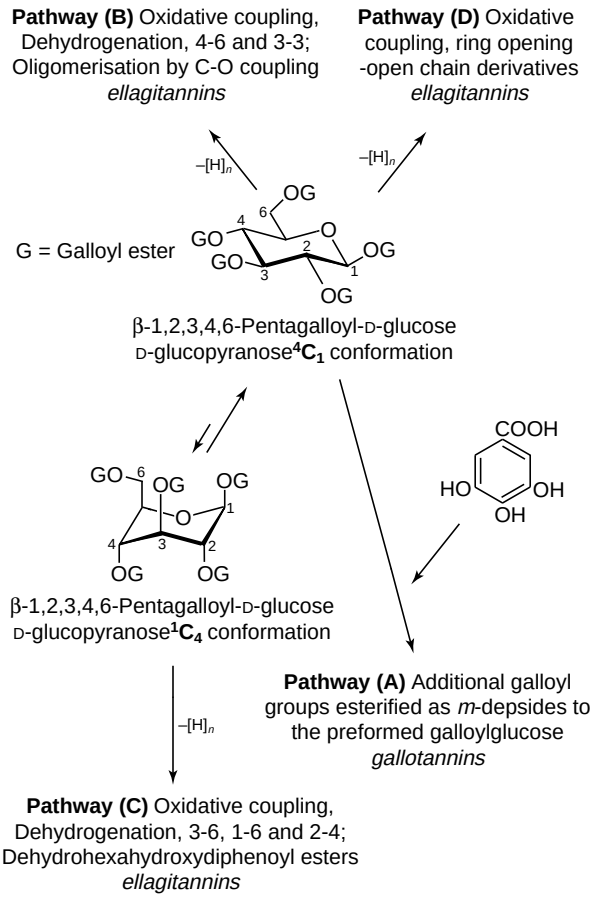


Figure 10. Biogenesis of the gallotannins and ellagitannins; the metabolic embellishment of β-1,2,3,4,6-pentagalloyl-D-glucose, principal pathways.

Haslam, E. (1982) *Prog. Chem. Org. Nat. Prod.*, **41**, 1.
Haslam, E. (1982) *Plant Polyphenols – Vegetable Tannins Revisited*, Cambridge University Press, Cambridge.
Haslam, E. (1994) *Nat. Prod. Rep.*, **11**, 41.
Okuda, T. *et al*. (1981) *Heterocycles*, **15**, 653.
Okuda, T. *et al*. (1989) *J. Nat. Prod.*, **52**, 1.
Okuda, T. *et al*. (1989) *Planta Med.*, **55**, 117.
Okuda, T. *et al*. (1990) *Heterocycles*, **30**, 1195.
Okuda, T. *et al*. (1993) *Phytochemistry*, **32**, 507.
OKuda, T. *et al*. (1995) *Prog. Chem. Org. Nat. Prod.*, **66**, 1.
Schmidt, O. Th. (1956) *Prog. Chem. Org. Nat. Prod.*, **13**, 570.

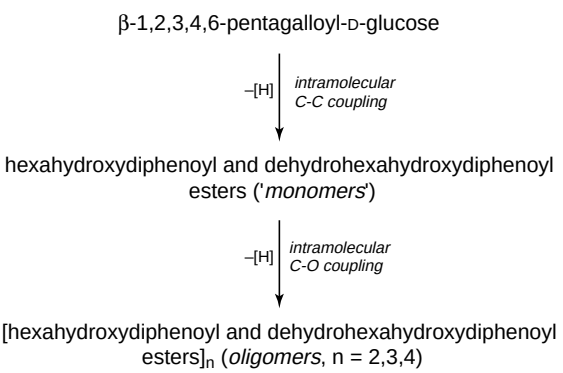


Figure 11. Overall patterns of oxidative metabolism of β-1,2,3,4,6-pentagalloyl-D-glucose in higher plants to yield ellagitannins.

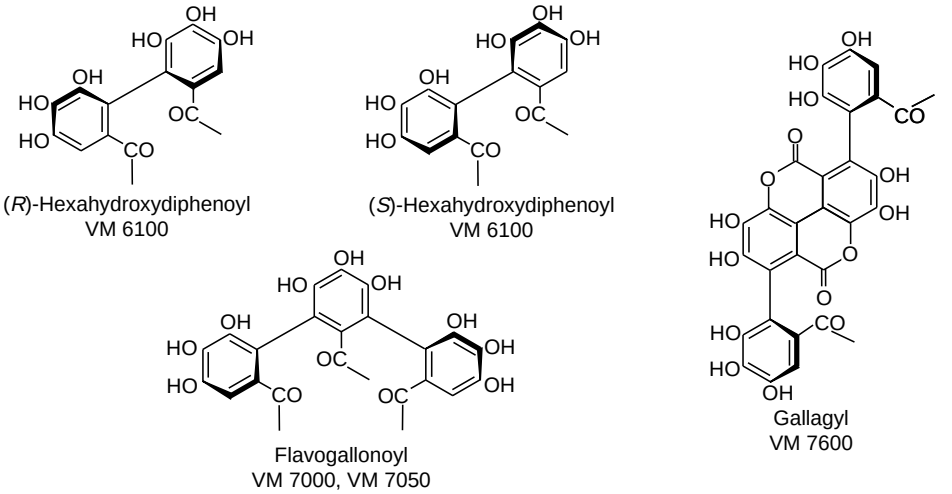


Figure 12. Principal derivatives of hexahydroxydiphenic acid formed by intramolecular C-C oxidative coupling.

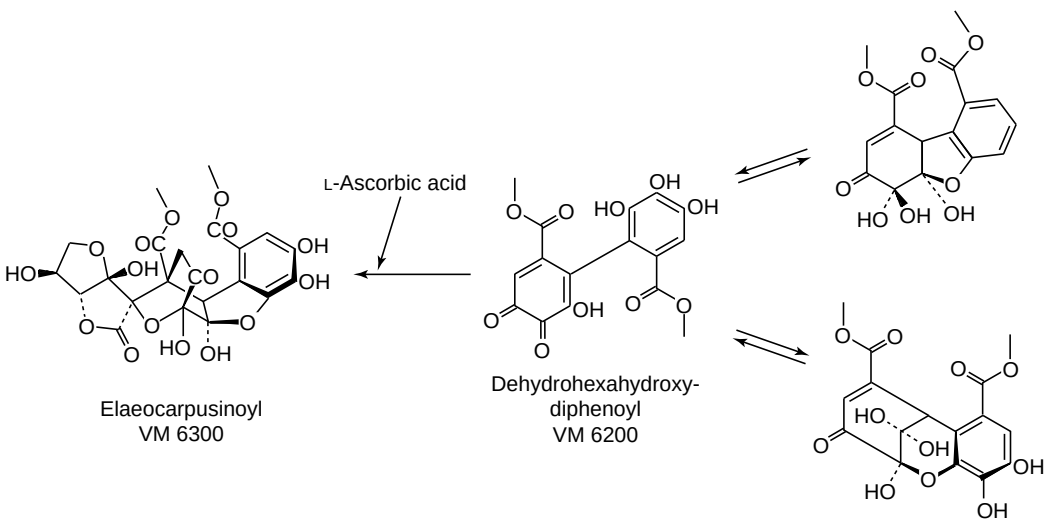


Figure 13. The dehydrohexahydroxydiphenoyl ester group and its derivatives.

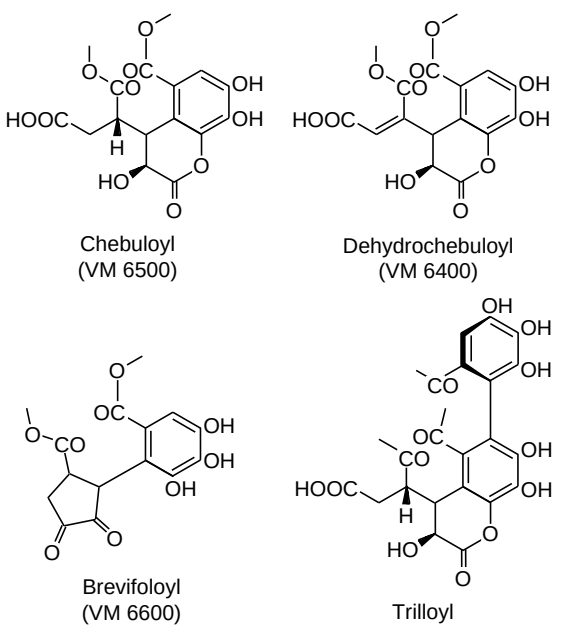
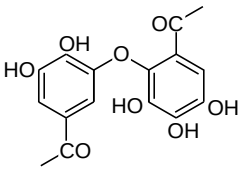
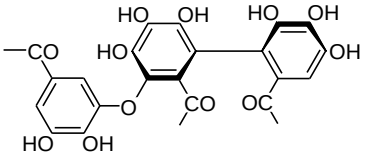


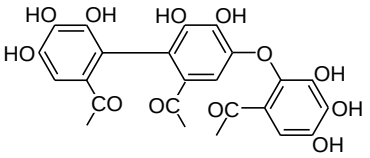
Figure 14. Ester derivatives of hexahydroxydiphenic acid in which one aromatic ring has undergone hydrolytic cleavage.



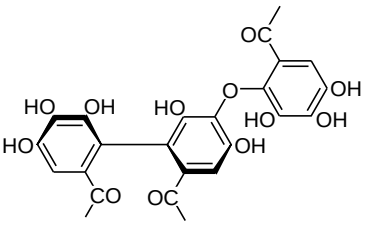
Dehydrodigalloyl
VM 6700



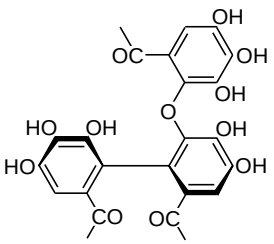
(*S*)-Sanguisorboyl
VM 6900



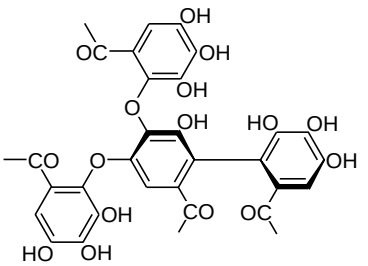
Valoneoyl
VM 6800, VM 6820, VM 6850



(*R*)-Tergalloyl
VM 7300, VM 7350



(*R*)-Macaranoyl
VM 7200

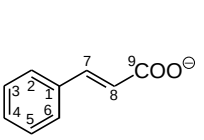


(*R*)-Euphorbinoyl
VM 7500

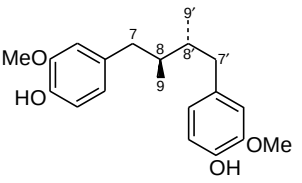
Figure 15. Principal ester groups formed by intermolecular C-O oxidative coupling of galloyl and hexahydroxydiphenoyl esters.

Lignans (VO)

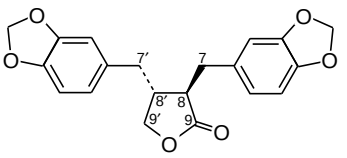
The lignans are a group of plant phenols whose structures are determined by the union of cinnamic monomers or their biogenetic equivalents. The lignan dimers linked at β-positions in the side chain were first defined in 1936 by R.D. Haworth and these are discussed first. This group of about 900 compounds has since been extended to include some 500 neolignans, which are dimerised in other ways, and higher oligomers of cinnamyl alcohols. Higher lignoid polymers occur in wood as lignin and are degraded during papermaking processes.



Cinnamate monomer



Dihydroguaiaretic acid

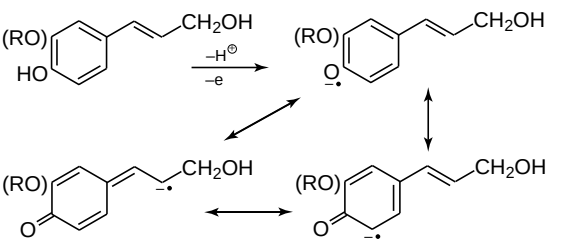


Hinokinin

The rings may bear different oxysubstituents and trioxygenation is common. **Dihydroguaiaretic acid** illustrates one typical aromatic substitution pattern and **Hinokinin** another. A single *para*-phenolic substituent is rare but this position is always occupied since the biosynthesis depends upon it. Free carboxylate groups are rare; ketals and ketones are found occasionally. The presence of an oxysubstituent *para* to the side chain in lignin and all lignans is consistent with a biosynthesis *via* coupling of mesomeric cinnamoyl radicals derived from a coumaryl alcohol.

The same net result would be obtained by two stage process initiated through the attack of such a radical upon its phenolic precursor. Growth of polymeric lignin can be rationalised through radical sites located on oxygen, on the ring system or in the side chain, whilst dimerisation of the latter leads to structures such as Dihydroguaiaretic acid and Hinokinin.

Lignans have been obtained as fragments from the degradation of lignin and direct evidence of lignan biosynthesis which substantiated the lignin analogy was obtained in 1969, when it was shown that cinnamyl precursors were incorporated into **Podophyllotoxin**. This work was subsequently extended by other workers who established this route in butyrolactones and furofurans.



The oxidation level of the cinnamyl residues when coupling occurs is uncertain and related studies are complicated by interaction of the matrix of lignan precursors with the larger lignin pool. It is commonly believed that lignin formation is not under enzymic control because lignin, in contrast to lignans, is

not optically active. However, it is possible that in lignin internal compensation occurs between the many chiral centres.

Dihydroguaiaretic acid and Hinokinin are representatives of major classes which are sometimes loosely referred to as ‘acyclic’; others may be defined as either carbocyclic or oxy-heterocyclic derivatives of these two types of parent.

The lignans are classified in the Type of Compound Index essentially according to Figure 16 which illustrates the principal lignan types including the unusual 7′,8-structure found in **Magnosalicin**.

The systematic nomenclature of such a structurally diverse, though biogenetically related, group of compounds is of limited utility and easily disguises structural similarities. The CA names of most lignans are given as synonyms within the entries, but the entry name is usually either a trivial name or a semisystematic name using the system originally proposed by Freudenberg and Weinges which has been extended by Moss and now accepted by IUPAC. This names lignans according to a ‘lign-’ scheme which more accurately reflects their biosynthetic origin. The cinnamyl unit of higher priority is numbered from 1–9 while that of lower priority bears primed numbers.

For Dihydroguaiaretic acid the two cinnamyl residues are equivalent so that no distinction need be made between them and the semi-systematic name is: 4,4′-dihydroxy-3,3′-dimethoxylignan. Note that under the CA system this compound would be named as a phenol. In Hinokinin, the sequence rules require that priority be given to the lactone carbonyl (C9) and this leads to the lignan name 3,3′,4,4′-bis (methylenedioxy)lignan-9,9′-olide.

Natural lignans are optically active although a few *meso*-compounds are known. Important physiological properties may be associated with a particular absolute configuration, as for example with the antitumour agent

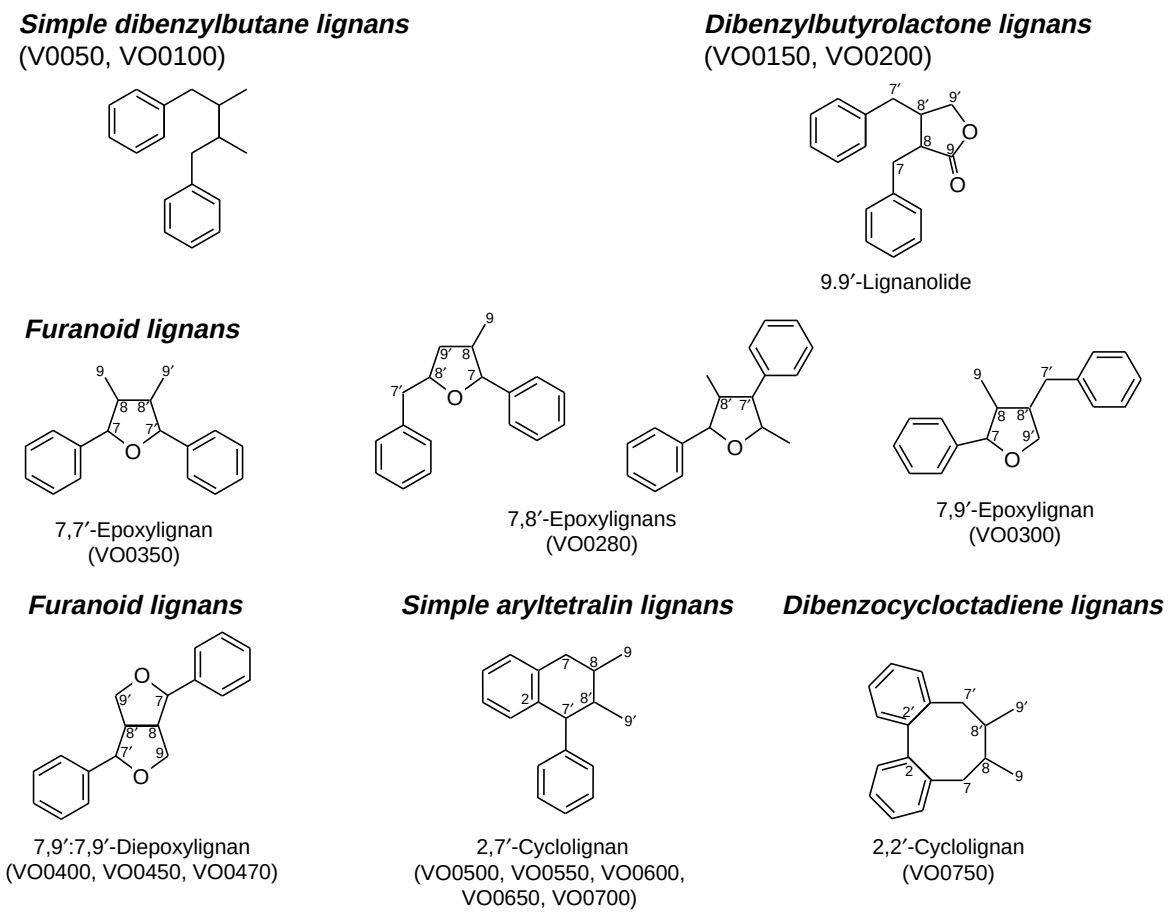


Figure 16. Lignan structural types.

Podophyllotoxin. The application of the Cahn-Ingold-Prelog sequence rules to lignans needs to be done with care as apparent inversions of configuration between closely related compounds owing to different substitution patterns are common.

Inherent dissymmetry is shown by some semisynthetic derivatives of Podophyllotoxin and in the dibenzocyclooctadienes. The absolute configuration of the latter major group is securely based on the X-ray analysis of **Gomisin D**. This is an atropoisomer with the *S*-configuration.

Neolignans (VO1500)

These compounds are also dimers of cinnamyl units but their structures are obtained by coupling of mesomeric radicals other than the β-β link typical of the lignans.

As the range of lignoids and their plant sources has widened so the distinction between lignans and neolignans has become less significant. Thus the neolignans were long – identified as more typical of plants of the family Lauraceae, but in recent years they have been isolated from the Piperacae, Magnoliaceae, Phytolaccacae, Rutaceae, Pinaceae and Berberidaceae amongst others. Further, lignans and neolignans frequently occur in plants of the same family, for example in *Piper wightii*. Significant developments were the isolation of the classical lignan **Megaceratonic acid** from a non-vascular plant – the hornwort *Megaceros flagellaris*, and of the pyranonyl hybrid **Scapaniapyrone** and its analogue from the liverworts *Scapania undulata* and *Jamesoniella autumnalis* respectively: at present only 5% of liverworts have been investigated.

The system of nomenclature illustrated above extends logically to include both neolignans and oligomeric lignoids since all are produced by the union of mesomeric cinnamate radicals. The structures (Figure 16) show the permutations of radical couplings at 8,8′-positions in lignans and some of the more varied combinations based on C-C and C-O linkages in neolignans are shown in Figure 17.

Sesquilignans and dilignans

In recent years a significant number of these oligomers has been identified largely owing to the use of HPLC allied to mass spectroscopy. Typical of the

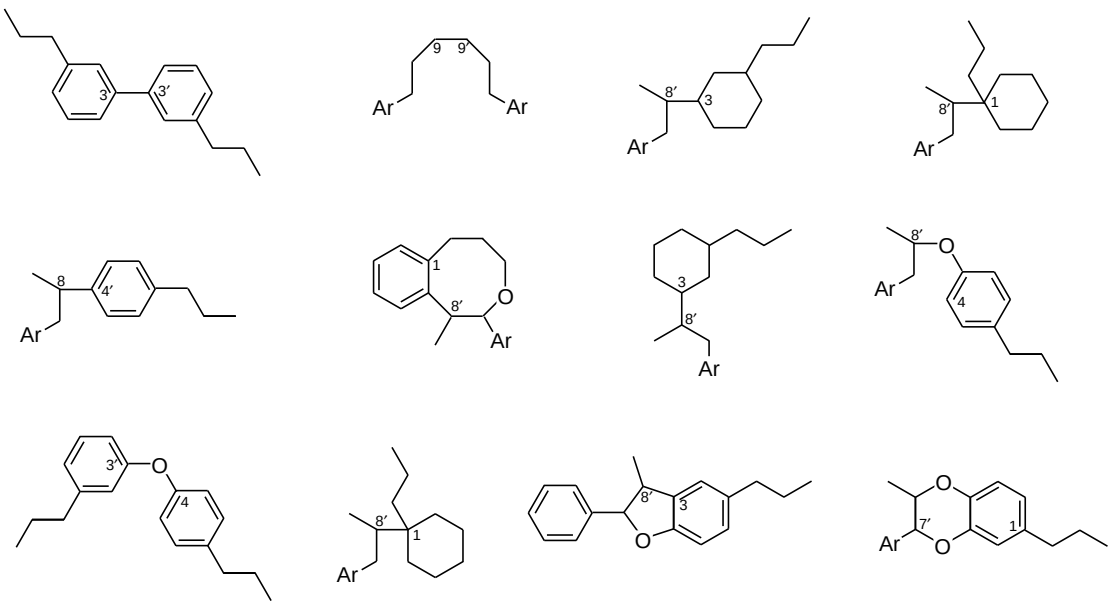


Figure 17. Skeletons of some neolignans.

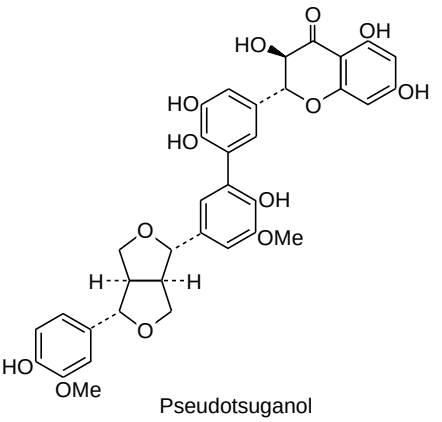
former group are the lappaols including **Lappaol A**, with three linked cinnamate residues. **Manassantin B** including four such residues is typical of the dilignans.

Note the variation in bond type between the classical C – C bonding in **Dunnianol** and that in **Salvianolic acid**, which is regarded as a member of this group although on hydrolysis of the enol ether and ester linkages the lignoid character will be lost altogether. A similar distinction needs to be drawn between the dilignans **Arctigenin E** and **Buddlenol E**, where four residues may be stably linked through C – C bonds or through a combination of C – C and ether linkages, and **Cannabisin G**. The latter falls within the definition of a dilignan yet on hydrolysis yields the arylethylamine and a ligna-dienoic acid.

Hybrid lignans

Here the lignoid includes a structure typical of another class of natural product: with the increased sensitivity of modern isolation methods the list of both members and known structural types is expanding. **Silybin** is a flavonolignan and in 1968 was the first known hybrid to be described; the coumarinolignan **Cleomiscosin** followed in 1979. Pseudotsuganol is illustrative of another type of flavonolignan.

More recently the stilbenolignan **Maackolin** and the xantholignan **Cadensin G** have been identified. Hybrid terpenes are typified by the C – C linked **Piperitylmagnolol** and by the C – O linked **Eudesmagnolol**. In **Chilianthin A** the lignan acid is esterified with an oleanyl triterpene. The hydrolysable macrocycle **Pelliatin** is of interest structurally and also because it is a constituent of a non-vascular plant, the liverwort *Pellia epiphylla*.



Pseudotsuganol

Ayres, D.C. and Loike, J.D. (1990) *Lignans, Chemical, Biological and Clinical Properties*, Cambridge Univ. Press, Cambridge.

Fengel, D. and Wegener, G. (1984) *Wood: Chemistry, Ultrastructure and Reactions*, de Gyten, Berlin.

Freudenberg, K. and Weinges, K. (1961) *Tetrahedron*, **15**, 115 (*nomenclature*).

Gottlieb, O.R. (1978) *Prog. Chem. Org. Nat. Prod.*, **35**, 1 (*neolignans*).

Gottlieb, O.R. and Yoshida, M. (1989) in *Natural Products Extraneous to the Lignocellulosic Cell Wall of Woody Plants* (eds J.W. Rowe and C.H. Kirk) Springer, Berlin (*neolignans*).

Haworth, R.D. (1936) *Ann. Rep. Progr. Chem.*, **33**, 266.

Lin, L-J. and Cordell, G.A. (1984) *Chem. Commun.*, 160 (*coumarinolignans*).

MacRae, W.D. and Towers, G.H.N. (1984) *Phytochemistry*, **23**, 1207 (*activity*).

Marston, A. and Hostettman, K. (1991) *Nat. Prod. Rep.*, **8**, 392 (*sesquilignans, dilignans*).

Sakakibara, A. *et al.* (1987) *Holzforschung*, **41**, 1.

Ward, R.S. (1993) *Nat. Prod. Rep.*, **10**, 1; 1995, **12**, 183; 1997, **14**, 43 (rev).

Polycyclic aromatic natural products (VQ)

A large proportion of the polycyclic aromatic compounds encountered in nature are quinonoid.

Quinone and quinonoid compounds are widely distributed and exist in all living organisms, often playing an important role in redox systems.

The quinones isolated from higher plants are usually relatively simple and frequently present as glycosides. The microbial quinones, on the other hand, are often more complex and exhibit greater structural diversity as well as wider variations in biological activity.

Eckardt, K. (1981) in *Antitumour compounds of natural origin, Chemistry and Biochemistry* (ed. A. Aszalos) CRC Press, Boca Raton, **II**, 28.

Gill, M. (1993) in *The Chemistry of Natural Products*, 2nd edn (ed. R.H. Thomson), Blackie, Glasgow, pp. 60.

Simpson, T.J. (1984) in *The Chemistry of Natural Products* (ed. R.H. Thomson), Blackie, Glasgow, pp. 107.

Thomson, R.H. (1971) *Naturally Occurring Quinones*, 2nd edn. Academic Press, London.

Thomson, R.H. (1987) *Naturally Occurring Quinones III*, Chapman & Hall, London.

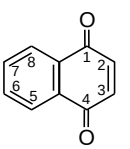
Turner, W.B. *et al.* (1983) *Fungal Metabolites II*, Academic Press, London.

Naphthalenes and naphthoquinones (VQ2000, VQ3000–VQ3060)

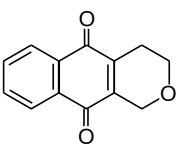
Most of the naphthalene derivatives found in nature are naphthoquinones, which are listed in the Type of Compound Index subdivided according to their number of oxygen substituents. They may arise by polyketide, terpenoid, shikimate pathways or a mixture of these. The purely terpenoid naphthalene derivatives are listed in the Terpenoid section under their biogenetic class.

Benzoisochromanquinones (VQ3100)

These are of polyketide origin and are found in fungi.



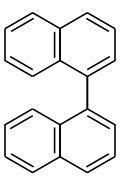
Naphthoquinone
1,4-Napthalenedione, 9CI



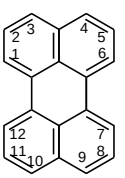
Benzoisochromanquinone

Binaphthyls, perylenes and the duclauxin group (VQ2500, VQ2600, VQ2700)

Binaphthyls and perylenes arise by radical coupling of naphthylenes. The **Duclauxin** group arises from radical coupling of a naphthopyran derivative.



Binaphthyl
[1,1′-Binaphthalene], 9CI

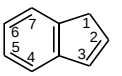


Perylene, 9CI

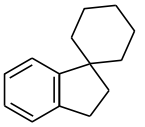
Turner, W.B. (1971) *Fungal Metabolites*, Academic Press, London.
Turner, W.B. *et al.* (1983) *Fungal Metabolites II*, Academic Press, London.
Weiss, U. *et al.* (1987) *Prog. Chem. Org. Nat. Prod.*, **52**, 1.

Indenes and Indan-1-spirocyclohexanes (VQ3300, VQ3400)

Simple natural indenes are listed under this heading apart from those of terpenoid origin that are listed in the appropriate Terpenoid section. The indan-1-spirocyclohexanes are found in cannabis and are probably derived biosynthetically from dibenzyls.



Indene

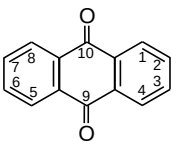


Indan-1-spirocyclohexane

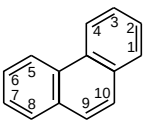
Crombie, L. *et al.* (1982) *J. Chem. Soc., Perkin Trans 1*, 1455.

Anthracenes and phenanthrenes (VQ3450, VQ4000–VQ4200, VQ4800–VQ5100)

Anthraquinones are the largest class of natural quinones. They are generally of polyketide origin. 9,10-Anthaquinones, which are listed subdivided according to their number of oxygen substituents, predominate with a small number of 1,2- and 1,4-anthraquinones (VQ4100). Bianthracenes presumably arise by radical coupling mechanisms. A number of non-quinonoid oxygenated phenanthrenes and 9,10-dihydrophenanthrenes are found in higher plants. 1,2-, 1,4- and 9,10-phenanthraquinones are also found as natural products in higher plants and fungi.



9,10-Anthraquinone
9,10-Anthracenedione, 9CI



Phenanthrene

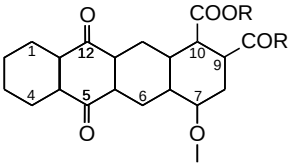
Bolton, R. (1988) in Rodd's *Chemistry of Carbon Compounds*, Suppl., (ed. M.F. Ansell), Elsevier, Amsterdam, Vol. IIIH, 1.
Gill, M. *et al.* (1987) *Prog. Chem. Org. Nat. Prod.*, **51**, 1.
Sainsbury, M. (1979) in *Rodd's Chemistry of Carbon Compounds*, 2nd edn, (ed. S. Coffey), Elsevier, Amsterdam, Vol. IIIH, 1.
Sargent, M.V. (1984) *Prog. Chem. Org. Nat. Prod.*, **45**, 103.
Thomson, R.H. (1971) *Naturally Occurring Quinones*, 2nd edn, Academic Press, London.
Thomson, R.H. (1987) *Naturally Occurring Quinones III*, Chapman & Hall, London.
Turner, W.B. *et al.* (1983) *Fungal Metabolites II*, Academic Press, London.
Wijnsma, R. *et al.* (1986) *Prog. Chem. Org. Nat. Prod.*, **49**, 79.

Anthracyclinones (VQ4300)

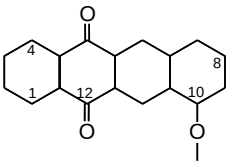
The anthracyclines produced by *Streptomyces* form a group of clinically useful antitumour agents. They also show potent activity against gram-positive bacteria but are generally too toxic to be of value. In addition to the 200 or so naturally occurring metabolites, many semisynthetic anthracyclines have been developed in the search for improved antitumour activity and lower toxicity. Biosynthesis

is from a decaketide precursor formed from nine acetates and one propanoate unit.

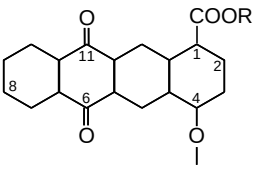
Numbering of anthracycline derivatives is confusing. The IUPAC numbering system is commonly used. However CA names them systematically as substituted naphthacenediones and in these cases the numbering depends on the hierarchy of the attached functional groups.



IUPAC



CA, e.g. Daunomycin



CA, e.g. Pyrromycin

Behal, V. *et al.* (1983) in *Biochemistry and genetic regulation of commercially important antibiotics* (ed. L.C. Vining) Addison-Wesley, pp. 255.

Eckardt, K. *et al.* (1988) *J. Basic Microbiol*, **28**, 137 (*biosynth*).

Fujiwara, A. *et al.* (1986) *CRC Crit. Rev. Biotechnol*. CRC Press, Boca Raton, **3**, 133 (*struct, biosynth*).

Kelly, T.R. (1984) *Tetrahedron*, **40,** 4537 (*synth*)

Krohn, K. (1986) *Angew. Chem. Int. Ed.*, **25,** 700 (*synth*).

Lown, J.W. (ed.) (1988) *Bioactive Reviews Vol. 6*, Elsevier, Amsterdam.

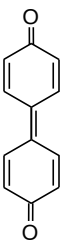
Thomas, G.J. (1990) *Recent Prog. Chem. Synth. Antibiot.*, 467 (*synth*)

Vigevani, A. *et al.* (1985) *Mag. Reson. Chem.*, **23**, 344 (*pmr, ir*).

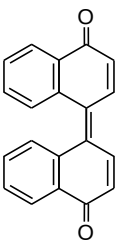
Wagner, C. *et al.* (1991) *J. Basic Microbiol.*, **31**, 223 (*biosynth*).

Extended quinones (VQ6000)

Extended quinones arise from simpler precursors by oxidative phenol coupling. Natural diphenanthroquinones have not been found. There are only a few extended *o*-quinones. The dinaphthoquinones and dianthroquinones may be further coupled to give perylene and similar ring systems.



Dibenzoquinone



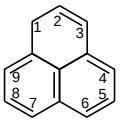
Dinaphthoquinone

Thomson, R.H. (1971) *Naturally Occurring Quinones*, 2nd edn. Academic Press, London.

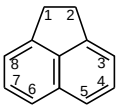
Thomson, R.H. (1987) *Naturally Occurring Quinones III*, Chapman & Hall, London.

Phenalenes and fluorenes (VQ7500, VQ7700)

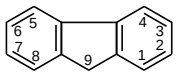
There are only a small number of natural products with the phenalene and fully aromatic fluorene skeletons.



Phenalene



Acenaphthene



Fluorene

Andrew, H.F. (1979) in *Rodd's Chemistry of Carbon Compounds*, 2nd edn. (ed. S. Coffey), Elsevier, Amsterdam, Vol. IIIH, 138.
Andrew, H.F. (1988) in *Rodd's Chemistry of Carbon Compounds*, Suppl., (ed. M.F. Ansell), Elsevier, Amsterdam, Vol IIIH, 27.
Cooke, R.G. *et al.* (1981) *Prog. Chem. Org. Nat. Prod.*, **40**, 153.
Turner, W.B. *et al.* (1983) *Fungal Metabolites II*, Academic Press, London.

Miscellaneous polycyclic aromatics (VQ9000)

Miscellaneous polycyclic aromatic systems that occur naturally probably arise by radical coupling of simpler systems.

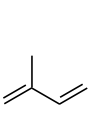
Andrew, H.F. (1988) in *Rodd's Chemistry of Carbon Compounds*, Suppl., (ed. M.F. Ansell), Elsevier, Amsterdam, Vol. IIIH, 71.
Campbell, N. *et al.* (1979) in *Rodd's Chemistry of Carbon Compounds*, 2nd edn. (ed. S. Coffey), Elsevier, Amsterdam, Vol. IIIH, 211.

Terpenoids (VS)

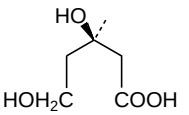
Classification of terpenoids

The immense variety of structural types found in the terpenoids was rationalised by the isoprene rule of Ruzicka. However, the number of exceptions to the regular arrangement of isoprene units led to the biogenetic isoprene rule which encompassed the possibility of rearrangements during biosynthesis. Terpenoids are thus seen as being formed from linear arrangements of isoprene units followed by various cyclisations and rearrangements of the carbon skeleton. They can also be biosynthetically modified by the loss or addition of carbon atoms. It is useful to classify terpenoids according to the number of isoprene units from which they are biogenetically derived, even though some carbons may have been added or lost. (This sometimes causes some uncertainty if it is believed that more than five carbons have been lost; only a biosynthetic study can resolve this issue. For example the irones (C_{15}) are derived biosynthetically from the iridial group (C_{31})).

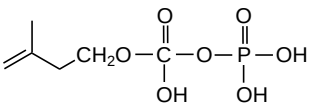
The biogenetic isoprene rule implies the involvement of a branched five carbon unit in the biosynthesis of terpenoids. Isoprene, although a natural product, is not a precursor of the terpenoids. The biosynthetic origin of this five carbon unit is well established. The pathway involves mevalonic acid which is converted into isopentenyl diphosphate, the five-carbon precursor of the terpenoids. However alternative biosynthetic pathways to isopentenyl diphosphate are now becoming common.



Isoprene



Mevalonic acid



Isopentenyl diphosphate

There are a few naturally-occurring branched C_5 compounds. These are listed in the Type of Compound Index under Hemiterpenoids (VS0050).

Chappell, J. (1995), *Ann. Rev. Plant Physiol. Plant Mol. Biol.*, **46**, 521.
Dewick, P.M. (1997), *Nat. Prod. Rep.*, **14**, 111
Hanson, J.R. (1997) in *Comprehensive Organic Chemistry* (eds D.H.R. Barton *et al.*), Pergamon, Oxford, Vol. 5, p. 989.
Loomis, W.D. *et al.* (1973) *Recent Adv. Phytochem.*, **6**, 147.
Ramos-Valdivia, A.C. *et al.* (1997) *Nat. Prod. Rep.*, **14**, 591.
Ruzicka, L. *et al.* (1921) *Helv. Chim. Acta*, **4**, 505.
Ruzicka, L. *et al.* (1953) *Experientia*, **9**, 357.
Ruzicka, L. (1959) *Proc. Chem. Soc.*, 341.

Nomenclature

The systems used for the nomenclature of terpenoids have evolved over a long period. For many terpenoid classes more than one name has been proposed for the carbon skeleton and in a large number of cases several numbering systems are in use. DNP has used the most accepted numbering system for most skeletal types. In cases for which no numbering system has been proposed or where several are in use, preference has been given to the biogenetic system.

Many terpenoids are named as formal variants of steroid structures and their nomenclature and numbering therefore follows on from that of steroids, which is described more fully in a subsequent section.

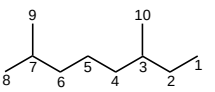
Monoterpenoids (VS0100–VS1200)

Monoterpenoids have been isolated from the fragrant oils of many plants and are important in the perfumery and flavour industries. Monoterpenoids are also found in many marine organisms, where they are generally halogenated, and as insect pheromones and defence secretions. The biosynthetic pathways of the main classes of monoterpenes have been well studied.

Banthorpe, D.V. *et al.* (1972) *Chem. Rev.*, **72**, 115.
Croteau, R. (1981) in *Biosynthesis of Isoprenoid Compounds*, (eds J.W. Porter *et al.*), Wiley, New York, Vol. 1, p. 225.
Croteau, R. (1987) *Chem. Rev.*, **87**, 929.
Croteau, R. *et al.* (1994), *Recent Adv. Phytochem.*, **28**, 193.
Grayson, D.H. *et al.* (1997) *Nat. Prod. Rep.*, **14**, 477.
Hill, R.A. (1993) in *The Chemistry of Natural Products*, 2nd edn (Ed. R.H. Thomson), Blackie, Glasgow, pp. 107.
Lerdau, M. (1994), *Trends Ecol. Evol.*, **9**, 58.

Acyclic monoterpenoids (VS0100)

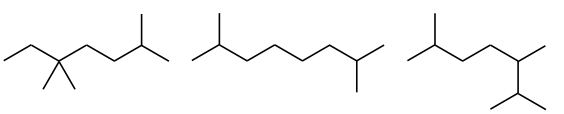
In this section are grouped the regular linear monoterpenoids, that is those formed by a head to tail arrangement of the isoprene units. They are principally found in plants and in insect exudates. No semi-systematic name has been ascribed to this carbon skeleton probably because the systematic 2,6-dimethyloctane naming is straightforward. The numbering system shown below is in line with that used with other acyclic terpenoids.



Regular acyclic monoterpenoid skeleton
2,6-Dimethyloctane, 9CI, 8CI

Irregular acyclic monoterpenoids (VS0150)

Some acyclic monoterpenoids arise from other arrangements of the isoprene units. These may arise by alternative linkages of the units, e.g. head to head, by rearrangement of regular acyclic monoterpenoids or by cleavage of cyclic monoterpenoids.



Irregular acyclic monoterpenoid skeletons

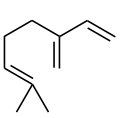
Halogenated dimethyloctane monoterpenoids (VS0200)

This group is obtained principally from marine organisms. They are all regular acyclic monoterpenoids and the numbering system follows the accepted pattern.

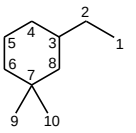
Naylor, S. *et al.* (1983) *Prog. Chem. Org. Nat. Prod.*, **44**, 189.

Ochtodane monoterpenoids (VS0220)

The ochtodanes are also principally from marine organisms particularly *Ochtodes* spp. and presumably arise by cyclisation of myrcene.



Myrcene

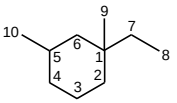


Ochtodane
3-Ethyl-1,1-dimethylcyclohexane, 9CI

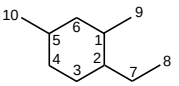
Naylor, S. *et al.* (1983) *Prog. Chem. Org. Nat. Prod.*, **44**, 189.

1-Ethyl-1,3-dimethylcyclohexane and 1-ethyl-2,4-dimethylcyclohexane monoterpenoids (VS0240, VS0260)

These two terpenoid skeletons are only found in marine organisms and presumably arise by cyclisation of a regular acyclic monoterpene skeleton followed in the latter case by an ethyl migration. The numbering systems reflect their probable biogenetic relationship.



1-Ethyl-1,3-dimethyl-
cyclohexane



1-Ethyl-2,4-dimethyl-
cyclohexane

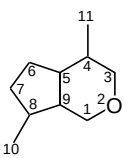
Naylor, S. *et al.* (1983) *Prog. Chem. Org. Nat. Prod.*, **44**, 189.

Cyclopropane and cyclobutane monoterpenoids (VS0300, VS0350)

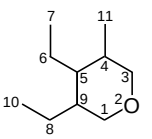
These include the pyrethrin terpenoids such as **Chrysamthemic acid**, and **Grandisol** the pheromone of the male boll weevil.

Iridoid, 10-alkyliridoid and secoiridoid monoterpenoids (VS0400, VS0420, VS0440)

The iridoids and secoiridoids form a large group of plant constituents that are found usually, but not invariably, as glycosides. Their biosynthesis has been well established. C-11 is missing in some iridoids. The nitrogen-containing iridoids and the large and important group of alkaloids derived from Secologanin are described in the alkaloid section below.



Iridoid skeleton
4,7-Dimethylcyclopenta[*c*]pyran, 9CI

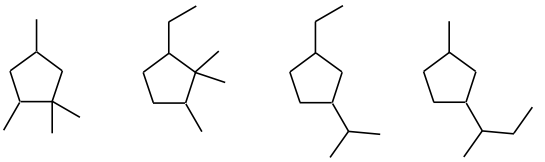


Secoiridoid skeleton

Boros, C.A. *et al.* (1990) *J. Nat. Prod.*, **53**, 1055.
El-Naggar, L.J. *et al.* (1980) *J. Nat. Prod.*, **43**, 649.
Inouye, H. *et al.* (1986) *Prog. Chem. Org. Nat. Prod.*, **50**, 169.
Inouye, H. *et al.* (1991) *Methods Plant Biochem.*, **7**, 99.

Other cyclopentane monoterpenoids (VS0450)

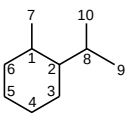
This is a large group containing several carbon skeletons that probably arise biogenetically by cleavage of bicyclic monoterpenoids.



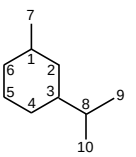
Some cyclopentane monoterpenoid skeletons

Menthane monoterpenoids (VS0500, VS0520, VS0540)

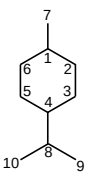
The menthane group comprises three isomeric types, *o*-, *m*- and *p*-menthanes. The *p*-menthanes are the most widespread and arise by a cyclisation of a regular acyclic monoterpenoid. The *o*- and *m*- menthanes are much rarer, and presumably arise by alkyl migration of *p*-menthanes. The numbering systems of the menthanes reflect their biogenetic relationship. Since *p*-menthane has a plane of symmetry the numbering of ring substituents is chosen to give the lowest numbers consistent with the avoidance of compound locants for double bonds when possible. For example the name *p*-menth-1-en-6-one is preferred to *p*-menth-1(6)-en-2-one.



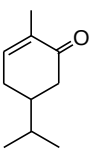
o-Menthane



m-Menthane



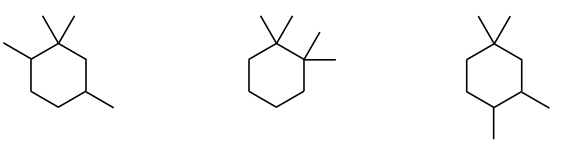
p-Menthane, 8CI
1-Methyl-4-(1-methylethyl)-cyclohexane, 9CI



p-Menth-1-en-6-one

Other cyclohexane monoterpenoids (VS0600)

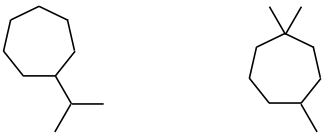
Cyclohexane skeletons that are not included in previous groups are collected here. They probably arise by cleavage of bicyclic monoterpenoids.



Other cyclohexane monoterpenoid skeletons

Cycloheptane monoterpenoids (VS0700)

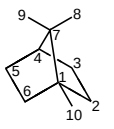
This small group of compounds may arise by ring expansion of the *p*-menthane skeleton.



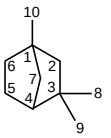
Cycloheptane monoterpenoid skeletons

Bicyclic monoterpenoids (VS0800–VS1050)

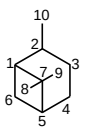
The bicyclic monoterpenoids arise biogenetically by further cyclisation of monocyclic terpenoids followed by various rearrangements. They easily undergo a wide variety of skeletal rearrangements and many may be artifacts produced during isolation procedures. The numbering systems given below are those most commonly used and follow the systematic (Von Baeyer) numbering scheme, although particularly in the older literature several numbering systems can be found, e.g. for the pinanes.



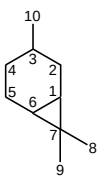
Camphane
1,7,7-Trimethylbicyclo-[2.2.1]heptane, 9CI
VS0800



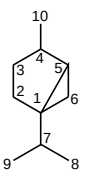
Fenchane
1,3,3-Trimethylbicyclo-[2.2.1]heptane, 9CI
VS0850



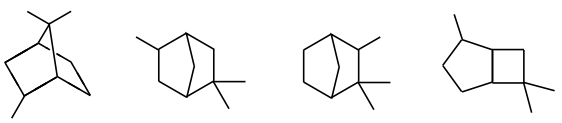
Pinane
2,6,6-Trimethylbicyclo[3.1.1]heptane, 9CI
VS0900



Carane
3,7,7-Trimethylbicyclo-[4.1.0]heptane, 9CI
VS0950



Thujane
4-Methyl-1-(1-methylethyl)-bicyclo[3.1.0]hexane, 9CI
VS1000



Miscellaneous bicyclic monoterpenoids VS1050

Pelter, A. *et al.* (1969) in *Rodd's Chemistry of Carbon Compounds*, 2nd edn. Elsevier, Amsterdam, IIC, 136.

Tricyclic monoterpenoids (VS1200)

Tricyclene and **Teresantalol** are two examples of this small group of natural compounds.

Tricyclene
1,7,7-Trimethyltricyclo[2.2.1.$0^{2,6}$]heptane, 9CI

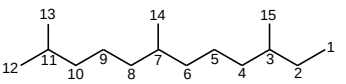
Sesquiterpenoids (VS1300–VS5320)

The sesquiterpenoids are C_{15} compounds formed by the assembly of three isoprenoid units. They are found in many living systems but particularly in higher plants. There is a large number of sesquiterpenoid carbon skeletons, which all however arise from the common precursor, farnesyl pyrophosphate, by various modes of cyclisations followed, in many cases, by skeletal rearrangement.

Bryant, R. (1969) in *Rodd's Chemistry of Carbon Compounds*, 2nd edn. Elsevier, Amsterdam, IIC, 256.
Cane, D.E. (1981) in *Biosynthesis of Isoprenoid Compounds*, (eds J.W. Porter *et al.*), Wiley, New York, Vol. 1, p. 283.
Cane, D.E. (1990) *Chem. Rev.*, **90**, 1089.
Cordell, G.A. (1976) *Chem. Rev.*, **76**, 425.
Fraga, B.M. (1998) *Nat. Prod. Rep.*, **15**, 73
Hill, R.A. (1993) in *The Chemistry of Natural Products*, 2nd edn (ed. R.H. Thomson), Blackie, Glasgow, pp. 117.
Parker, W. *et al.* (1967) *Quart. Rev. Chem. Soc.*, **21**, 331.
Roberts, J.S. (1972) in *Chemistry of Terpenes and Terpenoids*, (ed. A.A. Newman) Academic Press, London, p. 88.

Simple farnesane sesquiterpenoids (VS1300)

The simple farnesanes are named semi-systematically in this Dictionary although the systematic trimethyldodecane naming is used extensively in the literature. The farnesane numbering system is used as a biogenetic numbering system for many sesquiterpenoid skeletons.



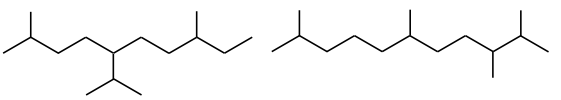
Farnesane
2,6,10-Trimethyldodecane, 9CI

Furanoid farnesane sesquiterpenoids (VS1320)

Although many numbering systems have been used for furanoid farnesanes, such as **Ngaione**, it is logical to use the farnesane numbering system for this group.

Irregular acyclic sesquiterpenoids (VS1400)

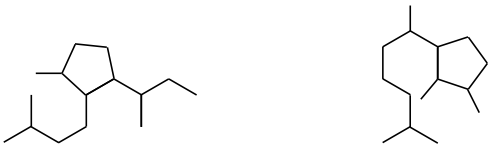
The various skeletons that are not clearly head to tail arrangements of isoprenoid units may arise by rearrangement of a regular acyclic precursor or by cleavage of a cyclic terpenoid structure such as a cembrane.



Some irregular acyclic sesquiterpenoid skeletons

Miscellaneous cyclobutane and cyclopentane sesquiterpenoids (VS1420, VS1430)

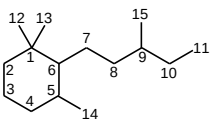
This group contains a variety of miscellaneous cyclobutane and cyclopentane sesquiterpenoids



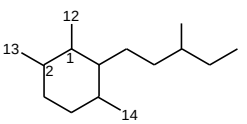
Some miscellaneous cyclopentane sesquiterpenoid skeletons

Cyclofarnesane and rearranged cyclofarnesane sesquiterpenoids (VS1450, VS1460, VS1470)

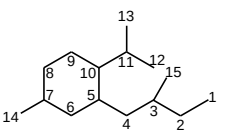
Cyclofarnesanes arise by formation of a six-membered ring between carbons 6 and 11 of farnesane. They include **Abscisic acid** whose accepted numbering system is 'non-farnesane'. Methyl group migration gives the rearranged cyclofarnesane skeleton. The herbertianes, included in this group, (not to be confused with herbertanes) are 5,10-cyclofarnesanes.



Cyclofarnesane
1,1,3-Trimethyl-2-(3-methylpentyl)cyclohexane



Rearranged cyclofarnesane skeleton
1,2,4-Trimethyl-3-(3-methylpentyl)cyclohexane

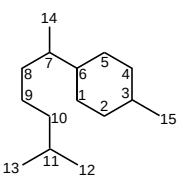


Herbertiane
4-Methyl-2-(2-methylbutyl)-1-(1-methylethyl)cyclohexane

Milborow, B.V. *et al.* (1991) *Methods in Plant Biochem.*, **7**, 213.
Weyerstahl, P. (1992) *Annalen*, 1325

Bisabolane sesquiterpenoids (VS1500)

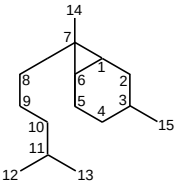
The bisabolanes are a fairly large group mainly found as constituents of higher plants. The numbering system used for bisabolanes is the same as the farnesane system. Since the cyclohexane ring has a plane of symmetry, substituents in this ring should be numbered where possible avoiding the compound locant, 1(6), for a double bond and keeping the numbers for the substituents in the cyclohexane ring as low as possible.



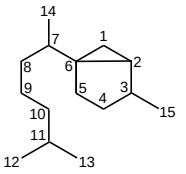
Bisabolane
1-(1,5-Dimethylhexyl)-4-methylcyclohexane, 9CI

Cyclobisabolane sesquiterpenoids (VS1550)

This section includes the sesquicarane and sesquisabinane carbon skeletons numbered in accordance with bisabolane.



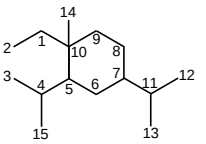
Sesquicarane
3,7-Dimethyl-7-(4-methyl-pentyl)bicyclo-[4.1.0]heptane, 9CI



Sesquisabinane
1-(1,5-Dimethylhexyl)-4-methylbicyclo[3.1.0]-hexane, 9CI

Elemane sesquiterpenoids (VS1600)

Elemanes are numbered consistently with eudesmanes (see below) and germacranes. They are rapidly formed *in vitro* by Cope rearrangement of the corresponding 1(10), 4-germacradienes and it is possible that they are artifacts produced during the isolation procedures. Some elemanes are oxidatively modified, e.g. **Vernolepin** which presumably is not formed by a Cope rearrangement during isolation.

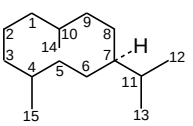


Elemane
1-Ethyl-1-methyl-2,4-bis-(1-methylethyl)cyclohexane, 9CI

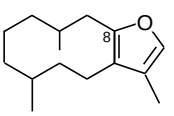
Germacranes (VS1650–VS1700)

The numbering of the germacrane skeleton poses a problem since there is a plane of symmetry through carbons 2 and 7. Germacranes are normally drawn in a conventional way as shown below with H-7 in the α-configuration. Care should be taken with the small number of germacranes with a double bond at C-7 as the ring can be numbered in either direction. Germacranes frequently have double bonds in the 1(10) and 4 positions. There have been proposals to give different names to the skeletons with (1(10)*Z*, 4*E*) (melampolides) and (1(10)*E*, 4*Z*) (heliangolides) configurations. However this is confusing and in DNP all compounds are named as germacranes. A further problem with the representation of germacranes arises from substituents at carbons drawn as reentrant angles. Wherever possible germacranes should be drawn without substituents at reentrant centres as in this Dictionary, and care should be exercised when reading the literature.

The large germacrane group is divided into simple germacranes, that is those without a lactone or furan ring (VS1650), 12,6-germacranolides (VS1660), 12,8-germacranolides (VS1670), furanogermacranes, nor- and homo-germacranes (VS1680), secogermacranes (VS1690) and cyclogermacranes (VS1700).



Germacrane
1,7-Dimethyl-4-(1-methyl-ethyl)cyclodecane, 9CI

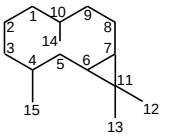


Furanogermacrane
4,5,6,7,8,9,10,11-Octahydro-3,6,10-trimethylcyclodeca[*b*]-furan, 9CI

Brown, D.S. *et al.* (1992) *Heterocycles*, **34**, 807.
Fischer, N.H. *et al.* (1979) *Prog. Chem. Org. Nat. Prod.*, **38**, 47.
Fischer, N.H. (1990) *Recent Adv. Phytochem.*, **24**, 161.

Lepidozanes and bicyclogermacrane sesquiterpenoids (VS1710)

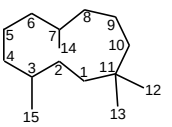
Bicyclogermacranes, found in higher plants, have a *cis*-fused cyclopropane ring junction whereas the stereoisomeric lepidozanes from liverworts have a *trans*-fused ring junction.



Bicyclogermacrane
3,7,11,11-Tetramethylbicyclo[8.1.0]undecane, 9CI

Humulane sesquiterpenoids (VS1720)

At least three numbering systems are in use for the humulane skeleton. In DNP the farnesane numbering system has been chosen rather than those based on germacrane numbering since the biosynthetic pathway to humulane does not involve germacrane.

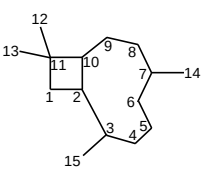


Humulane
1,1,4,8-Tetramethylcycloundecane, 9CI

Gonzaléz, A.G. *et al.*, (1995), *Prog. Chem. Org. Nat. Prod.*, **64**, 1

Caryophyllane sesquiterpenoids (VS1730)

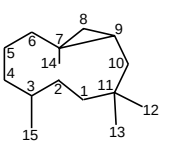
Further cyclisation of the humulane skeleton between carbons 2 and 10 produces the caryophyllane skeleton. Several numbering systems are in use for caryophyllanes; the one chosen for DNP is that based on farnesane.



Caryophyllane
2,6,10,10-Tetramethylbicyclo[7.2.0]undecane, 9CI

Bicyclohumulane sesquiterpenoids (VS1740)

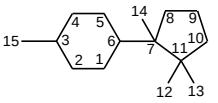
A small group of 7,9-cyclohumulanes is named as bicyclohumulanes.



Bicyclohumulane
1,5,8,8-Tetramethylbicyclo[8.1.0]undecane, 9CI

Cuparane sesquiterpenoids (VS1750)

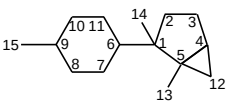
Cuparane is formed by cyclisation between carbons 6 and 11 of the bisabolane skeleton and the numbering system used here takes account of this fact. Cuparanes are found in liverworts, higher plants and marine organisms.



Cuparane
(Most have an aromatic ring and are named in CA
as substituted benzenes)

Cyclolaurane sesquiterpenoids (VS1760)

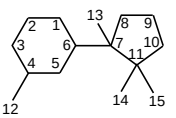
Cyclolauranes found in marine organisms may be considered as cyclocuparanes but as they co-occur with lauranes, the numbering system has been chosen to agree with the accepted laurane system.



Cyclolaurane
1,2-Dimethyl-2-(4-methylcyclohexyl)bicyclo[3.1.0]hexane, 9CI

Herbertane sesquiterpenoids (VS1800)

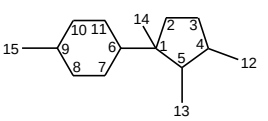
Herbertanes are a small group of compounds isolated from liverworts and fungi. As the biosynthesis of this skeleton is not known the numbering system proposed in the literature is used.



Herbertane
(Mostly named as substituted benzenes in 9CI)

Laurane sesquiterpenoids (VS1850)

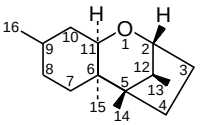
Lauranes are found in marine organisms, particularly *Laurencia* spp.



Laurane
(Mostly named as substituted benzenes in 9CI)

Trichothecane sesquiterpenoids (VS1900)

The biologically active trichothecanes are produced by various species of imperfect fungi. Probably their most important biological activity is their cytostatic action. Most trichothecanes contain a 12,13-epoxide grouping (Scirpane) and for convenience many trichothecanes are named as derivatives of the parent Scirpane in DNP.



Trichothecane, 9CI

A group of naturally occurring trichothecanes are macrocyclic lactones, incorporating a bridge derived from mevalonic acid and acetate. An example is **Baccharinol**.

Betina, V. (ed.) (1984) *Mycotoxins: Production, Isolation, Separation and Purification*, Elsevier, Amsterdam, Vol. 8.

Grove, J.F. (1996), *Prog. Chem. Org. Nat. Prod.*, **69**, 1

Johann Wolfgang Goethe-universitaet (1984) *Synform*, 2 (*synth*).

Lacey, J. (1987) *Trichothecenes and other Mycotoxins*, Wiley, New York.

McDougal, P.G. *et al.* (1985) *Prog. Chem. Org. Nat. Prod.*, **47**, 153.

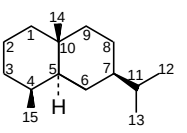
Steyn, P.S. (ed.) (1986) *Mycotoxins and Phycotoxins*, Elsevier, Amsterdam.

Ueno, Y. (ed.) (1983) *Trichothecenes, Chemical, Biological and Toxicological Aspects*, Elsevier, Amsterdam.

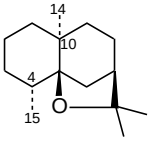
Wylie, T. *et al.* (1977) *Mycotoxic fungi, Mycotoxins and Mycotoxicoses*, Dekker, New York (3 vols).

Eudesmane sesquiterpenoids (VS1950–VS2000)

Eudesmanes are called selinanes in the older literature. The eudesmanes found in higher plants generally have the stereochemistry shown below. *ent*-eudesmanes are found in some species of liverworts. As with the germacrane group, the eudesmanes are divided into groups comprising simple eudesmanes (VS1950), eudesman-12,6-olides (VS1970), eudesman-12,8-olides and furanoeudesmanes (VS1975), secoeudesmanes (VS1990), and noreudesmanes (VS2000). There is also a large group of esters based on the dihydro-*β*-agarofuran skeleton which are grouped separately (VS1980). Within the eudesmane group, particularly with dihydro-*β*-agarofuran derivatives, there is some confusion concerning the numbering of carbons 14 and 15. The numbering given here should be adopted.



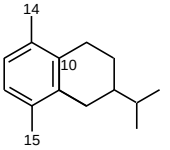
Eudesmane
Decahydro-1,4*a*-dimethyl-7-(1-methylethyl)-naphthalene, 9CI



Dihydro-β-agarofuran
2,2,5,9-Tetramethyl-2*H*-3,9*a*-methano-1-benzoxepin, 9CI

Emmotin sesquiterpenoids (VS2010)

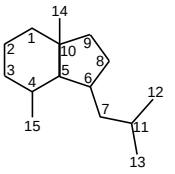
The emmotin group are 14(10 → 1)-abeoeudesmanes which have an aromatic ring A and the methyl at C-10 migrated to C-1.



Emmotin skeleton

Oppositane sesquiterpenoids (VS2020)

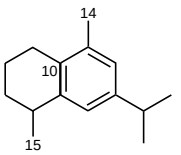
The oppositanes are 8(7 → 6)-abeoeudesmanes and are found in plants and marine organisms.



Oppositane
Octahydro-3*a*,7-dimethyl-1-(2-methylpropyl)-1*H*-indene, 9CI

Farfugin sesquiterpenoids (VS2040)

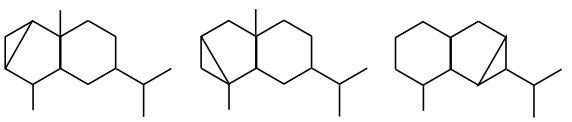
The farfugin group are 14(10 → 9)-abeoeudesmanes which have an aromatic ring B and the methyl at C-10 migrated to C-9.



Farfugin skeleton

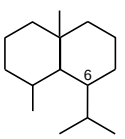
Cycloeudesmane sesquiterpenoids (VS2050)

Various cycloeudesmanes are included in this section.



Gorgonane sesquiterpenoids (VS2060)

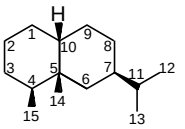
The gorgonanes are derived from eudesmanes by isopropyl group migration to C-6.



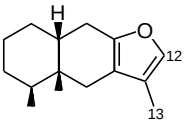
Gorgonane
Decahydro-1,4*a*-dimethyl-8-(1-methylethyl)naphthalene, 9CI

Eremophilane sesquiterpenoids (VS2100–VS2130)

The eremophilanes have been shown to be derived from eudesmanes by migration of the methyl group at C-10 to C-5. There is confusion in the literature about the numbering of carbons 14 and 15; the biogenetic numbering given below should be used. The normal stereochemistry is shown, although there are several exceptions to this. As with the other larger categories, the eremophilanes are separated into simple eremophilanes (VS2100), eremophilanolides and furanoeremophilanes (VS2110), seco- and abeoeremophilanes (VS2120) and noreremophilanes (VS2130).



Eremophilane
Decahydro-1,8*a*-dimethyl-
7-(1-methylethyl)-
naphthalene, 9CI

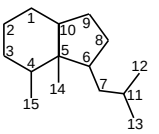


Furanoeremophilane
4,4*a*,5,6,7,8,8*a*,9-Octahydro-
3,4*a*,5-trimethylnaphtho-
[2,3-*b*]furan, 9CI

Pinder, A.R. (1977) *Prog. Chem. Org. Nat. Prod.*, **34**, 82.

Chiloscyphane sesquiterpenoids (VS2140)

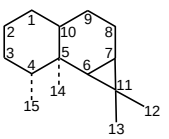
The chiloscyphanes, a small group isolated from liverworts, are 8(7 → 6)-abeoeremophilanes.



Chiloscyphane
Octahydro-7,7*a*-dimethyl-1-(2-methylpropyl)-1*H*-indene, 9CI

Aristolane sesquiterpenoids (VS2150)

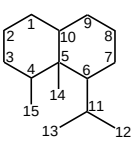
The aristolanes, isolated *inter alia* from *Aristolochia* spp., are 6,11-cycloeremophilanes.



Aristolane
Decahydro-1,1,7,7*a*-tetramethyl-1*H*-cyclopropa-
[*a*]naphthalene, 9CI

Nardosinane sesquiterpenoids (VS2160)

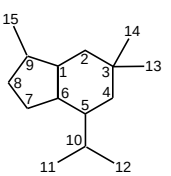
The nardosinanes, isolated from marine organisms, are eremophilanes in which the isopropyl group has migrated to carbon 6.



Nardosinane
Decahydro-1,8*a*-dimethyl-8-(1-methylethyl)naphthalene

Brasilane sesquiterpenoids (VS2170)

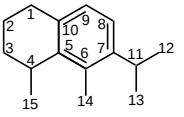
No biosynthetic scheme has been proposed for this skeleton.



Brasilane
Octahydro-1,6,6-trimethyl-4-(1-methylethyl)-1*H*-indene

Cacalol sesquiterpenoids (VS2180)

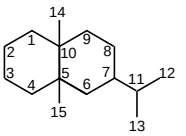
The cacalol sesquiterpenoids occur, *inter alia*, in *Cacalia* spp. and are eremophilanes, typically with an aromatic ring B, in which carbon-14 has further migrated to C-6.



Cacalol skeleton

Valerane sesquiterpenoids (VS2200)

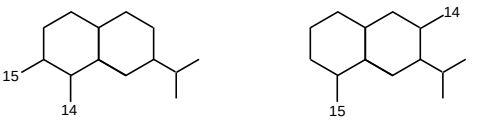
The valeranes (also called jatamansanes), mainly from *Valeriana* spp., are eudesmanes where the methyl group at C-4 has migrated to C-5. There is normally a carbonyl group at C-4.



Valerane
Decahydro-4*a*,8*a*-dimethyl-2-(1-methylethyl)naphthalene

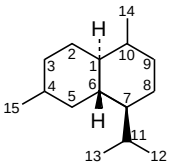
Miscellaneous rearranged eudesmane sesquiterpenoids (VS2220)

Various methyl-migrated eudesmanes are grouped together here. The methyl groups should retain their numbering on migration.

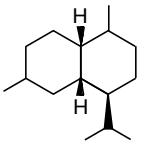


Cadinane sesquiterpenoids (VS2250, VS2260)

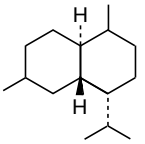
The nomenclature, numbering and absolute stereochemistry of this group is somewhat confused. Biogenetic (germacrane) numbering is used in DNP, but many other numbering systems have been used in the literature. The names of the skeletons depend on the relative stereochemistries at carbons 1, 6 and 7 as indicated. Moreover, the aromatised skeletons are given different names, calamenene and cadalene, and these are often given different numbering systems. Nor-and seco-cadinanes are grouped separately (VS2260).



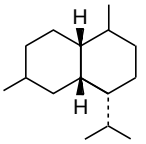
Cadinane
Decahydro-1,6-dimethyl-
4-(1-methylethyl)naphthalene, 9CI



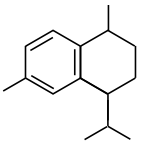
Muurolane



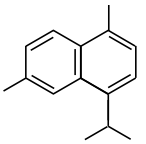
Bulgarane



Amorphane



Calamenene

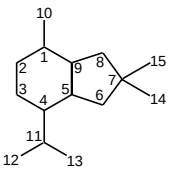


Cadalene

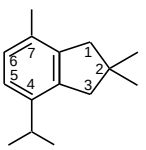
Bordoloi, M. *et al.* (1989) *Phytochemistry*, **28**, 2007.

Alliacane sesquiterpenoids (VS2270)

The alliacanes from *Marasmius alliacus* and the aromatic primnatrienes from a *Primnoeides* spp. have unfortunately been assigned different numbering systems. The biosynthesis of this skeleton has not been established.



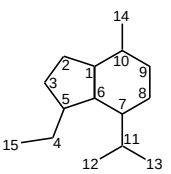
Alliacane
Octahydro-2,2,4-trimethyl-
7-(1-methylethyl)-
1*H*-indene, 9CI



Primnatriene

Oplopane sesquiterpenoids (VS2280)

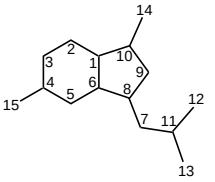
Oplopanes, from higher plants, are 3(4 → 5)-abeocadinanes and the numbering system used here is biogenetic.



Oplopane
1-Ethyloctahydro-4-methyl-7-(1-methylethyl)-1*H*-indene

Mutisianthol sesquiterpenoids (VS2290)

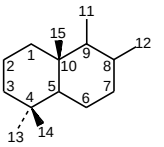
This is a small group of 6(7 → 8)-abeocadinanes from *Jungia* and *Mutisia* spp. Again the numbering system is biogenetic.



Mutisianthol group skeleton
Octahydro-1,5-dimethyl-3-(2-methylpropyl)-1*H*-indene

Drimane sesquiterpenoids (VS2300)

The drimanes, from fungi and higher plants, arise by direct cyclisation of a farnesane derivative. The accepted numbering system is shown. Compounds of the *ent*- series such as **Iresin** were isolated earlier from *Iresine* spp. Nor- and secodrimanes are grouped separately (VS2320).

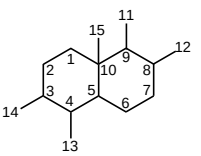


Drimane
Decahydro-1,1,4*a*,5,6-pentamethylnaphthalene

Cordell, G.A. (1976) *Chem. Rev.*, **76**, 425.
Djerassi, C. *et al*. (1958) *J. Am. Chem. Soc.*, **80**, 2593.
Jansen, B.J.N. *et al.* (1991) *Nat. Prod. Rep.*, **8**, 309;319.

Coloratane sesquiterpenoids (VS2310)

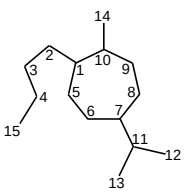
The coloratanes, few in number, are related to the drimanes. A methyl migration from C-4 to C-3 has occurred.



Coloratane
Decahydro-1,2,4*a*,5,6-pentamethylnaphthalene

Xanthane sesquiterpenoids (VS2380)

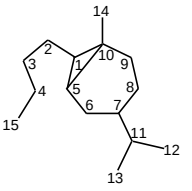
The xanthanes, originally isolated from *Xanthium* spp., are 4,5-secoguaianes.



Xanthane
1-Butyl-2-methyl-5-(1-methylethyl)cycloheptane

Carabrane sesquiterpenoids (VS2390)

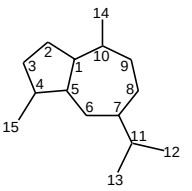
The carabranes are a small group of 5,10-cycloxanthanes.



Carabrane
7-Butyl-1-methyl-4-(1-methylethyl)bicyclo[4.1.0]heptane

Guaiane sesquiterpenoids (VS2400–VS2440)

This large group is divided into simple guaianes (VS2400), 12,6-guaianolides (VS2410), 12,8-guaianolides (VS2420), guaiane dimers (VS2430), and seco-, cyclo- and abeoguaianes (VS2440).

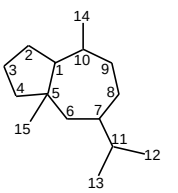


Guaiane
Decahydro-1,4-dimethyl-7-(1-methylethyl)azulene, 9CI

Fischer, N.H. *et al.* (1979) *Prog. Chem. Org. Nat. Prod.*, **38**, 47.
Fischer, N.H. *et al.* (1990) *Recent Adv. Phytochem.*, **24**, 161.

Pseudoguaiane sesquiterpenoids (VS2450, VS2470)

The migration of a methyl group of the guaiane skeleton from C-4 to C-5 produces the pseudoguaiane skeleton. There is often an oxygen function at C-4.

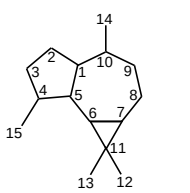


Pseudoguaiane
Decahydro-4,8*a*-dimethyl-7-(1-methylethyl)azulene

Fischer, N.H. *et al.* (1974) *Prog. Chem. Org. Nat. Prod.*, **38**, 47.

Aromadendrane sesquiterpenoids (VS2500)

The aromadendranes are 6,11-cycloguaianes. The smaller groups of 5,10-cycloaromadendranes and seco-aromadendranes, including the plagiochilins which are 2,3-secoaromadendranes, are listed separately (VS2520, VS2540).

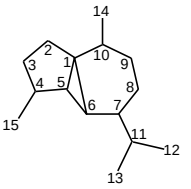


Aromadendrane
Decahydro-1,1,4,7-tetramethyl-1*H*-cycloprop[*e*]azulene, 9CI

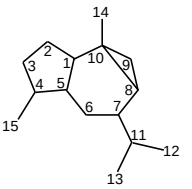
Gijsen, H.J.M. *et al.* (1995), *Prog. Chem. Org. Nat. Prod.*, **64**, 149

Cubebane and ivaxillarane sesquiterpenoids (VS2600, VS2620)

The small groups of cubebanes and ivaxillaranes are 1,6- and 8,10-cycloguaianes respectively.



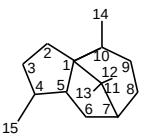
Cubebane
Octahydro-3,7-dimethyl-4-(1-methylethyl)-1*H*-cyclopenta[1,3]cyclopropa[1,2]-benzene, 9CI



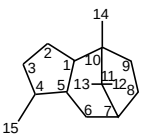
Ivaxillarane

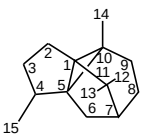
Patchoulane and rearranged patchoulane sesquiterpenoids (VS2650, VS2660)

Patchoulanes are 1,11-cycloguaianes. Various rearranged patchoulanes are also found, in for example patchouli oil. Biogenetic numbering has been used here.



Patchoulane

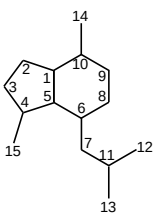




Rearranged patchoulanes

Valerenane sesquiterpenoids (VS2710)

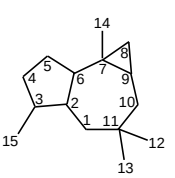
Valerenanes are 8(7 → 6)-abeoguaianes. Only a few representatives have been reported. They are not to be confused with the valeranes (see above).



Valerenane
Octahydro-1,4-dimethyl-7-(2-methylpropyl)-1*H*-indene

Africanane sesquiterpenoids (VS2750)

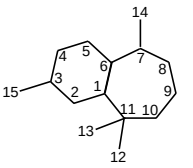
The farnesane numbering system is used for the africanane skeleton although the biosynthesis has not been established conclusively. Some compounds (e.g. **Africanone**) have been named as africananes but have since been shown to have a different skeleton.



Africanane
Decahydro-3,3,5,7*b*-tetramethyl-1*H*-cycloprop[*e*]azulene

Lippifoliane and himachalane sesquiterpenoids (VS2760, VS2780)

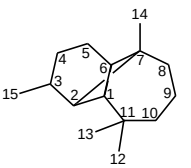
Many numbering systems have been used for the himachalane skeleton. The farnesane system has been used here. The small group of lippifolianes are 7,9-cyclohimachalanes.



Himachalane
Decahydro-2,5,9,9-tetramethyl-1*H*-benzocycloheptene, 9CI

Longipinane sesquiterpenoids (VS2800)

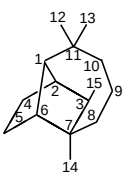
Longipinanes are 2,7-cyclohimachalanes. Several numbering systems have been used for this skeleton, the farnesane system being used for consistency here. The absolute configuration of this group is in some doubt.



Longipinane
2,6,6,9-Tetramethyltricyclo[5.4.0.0^{2,8}]undecane, 9CI

Longifolane sesquiterpenoids (VS2850)

Several numbering systems have been used in the older literature but as the biosynthesis has been established, the biogenetic farnesane system is used here.

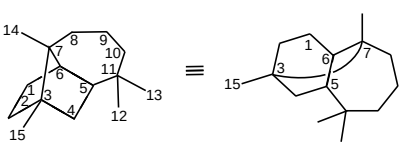


Longifolane
Decahydro-4,8,8,9-tetramethyl-1,4-methanoazulene, 9CI

Dev, S. (1981) *Acc. Chem. Res.*, **14**, 82.
Dev, S. (1981) *Prog. Chem. Org. Nat. Prod.*, **40**, 49.

Longibornane sesquiterpenoids (VS2900)

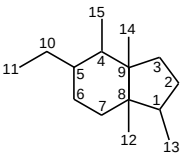
Longibornanes are 3,7-cyclohimachalanes.



Longibornane
Decahydro-4,5,5,8*a*-tetramethyl-1,4-methanoazulene, 9CI

Pinguisane sesquiterpenoids (VS3000)

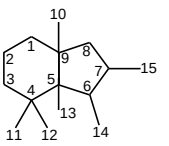
The structures of several pinguisanes from liverworts have recently been revised. Many of the trivial names in this group are confusing. The commonly used numbering system is shown.



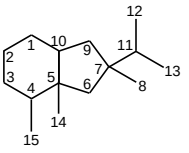
Pinguisane
5-Ethyloctahydro-1,3*a*,4,7*a*-tetramethyl-1*H*-indene

Thapsane and fukinane sesquiterpenoids (VS3050, VS3080)

The thapsanes occur in *Thapsia* spp. while fukinanes are found in *Petasites japonicus*. Both types are relatively uncommon. One group of workers has defined the thapsane skeleton as including an oxygen bridge, but that is not followed here.



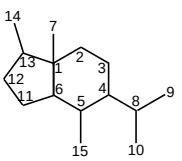
Thapsane
Octahydro-1,2,3*a*,7,7,7*a*-hexamethyl-1*H*-indene



Fukinane
Octahydro-2,3*a*,4-trimethyl-2-(1-methylethyl)-1*H*-indene

Picrotoxane sesquiterpenoids (VS3100)

The picrotoxanes are bitter, toxic constituents of the Orchidaceae. They are generally highly oxygenated. The numbering system in general use seems to be based on the menthane system and is inconsistent with the majority of other schemes.

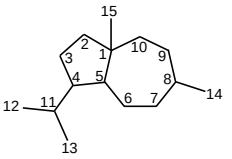


Picrotoxane
Octahydro-1,4,7*a*-trimethyl-5-(1-methylethyl)-1*H*-indene

Fischer, N.H. *et al.* (1979) *Prog. Chem. Org. Nat. Prod.*, **38**, 47.

Daucane sesquiterpenoids (VS3180)

Many numbering systems have been used for the daucanes; that chosen here is related to the guaiane system. Daucanes are also called carotanes but this name is not recommended because of possible confusion with the carotenoids.

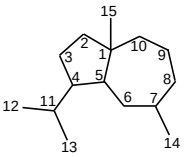


Daucane
Decahydro-3*a*,6-dimethyl-1-(1-methylethyl)azulene

Fraga, B.M. (1989) in *Studies in Natural Products Chemistry*, (ed. Atta-ur-Rahman) Elsevier, Amsterdam, p. 721.
Ghisalberti, E.L. (1994), *Phylochemistry*, **37**, 597.
Gonzaléz, A.G. *et al.* (1995), *Prog. Chem. Org. Nat. Prod.*, **64**, 1.

Isodaucane sesquiterpenoids (VS3190)

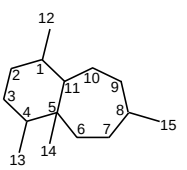
The isodaucanes (also called salviolanes) are clearly related to daucanes. The numbering system used here is again related to the guaiane system.



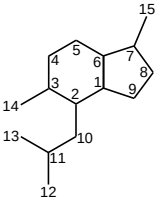
Isodaucane
Decahydro-3*a*,7-dimethyl-1-(1-methylethyl)azulene

Perforane and pacifigorgiane sesquiterpenoids (VS3200, VS3350)

The perforanes form a small group found in *Laurencia* spp. Pacifigorgianes are found in liverworts, higher plants and marine organisms.



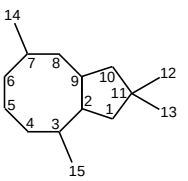
Perforane
Decahydro-1,4,7,9*a*-tetra-methyl-1*H*-benzocyclo-heptane



Pacifigorgiane
Octahydro-1,5-dimethyl-4-(2-methylpropyl)-1*H*-indene

Asteriscane sesquiterpenoids (VS3380)

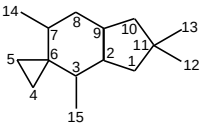
The asteriscanes form a small group isolated from *Asteriscus* spp. The farnesane numbering system is used here.



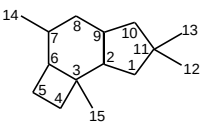
Asteriscane
Decahydro-2,2,4,8-tetramethyl-1*H*-cyclopentacyclooctene, 9Cl

Illudane and protoilludane sesquiterpenoids (VS3400, VS3420)

Although historically different numbering systems have been proposed for the illudane skeleton and the related groups, the biogenetic farnesane numbering is shown here.



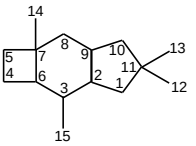
Illudane
Decahydro-2′,2′,4′,6′-tetramethylspiro[cyclopropane-1,5′-[5*H*]indene], 9CI



Protoilludane
Decahydro-3,6,6,7*b*-tetramethyl-1*H*-cyclobut[*e*]indene, 9CI

Sterpurane sesquiterpenoids (VS3430)

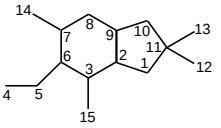
Sterpuranes are biogenetically related to the illudanes. The farnesane numbering system is used in preference to others.



Sterpurane
Decahydro-2*a*,5,5,7-trimethyl-1*H*-cyclobut[*f*]indene, 9CI

Illudalane sesquiterpenoids (VS3470)

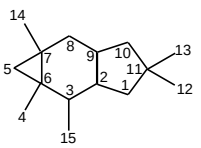
The illudalanes are 4,6-secoilludanes and include the pterosins from bracken, most of which have an aromatised ring.



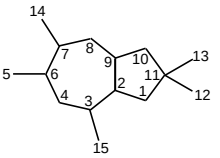
Illudalane
5-Ethyloctahydro-2,2,4,6-tetramethyl-1*H*-indene, 9CI

Isolactarane, merulane, lactarane and marasmane sesquiterpenoids (VS3470, VS3475, VS3480, VS3500)

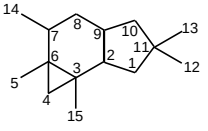
These four groups are biogenetically related to the illudanes and the numbering system used here reflects this fact. Numbering systems in the literature are similar but care should be exercised with the numbering of the methyl groups.



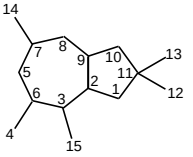
Isolactarane



Lactarane



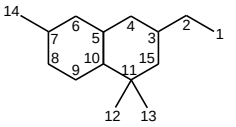
Marasmane



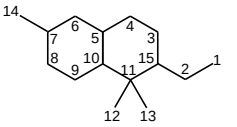
Merulane

Furodysin and furodysinin sesquiterpenoids (VS3550, VS3560)

A farnesane numbering system is used for the furodysin and the rearranged furodysinin groups from *Dysidea* spp.



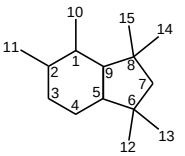
Furodysin skeleton
3-Ethyldecahydro-
1,1,6-trimethylnaphthalene



Furodysinsin skeleton
2-Ethyldecahydro-
1,1,6-trimethylnaphthalene

Botrydial sesquiterpenoids (VS3600)

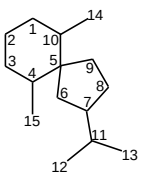
A non-biogenetic numbering system has been adopted for the botrydial group from *Botrytus* spp.



Botrydial skeleton
Octahydro-1,1,3,3,4,5-hexamethyl-1*H*-indene

Spirovetivane sesquiterpenoids (VS3700)

Spirovetivanes (also known as vetispiranes) are found in vetiver oil and also occur as phytoalexins in infected potatoes. The numbering system reflects their eudesmane derivation.

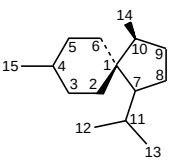


Spirovetivane
6,10-Dimethyl-2-(1-methylethyl)spiro[4.5]decane, 9CI

Marshall, J.A. *et al.* (1974) *Prog. Chem. Org. Nat. Prod.*, **31**, 283.

Acorane sesquiterpenoids (VS3750)

The acoranes and the enantiomeric alaskanes have a symmetrical six-membered ring. It has been suggested that C-2 should be chosen to be *syn*- to C-14.

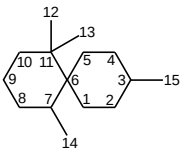


Acorane
1,8-Dimethyl-4-(1-methylethyl)spiro[4.5]decane, 9CI

Marshall, J.A. *et al.* (1974) *Prog. Chem. Org. Nat. Prod.*, **31**, 283.

Chamigrane sesquiterpenoids (VS3800)

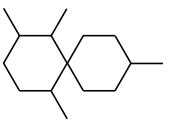
The chamigranes are a group of mainly marine natural products, mostly from *Laurencia* and *Aplysia* spp. The numbering system is based on farnesane. There are also some secochamigranes known (VS 3810).



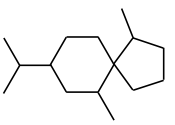
Chamigrane
1,1,5,9-Tetramethylspiro[5.5]undecane, 9CI

Miscellaneous spirosesquiterpenoids (VS3850)

Some miscellaneous spirosesquiterpenoid skeletons are collected in this section.



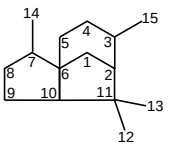
1,2,5,9-Tetramethylspiro-
[5.5]undecane



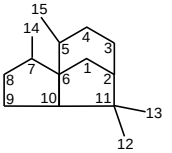
1,6-Dimethyl-8-(1-methyl-
ethyl)spiro[4.5]decane

Cedrane and isocedrane sesquiterpenoids (VS3900, VS3920)

The cedranes occur in wood oils and lac resins. Several numbering systems have been used. The farnesane system is used here. Isocedranes are rearranged cedranes; a related numbering system is in use.



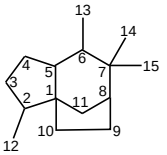
Cedrane
Octahydro-3,6,8,8-tetra-
methyl-1*H*-3*a*,7-methano-
azulene



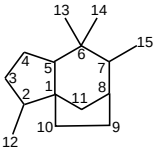
Isocedrane
Octahydro-3,4,8,8-tetra-
methyl-1*H*-3*a*,7-methano-
azulene

Zizaane and prezizaane sesquiterpenoids (VS3950, VS3960)

The zizaanes and prezizaanes are found in wood oils including vetiver oil.



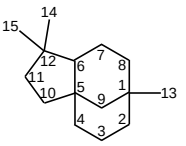
Zizaane
Octahydro-3,7,8,8-tetra-
methyl-1*H*-3*a*,6-methano-
azulene, 9CI



Prezizaane
Octahydro-3,7,7,8-tetra-
methyl-1*H*-3*a*,6-methano-
azulene

Clovane sesquiterpenoids (VS4000)

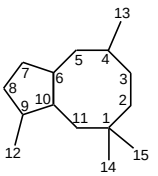
Only one clovane is natural and it is possibly an artifact from the acid catalysed rearrangement of a caryophyllene.



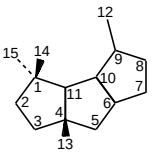
Clovane
Decahydro-1,1,7-trimethyl-3*a*,7-methano-3*aH*-cyclopentacyclooctene, 9CI

Precapnellane and capnellane sesquiterpenoids (VS4200, VS4250)

Precapnellanes and capnellanes are of marine origin. Capnellanes are 4,11-cycloprecapnellanes.



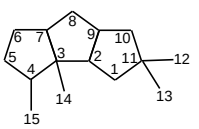
Precapnellane
Decahydro-1,5,8,8-tetra-methyl-1*H*-cyclopenta-cyclooctene



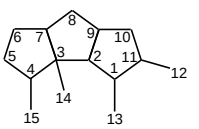
Capnellane
Decahydro-3,3,4,7*a*-tetra-methyl-1*H*-cyclopenta[*a*]-pentalene, 9CI

Hirsutane and rearranged hisutane sesquiterpenoids (VS4300, VS4310)

The hirsutanes are antibiotics from various fungi. Methyl migrated analogues are also found.



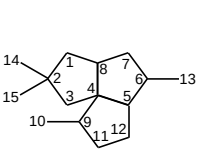
Hirsutane
Decahydro-2,2,3*b*,4-tetra-methyl-1*H*-cyclopenta[*a*]-pentalene



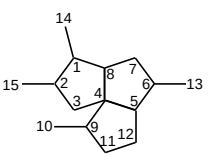
Rearranged hirsutane
Decahydro-2,3,3*b*,4-tetra-methyl-1*H*-cyclopenta[*a*]-pentalene

Pentalenane sesquiterpenoids (VS4400)

The pentalenanes are antibiotics from *Streptomyces* spp. Also included in this section are the rearranged and seco-pentalenanes.



Pentalenane
Decahydro-1,4,7,7-tetra-methylcyclopenta[*c*]-pentalene

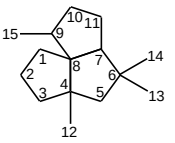


Rearranged pentalenane
Decahydro-1,4,6,7-tetra-methylcyclopenta[*c*]-pentalene

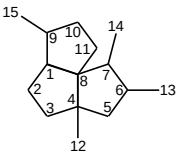
Cane, D.E. *et al.* (1992) *J. Org. Chem.*, **57**, 844.

Silphinane, silphiperfoliane and presilphiperfoliane sesquiterpenoids (VS4450, VS4460, VS4470)

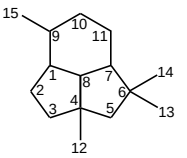
These three groups are related biogenetically. The absolute configurations have not been established unambiguously.



Silphinane
Decahydro-1,4,4,5*a*-tetra-methylcyclopenta[*c*]-pentalene



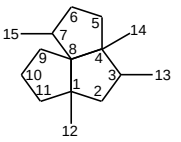
Silphiperfoliane
Decahydro-1,4,5,6*a*-tetra-methylcyclopenta[*c*]-pentalene



Presilphiperfoliane

Isocomane sesquiterpenoids (VS4500)

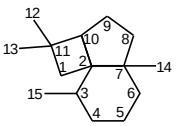
Several numbering systems have been used for this skeleton, including four from one author. The originally proposed system is used here.



Isocomane
Decahydro-1,3*a*,4,6-tetramethylcyclopenta[*c*]pentalene

Panasinsane sesquiterpenoids (VS4630)

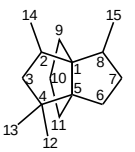
The panasinsanes from *Panax ginseng* are 2,7-cyclocaryophyllanes. The farnesane numbering system is used here.



Panasinane
Decahydro-2,2,4*a*,8-tetramethylcyclobut[*c*]indene, 9CI

Modhephane sesquiterpenoids (VS4700)

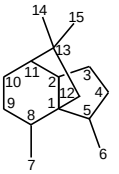
The modhephanes form a small group of bridged tricyclic sesquiterpenes from *Isocoma wrightii*. The numbering system originally proposed is used here.



Modhephane
Tetrahydro-1,1,3,4-tetramethyl-1*H*,4*H*-propanopentalene

Quadrane sesquiterpenoids (VS4750)

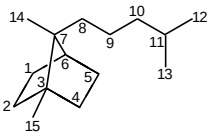
The quadranes are antibiotics from *Aspergillus terreus*.



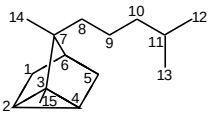
Quadrane

Campherenane, α-santalane and β-santalane sesquiterpenoids (VS4770, VS4780, VS4790)

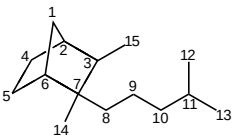
The campherenanes from *Cinnamomum camphora* and the α- and β-santalanes from sandalwood oil (*Santalum album*) are biogenetically related. The farnesane numbering system is used.



Campherenane
1,7-Dimethyl-7-(4-methyl-pentyl)bicyclo[2.2.1]-heptane, 9CI



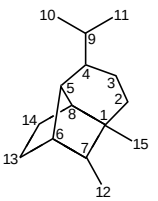
α-Santalane
2,3-Dimethyl-2-(4-methyl-pentyl)tricyclo[2.2.1.0^{2,6}]-heptane



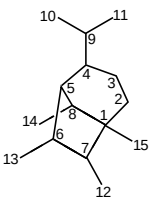
β-Santalane
2,3-Dimethyl-2-(4-methylpentyl)bicyclo[2.2.1]heptane

Sativane sesquiterpenoids (VS4800)

The sativanes and their 13,14-seco-derivatives, the helminthosporanes, are fungal metabolites from *Helminthosporum sativum*.



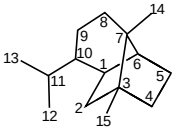
Sativane
Octahydro-4,8-dimethyl-7-(1-methylethyl)-1,4-methano-1*H*-indene



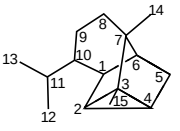
Helminthosporane
1,6,7,8-Tetramethyl-4-(1-methylethyl)bicyclo-[3.2.1]octane

Copacamphane and sinularane sesquiterpenoids (VS4820, VS4850)

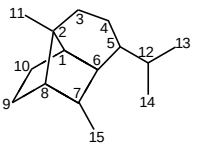
Copacamphanes and cyclocopacamphanes have regular farnesane skeletons and a farnesane numbering system has been used. Sinularanes and cyclocopacamphanes are rearrangement products.



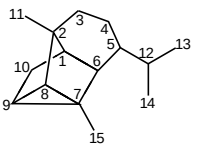
Copacamphane
Octahydro-4,8-dimethyl-7-(1-methylethyl)-1,4-methano-1*H*-indene, 9CI



Cyclocopacamphane
Octahydro-1,7*a*-dimethyl-5-(1-methylethyl)-1,2,4-metheno-1*H*-indene, 9CI



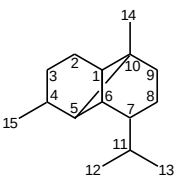
Sinularane
Octahydro-7*a*,8-dimethyl-5-(1-methylethyl)-1,4-methano-1*H*-indene, 9CI



Cyclosinularane
Octahydro-1,4-dimethyl-7-(1-methylethyl)-1,2,4-metheno-1*H*-indene, 9CI

Copaane sesquiterpenoids (VS4960)

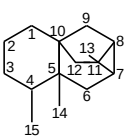
Copaanes can either be considered as 5,10-cyclocadinanes or 1,6-cycloeudesmanes.



Copaane
1,3-Dimethyl-8-(1-methylethyl)tricyclo[4.4.0.0^{2,7}]decane, 9CI

Ishwarane sesquiterpenoids (VS5000)

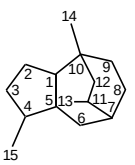
The ishwaranes form a small group of 10,12-cycloeremophilanes from *Aristolochia* spp; 7,11- and 8,11-seco derivatives are known and are included in this group.



Ishwarane
Decahydro-1,6,6*a*-trimethyl-1,2*a*-methano-2*aH*-cyclopropa[*b*]naphthalene, 9CI

Rotundane sesquiterpenoids (VS5020)

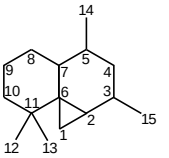
The rotundanes are 10,12-cycloguaianes.



Rotundane
Decahydro-1,4,6-trimethyl-4,7-ethanoazulene, 9CI

Thujopsane sesquiterpenoids (VS5040)

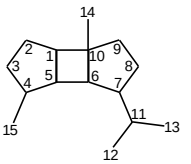
Thujopsanes are found in higher plants whilst *ent*-thujopsanes are found in liverworts. The biogenetic farnesane numbering system is used.



Thujopsane
Decahydro-2,4*a*,8,8-tetramethylcyclopropa[*d*]naphthalene, 9CI

Bourbonane sesquiterpenoids (VS5050)

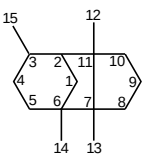
Bourbonanes are 6,10-cycloguaianes. A guaiane numbering system is used.



Bourbonane
Decahydro-3*a*,6-dimethyl-1-(1-methylethyl)cyclobuta[1,3:3,4]-dicyclopentene, 9CI

Gymnomitrane sesquiterpenoids (VS5070)

Gymnomitranes are biogenetically related to cuparanes. A farnesane numbering system is used here.



Gymnomitrane
Decahydro-3*a*,4,7,8*a*-tetramethyl-4,8-methanoazulene, 9CI

Miscellaneous sesquiterpenoids (VS5090–VS5320)

Sesquiterpenoid skeletons that do not fit easily into the above categories are collected here. They are divided into monocyclic, bicyclic, tricyclic and tetracyclic sesquiterpenoids.

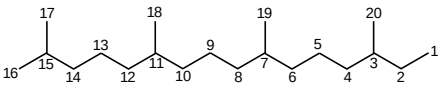
Diterpenoids (VS5350–VS7340)

The diterpenoids constitute a large group of compounds derived from geranylgeranyl pyrophosphate. They are found in higher plants, fungi, insects and marine organisms.

Hanson, J.R. (1971) *Prog. Chem. Org. Nat. Prod.*, **29**, 395.
Hanson, J.R. (1972) in *Chemistry of Terpenes and Terpenoids*, (ed. A.A. Newman) Academic Press, London, p. 155.
Hanson, J.R. (1998) *Nat. Prod. Rep.*, **15**, 93.
Hill, R.A. (1993) in *The Chemistry of Natural Products*, 2nd edn (ed. R.H. Thomson), Blackie, Glasgow, pp. 124.
West, C.A. (1981) in *Biosynthesis of Isoprenoid Compounds*, (eds J. W. Porter *et al.*) Wiley, New York, Vol. 1, p. 375.

Phytane diterpenoids (VS5350)

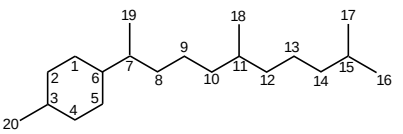
Phytanes are regular acyclic diterpenoids. The phytane numbering system is shown here.



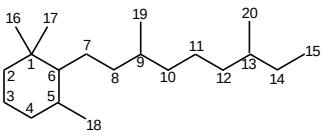
Phytane
2,6,10,14-Tetramethylhexadecane

Prenylbisabolane and 10,15-cyclophytane diterpenoids (VS5380, VS5390)

The prenylbisabolanes arise by cyclisation between carbons 1 and 6 of the phytane skeleton. They retain their phytane numbering system. The 10,15-cyclophytanes are important compounds including the retinal group. Since 10,15-cyclophytanes resemble carotenoids, a carotenoid-like numbering system has usually been adopted. It is possible to view 10,15-cyclophytanes as 9,10-secolabdanes and some are named and numbered as such in the literature.



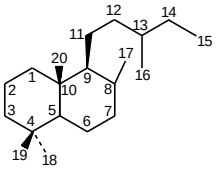
Prenylbisabolane
1-Methyl-4-(1,5,9-trimethyldecyl)cyclohexane



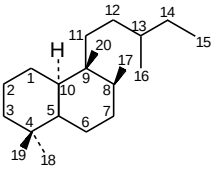
10,15-Cyclophytane
1,1,3-Trimethyl-2-(3,7-dimethylnonyl)cyclohexane

Labdane and halimane diterpenoids (VS5400–VS5470)

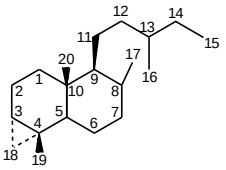
Labdanes form a large group and occur in both enantiomeric series. The halimanes are derived from labdanes by migration of the C-20 methyl group to C-9. Nor-, seco- and rearranged labdanes, including the gnaphalanes, are presented in separate sections.



Labdane
Decahydro-1,1,4*a*,6-tetra-methyl-5-(3-methyl-pentyl)naphthalene, 9CI



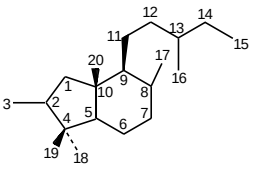
Halimane
Decahydro-1,1,5,6-tetra-methyl-5-(3-methyl-pentyl)naphthalene



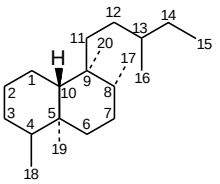
Gnaphalane

Colensane and clerodane diterpenoids (VS5480–VS5530)

Colensanes are a small group of 4(3 → 2) abeolabdanes. Clerodanes arise from labdanes by two methyl migrations. They are abundant and are found particularly in *Teucrium* spp. where they are highly oxygenated. In the past *ent*-clerodanes have been named as neoclerodanes and kolavanes but these names are not widely used.



Colensane
Octahydro-1,1,2,3*a*,5-pentamethyl-4-(3-methyl-pentyl)-1*H*-indene

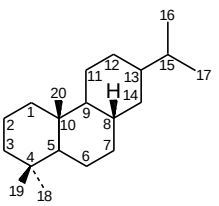


ent-Clerodane
Decahydro-1,2,4*a*,5-tetramethyl-1-(3-methyl-pentyl)naphthalene, 9CI

Merritt, A.T. *et al.* (1992) *Nat. Prod. Rep.*, **9**, 24.
Piozzi, F. *et al.* (1987) *Heterocycles*, **25**, 807.
Rodriguez-Hahn, L. *et al.* (1994), *Prg. Chem. Org. Nat. Prod.*, **63**, 107.
Rodriguez-Hahn, L. *et al.* (1995), *Recent Adv. Phytochem.*, **29**, 311.

Abietane diterpenoids (VS5550–VS5600)

Abietanes may arise from pimaranes by migration of the methyl group at C-13. This large group is divided into sections including furanoabietanes (VS5560), secoabietanes and secofriedoabietanes (VS5570), nor- and homoabietanes (VS5580), abeoabietanes (VS5590) and abietane dimers (VS5600).

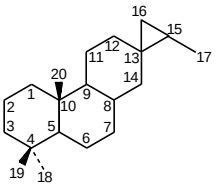


Abietane
Tetradecahydro-1,1,4*a*-trimethyl-7-(1-methylethyl)-phenanthrene, 9CI

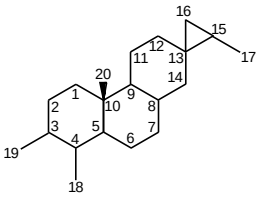
San Feliciano, A. (1993), *Planta Med.*, **59**, 485.

13,16-Cycloabietane and 17(15 → 16)-abeoabietane diterpenoids (VS5620, VS5630)

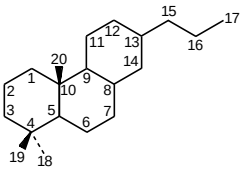
These groups include the coleons, lanugones and plectranthones from *Coleus* and *Plectranthus* spp. Included in each group are the corresponding 19(4 → 3)-abeo derivatives.



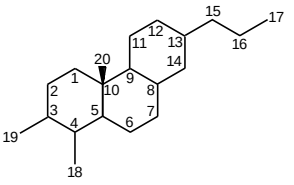
13,16-Cycloabietane



19(4→3)-Abeo-13,16-cyclo-abietane



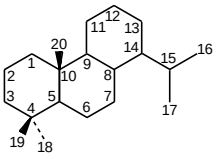
17(15→16)-Abeoabietane
Tetradecahydro-1,1,4*a*-trimethyl-7-propyl-phenanthrene



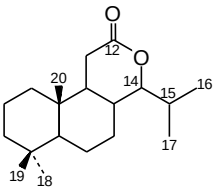
17(15→16),19(4→3)-Bisabeoabietane
Tetradecahydro-1,2,4*a*-trimethyl-7-propyl-phenanthrene

Totarane and nagilactone diterpenoids (VS5650, VS5660)

The totaranes may arise from abietanes by migration of the isopropyl group from C-13 to C-14. They normally have an aromatic ring C and are found in several species of higher plants. The nagilactone group are seconortotaranes found in *Podocarpus* spp.



Totarane
Tetradecahydro-1,1,4*a*-trimethyl-8-(1-methylethyl)-phenanthrene

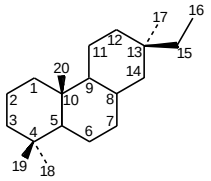


Nagilactone skeleton

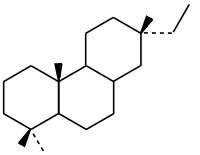
Bendall, J.G. *et al.* (1995) *Aust. J. Chem.*, **48**, 883.

Pimarane, rosane, erythroxylane, parguarane, devadarane and isopimarane diterpenoids (VS5700–VS5770)

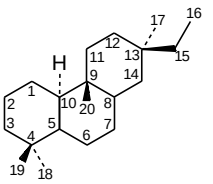
The pimaranes (VS5700) and isopimaranes (VS5750) (formerly called sandaracopimaranes) arise by further cyclisation of the labdane skeleton. Pimaranes have the ethyl group at C-13 *syn*- to the methyl group at C-10 whereas in the isopimaranes they are *anti*-. Both pimaranes and isopimaranes occur in both enantiomeric series. Rosanes (VS5710) arise by migration of the C-10 methyl group of pimaranes to C-9. 13-*epi*-Rosanes, e.g. **Rimuene** arise in a similar manner from isopimaranes. Erythoxylanes (VS5720), parguaranes (VS 5730) and devadaranes (VS5740) represent a further degree of rearrangement of pimaranes and isopimaranes.



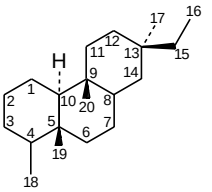
Pimarane
7-Ethyltetradecahydro-1,1,4*a*,7-tetramethyl-phenanthrene



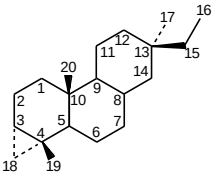
Isopimarane



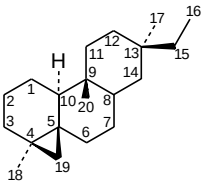
Rosane
7-Ethyltetradecahydro-1,1,4*b*,7-tetramethyl-phenanthrene



Erythroxylane
7-Ethyltetradecahydro-
1,4*b*,7,10*a*-tetramethyl-
phenanthrene



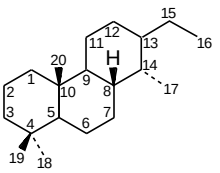
Parguarane
5-Ethyltetradecahydro-
1*a*,5,7*b*-trimethyl-1*H*-
cyclopropa[*a*]phenanthrene



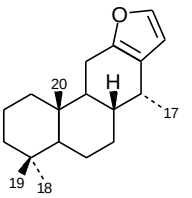
Devadarane
8-Ethyltetradecahydro-
3*a*,8,10*a*-trimethyl-
cyclopropa[*j*]phenanthrene

Cassane and vouacapane diterpenoids (VS5800)

The cassanes presumably arise by methyl migration in the pimarane skeleton from C-13 to C-14. The *Erythrophleum* alkaloids are simple derivatives of cassanes and are listed in this section as well as in the Alkaloid section. Furanocassanes are called vouacapanes.



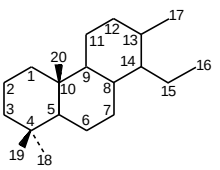
Cassane
7-Ethyltetradecahydro-
1,1,4*a*,8-tetramethyl-
phenanthrene



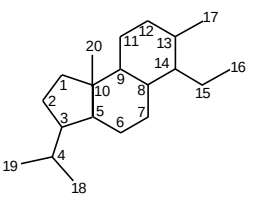
Vouacapane skeleton

Cleistanthane and isocleistanthane diterpenoids (VS5850)

The cleistanthane skeleton arises from pimaranes or isopimaranes by the migration of the ethyl group from C-13 to C-14. Isocleistanthane is a 2(4 → 3)-abeocleistanthane.



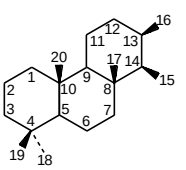
Cleistanthane
7-Ethyltetradecahydro-
1,4,4*a*,7-tetramethyl-
phenanthrene



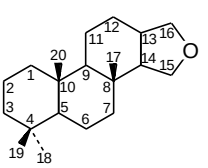
Isocleistanthane skeleton

Isocopalane and spongiane diterpenoids (VS5950)

Isocopalanes and spongianes are of marine origin and both have the same carbon skeleton. A spongiane or spongian is a 15,16-epoxyisocopalane.



Isocopalane
Tetradecahydro-1,1,4*a*,7,8,8*a*-
hexamethylphenanthrene



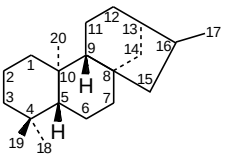
Spongiane skeleton

Podocarpane diterpenoids (VS5980)

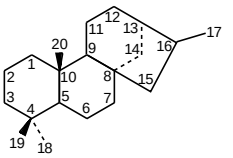
Miscellaneous podocarpane derivatives that cannot be easily classified are collected in this section.

Kaurane and phyllocladane diterpenoids (VS6000–VS6040)

The kauranes arise by further cyclisation of a pimarane cation intermediate followed by rearrangement. Most kauranes occur in the *ent*-series. The less common phyllocladanes have the methylene bridge on the opposite side from the methyl at C-10. Nor-, seco- and rearranged kauranes are placed in separate sections. The seco-kaurane group includes the *Rabdosia* constituents, e.g. **Enmein**. Care should be taken when using *Chemical Abstracts* for this skeleton as *ent*-kaurane is taken as the stereoparent and is named kaurane; 19-substituted *ent*-kauranes are named as having a (4α)-18-substituent and phyllocladanes are named as (5α,9α,10β)-kauranes. CA also uses Enmein as a basis for naming some of the secokauranes.



ent-Kaurane
≡ Kaurane, 9CI

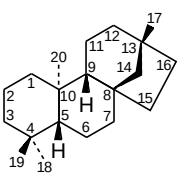


Phyllocladane
≡ 5α,9α,10β-Kaurane, 9CI

Alhazimi, H.M.G. *et al.* (1994), *J. Chem. Soc. Pak.*, **16**, 193.
Fujita, E. *et al.* (1984) *Prog. Chem. Org. Nat. Prod.*, **46**, 77.

Beyerane diterpenoids (VS6050)

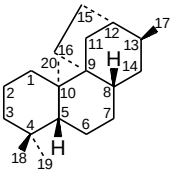
Beyeranes are formed by cyclisation of a pimarane cation intermediate without rearrangement. Most beyeranes belong to the *ent*-series. *ent*-Beyerane is also known as stachane. *Chemical Abstracts* names *ent*-beyeranes as 13-methyl-17-norkauranes.



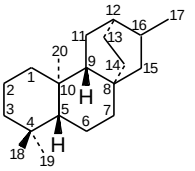
ent-Beyerane

Villanovane, atisane, trachylobane and helvifulvane diterpenoids (VS6080, VS6100, VS6150, VS6160)

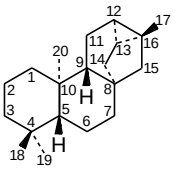
Rearrangement of the beyerane cation intermediate produces the villanovane and atisane skeletons. Rearrangement of the beyerane cation intermediate followed by cyclisation leads to the trachylobane and helvifulvane skeletons. These skeletons are mostly found in the *ent*-series. *ent*-Atisane and *ent*-trachylobane are named in CA as atisane and trachylobane (see remarks under kaurane, above).



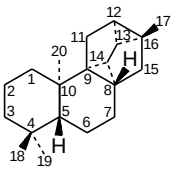
ent-Villanovane
Tetradecahydro-4,4,8,11*b*-tetramethyl-9,11*a*-methano-11*aH*-cyclohepta[*a*]-naphthalene, 9CI



ent-Atisane
≡ Atisane, 9CI



ent-Trachylobane
≡ Trachylobane, 9CI

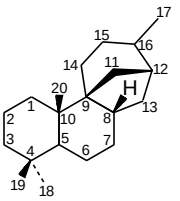


ent-Helvifulvane
Tetradecahydro-4,4,7*a*,9*b*-tetramethyl-8,8*a*-methano-9*aH*-cyclopropa[*b*]-phenanthrene, 9CI

Fraga, B.M. (1994), *Phytochem. Anal.*, **5**, 49 (Trachylobanes)

Aphidicolane diterpenoids (VS6180)

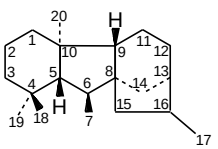
This small group includes the biologically active **Aphidicolin** and the **Stemodin** type from *Stemodia maritima*. The methano-bridge stereochemistry is opposite in Aphidicolin and Stemodin.



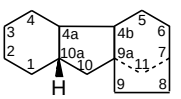
Aphidicolane
4,4,17-Trimethyl-9,15-cyclo-*C*,18-dinor-14,15-secoandrostane

Gibberellins (VS6200)

The gibberellins, important plant hormones, are based on the *ent*-gibberellane skeleton. They are produced by higher plants and the rice plant infecting fungus *Gibberella fujikuroi*. They have also been isolated from red and green algae. The biosynthesis of the gibberellins has been well studied and it is clear that they are derived from *ent*-kaurene. In CA they are named as derivatives of the stereoparent gibbane, which is a C_{15} skeleton and has a completely different numbering scheme. Many of the natural gibberellins are C_{19} norditerpenes.



ent-Gibberellane



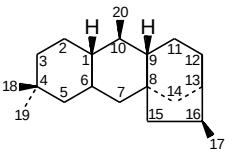
Gibbane, 9CI

Bakker, H.S. *et al.* (1974) *Tetrahedron*, **30**, 3631 (*biosynth*).
Beale, M.H. *et al.* (1990) *Methods Plant Biochem.*, **24**, 203 (*rev*).
Bearder, J.R. *et al.* (1976) *J. Chem. Soc., Chem. Commun.*, 834 (*biosynth*).

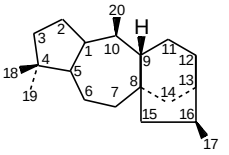
Bearder, J.R. *et al.* (1977) *Biochem. Soc. Trans.* **5**, 569 (*rev*).
Binks, R. *et al.* (1969) *Phytochemistry*, **8**, 271 (*ms*).
Ceccarelli, N. *et al.* (1983) *Phytochemistry* **22**, 2203 (*biosynth*).
Crosier, A. *et al.* (1970) *Can. J. Bot.*, **48**, 867 (*biochem*).
Evans, R. *et al.* (1970) *J. Chem. Soc. C*, 2601 (*biosynth*).
Evans, R. *et al.* (1975) *J. Chem. Soc., Perkin Trans. 1*, 1514 (*cmr*).
Hanson, J.R. (1965) *J. Chem. Soc.*, 5036 (*pmr*).
Hanson, J.R. (1990) *Nat. Prod. Rep.*, **7**, 41 (*rev*).
Kamiya, Y. *et al.* (1983) *Phytochemistry*, **22**, 681 (*biosynth*).
Lang, A. *et al.* (1970) *Annu. Rev. Plant Physiol.*, **21**, 537 (*rev*).
MacMillan, J. *et al.* (1970) *Aspects Terpenoid Chem. Biochem., Proc. Phytochem. Soc. Symp., 2nd*, 153 (*rev*).
Mander, L.N. *et al.* (1988) *Nat. Prod. Rep.*, **6**, 541 (*synth*).
Mander, L.N. *et al.* (1992) *Chem. Rev.*, **92**, 573 (*synth*).
Phinney, B.O. *et al.* (1990) *Recent Adv. Phytochem.*, **24**, 203 (*rev*).
Takahashi, N. *et al.* (1969) *Org. Mass Spectrom.*, **2**, 711 (*ms*).
West, C. (1969) *Biochem. J.*, **114**, 3P (*rev*).
Yamaguchi, I. *et al.* (1975) *J. Chem. Soc., Perkin Trans. 1*, 992 (*cmr*).
Yamane, H. *et al.* (1977) *Phytochemistry*, **16**, 831 (*biosynth*).

Leucothol and grayanotoxane diterpenoids (VS6220, VS6300)

The leucothol group and the grayanotoxins are also rearrangement products of the *ent*-kaurane skeleton.



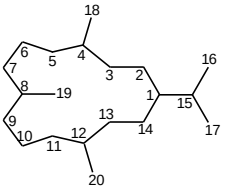
Leucothol skeleton



Grayanotoxane, 9CI

Cembrane diterpenoids (VS6400–VS6410)

The cembranes form a large group of diterpenoids found in higher plants (e.g. tobacco and conifers), insects and marine organisms. The cembrane nucleus has a plane of symmetry and is conventionally drawn with C-7 at the top as defined by the C-1, C-8 axis, C-7 being chosen as bearing a double bond or equivalent. The numbering system shown is generally accepted. Many polycyclic diterpenoids can be regarded as formally arising by cyclisation of the cembrane skeleton (or the related casbane skeleton – see below). As with germacranes, care is necessary in interpreting published configurations at centres involving reentrant angles.

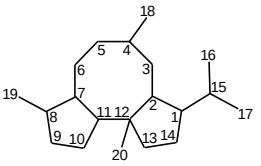


Cembrane
1,7,11-Trimethyl-4-(1-methylethyl)cyclotetradecane, 9CI

Tius, M.A. (1988) *Chem. Rev.*, **88**, 719.
Wahlberg, I. *et al.* (1992) *Prog. Chem. Org. Nat. Prod.*, **59**, 141; **60**, 1.
Weinheimer, A.J. *et al.* (1979) *Prog. Chem. Org. Nat. Prod.*, **36**, 285.

Rearranged cembrane diterpenoids (VS6420)

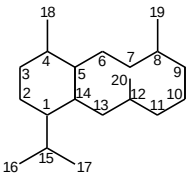
This group contains assorted, unusual macrocyclic diterpenoids including a basmane derivative which is formally a (2,12 : 7,11) cyclised cembrane.



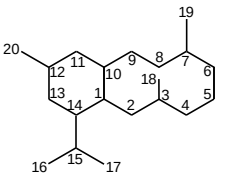
Basmane

Eunicellane and asbestinane diterpenoids (VS6440, VS6450)

These are marine natural products. The eunicellane (cladiellane) skeleton is formally a 5,14-cyclocembrane and the cembrane numbering system is preferred. The closely related asbestinane group has been assigned a different numbering system.



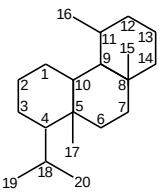
Eunicellane



Abestinane

Sphaerane diterpenoids (VS6460)

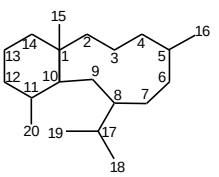
The bromosphaerol group of marine natural products contains an unusual carbon skeleton. The numbering system is as shown. Bicyclic (lacking the 1, 10-bond) and tetracyclic (with a 2,17-bond) derivatives are known. (N.B. Sphaeranes are not to be confused with sphaeroanes, see below).



Sphaerane
Tetradecahydro-5,8*a*,10*a*-trimethyl-1-(1-methylethyl)-phenanthrene

Briarane diterpenoids (VS6470)

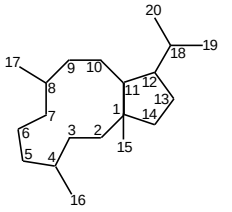
The briaranes are marine diterpenoids with the numbering system as shown. The carbon skeleton is formerly a 3,8-cyclocembrane.



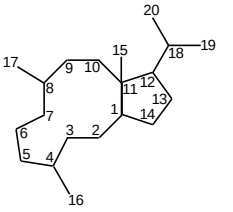
Briarane

Dolabellane and modified dolabellane diterpenoids (VS6500, VS6510)

Dolabellanes are found in marine organisms and in liverworts. Several numbering systems have been used in the literature. We have used the one shown. The modified dolabellane group includes the neodolabellanes in which a methyl has migrated from C-1 to C-11. A rare 3,9-cyclodolabellane is also included in this group.



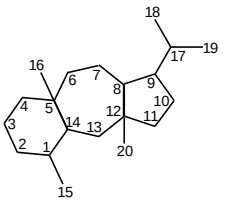
Dolabellane



Neodolabellane

Dolastane and modified dolastane diterpenoids (VS6540, VS6550)

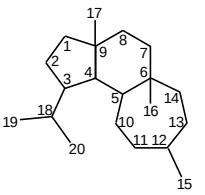
The name clavularane was originally used for this group of marine natural products but now dolastane appears to be widely accepted. Dolastane is a 3,8-cyclodolabellane but a different numbering system is used. The modified dolastane group contains 8,9-secodolastanes and a chromophycane, a new skeletal type related to dolastane by migration of the methyl C-20 to C-13.



Dolastane
Tetradecahydro-3*a*,5,8*a*-trimethyl-1-(1-methylethyl)-benz[*f*]azulene, 9CI

Cyathane diterpenoids (VS6560)

The cyathanes are fungal metabolites. The biosynthesis of this unusual skeleton has been studied. The accepted numbering system is as shown.

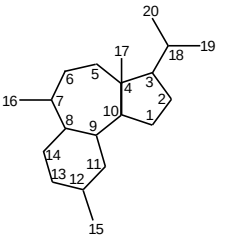


Cyathane
Tetradecahydro-3*a*,5*a*,8-trimethyl-1-(1-methylethyl)-cyclohept[*e*]indene

Turner W.B. *et al.* (1983) *Fungal Metabolites II*, Academic Press, London.

Sphaeroane diterpenoids (VS6570)

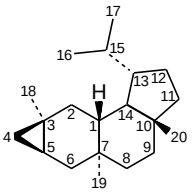
The sphaeroanes are marine algal products with a skeleton which is formally a 2,7-cyclodolabellane though the numbering system is different from that of dolabellanes.



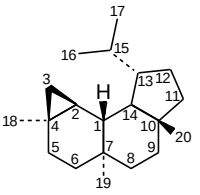
Sphaeroane
Dodecahydro-3*a*,6,9-trimethyl-3-(1-methylethyl)-
benz[*e*]azulene, 9CI

Verrucosane and modified verrucosane diterpenoids (VS6580, VS6590)

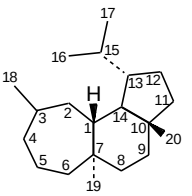
Verrucosanes and their modifications, including neoverrucosanes, homoverrucosanes (3,5-secoverrucanes) and homoneoverrucosanes (2,4-seconeoverrucosanes) are liverwort natural products. The tetracyclic verrucosane skeleton is formally related to dolabellane by 4,10- and 6,8-bond formation. A different numbering system from that of dolabellane is used. The isomeric neoverrucosane has the cyclopropane bridging C-2 and C-3.



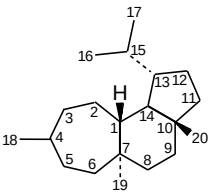
Verrucosane
Tetradecahydro-3*a*,5*a*,7*a*-trimethyl-
1-(1-methylethyl)cyclopenta-
[*a*]cyclopropa[*g*]naphthalene, 9CI



Neoverrucosane



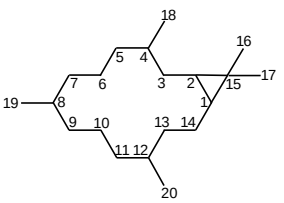
Homoverrucosane



Homoneoverrucosane

Casbane diterpenoids (VS6600)

The casbane skeleton is closely related to cembrane and it seems logical to use the cembrane numbering system although many others have been used in the literature. Casbane, like cembrane, is the formal parent of several important groups of diterpenoids especially from the Euphorbiaceae and Thymelaeceae. Berdimerane has a rearranged casbane skeleton.

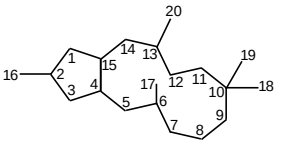


Casbane
3,7,11,15,15-Pentamethylbicyclo[12.1.0]pentadecane, 9CI

Evans, F.J. *et al.* (1983) *Prog. Chem. Org. Nat. Prod.*, **44**, 1.

Jatrophane and 9,13-cyclojatrophane diterpenoids (VS6610, VS6620)

Jatrophane is the parent skeleton of a group of macrocyclic diterpenoids from *Euphorbia* species. Formally it can be derived from casbane by 6,10-cyclisation and opening of the cyclopropane. The numbering system shown is used fairly consistently in all the related diterpenoids of this series. Two examples of the 9,13-cyclojatrophane skeleton have been isolated.

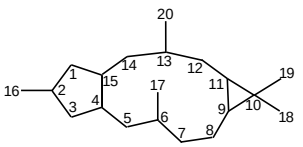


Jatrophane

Evans, F.J. *et al.* (1983) *Prog. Chem. Org. Nat. Prod.*, **44**, 1.

Lathyrane diterpenoids (VS6650)

Derivatives of the lathyrane skeleton occur widely in the Euphorbiaceae as mixed esters. Lathyrane is formally a 6,10-cyclocasbane. There is some confusion in the literature over the configuration of the methyl group attached to C-13. Reentrant angles should be avoided if possible.

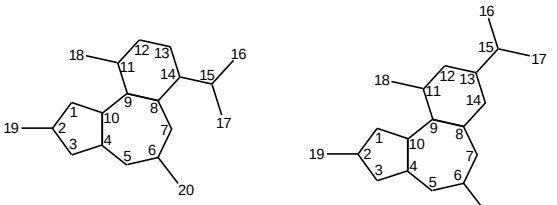


Lathyrane
Tetradecahydro-1,1,3,6,9-pentamethyl-1*H*-cyclopenta[*a*]-cyclopropa[*f*]cycloundecene

Evans, F.J. *et al.* (1983) *Prog. Chem. Org. Nat. Prod.*, **44**, 1.

Rhamnofolane and daphnane diterpenoids (VS6660, VS6680)

The rhamnofolane and daphnane skeletons are closely related, being formally derived from casbane by two cyclisations (6,10 and 5,14) followed by cleavage of the 1,15 (daphnane) or 2,15 (rhamnofolane) cyclopropane bonds. Note that in *Chemical Abstracts*, Daphnane is an alkaloidal stereoparent. Terpenoid daphnane derivatives are named as derivatives of Daphnetoxin.



Rhamnofolane

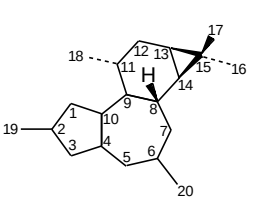
Daphnane

Evans, F.J. *et al.* (1983) *Prog. Chem. Org. Nat. Prod.*, **44**, 1.

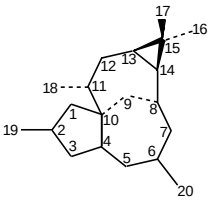
Tigliane and ingenane diterpenoids (VS6700, VS6710)

The tigliane nucleus is formerly derived from casbane by 6,10- and 5,14-cyclisations. This is the carbon framework of phorbol whose derivatives occur

widely in the Euphorbiaceae and are renowned for their tumour promoting and irritant activity. The ingenane skeleton is derived by rearrangement of tigliane. Ingenane esters also have irritant properties.



Tigliane
Tetradecahydro-1,1,3,6,8-pentamethyl-1*H*-cyclopropa[3.4]benz[1,2-*e*]azulene, 9CI

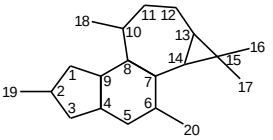


Ingenane
Dodecahydro-1,1,4,7,9-pentamethyl-1*H*-2,8*a*-methanocyclopenta[*e*]cyclopropa[*e*]cyclodecene, 9CI

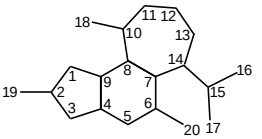
Evans, F.J. *et al.* (1983) *Prog. Chem. Org. Nat. Prod.*, **44**, 1.

Jatropholane and secojatropholane diterpenoids (VS6720)

Jatropholanes arise by 5,13- and 6,10-cyclisation of casbane. Subsequent cleavage of the 2,15-cyclopropane bond affords crotofolane. Only a few examples of these skeletons have been reported.



Jatropholane
Hexadecahydro-1,1,2,5,7-pentamethyl-6*H*-cyclohept-[1,2-*e*]indene

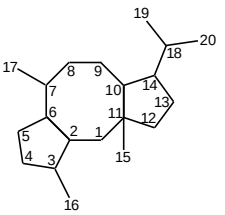


Crotofolane
Tetradecahydro-2,5,10-trimethyl-6-(1-methylethyl)cyclohept[*e*]-indene, 9CI

Evans, F.J. *et al.* (1983) *Prog. Chem. Org. Nat. Prod.*, **44**, 1.

Fusicoccane diterpenoids (VS6750)

Fusicoccanes occur in fungi and liverworts. Biosynthetic evidence favours the involvement of a dolabellane-like precursor. The accepted numbering system differs from that of dolabellane.

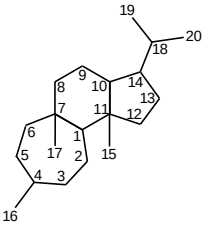


Fusicoccane
Tetradecahydro-1,4,9*a*-trimethyl-7-(1-methylethyl)dicyclopenta[*a,d*]-cyclooctene, 9CI

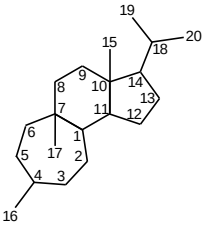
Krasnopolskaya, L.M. (1994), *J. Plant Growth Regulation*, **13**, 39.
Turner, W.B. *et al.* (1983) *Fungal Metabolites II*, Academic Press, London.

Valparane and mulinane diterpenoids (VS6770, VS6780)

The valparanes and mulinanes are related to the fusicoccanes and are numbered accordingly.



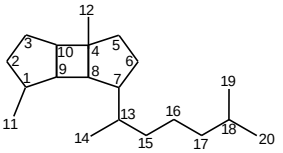
Valparane
Tetradecahydro-5*a*,8,10*b*-trimethyl-3-(1-methylethyl)cyclohept[*e*]indene, 9CI



Mulinane

Spatane diterpenoids (VS6800, VS6810)

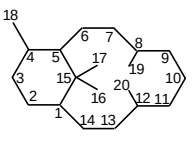
The spatane skeleton is formally derived from a prenylgermacrane by 1,5- and 6,10-cyclisation. The numbering system unfortunately does not reflect this derivation. Spatanes and the related 4,10-secospatanes are marine natural products.



Spatane
Decahydro-3*a*,6-dimethyl-1-(1,5-dimethylhexyl)cyclobuta[1,2:3,4]dicyclopentene, 9CI

Verticillane diterpenoids (VS6880)

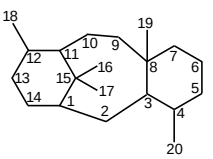
The verticillane group is formally derivable from cembrane by an 11,15-cyclisation. A non-cembrane numbering system is used. Only a few members of this group have been reported.



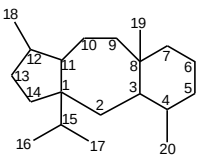
Verticillane
4,8,12,15,15-Pentamethylbicyclo[9.3.1]pentadecane, 9CI

Taxane and 11(15→1)-Abeotaxane diterpenoids (VS6900, VS6950)

The taxanes form an important group of biologically active diterpenoids and alkaloids from *Taxus* spp. The skeleton is related to verticillane by a further cyclisation. The accepted numbering system is as shown. A large number of 11(15→1)-Abeotaxanes have been isolated recently.



Taxane
Tetradecahydro-4,9,12*a*,13,13-pentamethyl-6,10-methanobenzocyclodecene, 9CI



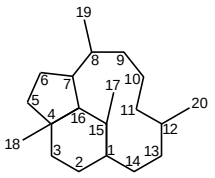
11(15→1)-Abeotaxane

Appendino, G. (1995) *Nat. Prod. Rep.*, **12**, 349.
Das, B. (1996) *Indian J. Chem., Sect. B.*, **35**, 883.
Das, B. (1995) *Planta Med.*, **61**, 393.

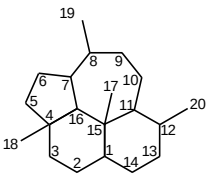
Kingston, D.G.I *et al.* (1993) *Prog. Chem. Org. Nat. Prod.*, **61**, 1.
Swindell, C.S. (1993) in *Studies in Natural Product Chemistry*, Vol 12, (ed. Atta-ur-Rahman), Elsevier, p. 170.

Trinervitane and kempane diterpenoids (VS7000, VS7010)

The trinervitanes, 7,16-secotrinervitanes, 17-methyltrinervitanes and kempanes are constituents of the defence secretions of termites. Kempanes are 11,15-cyclotrinervitanes. The numbering system of these unusual diterpenoids is as shown.



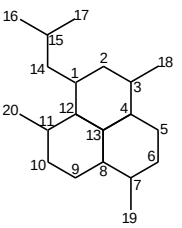
Trinervitane
Hexadecahydro-1,4,8,12-tetramethyl-1,11-ethanocyclopenta-cycloundecane, 9CI



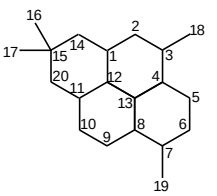
Kempane
Hexadecahydro-2*a*,7,10,10*c*-tetramethyl-napth-[2,1,8-*cde*]-azulene, 9CI

Amphilectane, cycloamphilectane, adociane and neoamphilectane diterpenoids (VS7020, VS7030, VS7040)

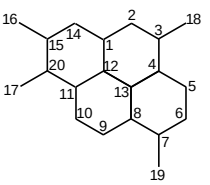
The amphilectanes (including the pseudopterosins), the cycloamphilectanes and adocianes (also called is isocycloamphilectanes) and neoamphilectanes are marine products. They are found with serrulatane derivatives from which amphilectanes are presumably derived by cyclisation. Cycloamphilectanes represent a further cyclisation and adocianes have undergone a methyl migration. Neoamphilectanes are 2(1 → 12) abeoamphilectanes



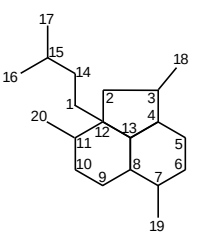
Amphilectane
Dodecahydro-1,4,7-trimethyl-3-(2-methylpropyl)-1*H*-phenalene, 9CI



Cycloamphilectane
Hexadecahydro-1,4,7,7-tetramethylpyrene



Adociane
Hexadecahydro-1,2,5,8-tetramethylpyrene, 9CI

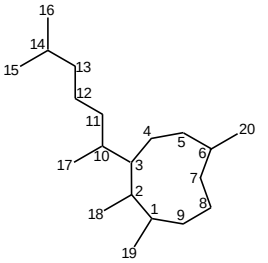


Neoamphilectane

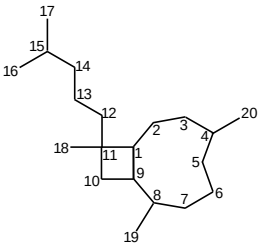
König, G.M. (1996) *J. Org. Chem*, **61**, 3259.

Xenicane and xeniaphyllane diterpenoids (VS7100, VS7110, VS7150)

Xenicanes and xeniaphyllanes are marine natural products. Various nor-, seco- and cyclo-xenicanes are listed in a separate section (VS7110). Xeniaphyllanes are the diterpenoid equivalent of the caryophyllane skeleton. Xenicanes are cleaved xeniaphyllanes.



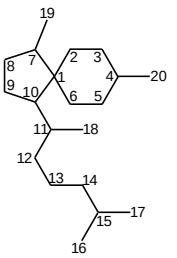
Xenicane
1,2,6-Trimethyl-3-(1,5-dimethylhexyl)-cyclononane



Xeniaphyllane
2,6,10-Trimethyl-10-(4-methylpentyl)-bicyclo[7.2.0]undecane

Viscidane diterpenoids (VS7160)

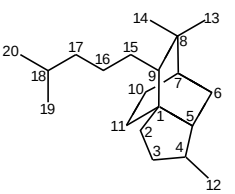
Viscidanes, from *Eremophila* spp., are the diterpenoid equivalents of the acoranes.



Viscidane
1,8-Dimethyl-4-(1,5-dimethylhexyl)spiro[4.5]decane

Eremane diterpenoids (VS7180)

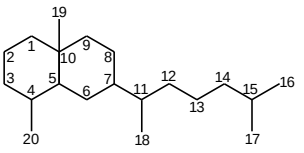
Eremanes have an unusual carbon skeleton. They are isolated from *Eremophila* spp.



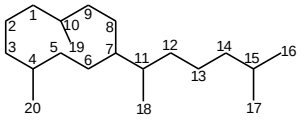
Eremane
Decahydro-1,5,5-trimethyl-4-(4-methylpentyl)-3*a*,6-ethano-3*aH*-indene

Prenyleudesmane, prenylgermacrane and prenylbicyclogermacrane diterpenoids (VS7190, VS7200, VS7210)

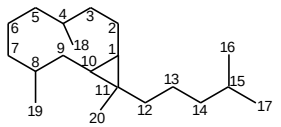
These three groups of ‘extended’ sesquiterpenoid skeletons are largely of marine origin.



Prenyleudesmane
Decahydro-7-(1,5-methylhexyl)-1,4*a*-dimethylnaphthalene



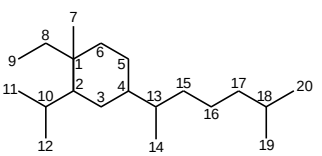
Prenylgermacrane
4-(1,5-dimethylhexyl)-1,7-dimethylcyclodecane



Prenybicyclogermacrane
3,7,11-Trimethyl-11-(4-methylpentyl)bicyclo-
[8.1.0]undecane

Lobane diterpenoids (VS7220)

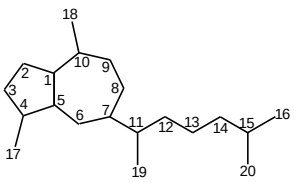
Lobanes are of marine origin and are ‘extended’ elemanes. A most unusual non-standard numbering system is used in the literature.



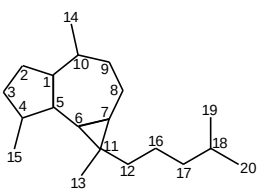
Lobane
4-(1,5-dimethylhexyl)-1-ethyl-1-methyl-2-
(1-methylethyl)cyclohexane

Pachydictyane and cneorubin diterpenoids (VS7230, VS7240)

These two groups are also ‘extended’ sesquiterpenoids. The pachydictyanes are prenylguaianes from marine organisms and the cneorubin group are prenylaromadendranes found in leaves of *Cneorum tricoccon*.



Pachydictyane
Decahydro-7-(1,5-dimethyl-
hexyl)-1,4-dimethyl-
azulene

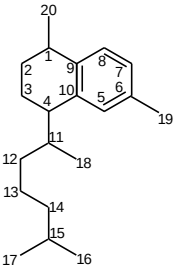


Cneorubin skeleton
Decahydro-1,4,7-trimethyl-1-
(4-methylpentyl)-1*H*-
cycloprop[*e*]azulene

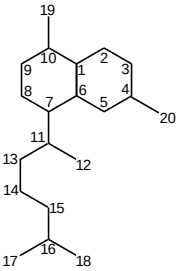
Hardt, I.H. (1996), *Phyhochemistry*, **43**, 71.

Serrulatane and biflorane diterpenoids (VS7250)

The biflorane skeleton is found in marine organisms, insects and *Eremophila* spp. The skeleton is an ‘extended’ cadinane. The serrulatane name is given to the aromatic analogue. Unfortunately different numbering systems have been given to serrulatanes and bifloranes.



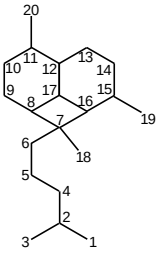
Serrulatane



Biflorane
Decahydro-4-(1,5-dimethyl-hexyl)-1,6-dimethyl-naphthalene

Decipiane diterpenoids (VS7260)

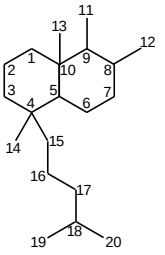
Decipianes from *Eremophila* spp. are cyclised bifloranes. Yet another numbering system is used for this skeleton.



Decipiane
Decahydro-1,2,5-trimethyl-1-(4-methylpentyl)-1*H*-cyclobuta[*de*]naphthalene, 9CI

Sacculatane diterpenoids (VS7270)

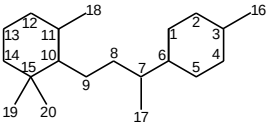
Sacculatanes are ‘extended’ drimanes and are found in liverworts.



Sacculatane
Decahydro-1,4*a*,5,6-tetramethyl-1-(4-methylpentyl)-naphthalene

Obtusane diterpenoids (VS7280)

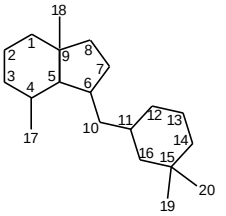
The obtusanes, of marine origin, are bicyclic phytanes. The numbering system is almost the same as for phytane. (Note that the terpenoid **Obtusane** itself is a chamigrane).



Obtusane
1,1,3-Trimethyl-2-[3-(4-methylcyclohexyl)butyl-
cyclohexane

Irieol diterpenoids (VS7290)

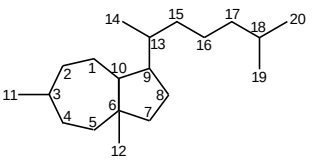
The irieol group, also of marine origin, represents an unusual diterpenoid skeleton.



Irieol skeleton
Octahydro-1-[(3,3-dimethylcyclohexyl)methyl]-
3*a*,7-dimethyl-1*H*-indene

Sphenolobane diterpenoids (VS7300)

The sphenolobane skeleton is an ‘extended’ daucane skeleton.



Sphenolobane
Decahydro-1-(1,5-dimethylhexyl)-3*a*,6-dimethylazulene

Miscellaneous diterpenoids (VS7310–VS7340)

Diterpenoids that do not easily fit into the other categories are collected here. Mono-, bi-, tri- and tetracyclic diterpenoids are listed separately in the Type of Compound Index.

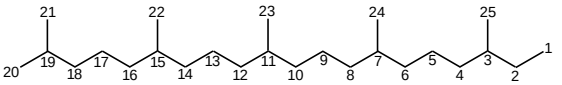
Sesterterpenoids (VS7400–VS7580)

Sesterterpenoids are a small group of natural products that arise from five isoprene units. Although sesterterpenoids strictly have 25 carbons, there are many nor- and alkylated members. Also included here are the C_{21} acyclic terpenoids although their biosynthetic relationship with the sesterterpenoids has not been established with certainty. Sesterterpenoids are found in fungi, higher plants, insects and marine organisms.

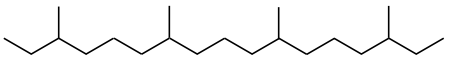
Cordell, G.A. (1974) *Phytochemistry*, **13**, 2343.
Crews, P. *et al.* (1985) *Prog. Chem. Org. Nat. Prod.*, **48**, 203.
González, A.G. *et al.* (1983) *J. Nat. Prod.*, **46**, 256.
Hanson, J.R. *et al.* (1996) *Nat. Prod. Rep.*, **13**, 529.
Hill, R.A. (1993) in *The Chemistry of Natural Products*, 2nd edn (ed. R.H. Thomson), Blackie, Glasgow, pp. 129.

Acyclic and noracyclic sesterterpenoids (VS7400)

The acyclic sesterterpenoids arise by a head to tail fusion of isoprene units. The accepted numbering system is used here. The noracyclic sesterterpenoids (VS7410) are numbered in a similar way; however, a problem arises with the symmetry of the C_{21} compounds as they may be numbered from either end. The acyclic sesterterpenoids frequently contain furanoid rings.



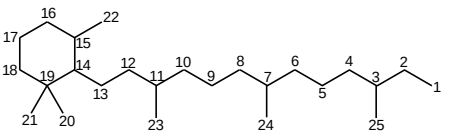
Acyclic sesterterpenoid skeleton
2,6,10,14,18-Pentamethyleicosane



C_{21} sesterterpenoid skeleton
3,7,11,15-Tetramethylheptadecane

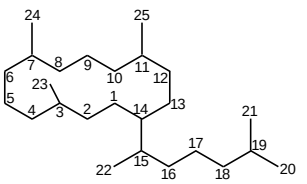
Cyclohexane sesterterpenoids (VS7420)

Most of the cyclohexane sesterterpenoids arise by cyclisation of the acyclic skeleton between carbons 14 and 19.



Cericerane sesterterpenoids (VS7440)

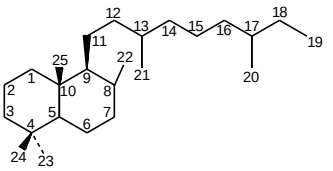
Cericeranes arise by cyclisation between carbons 1 and 14 of the acyclic skeleton, retaining the numbering system. The symmetry of the cyclotetradecane ring leads to some ambiguity of numbering (cf. cembranes).

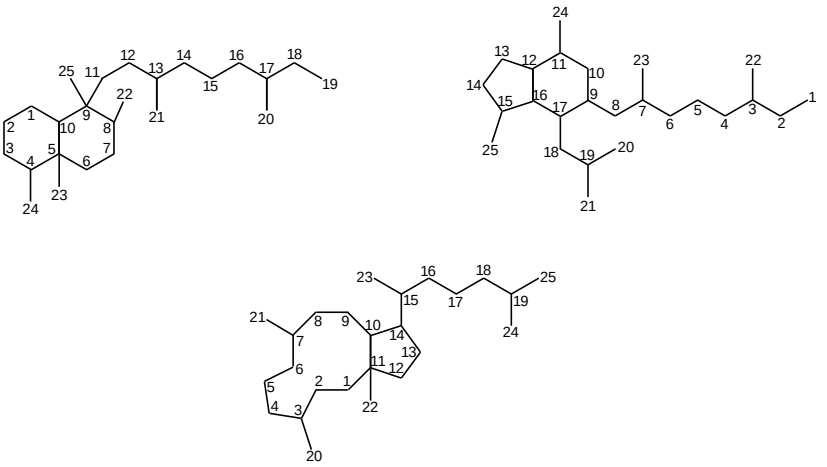


Cericerane
1-(1,5-Dimethylhexyl)-4,8,12-trimethylcyclolotetradecane

Bicyclic sesterterpenoids (VS7460)

Various bicyclic sesterterpenoids are known. Some are prenylated analogues of diterpene skeletons and the numbering systems are related to the corresponding diterpenoid systems. Others have biogenetic numbering systems.

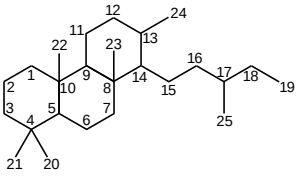




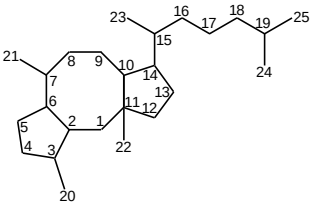
Some bicyclic sesterterpenoid skeletons

Cheilanthane and ophiobolane sesterterpenoids (VS7500, VS7520)

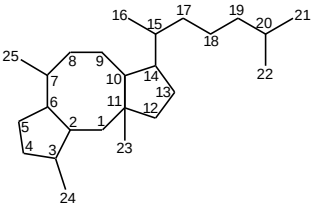
The accepted numbering systems for the cheilanthanes and ophiobolanes are shown here. *Chemical Abstracts* uses ophiobolane as a stereoparent; however it uses a different numbering system for the non-ring carbons.



Cheilanthane
4,4,8-Trimethyl-*D*(15),24-dinor-13,17-secocholane, 9CI



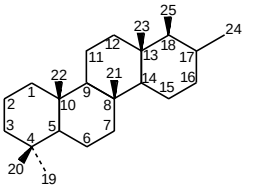
Ophiobolane



Ophiobolane, CA numbering

Scalarane sesterterpenoids (VS7540)

The scalarane numbering system is shown here. Carbons 12, 24 and 25 are generally oxygenated in this skeleton. Several methyl and dimethyl scalaranes are found in marine organisms. The additional methyl groups attached to C-24 and C-20 are numbered 26 and 27 respectively.

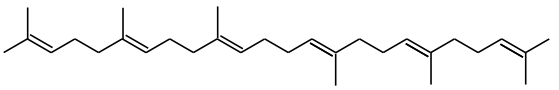


Scalarane
4,4,8,17,17*a*-Pentamethyl-*D*-homoandrostane, 9CI

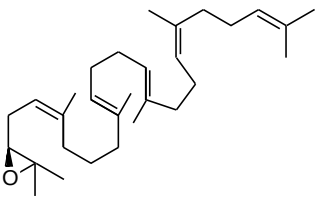
Bowden, B.F. (1992), *J. Nat. Prod.*, **55**, 1234.

Triterpenoids (VS7600–VS9450)

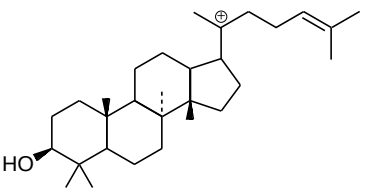
The triterpenoids constitute a large diverse group of natural products derived from squalene or, in the case of 3*β*-hydroxytriterpenoids, the 3*S*-isomer of squalene 2,3-epoxide. The conformation that *all-trans*-squalene 2,3-epoxide adopts when the initial cyclisation takes place determines the stereochemistry of the ring junctions in the triterpenoid produced. Thus cyclisation of the chair-boat-chair-boat conformation gives the protostane cation and cyclisation of the chair-chair-chair-boat conformation leads to the dammarane cation. The initially formed cation intermediate may undergo a series of 1,2-hydride and methyl migrations, commonly called backbone rearrangements, to give a variety of skeletal types.



Squalene

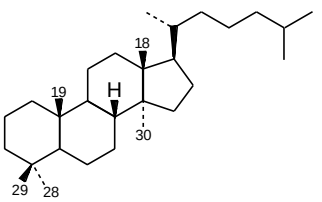


Squalene 2,3-epoxide

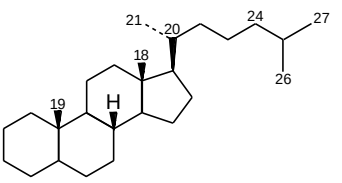


Protostane cation intermediate

Backbone rearrangement of the protostane cation gives the lanostane skeleton; **Lanosterol** is the biogenetic precursor of the steroids in animals. The methyl groups at carbons 4 and 14 are removed during steroid biosynthesis. The steroid numbering system is adopted for lanostane and related tetracyclic triterpenoids. The three methyl groups that were removed during the biosynthesis of steroids are currently numbered 28, 29 and 30 as shown. However, older literature uses the numbers 31, 30 and 32, respectively. This was based on the assignment of carbon numbers 28 and 29 to the stigmastane ethyl group, even though most lanostanes do not have such an ethyl group. The numbering in DNP follows the currently accepted convention. (See also Steroid section following).



Lanostane numbering



Steroid numbering

Abe, I. *et al.* (1993) *Chem. Rev.*, **93**, 2189.
Connolly, J.D. *et al.* (1972) in *Chemistry of Terpenes and Terpenoids*, (ed. A.A. Newman) Academic Press, London, p. 207.

Connolly, J.D. *et al.* (1991) *Methods Plant Biochem.*, **7**, 331.
Connolly, J.D. *et al.* (1997) *Nat. Prod. Rep.*, **14**, 661.
Goodwin, T.W. (1981) in *Biosynthesis of Isoprenoid Compounds*, (eds J.W. Porter *et al.*) Wiley, New York, Vol. 1, p. 443.
Hill, R.A. (1993) in *The Chemistry of Natural Products*, 2nd edn (ed. R.H. Thomson), Blackie, Glasgow, pp. 131.
Mahato, S.B. *et al.* (1997) *Phytochemistry*, **44**, 1185.
Spencer, T.A. (1994) *Acc. Chem. Res.*, **27**, 83.

The main tetracyclic triterpenoid skeletons have the steroid numbering for the skeleton including the side chain and only the methyl groups will be numbered in the structures that follow. As a general rule the methyls which migrate during terpenoid biosynthesis retain their numbering.

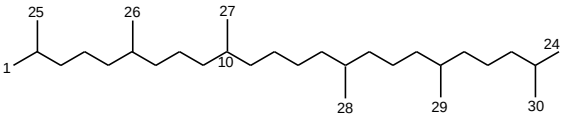
CA names most tetracyclic triterpenoids as derivatives of the steroid stereoparents. This has the disadvantage that some are assigned different names from those commonly used. The CA names for some common skeletons are listed below.

	Chemical Abstracts name
Protostane	Dammarane, (8α, 9β, 13α, 14β)-
Fusidane	29-Nordammarane, (4α, 8α, 13α, 14β)-
Cycloartane	9,19-Cyclolanostane
Cucurbitane	19-Norlanostane, 9-methyl-, (9β, 10α)-
Euphane	Lanostane, (13α, 14β, 17α)-
Tirucallane	Lanostane, (13α, 14β, 17α, 20*S*)-
Apotirucallane	Cholestane, 4,4,8-trimethyl-, (13α, 17α, 20*S*)

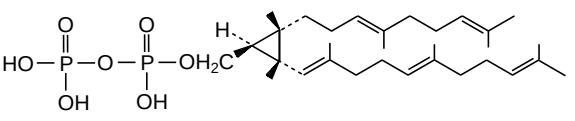
The CA nomenclature of some other triterpenoids is idiosyncratic, e.g. Malabaricanes.

Linear triterpenoids (VS7600)

This group contains simple derivatives of squalene. The preferred numbering system is shown and is used for the related polyether derivatives found in marine algae, e.g. *Laurencia* spp. Also included are C_{30} polyprenols, and some homo- and nor-squalenes. Squalene is formed biosynthetically from farnesyl pyrophosphate *via* presqualene pyrophosphate.



Squalane
2,6,10,15,19,23-Hexamethyltetracosane, 9CI

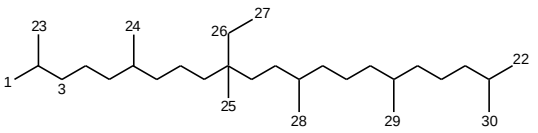


Presqualene pyrophosphate

Julia, M.Y. (1991) *Chem. Soc. Rev.*, **20**, 129.
Poulter, C.D. *et al.* (1981) in *Biosynthesis of Isoprenoid Compounds*, (eds J.W. Porter *et al.*) Vol. 1, Wiley, New York, p. 413.

Botryococcene triterpenoids (VS7620)

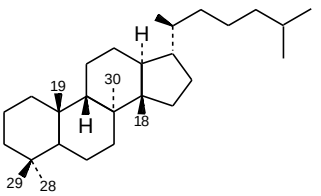
The alga *Botryococcus braunii* produces a series of branched alkylated isoprenoid hydrocarbons based on botryococcane. The names of individual compounds indicate the number of carbons, e.g. ***C_{34}*-Botryococcene**. Of the several numbering systems that have been used, the one below is preferred. Cyclised botryococcenes are also listed in this section.



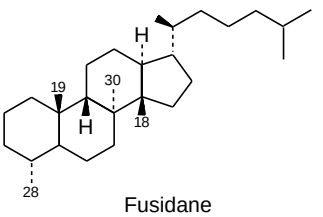
C_{30}-Botryococcane
10-Ethyl-2,6.10,13,17,21-hexamethyldocosane

Prostostane and fusidane triterpenoids (VS7700)

The protostanes, e.g. Protosterol, form a small group that arise from cyclisation of squalene or its 2,3-epoxide without backbone rearrangement. The fusidanes form a small but important group of antibiotics, e.g. **Fusidic acid**, which lack one of the methyl groups at carbon 4. The numbering and stereochemistry of these skeletons is indicated below.



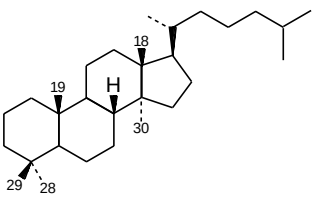
Protostane



Fusidane

Lanostane triterpenoids (VS7750)

Backbone rearrangement of the protostane cation leads to the lanostanes, a large group. **Lanosterol** is a key intermediate in steroid biosynthesis. They are uncommon in plants but some fungi, e.g. *Ganoderma lucidum*, are a prolific source. Some rearranged lanostanes are also included in this section.

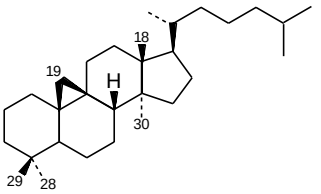


Lanostane

The nomenclature of the numerous *G. lucidum* products is highly confused owing to the application of identical trivial names to compounds of different structure by several different groups working simultaneously (for full details, see individual entries). It is recommended that systematic nomenclature be used.

Cycloartane triterpenoids (VS7800)

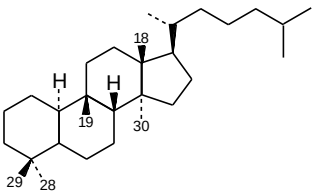
Backbone rearrangement of the protostane cation including cyclisation to form a 9,19 bond produces the cycloartane skeleton. Plants use **Cycloartenol** and not Lanosterol for the biosynthesis of phytosterols. Cycloartanes are often named as 9,19-Cyclolanostanes in the literature.



Cycloartane

Cucurbitane triterpenoids (VS7900)

More extensive backbone rearrangement of the protostane cation affords the cucurbitane skeleton. The cucurbitacins, e.g. **Cucurbitacin A**, are found in the Cucurbitaceae and are of interest because of their biological activity. Many occur as glycosides.

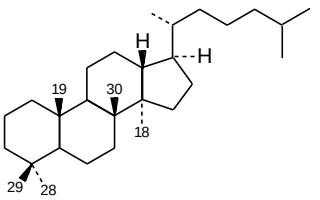


Cucurbitane

Lavie, D. *et al.* (1971) *Prog. Chem. Org. Nat. Prod.*, **29**, 307.
Miro, M (1995), *Phytother. Res.*, **9**, 159.

Dammarane triterpenoids (VS7950)

Collapse of the dammarane cation without backbone rearrangement leads to the dammarane skeleton (a stereoisomer of protostane). Dammaranes often occur as glycosides and are commonly found among the much-studied saponins of ginseng.

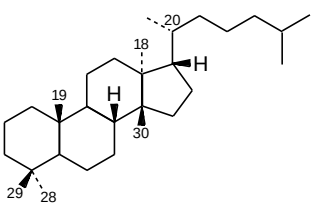


Dammarane

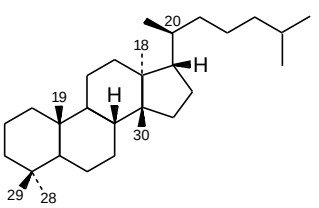
Tanaka, O. *et al.* (1984) *Prog. Chem. Org. Nat. Prod.*, **46**, 1.

Tirucallane/euphane triterpenoids (VS8000)

Backbone rearrangement of the dammarane cation (analogous to the protostane-lanostane rearrangement) yields the euphane skeleton and its 20-epimer, the tirucallane skeleton. There is frequent confusion in the literature about the stereochemistry at carbon-20 in these compounds.



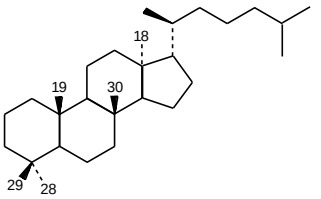
Euphane



Tirucallane

Apotirucallane triterpenoids (VS8050)

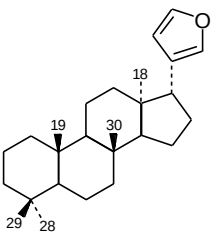
Further rearrangement of the tirucallane skeleton, the *apo*-rearrangement, affords the apotirucallane skeleton. Apotirucallanes are the notional parents of the tetranortriterpenoids (limonoids) and the quassinoids.



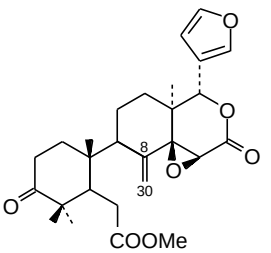
Apotirucallane

Nortriterpenoids (VS8100–VS8130)

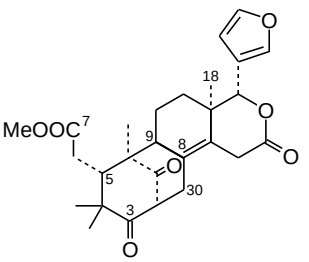
The tetranotriterpenoids (limonoids) are formed by loss of four terminal carbons of the apotirucallane skeleton. The side chain is typically a β-substituted furan although other oxidation levels are found to a lesser extent. A series of ring-cleavages and rearrangements can lead to a wide range of structures. For example cleavage of rings B and C may be followed by cyclisation to form a bicyclo[3.3.1]nonanolide system. In the Type of Compound Index, these compounds are presented in three groups – intact tetranortriterpenoids (VS8100), ring cleaved derivatives (VS8120) and rearranged derivatives (VS8130). The last group contains the bitter principles of the Cneoraceae, e.g. **Cneorin C**.



Tetranortriterpenoid skeleton



A tetranortriterpenoid with rings B and D cleaved



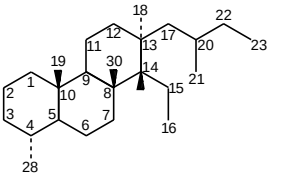
A bicyclo[3.3.1]nonanolide tetranortriterpenoid

Isman, M.B. *et al.* (1996), *Recent Adv. Phytochem.*, **30**, 155.
Mondon, A. *et al.* (1983) *Prog. Chem. Org. Nat. Prod.*, **44**, 101.
Taylor, D.A.H. (1984) *Prog. Chem. Org. Nat. Prod.*, **45**, 1.

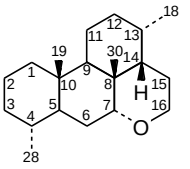
Quassinoid nortriterpenoids (VS8200, VS8205)

The quassinoids are found in the Simaroubaceae family and are closely related to the tetranortriterpenoids. This relationship is emphasised by the presence of several C_{25} precursors (e.g. **Simarolide**) which lose a further five carbon atoms to give the C_{20} picrasane skeleton. C_{18} and C_{19} quassinoids are also known but are less common. The quassinoids have attracted much synthetic effort because of their cytotoxic activity. *Chemical Abstracts* uses the picrasane skeleton as a

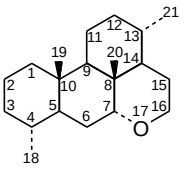
stereoparent; however the numbering system used by *Chemical Abstracts* differs from the accepted system; the oxygen atom is numbered.



C_{25} Quassinoid skeleton



Picrasane (C_{20})-usual numbering

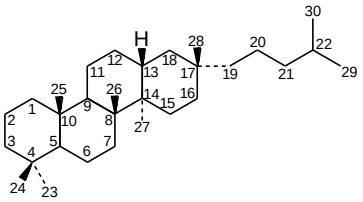


Picrasane, *CA* numbering

Kawada, K. *et al.* (1989) *Org. Prep. Proced. Int.*, **21**, 521 (*synth*).
Polonsky, J. (1973) *Prog. Chem. Org. Nat. Prod.*, **30**, 101; (1985) **47**, 221.

Baccharane triterpenoids (VS8230)

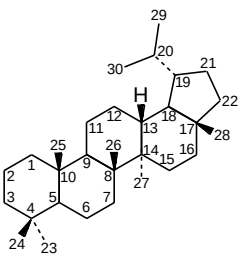
Cyclisation of squalene or squalene 2,3-epoxide in the chair-chair-chair-boat conformation leads initially to the baccharane skeleton. Backbone rearrangement of this skeleton leads to the lemnaphyllane and shionane skeletons (listed in the miscellaneous triterpenoid group). *Chemical Abstracts* treats this group as 18,19-secolupanes. The numbering system parallels that of lupane apart from the carbon atoms of ring D. Occasionally a steroid-like numbering system is used. This cyclisation of squalene provides an entry into the pentacyclic triterpenoids.



Baccharane
D : *B*-Friedo-18,19-secolupane, 9CI

Lupane triterpenoids (VS8250)

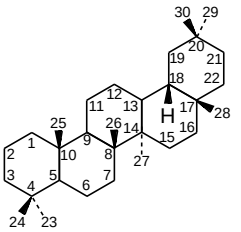
Formation of a five membered ring E from the baccharane precursor affords the lupane skeleton whose numbering system is as shown. 3(2 → 1)-Abeolupanes (VS8260) and assorted nor-, friedo- and seco-lupanes (VS8270) are listed separately.



Lupane, 9CI
(23/24 substituents specified in CA as (4α)23- and (4β)23- respectively)

Oleanane triterpenoids (VS8300)

Formation of a six-membered ring E from the baccharane precursor leads to the oleanane group. Oleananes form the largest group of triterpenoids and occur widely in the plant kingdom often as glycosides. The fairly numerous nor-, seco- and abeooleananes are listed separately (VS8310).

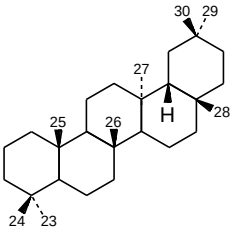


Oleanane, 9CI
(23/24 substituents specified in CA as (4α)23- and (4β)23- respectively; and 29/30 substituents as (20α)29 and (20β)29 respectively)

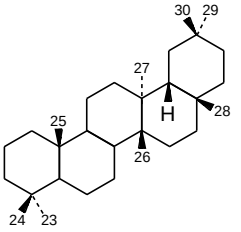
Agrawal, P.K. *et al.*, (1992), *Prog. Nuclear Magn. Reson. Spect.*, **24**, 1.
Mahato, S.B. *et al.* (1988) *Phytochemistry*, **27**, 3037.
Tschesche, R. *et al.* (1973) *Prog. Chem. Org. Nat. Prod.*, **30**, 461.

Taraxerane, multiflorane, glutinane and friedelane triterpenoids (VS8350–VS8510)

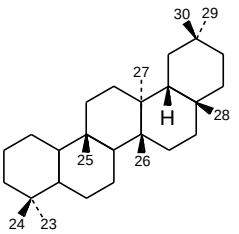
These triterpenoids arise by increasing degrees of backbone rearrangement of the oleanane skeleton.



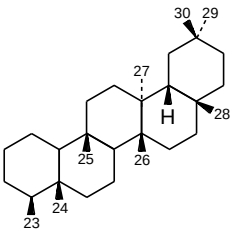
Taraxerane
D -Friedooleanane, 9CI



Multiflorane
D : *C*-Friedooleanane, 9CI



Glutinane
D : *B*-Friedooleanane, 9CI

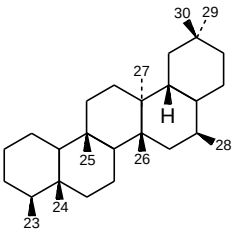


Friedelane
D : *A*-Friedooleanane, 9CI

Chandler, R.F. *et al.* (1979) *Phytochemistry*, **18**, 711.
Gunatikala, A.A.L. (1996), *Prog. Chem. Org. Nat. Prod.*, **67**, 1.

Pachysanane triterpenoids (VS8520)

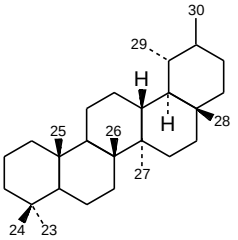
The pachysananes are friedelane derivatives which have the C-28 methyl group attached to C-16.



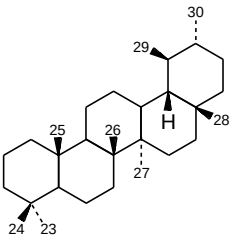
Pachysanane
16-Methyl-*D* : *A*-friedo-28-noroleanane, 9CI

Taraxastane, ursane and bauerane triterpenoids (VS8550–VS8700)

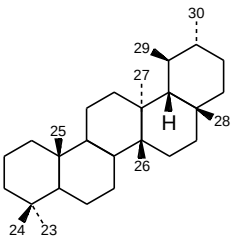
Methyl migration in ring E of the oleanane precursor leads to the taraxastane skeleton (following proton loss) or to the stereoisomeric ursane skeleton (following a series of hydride shifts). These two systems are often confused in the literature. The bauerane skeleton is related to ursane by backbone rearrangement.



Taraxastane



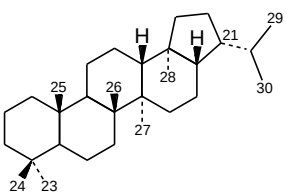
Ursane, 9CI



Bauerane
D : *C*-Friedoursane, 9CI

Hopane triterpenoids (VS8720, VS8730)

Cyclisation of squalene in the chair-chair-chair-chair-chair conformation affords the hopane skeleton and following a ring E expansion step, the gammacerane skeleton (see below). Degraded and extended hopanes occur widely in natural sediments. 21α*H*-Hopanes (moretanes) arise by cyclisation of squalene in the chair-chair-chair-chair-boat conformation.

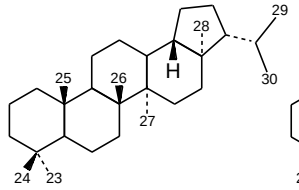


Hopane
A′-Neogammacerane, 9CI

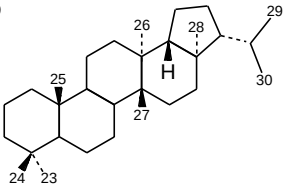
Ourisson, G. *et al.* (1992) *Acc. Chem. Res.*, **25**, 403.

Neohopane, fernane, adianane and filicane triterpenoids (VS8770, VS8800, VS8850, VS8870)

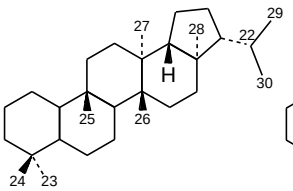
Backbone rearrangement of the moretane skeleton leads in turn to the neohopanes (neomotianes), fernanes, adiananes and filicanes.



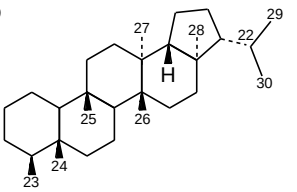
Neohopane
B′ : *A′*-Neogammacerane, 9CI



Fernane
D : *C*-Friedo-*B′* : *A′*-neogammacerane, 9CI



Adianane
D : *B*-Friedo-*B′* : *A′*-neogammacerane, 9CI

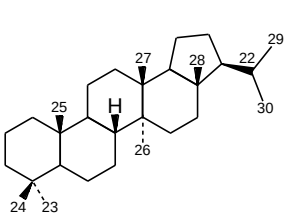


Filicane
D : *A*-Friedo-*B′* : *A′*-neogammacerane, 9CI

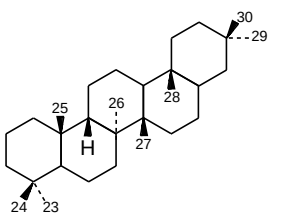
Murakami, T. *et al.* (1988) *Prog. Chem. Org. Nat. Prod.*, **54**, 1.

Arborinane and stictane triterpenoids (VS8850, VS8900)

Cyclisation of squalene, or its 2,3-epoxide, in the chair-boat-chair-chair-boat conformation followed by ring expansion of ring E yields the stictane skeleton. Members of this group occur in lichens, e.g. *Sticta* spp. Backbone rearrangement of this initial cyclisation product gives the arborinane skeleton. Two variants to this series **Boehmerol** and **Boehemerone** have undergone partial backbone rearrangement. (See under Miscellaneous triterpenoids).



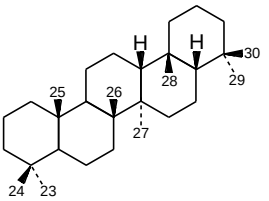
Aborinane
D : *C*-Friedo-*B′* : *A′*-neogammacerane, 9CI



Stictane
21,21-Dimethyl-29,30-dinorgammacerane, 9CI

Gammacerane triterpenoids (VS8950)

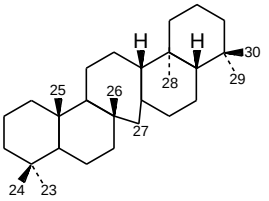
The gammacerane skeleton arises from the same cyclisation as hopane. The most notable gammacerane derivative is **Tetrahymenol**, a metabolite of the protozoon *Tetrahymena pyriformis*, whose hydroxyl function is derived from water and not from squalene oxide.



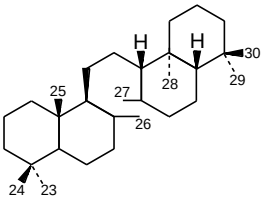
Gammacerane, 9CI

Serratane and onocerane triterpenoids (VS9000, VS9050)

Cyclisation of squalene, or more likely, its bisepoxide, from both ends affords the onocerane skeleton. Further cyclisation leads to the serratane skeleton.



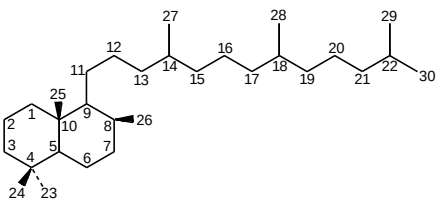
Serratane
C(14*a*)-Homo-27-norgammacerane, 9CI



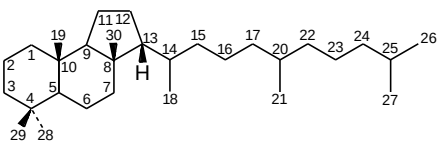
Onocerane
8,4-Secogammacerane, 9CI

Polypodane, malabaricane and podiodane triterpenoids (VS9080, VS9100)

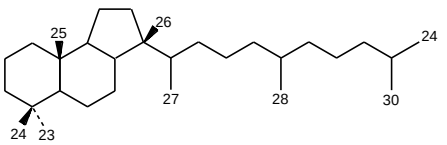
Partial cyclisation of squalene 2,3-epoxide from one end leads to the polypodane and the malabaricane groups. The podiodanes are malabaricanes which have undergone a methyl migration. The isomalabaricane skeleton (8,9-diepimer of malabaricane) has recently emerged. The most widely used numbering systems are given below.



Polypodane



Malabaricane
15-Methyl-*D*-homo-*C*,30-dinor-13,17*a*-secodammarane, 9CI



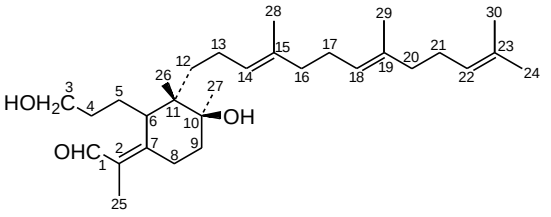
Podiodane

Miscellaneous triterpenoids (VS9300)

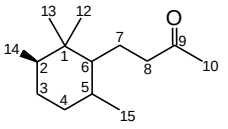
This group contains assorted triterpenoid skeletons which are less easily classified. It includes intriguing compounds such as **Sipholenone C** from the sponge *Siphochalina siphonella*.

Iridal group norterpenoids (VS9350)

The iridals are constituents of *Iris* spp. which serve as the precursors of the important perfumery chemicals, the irones. The numbering system of iridal is based on that of squalene. The irones are also included in this section. The numbering system of the irone skeleton is based on the carotenoids.



Iridal



Irone

Jaenicke, L. *et al.* (1986) *Prog. Chem. Org. Nat. Prod.*, **50**, 1.
Jaenicke, L. *et al.* (1990) *Pure Appl. Chem.*, **62**, 1365.

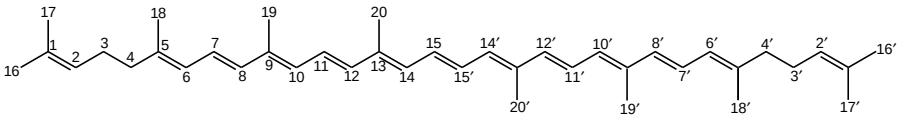
Tetraterpenoids (VS9400)

The tetraterpenes arise by head to head coupling of two geranylgeranylpyrophosphate molecules.

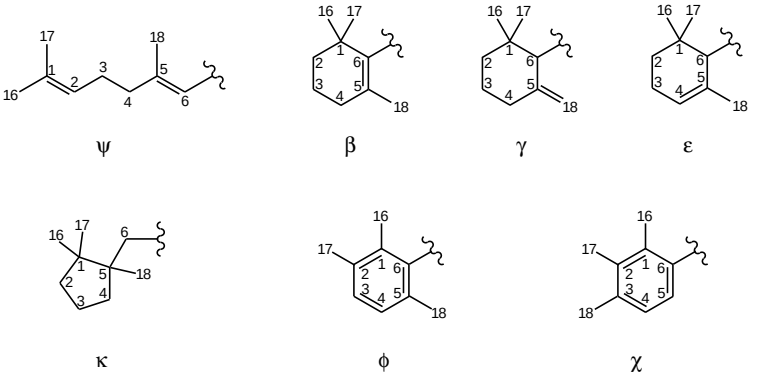
Spurgeon, S.L. *et al.* (1981) in *Biosynthesis of Isoprenoid Compounds*, (eds J.W. Porter *et al.*), Wiley, New York, Vol. 2, p. 1.

Carotenoids

These include the hydrocarbons (carotenes) and their oxygenated derivatives (xanthophylls). Carotenoid nomenclature is based on a stem name, carotene, and two Greek letters as prefixes to define the two end groups. The numbering system and end groups are given below.



ψ,ψ-Carotene



Carotenoid end-groups

IUPAC treats ‘hydro’ prefixes in carotenoid names as non-detachable. This Dictionary follows IUPAC recommendations for nomenclature except that the ‘hydro’ prefix is treated as detachable and is placed alphabetically with the other prefixes. CA also uses a detachable ‘hydro’ prefix but it does not use hypothetical parents such as *β*-caroten-6-ols which are incapable of existence (see current *Chemical Abstracts Index Guide*). The following examples illustrate this point.

IUPAC name	*Chemical Abstracts* name
5,6-Dihydro-*β*,*β*-caroten-3-ol	5,6-Dihydro-*β*,*β*-caroten-3-ol
5,6-Dihydro-*β*,*β*-caroten-6-ol	5,6-Dihydro-6-hydroxy-*β*,*β*-carotene

Britton, G. (1991) *Methods Plant Biochem.*, **7**, 473.

Britton, G. (1991) *Nat. Prod. Rep.*, **8**, 223.

Goodwin, T.W. (1980) *Biochemistry of the Carotenoids*, 2nd edn. Chapman & Hall, London.

Goodwin, T.W. (1992) *Methods Enzymol.*, **213**, 167.

Hill, R.A. (1993) in *The Chemistry of Natural Products*, 2nd edn (ed. R.H. Thomson), Blackie, Glasgow, pp. 135.

IUPAC (1975) *Pure Appl. Chem.*, **41**, 407.

Pfander, H. (1981) *Key to Carotenoids*, Birkhäuser, Basel.

Pfander, H. (1992) *Methods Enzymol.*, **213**, 3.

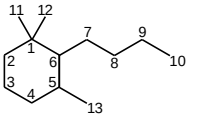
Ramage, R. (1972) in *Chemistry of Terpenes and Terpenoids*, (ed. A.A. Newman), Academic Press, London, p. 288.

Sandmann, G. (1994) *Eur. J. Biochem.*, **223**, 7.

Miscellaneous terpenoids (VS9450–VS9910)

Megastigmane norterpenoids (VS9450)

This is a fairly large group of C_{13} compounds generally thought to be degraded carotenoids or catabolites of abscisic acid.



Megastigmane
2-Butyl-1,1,3-trimethylcyclohexane

Izoe, S. *et al.* (1969) *Tetrahedron Lett.*, 279.
Powell, R.G. *et al.* (1986) *J. Org. Chem.*, **51**, 1074.

Apocarotenoids (VS9700)

Apocarotenoids are carotenoids in which the carbon skeleton has been shortened by the formal removal of fragments from one or both ends. A locant is used to indicate that all the molecule beyond the carbon with that locant has been removed. It is not necessary to give a Greek-letter end group designation if the apo-locant is greater than 5.

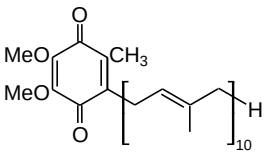
Polyterpenoids (VS9800)

This section includes the higher carotenoids and polyprenols with more than 40 carbons.

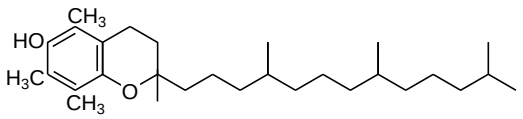
Britton, G. (1991) *Nat. Prod. Rep.*, **8**, 223.
Tanaka, Y. (1991) *Methods Plant Biochem.*, **7**, 519.

Meroterpenoids (VS9900)

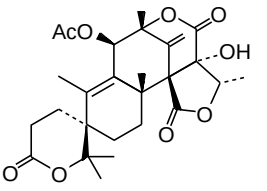
Meroterpenoids are of mixed biogenesis containing terpenoid and non-terpenoid derived fragments. This broad definition could include the vast number of simple prenylated phenolics but is normally reserved for compounds where the terpenoid fragment comprises a large part of the molecule. The polyprenylated quinones and chromanols typified by the ubiquinones and tocopherols are clearly of mixed biogenesis but the metabolites of *Aspergillus ustus* such as Austin could be mistaken for sesterterpenoids. In fact these metabolites have been shown to be derived from a sesquiterpenoid fragment and an aromatic polyketide fragment.



Ubiquinone 10



α-Tocopherol

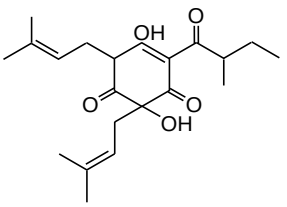


Austin

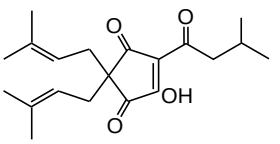
Ahmed, S.A. *et al.* (1989) *J. Chem. Soc. Perkin Trans 1*, 807.
Hubscher, M. *et al.* (1990) *Helv. Chim Acta*, **73**, 782; 1068.
Konishi, K. *et al.* (1987) *Chem. Pharm. Bull.*, **35**, 1531.

Hop meroterpenoids (VS9910)

The bitter hop constituents exemplified by Adhumulone and the ring contracted Hulupone have been shown to be derived by prenylation of a polyketide aromatic ring.



Adhumulone



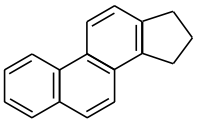
Hulupone

Drawert, F. *et al.* (1976) *Phytochemistry*, **15**, 1689; 1695.

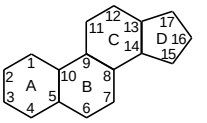
Steroids (VT)

For general information on the biogenesis of steroids, see the preceding terpenoid section.

The steroid structure is based on four carbocyclic rings arranged as in cyclopenta[*a*]phenanthrene, which is normally fully or partially reduced so that only limited unsaturation, if any, is present. The four steroid rings are labelled, and their carbon atoms are numbered according to the universal convention illustrated.

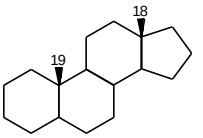


Cyclopenta[*a*]phenanthrene



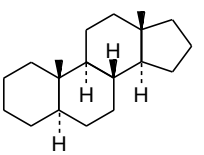
Basic steroid structure

The great majority of steroids also have one or two methyl groups present at the bridgehead positions C-10 and C-13; the methyl carbon atoms are numbered C-19 and C-18, respectively.



Methyl groups, hydrogen atoms, or substituents at the bridgehead positions C-8,9,10,13, and 14 are assumed to have the 8β, 9α, 10β, 13β, 14α configurations unless otherwise specified. C-5 is a special case, as there are many steroids of each of the 5α and 5β configurations, and it is therefore necessary to specify the C-5 configuration for any steroid which is saturated at C-5. (e.g. 5α-Androstane or 5β-Androstane).

It is worth noting here some changes in *Chemical Abstracts* indexing policy. Prior to the 8th Collective Index (1967), the indexing of steroid stereoisomers gave priority to the C-5 configuration which effectively led to a separation of 5α- and 5β-steroids. Users should be alert to this when searching the literature before 1967.

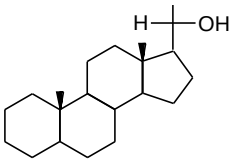


5α-Androstane

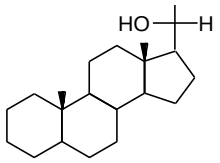
The hydrogen atoms at C-8,9, and 14 are generally omitted from formulae if they have the natural configurations shown here.

Any side-chain at C-17 is assumed to have the 17β-configuration unless otherwise indicated. This is shown either by using a wedge bond or, where there is any possibility of uncertainty owing to substitution at C-20, by drawing in the C-17 α-hydrogen atom.

Configurations of substituents in the side chain were formerly also indicated by α or β, according to a convention devised by Fieser, whereby the side-chain is drawn in Fischer projection, with the highest numbered locant at the top.



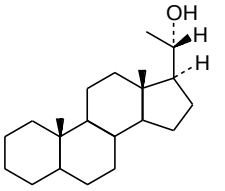
20α-form



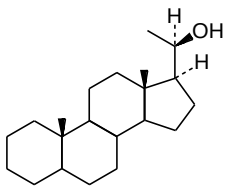
20β-form

The Fieser convention for pregnan-20-ols

However, use of this arbitrary system demands either a good memory or reference to printed texts. The unambiguous Cahn-Ingold-Prelog sequence rule descriptors (*R* or *S*) are now recommended for side-chain configurations. Designations according to Fieser's system are also given in DNP entries where these are or have been in widespread use.

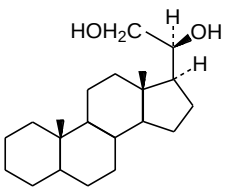


20*S* (≡ 20α)

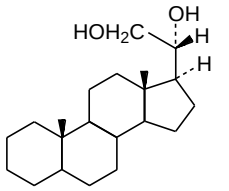


20*R* (≡ 20β)

The presence of substituents at C-17 or C-21 may change the priority of groups so that 20*S* is no longer equivalent to 20α. This happens for example in the pregnane-20,21-diols.



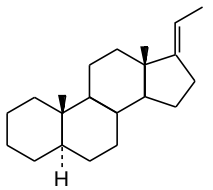
20*S* (≡ 20β)



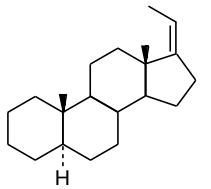
20*R* (≡ 20α)

Difficulty in labelling configurations unambiguously can also occur in secosteroids, in which a ring bond has been broken. Parts of the molecule which are normally constrained when the rings are intact become free to rotate in relation to each other so that α and β lose their defined meanings. The sequence rule is again recommended to overcome this problem. The compounds of the Vitamin D series (9,10-secosteroids) are the most important in this class.

The sequence rule descriptors (*E*-) and (*Z*-) are required for defining side-chain double bond configurations.



(*E*)-5α-Pregn-17(20)-ene
5α-Pregn-17(20)*E*-ene



(*Z*)-5α-Pregn-17(20)-ene
5α-Pregan-17(20)*Z*-ene

(Note that both locants for unsaturation are required when the numbers are non-consecutive.)

Bernstein, S. *et al.* (1968) *Physical Properties of Steroid Conjugates*, Springer-Verlag, Berlin.

Danielsson, H. and Sjövall, J. (1985) *Sterols and Bile Acids*, Elsevier, Amsterdam.

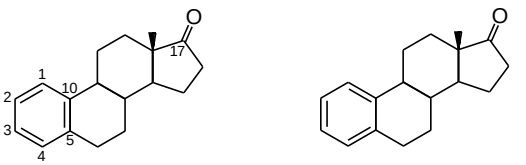
Duax, W.L. and Norton, D.A. (eds) (1975) *Atlas of Steroid Structure*, Vol. 1; Vol. 2 (1984) Plenum, New York.
Fieser, L.F. and Fieser, M. (1959) *Steroids*, Reinhold, New York.
Goad, L.F. and Akihisa, T. (1997) *Analysis of Sterols*, Blackie, London.
Hanson, J.R. (1997) *Nat. Prod. Rep.*, **14**, 373.
Hill, R.A., Kirk, D.N., Makin, H.L.J. and Murphy, G.M. (1991) *Dictionary of Steroids*, Chapman & Hall, London.
Makin, H.L.J. (ed.) (1984) *Biochemistry of Steroid Hormones*, 2nd edn, Blackwell, Oxford
Nair, P.P. and Kritchevsky, D. (eds) (1971) *The Bile Acids*, Vol. 1; Vol. 2 (1973); Vol. 3 (1976); Setchell, K.D.R., Kritchevsky, D. and Nair, P.P. (eds) (1988) Vol. 4, Plenum, New York.
Turner, A.B. (1993) in *The Chemistry of Natural Products*, 2nd edn (ed. R.H. Thomson), Blackie, Glasgow, pp. 140.
Zeelen, F.J. (1990) *Medicinal Chemistry of Steroids*, Elsevier, Amsterdam.
Zeelen, F.J. (1994) *Nat. Prod. Rep.*, **11**, 607 (*synth*).

In the notes which follow, the carbon numbers used to classify the different types of steroids refer only to those which constitute the parent steroid skeleton. They do not include any carbon atoms which may be present as substituents (e.g. 6-methyl), or in derivative groups such as the ethers or esters of steroid alcohols. Within the Type of Compound Index, the steroid groups are arranged in order of increasing carbon number, which may not correspond exactly with the order in which they are discussed in the following sections.

Estrane steroids (aromatic ring A, C_{18}) (VT0100)

Estrane (oestrane) is the parent hydrocarbon of the estrogens, the hormones responsible for development of the female reproductive organs and secondary sex characteristics. The original spellings ‘oestrogen’ and ‘oestrane’, although still used, are superseded by the form with the initial ‘o’ omitted.

The estrogens have an aromatic ring A. Two ‘Kekulé’ forms can be drawn, as for all benzene derivatives.



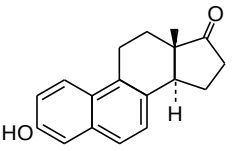
(*i*) [1,3,5(10)-triene] (*ii*) [1(10),2,4-triene]

Kekulé forms of Estrone

For purposes of naming and illustrating estrogens the form (*i*) is preferred; the natural estrogens are accordingly derivatives of estra-1,3,5(10)-triene. They have a hydroxyl group at C-3, and hydroxyl or carbonyl at C-17. Some metabolites have additional oxygen functions elsewhere.

The trivial names and abbreviations of **Estrone** (3-hydroxyestra-1,3,5(10)-trien-17-one), **Estradiol** (estra-1,3,5(10)-triene-3,17β-diol) and **Estriol** (estra-1,3,5(10)-triene-3,16α,17β-triol), are commonly used, especially in biochemical and medical contexts. These trivial names are sometimes incorporated into those of derivatives (e.g. **17α-Methylestradiol**).

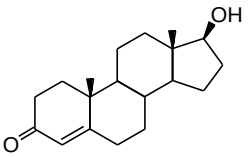
Equine estrogens, with unsaturation additionally in ring B, belong to the estra-1,3,5(10),7-tetraene and estra-1,3,5,7,9-pentaene series. Note the recommended change in numbering for unsaturation in the latter case, where it is possible to avoid the need for a compound locant [5(10)].



3-Hydroxyestra-1,3,5,7,9-pentaen-17-one

Androstane steroids (C_{19}) (VT0250)

Androstane is the parent hydrocarbon of the male hormone Testosterone (17β-hydroxyandrost-4-en-3-one) and its derivatives and metabolites. The androstane ring structure with the two bridgehead methyl groups is common to most other groups of steroids (pregnanes, cholanes, cholestanes, etc) which have side-chains at C-17 and are discussed individually below.

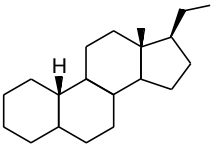


17β-Hydroxyandrost-4-en-3-one (Testosterone)

Zhou, Z.X. *et al.* (1994) *Recent. Prog. Hormone Res.*, **49**, 249.

C_{20} steroids (VT0400)

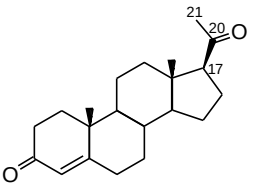
These are scarce among natural products, and are limited to a few 19-nor-pregnanes, with the pregnane skeleton (see below) but lacking the bridgehead methyl group (C-19). Alternative names based upon 17-ethylestrane are also often used for this series of compounds (see pregnanes).



19-Norpregnane

Pregnane steroids (C_{21}) (VT0450)

Pregnane is the parent hydrocarbon of the pregnancy hormone progesterone (pregn-4-ene-3,20-dione), and of the great majority of the corticosteroids and many other natural products, which together make the pregnanes the largest single group of steroids. The skeletal structure comprises androstane with a two-carbon (ethyl) chain at C-17. The chain is in the β-configuration unless 17α-pregnane is specified.



Pregn-4-ene-3,20-dione (Progesterone)

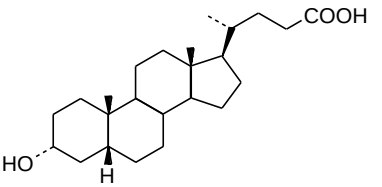
Many pregnane derivatives have hydroxyl or a related group at C-17. To avoid any ambiguity as to the configuration, epimeric forms of 17-substituted pregnanes are specified in DNP as 17αOH or 17βOH.

Norcholan-23-oic acid (C_{23}) and cholan-24-oic acid (C_{24}) steroids (VT0650, VT0800)

The largest single group in this class comprises the bile acids, the majority of which are cholan-24-oic acids. The shorter form 'cholanic acid' has been widely used but is not recommended.

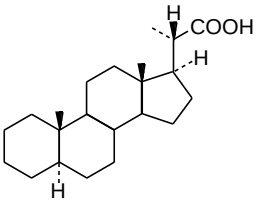
Naturally occurring bile acids are mainly 5β-cholan-24-oic acids with hydroxyl substitution at C-3, and variously at other sites. The orientation of the hydroxyl groups is usually α-, and bile acids have well-defined polar and non-polar regions. This amphipathic quality of bile acids is essential to their physiological functions.

Bile acids are biosynthesised from cholesterol; *primary* bile acids directly so (e.g. **Cholic acid, Chenodeoxycholic acid**); subsequent action by intestinal bacteria yields *secondary* bile acids (e.g. **Deoxycholic acid**, Lithocholic acid).



3α-Hydroxy-5β-cholan-24-oic acid (Lithocholic acid)

Omission of one of the side chain carbon atoms leads to the 24-nor series (24-norcholan-23-oic acids); loss of two carbon atoms gives the 23,24-dinorcholan-22-oic acids. The latter are sometimes named as pregnane-20-carboxylic acids, requiring a sequence-rule descriptor of the C-20 configuration.



23,24-Dinor-5α-cholan-22-oic acid
5α-Pregnane-20*S*-carboxylic acid

Whenever cholane nomenclature is used, the side chain has the C-20 configuration which is illustrated above for Lithocholic acid (20*R*, in the absence of substituents at or near C-20), unless the opposite is specified.

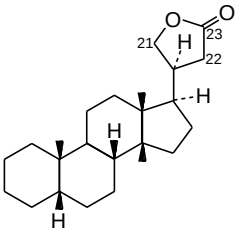
Haslewood, G.A.D. (1978) *The Biological Importance of Bile Salts*, North-Holland, Amsterdam.

Cardanolide steroids (C_{23}) (VT0750)

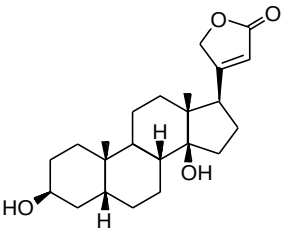
Cardanolide is the parent compound of the *Digitalis* glycosides and comprises the androstane skeleton with a γ-lactone ring attached at C-17. The configuration at C-5 must be stated, but is frequently β. Prior to the most recent IUPAC-IUB recommendations (1989), the 14α-configuration was assumed unless the 14β-configuration was indicated as an affix. Almost all natural products in these series, however, are of 14β type, and the convention for C-14 has been reversed so that the cardanolide name implies 14β-configuration. This is in contrast to the rule for all other steroid classes. The change from the older system, which has been in use for several decades, seems likely to lead to confusion and so the C-14 configuration is specified for *all* such compounds in

DNP. The formulae illustrated below also show the 17β and 20*R* configurations which are implied in the absence of a contrary indication.

The naturally occurring compounds generally have a 20(22)-double bond and are commonly called cardenolides, for example Digitoxigenin is 3β,14β-dihydroxy-5β-card-20(22)-enolide. The cardenolide glycosides are of particular interest because of their cardiac activity.



5β-Cardanolide

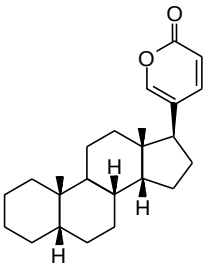


Digitoxigenin

Connolly, J.D. *et al.* (1991) *Methods Plant Biochem.*, **7**, 369.
Deepak, D. (1996), *Prog. Chem. Org. Nat. Prod.*, **69**, 71.
May, P.M. (1990) *Comprehensive Medicinal Chemistry*, Pergamon Press, Oxford, Vol. 2, p. 206.

Bufanolide steroids (C_{24}) (VT0900)

Bufanolide also has the androstane skeleton but in this case a δ-lactone ring is attached at C-17. It is the parent compound of the cardioactive constituents obtained from toad skin secretions and the sea onion or squill (*Scilla maritima*). As for the cardanolides, 14β-, 17β- and 20*R*-configurations are implied in the name. The naturally occurring compounds are generally doubly unsaturated in the lactone ring (bufa-20, 22-dienolide), and often occur as glycosides or as conjugates.



5β-Bufa-20,22-dienolide

The sterols

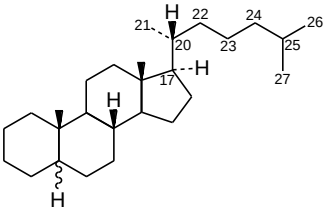
The sterols comprise several major groups of steroids characterised by having a hydroxyl group at C-3, normally in the β-configuration, and branching side chains of from eight to ten or more carbon atoms at C-17. They occur widely throughout the animal and particularly the plant kingdoms. They have both structural roles, as membrane constituents, and a key place in the biosynthetic sequences which lead to the steroid hormones and other biologically active steroidal species.

The following sections detail the main features of the various parent hydrocarbons which provide the structural basis and classification of the sterols.

Good, L.J. (1991) *Methods Plant Biochem.*, **7**, 369.
Kerr, R.G. *et al.* (1991) *Nat. Prod. Rep.*, **8**, 465.
Minale, L., (1993), *Prog. Chem. Org. Nat. Prod.*, **62**, 75.

Cholestane steroids (C_{27}) (VT1050, VT1100)

The cholestane skeleton, which derives its name from the longest-known and most familiar compound of its class, **Cholesterol**, can be regarded as the parent from which almost all other sterols are derived. This is true structurally, if not necessarily in terms of the detailed biosynthetic pathway.



Cholestane (5α- or 5β-Cholestanes)

Other classes of sterols are derived from cholestane by the addition of one or more carbon atoms at side-chain positions, most commonly C-24 (see ergostanes, stigmastanes, etc, below).

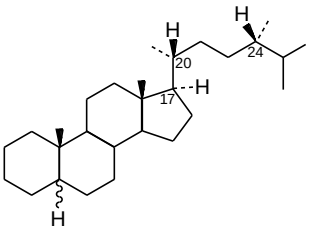
Several other steroid classes have structures based upon the C_{27} cholestane framework, although this is not always immediately apparent from the formulae as drawn. The ecdysteroids (insect moulting hormones) are highly oxygenated cholestanes. Many plant products that are commonly classified as tetracyclic triterpenes are cholestanes with altered stereochemistry and/or additional methyl substitution in the ring system, notably at C-4, -8, or-14. The dividing line between the sterols and the tetracyclic triterpenes is a matter mainly of origin and custom. Spirostans, furostans, and many of the steroidal alkaloids have structures which are formally derived from cholestanes by linking between two side-chain sites, or between a side-chain and a skeletal carbon, via an oxygen (epoxy) or nitrogen (epimino) bridge. **Vitamin D_3** and its analogues are 9,10-secocholestanes. All of these classes are described separately below.

Alkylated cholestanes of many types occur widely in plants, fungi, and marine organisms. The very large classes of 24-methylcholestanes (ergostanes and campestanes) and 24-ethylcholestanes (stigmastanes and poriferastanes) are sufficiently important that their parent hydrocarbons have been assigned these special systematic names (not used in *Chemical Abstracts* however). They are treated in separate sections below. The 4,4,14-trimethylcholestanes (lanostanes) are covered in the preceding terpenoid section. Many alkylcholestane derivatives, however, fall outside these major groups, and have not been dignified by special class names. They are treated in DNP as derivatives of cholestane. Others are homocholestanes, in which additional carbon atoms lengthen the side-chain, rather than branching off it. Many of these unusual sterols are best known by trivial names that reflect their biological origins.

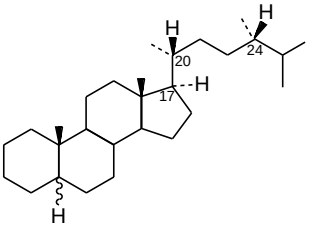
The cholestan-26-oic (or 27-oic) acids (VT1100) form a small but significant group of bile acids (see cholanes above).

Ergostane steroids (excluding withanolides and brassinolides) (VT1300)

The 24-methylcholestane structure is termed either ergostane or campestane, depending upon the configuration at C-24 although *Chemical Abstracts* indexes campestanes as 24*R*-ergostanes. The Fieser convention (see above) defines ergostanes as 24β-methylcholestanes; campestanes are 24α-methylcholestanes. These stereochemical labels have the advantage of being unaffected by adjacent substitution or unsaturation. While the saturated and Δ^{25}-unsaturated ergostane side-chanes have the 24*S* configuration, the altered priorities of groups around C-24 give ergost-22-ene the 24*R* configuration.



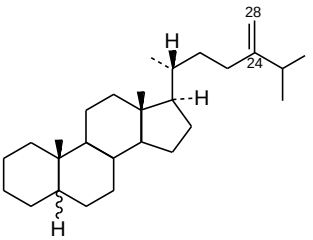
Ergostane



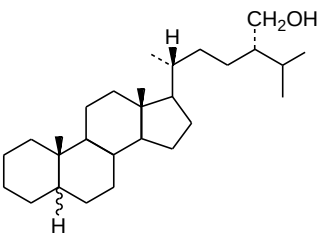
Campestane

For historical reasons, most of the compounds of these classes have become known as ergostane (or ergosterol) derivatives, even though, according to current nomenclature, some of them should strictly be named as campestanes. In DNP, therefore, ergostane-based nomenclature is generally given precedence, with campestane synonyms added where appropriate.

A further complication, firmly rooted in historical precedent, is the use of the locant C-28 for the carbon atom of the 24-methyl group. The latest IUPAC-IUB recommendation is that the locant C-28 be reserved for the 4α-methyl group in lanostanes, and in other 4,4-dimethylsterols of terpenoid type, with C-29 and C-30 allocated, respectively, to the 4β- and 14α-methyl groups.* The locant C-28 has therefore acquired two distinct meanings, according to context. In DNP the C-24 methyl group in ergostanes and campestanes retains its original locant as C-28, allowing the use of derivative names containing such expressions as ergost-24(28)-ene (for 24-methylenecholestanes) or ergostan-28-ol (for 24-hydroxymethylcholestanes). The cholestane-based synonyms favoured by IUPAC-IUB are also given, where necessary, for clarity.



24-Methylenecholestane
[Ergost-24(28)-ene]



24-Hydroxymethylcholestane
(Ergostan-28-ol)

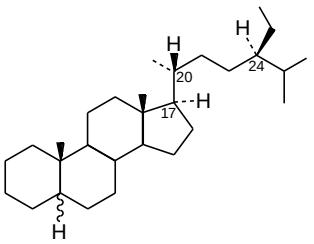
The C_{28} ergostane skeleton occurs in some other groups of compounds of steroidal type, notably the withanolides and the brassinolides, which are highly oxygenated ergostane derivatives (see below). Compounds of the Vitamin D_2 class are 9,10-secoergostane derivatives (see Vitamin D, below). Ergostanes with the 9β,10α-configuration, which are among the products of photochemical transformation of ergosta-5,7-dienes, have commonly been named as lumistanes, although this term is not recognised in the IUPAC-IUB rules for nomenclature.

Stigmastane steroids (C_{29}) (VT1550)

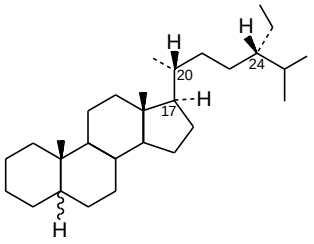
These are the 24-ethylcholestanes, stigmastanes and poriferastanes being epimeric at C-24. The long history of stigmastane-based nomenclature, derived from the common plant sterol Stigmasterol**,** has ensured that this is by far the

* The current IUPAC-IUB-recommended locant number for the 24-methyl carbon is 24[1], applicable, for example, in listing ^{13}C nmr assignments, but not recommended for use as a locant for unsaturation or further substitution. In such cases the entire C-24 substituent should be appropriately named, e.g. as a 24-methylene or 24-(hydroxymethyl) derivative of the cholestane series.

more widely used of the two names, a situation paralleling that described above for ergostanes and campestanes. In the Fieser system, stigmastanes have the 24α configuration, and poriferastanes are 24β. Again the sequence rule is now preferred, with 24*R* or 24*S* depending upon local substitution and/or unsaturation. *Chemical Abstracts* indexes poriferastanes as 24*S*-stigmastanes.

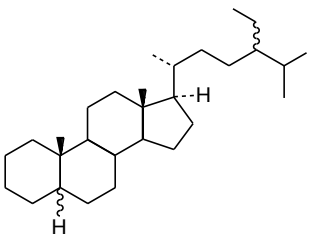


Stigmastane

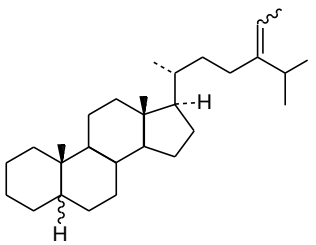


Poriferastane

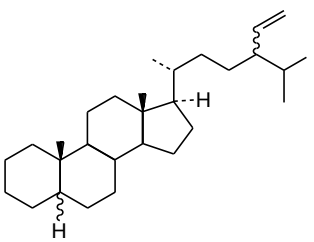
As with ergostanes, common usage over several decades has favoured the locants C-28 and C-29 for the two ethyl carbon atoms, and these are used here. The IUPAC-IUB recommendation is that the two ethyl carbon atoms be designated 24^1 and 24^2 whenever locants are needed. Synonyms based upon 24-ethylcholestane, 24-ethylidenecholestane, or 24-vinylcholestane are given in DNP where suitable.



24-Ethylcholestane



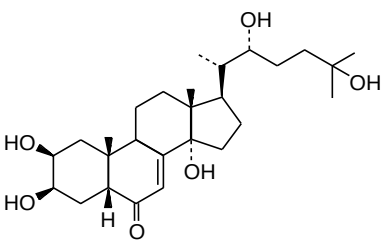
24-Ethylidenecholestane



24-Vinylcholestane

Ecdysteroids (C_{27}) (VT1150)

Ecdysteroids or ecdysones are moulting hormones of insects and crustaceans. They have also been isolated from many plants. The first ecdysone to be isolated was α-Ecdysone from the silkworm (*Bombyx mori*). Most ecdysteroids have a 2β,3β,14α,20,22-pentahydroxy-5β-cholest-7-en-6-one skeleton with further hydroxylation.



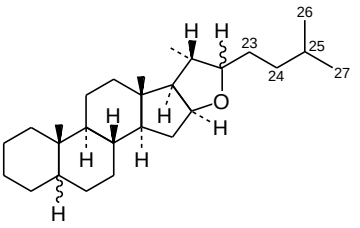
α-Ecdysone

Grieneisen, M.L. (1994) *Insect Biochem. Mol. Biol.*, **24**, 115
Nakanishi, K. (1971) *Pure Appl. Chem.*, **25**, 167.

Spirostan and furostan steroids (C_{27}) (VT1200, VT1250)

Many plant products belong to these related classes; a few of the spirostans, notably **Diosgenin** and **Hecogenin**, are sufficiently plentiful to have become major sources of steroidal intermediates for the synthesis of steroid hormones and pharmaceutical analogues.

The furostans are 16β, 22-epoxycholestanes, the extra ring being labelled as ring E. The parent structure furostan is defined as having the side-chain configuration illustrated. The configuration at C-22 (when saturated) is indicated according to the sequence rule. Those derivatives that are further substituted in the side chain also require sequence rule designations, including C-25 if C-26 is substituted. Some naturally-occurring furostan derivatives have additional epoxy rings between pairs of carbon atoms in the side chain. The spirostans (below) are a special case.



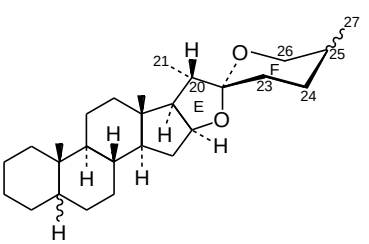
Furostan

Spirostans are 16β,22 : 22,26-diepoxycholestanes, or 22,26-epoxyfurostans. The sixth ring so formed is known as ring F. Chemically, the spiro centre at C-22 has the character of an internal acetal derived from a 16β, 26-dihydroxy-cholestan-22-one.

The parent name spirostan implies the configurations illustrated for C-20 and C-22, but that at C-25, and any other chiral locations if ring F is substituted, are given according to the sequence rule.

The omission of a terminal ‘e’ from the names furostan and spirostan recognises that they are not hydrocarbons. Derivative nomenclature for these classes of compounds, however, requires the addition of ‘e’ to the stem of the name if a consonant follows, e.g. 5α-spirostane-3β,12β-diol.

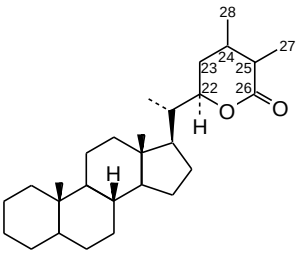
Tetrahedral geometry at the spiranic C-22 causes ring F to lie perpendicular to the general orientation of the other rings. Projection formulae onto the plane of the paper fail adequately to express this stereochemical relationship, and lead to difficulties in correctly illustrating the configurations of any substituents in ring F. The IUPAC-IUB-recommended way of drawing the formula avoids this problem by including a perspective representation of ring F, as shown below. The particular chair conformation illustrated is a matter of convention, and does not necessarily correspond to the preferred conformation in every case.



Spirostan

Withanolide and brassinolide steroids (C_{28}) (VT1400)

The withanolides are a group of naturally occurring plant steroids with an ergostane skeleton and a side-chain δ-lactone ring linking C-22 and C-26. The lactone ring is usually unsaturated at C-24, and there is a high level of oxygenation in the skeletal rings, frequently including a 2-en-1-one system and a 5,6-epoxide. The configurations are as shown below and the configuration at C-22 is usually *R*.

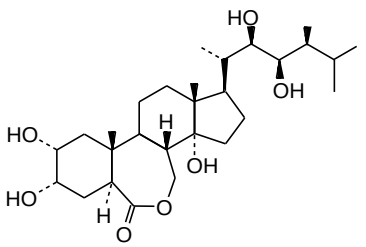


Withanolide

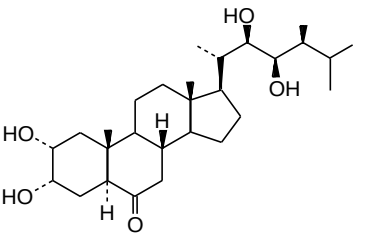
The Physalins are 13,14-secowithanolides with the formation of a 13,14- or 14,17-hemiacetal or acetal as in **Withaphysalin C** and **Physalin B**. Physalin B also has a new carbon-carbon bond between C-16 and C-24.

Brassinolides are a group of plant growth promoting substances originally isolated from rape pollen (*Brassica napus*) but now found to be widespread in plants. They are highly oxygenated ergostane derivatives, characterised by the expanded B-ring with incorporation of an oxygen atom to form an ε-lactone ring (B-homo-7-oxaergostan-6-one derivatives). The lactone is not essential for plant growth activity (Castasterone has an intact B-ring), but the 22*R*,23*R*-diol system is.

The oxygenation pattern bears some relationship to the ecdysteroids but the configurations at C-2,3 and 5 are α- in the brassinosteroids but are mostly β- in the ecdysteroids.



Brassinolide



Castasterone

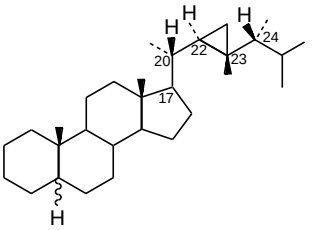
Fujioka, S. *et al.* (1997), *Nat. Prod. Rep.*, **14**, 1.
Glotter, E. *et al.* (1991) *Nat. Prod. Rep.*, **8**, 415.
Kirson, I. *et al.* (1981) *J. Nat. Prod.*, **44**, 633.
Ray, A.B. (1994), *Prog. Chem. Org. Nat. Prod.*, **63**, 1.

Gorgostane and other cyclopropacholestane steroids (C_{30}) (VT1700)

Gorgostane is the parent hydrocarbon of a widely-occurring group of sterols in marine organisms. Its skeleton comprises ergostane with an additional methyl group at C-23, and a methylene bridge between C-22 and C-23, forming a cyclopropane ring. Configurations in the side chain are as illustrated unless otherwise specified.

A wide variety of at least 100 diverse C_{30} and C_{31} marine sterols in the gorgostane and related structural classes are known. Sponges are the most prolific source. Sponge sterols are characterised by multiply alkylated side

chains, frequent presence of cyclopropane/cyclopropene functionality in the side chain, and wide variation in the steroid A–D ring skeleton, including many examples of A-nor and 19-nor variants.



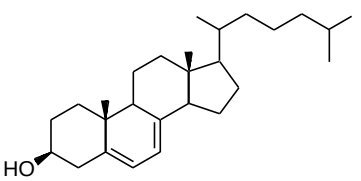
Gorgostane

D'Auria, M.V. *et al.* (1993) *Chem. Rev.*, **93**, 1839–1895.
Djerassi, C. *et al.* (1991) *Acc. Chem. Res.*, **24**, 371–378.
Faulkner, D.J. (1992) *Nat. Prod. Rep.*, **9**, 323–364.
Giner, J.L. (1993) *Chem. Rev.*, **93**, 1735–1752.

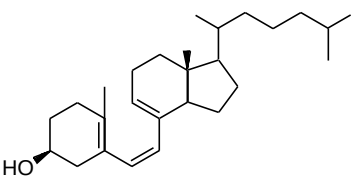
Vitamin D and related compounds (VT2850, VT2900)

The calcium-regulating vitamin D is found in two principal forms, vitamin D_2 and vitamin D_3, which differ only in the side chain. Vitamin D_2, sometimes called Ergocalciferol, is derived from the fungal sterol Ergosterol. Vitamin D_3, the natural mammalian form, is derived from cholesta-5,7-dien-3β-ol (7-dehydrocholesterol), and is accordingly known also as Cholecalciferol. Other compounds of the series are specified as belonging to either the ergostane or the cholestane series by use of the appropriate numerical subscript ($_2$ or $_3$). Both forms of vitamin D arise from photochemical ring-opening of the unsaturated ring B in the precursor sterol. The immediate products of ring-opening are known as previtamin D_2 or D_3, respectively. The previtamin has a (6*Z*)-9,10-seco-5(10),6,8-triene structure. Thermal rearrangement at physiological temperature shifts the unsaturation in the previtamin to form the vitamin itself, which has the (5*Z*,7*E*)-5,7,10(19)-triene structure. Metabolic changes in the liver and the kidney lead to introduction of hydroxyl substitution at C-25 and C-1, respectively, to give the active calcium-regulating hormones.

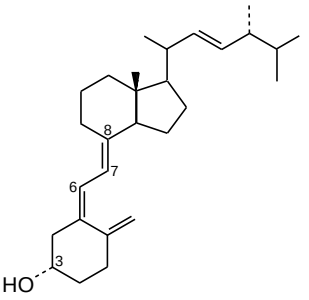
Formulae are usually drawn so as to represent the true elongated shape of the vitamin D molecule. To reach this conformation, the molecule has to undergo rotation around the 6,7-single bond within the triene system. This twisting reverses the orientation of ring A with respect to the remaining rings, so the normal meanings of α and β as applied to substituents in ring A become confused. Unambiguous sequence rule descriptors are usually preferred (but not by *Chemical Abstracts*), and are used in DNP. Thus the original 3β-hydroxy group becomes 3*S* in the vitamin, whereas the important '1α'-hydroxylated metabolites have the (1*S*,3*R*)-configuration. The sequence rule is also used, when necessary, to describe configurations at any other chiral centres in ring A, and at C-6 or C-7 in various reduced or oxidised derivatives of the triene system, as well as for side chain substituents.



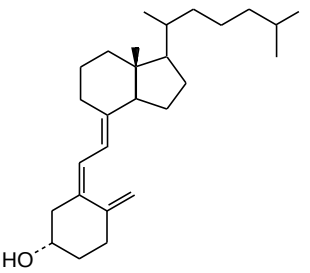
7-Dehydrocholesterol
(Cholesta-5,7-dien-3β-ol)



Previtamin D_3
(9,10-Secocholesta-5(10),6*Z*,8-trien-3*S*-ol)



Ergocalciferol; Vitamin D_2
(9,10-Secocholesta-5*Z*,7*E*,
10(19)22*E*-tetraen-3*S*-ol)



Cholecalciferol; Vitamin D_3
(9,10-Secocholesta-5*Z*,7*E*,
10(19)-trien-3*S*-ol)

Anon (1985) *Synform*, **3**, 75 (*synth*).
Coldwell, R.D. *et al.* (1990) *Steroids*, **55**, 418.
Okuda, K.I. *et al.* (1995), *J. Lipid. Res.*, **36**, 1641.
Zhu, G.D. *et al.* (1995), *Chem. Rev.*, **95**, 1877 (*synth*).

Aminoacids and peptides (VV)

Aminoacids (VV0050–VV0140)

Protein α-aminoacids (VV0050)

The common α-aminoacids are characterised by the structure $RCH(NH_2)COOH$, where R is an aliphatic (including hydrogen), aromatic or heterocyclic group. The exception is **Proline**, strictly an iminoacid, in which the N atom is incorporated into a 5-membered pyrrolidine ring.

They are the primary products of nitrogen anabolism in plants, where they are produced from ammonia (derived *ab initio* by nitrate reduction or nitrogen fixation) by a process called the glutamate synthetase cycle. This produces glutamate which is then transformed into the other aminoacids by a variety of processes.

The aminoacids thus represent the most important nitrogenous component (in terms of volume and accessibility) of the chiral pool produced by living organisms and are of great importance in chiral synthesis.

Several hundred plant aminoacids are known. Of these, 20 only (known as the primary protein aminoacids) are incorporated by all organisms into peptides and proteins (not all of these 20 aminoacids can be biosynthesised by animals). This protein synthesis occurs in the ribosomes by a process involving ribonucleic acid (RNA), the nucleoside chain of which transmits the template instructions of the DNA genetic material to the protein sequences, each primary aminoacid in the chain being coded for by one or more nucleoside base triplets or codons.

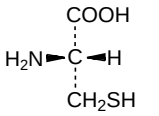
There is an IUPAC-IUB standard 3-letter code for each of the protein aminoacids (as well as for the common nonprotein aminoacids). For ease of computerised documentation of large peptide structures, one-letter codes have more recently been introduced but these are not used in DNP as the full structures of proteins and large peptides are not given in entries.

	IUPAC-IUB abbreviations	
1. Alanine	Ala	A
2. Arginine	Arg	R
3. Asparagine	Asn	N
4. Aspartic acid	Asp	D
5. Cysteine	Cys	C
6. Glutamic acid	Glu	E
7. Glutamine	Gln	Q
8. Glycine	Gly	G
9. Histidine	His	H
10. Isoleucine	Ile	I
11. Leucine	Leu	L
12. Lysine	Lys	K
13. Methionine	Met	M
14. Phenylalanine	Phe	F
15. Proline	Pro	P
16. Serine	Ser	S
17. Threonine	Thr	T
18. Tryptophan	Trp	W
19. Tyrosine	Tyr	Y
20. Valine	Val	V

Aminoacids and their corresponding 3-letter and 1-letter codes

Various posttranslational protein aminoacids known as secondary aminoacids may then arise in the protein by various processes such as conjugation of OH, SH or NH groups, *N*-methylation or hydroxylation (especially to produce **4-Hydroxyproline**). A special case of posttranslational change is the reversible oxidation of cysteine residues to produce the disulfide **Cystine** thus linking different parts of the peptide chain by disulfide bridges as part of the secondary structure of the protein.

With the exception of **Glycine**, all of the genetically coded protein aminoacids are chiral and belong to the L-series. In all cases except Cysteine, this corresponds to (*S*-) according to the Cahn-Ingold-Prelog convention. In Cysteine the higher priority of the —CH_2SH group over the —COOH group means that L- corresponds to (*R*-).



L-(≡*R*-)-Cysteine

Aminoacids of the opposite D-series can be detected in hydrolysates of aged proteins in which they arise by slow racemisation (they are also produced as artifacts of racemisation during acid or especially alkaline hydrolysis of polypeptides). D-Aminoacids are common constituents of antibiotics and bacterial proteins.

Barrett, G.C. (ed.) (1985) *Chemistry and Biochemistry of the Amino Acids*, Chapman & Hall, London.

Coppola, G.M. and Schuster, H.F. (1987) *Asymmetric Synthesis: Construction of Chiral Molecules Using Amino Acids*, Wiley, New York.

Hunt, S. (1991) in *Methods in Plant Biochemistry*, (ed. L.J. Rogers) Volume 5, Academic Press, New York, pp. 1–52.

Williams, R.M. (1989) *Synthesis of Optically Active Aminoacids*, Pergamon, Oxford.

Non-protein α-aminoacids (VV0100)

In addition to the proteinaceous aminoacids, plants produce several hundred known non-protein aminoacids which arise by a variety of metabolic routes. Some of these have demonstrated functions, for example as defence chemicals; the plant aminoacids probably perform a generalised nitrogen storage function.

A considerable number of atypical α-aminoacids have been isolated from microbial sources. They inhibit the growth of a range of microorganisms but their effects can be readily reversed by supplementing the growth medium by the requisite principal aminoacid.

Atypical aminoacids are encountered in the hydrolysates of microbial peptide antibiotics. These do not always occur in the free state but a number have been included in DNP since a given aminoacid may be present in a range of different peptides.

Hatanaka, S.I. *et al.* (1992) *Prog. Chem. Org. Nat. Prod.*, **59**, 1 (aminoacids from fungi).

Scannell, J.P. *et al.* (1974) in *Chemistry and Biochemistry of Amino Acids Peptides and Proteins*, Dekker, New York (*antimetabolites*).

β-Aminoacids (VV0120)

A number of β-aminoacids occur naturally, especially in peptide hormones and antibiotics. Of these the most widespread is **β-Alanine**.

Drey, C.N.C. (1985) in *Chemistry and Biochemistry of the Amino Acids*, (ed. G.C. Barrett) Chapman & Hall.

Peptides (VV0150–VV0600)

Peptides are oligomers and polymers notionally derived from aminoacids by condensation to produce amide linkages. The boundary between oligopeptides and polypeptides is arbitrary and in DNP has been set at 10 aminoacid residues. The configuration of aminoacid residues in polypeptides is assumed to be L- when not indicated otherwise.

There is evidence that in higher organisms small peptides (hormones) can arise only by cleavage of protein prohormones.

A large number of biologically-active atypical peptides have been isolated from bacteria, actinomycetes and fungi. Structurally they represent an extremely diverse group, encompassing those metabolites containing two or more aminoacid residues linked by a peptide bond, but possessing some additional features not characteristic of proteins. These may include unusual aminoacid residues, protein aminoacids with the D-configuration or raised to a higher oxidation level, or non-peptide linkages between residues (e.g. ester, lactone or a γ-glutamyl amide). In addition the molecules may be linear or cyclic, contain one or a combination of the above mentioned features, be modified by further interactions between the side chains of amino-acid units within the peptide, or conjugated with either lipids or sugar units.

Diketopiperazines (dipeptide anhydrides) (VV0150)

These are among the most numerous of all naturally occurring peptides. They range from simple cyclic dipeptides to highly complex fused ring systems such as the antiviral **Bicyclomycin** and the toxic 1,4-sulfur bridged **Sporidesmins** and related compounds. The ergot peptides (listed in the alkaloid section) can also be regarded as derivatives of cyclic dipeptides.

Nomenclature of the simple diketopiperazines is complicated by the proliferation of different ways of naming them. In DNP, systematic *Chemical Abstracts* names are used as their entry names, but the entries contain a full range of possible synonyms.

Cyclic oligo-and polypeptides (VV0500)

No cyclic homodetic tripeptides with or without biological activity have been observed to date. Cyclic peptides derived from 4–11 aminoacid residues linked by peptide bonds have been isolated from a variety of microorganisms. Their biological properties are diverse, ranging from antitumour activity for some cyclic tetrapeptides, through to iron complexation for some hexapeptides, the antibacterial properties of the **Gramicidin** and **Tyrocidin** decapeptides, and the immunosuppressant activity of the undecapeptides of the **Cyclosporin** family.

Depsipeptides (VV0600)

Cyclic heterodetic peptides or peptide lactones are those in which one or more of the peptide bonds have been replaced by ester linkages. **Valinomycin** and related antibiotics, though of no clinical value, are important biochemical tools in that they specifically complex with alkali metal ions. The **Actinomycin** family possess two peptide lactones attached to a common phenoxazine system and form stable complexes with DNA by intercalation; they are used clinically in the treatment of child leukaemia.

Large peptides and proteins (VV1000, VV2000)

Entries are given in DNP for the majority of bioactive peptides secreted by plants and animals for which reasonable structural information exists, including many insect neuropeptides which are an active field of research. Entries are presented for the most important non-enzyme proteins and for some enzymes, but full structures are not given in individual entries, the structures where known can be assessed *via* the cited references.

Large modified peptides

This is rather an arbitrary group including all those peptide antibiotics with a M_r greater than 1000. The development of sophisticated spectroscopic and analytical methodology over the past decade has led to the isolation and structural identification of a wide variety of highly modified peptides. The peptaibol group of linear peptides exemplified by **Alamethicin** are characterised by the presence of a large number of α-aminobutyric acid (aib) residues. These antibiotics, which form ion channels in biological and artificial membranes, are important biophysical tools. **Thiostrepton** and related antibiotics contain a central pyridine or reduced pyridine entity of unknown origin, together with a substantial number of cysteine derived thiazole units. The glycopeptides of the **Bleomycin** family which also possess similar thiazole units display remarkable antitumour activity. These molecules not only intercalate within double stranded DNA but also have an iron binding site for carrying singlet oxygen. This is positioned so as to effect oxidative cleavage of one of the DNA chains. Several of the Bleomycins are used clinically. Semi-synthesis has been employed to produce analogues for structure activity studies.

With such a diverse structural group it is impossible to provide an overview of the biosynthesis, but so far, for the majority of the larger bacterial peptide antibiotics investigated, such as the **Gramicidins**, **Bacitracins** and **Polymyxins**, it is evident that they are not synthesised on ribosomes but *via* the so-called multienzyme thiotemplate.

Bladon, C. (1993) in *The Chemistry of Natural Products*, 2nd edn (ed. R.H. Thomson) Blackie, Glasgow, p. 183 (*rev*).
Fusetani, N. *et al.* (1993), *Chem. Rev.*, **93**, 1793–1805 (*sponge peptides*)
Gross, E. (ed.) (1983) *The Peptides*, Academic Press, New York (*general*).
Lipmann, F. (1980) *Adv. Microbiol. Physiol.*, **21**, 227 (*biosynth*).
Sammes, P.G. (1975) *Prog. Chem. Org. Nat. Prod.*, **32**, 51 (*cyclodipeptides*).

β-Lactams

Penicillins and cephalosporins (VV0700, VV0800)

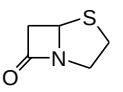
These are by far the most important group of the β-lactam antibiotics.

The naturally-occurring penicillins are a closely related group of antibacterial agents produced predominantly by fungi. They possess a common β-lactam-thiazolidine fused system. The *N*-acyl side-chain, in which variation can occur, is limited to a small number of aliphatic and aromatic groups.

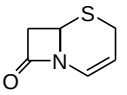
The naturally-occurring cephalosporins which are produced predominantly by *Acremonium/Cephalosporium* and *Streptomyces* spp. possess a common β-lactam-dihydrothiazine fused system, but in this case the side-chain is limited to an α-aminoadipoyl group.

All β-lactams act by inhibiting bacterial cell wall biosynthesis. They are, in varying degrees, susceptible to the inactivating β-lactam enzymes present in many pathogens. The penicillins and cephalosporins are biosynthetically related

to, and derived from, a common tripeptide precursor. The other groups appear to be produced by alternative pathways involving either peptide or aminoacid intermediates.



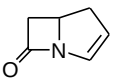
β-Lactam-thiazoline



β-Lactam-dihydrothiazine

Carbapenems (VV0900)

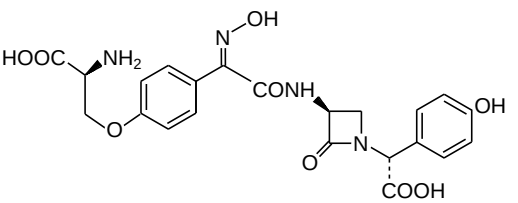
These are now a fairly substantial group of naturally-occurring bicyclic *β*-lactams. In terms of chemical stability they are highly sensitive compounds but nevertheless exhibit potent broad spectrum antibacterial activity. Due to the low titre and difficulties with isolation from microbial sources the most promising clinical candidate, Imipenem, is currently produced by total chemical synthesis.



Carbapenem nucleus

Monocyclic β-lactams (VV0920)

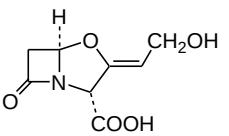
Monocyclic *β*-lactams, such as the monobactams and nocardicins, are bacterial products with limited antibacterial activity. However notable synthetics based on the natural system (e.g. Aztreonam) are potent antibiotics against gram negative organisms.



Nocardicin A

Clavams (VV0950)

The so-called clavams are predominantly produced by *Streptomyces* spp. The most important member of this group, Clavulanic acid, although exhibiting limited antibacterial properties is a potent *β*-lactamase inhibitor and is used clinically in combination with semisynthetic penicillins. The other naturally-occurring clavams, which have the opposite chirality at the ring junction, lack antibacterial properties but demonstrate some antifungal activity.



Clavulanic acid

Allen, J.D. *et al.* (1986) *Adv. Intern. Med.*, **31**, 119 (*β-lactams*).

Baldwin, J.E. (ed.) (1983) Tetrahedron Symposium in Print No. 10, 39 (*penicillins*).

Barrett, G.C. (ed.) (1985) *Chemistry and Biochemistry of the Amino Acids*, Chapman & Hall, London.

Bentley, P.H. *et al.* (eds), (1992) *Recent Adv. in the Chem. of β-Lactam Antibiotics, Proc. 4th Int. Symp.*, RSC, London.

Brennan, J. (1986) in *Amino Acids Peptides and Proteins*, specialist periodical reports, RSC, London, **17**, 171.

Brown, A.G. *et al.* (eds) (1985) *Recent Advances in Chemistry of β-lactam antibiotics*, RSC Special Publication no. 52.
Bycroft, B.W. *et al.* (1987) *Biotechnology Handbook*, Vol. 1, (ed. J. Peberdy) Plenum Press, New York, pp. 113.
Casy, A.F. *et al.* (1989) *J. Pharm. Biomed. Anal.*, **7**, 1121 (*ms*).
Demain, A.L. *et al.* (ed.) (1983) Antibiotics containing the β-lactam structure, *Handbook of Experimental Pharmacology*, p. 67.
Frydrych, C.H. (1991) *Amino Acids Pept.*, **22**, 294; (1992) **23**, 249.
Jensen, S.E. (1986) *CRC Crit. Rev. Biotechnol.*, CRC Press, Boca Raton, **3**, 277 (*biosynth*).
Kleinkauf, H. *et al.* (eds) (1990) Biochem. of Pept. Antibiot. Recent Adv. in *Biotechnol. of β-Lactams and Microbial Bioactive Pept.*, De Gruyter, Berlin.
Morin, R.B. *et al.* (1982) *Chemistry and Biology of β-lactam Antibiotics*, Vols 1–3, Academic Press, New York (*general*).
Ono, H. *et al.* (1990) *Biochem. Pept. Antibiot.*, 131.
O'Sullivan, J. *et al.* (1986) in *Biotechnology*, Vol 4, (ed. H. Page) VCH, Weinheim, Ger., p. 247.
Parker, W. *et al.* (1986) *Adv. Appl. Microbiol.*, **31**, 181 (*monobactams*).
Robinson, J.A. *et al.* (1985) *Nat. Prod. Rep.*, **2**, 293 (*biosynth*).
Rolinson, G.N. (1986) *J. Antimicrob. Chemother.*, **17**, 5.
Salton, M.R.J. *et al.* (eds) (1981) *β-Lactam antibiotics: Mode of Action, new developments and future prospects*, Academic Press, New York.
Stachulski, A.V. (1989) *Amino Acids Pept.*, **20**, 249; (1990) **21**, 248.
Walsh, T.F. (1988) *Annu. Rep. Med. Chem.*, **23**, 121.
Williamson, J.M. (1986) *CRC Crit. Rev. Biotechnol.*, **4**, 111 (*biosynth*).

Glycopeptides (VV3000)

The members of the **Vancomycin** family are the most significant within this relatively small and structurally self-evident category of antibiotics. Their activity is restricted to gram-positive organisms but they are particularly effective against the so-called multiresistant streptococcal and staphylococcal strains and for this reason have found significant clinical application.

Egge, H. *et al.* (1987) *Mass Spectrom. Rev.*, **6**, 331 (*ms*).
Lancini, G. *et al.* (1990) *Biochem. Pept. Antibiot.*, 159 (*Vancomycins*).
Williams, D.H. (1996) *Nat. Prod. Rep.*, **13**, 469 (rev).

Alkaloids (VX)

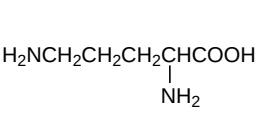
Alkaloids are a large group of nitrogen-containing secondary metabolites of plant, microbial or animal origin. The term originally implied pharmacologically active bases of plant origin, but the definition has subsequently been broadened so that it is now generally considered to include the majority of nitrogen-containing natural products with the exception of the simple aminoacids, proteins and nitrogen-containing substances of polyketide origin such as the aminoglycoside antibiotics. Basic properties may be weak or absent as in the various types of amide alkaloids. The class of microbial alkaloid overlaps considerably with that of the nitrogenous antibiotics, and substances such as the cytochalasans which show antibiotic properties are in DNP classified as alkaloidal, the definition being a matter of semantics.

Biogenetically and structurally the alkaloids are diverse and it is usual to discuss them in terms of biogenetic origin rather than purely on the basis of structural features. The organisation of alkaloid groups within the Type of Compound Index follows the order given below.

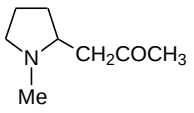
Alkaloids derived from ornithine

Simple ornithine alkaloids (VX0300)

Several simple alkaloids derived possibly from ornithine are known. These include Hygrine and **Stachydrine**. Condensation of two ornithine units with acetoacetate gives **Cuscohygrine**. Other alkaloids containing a pyrrolidine ring include **Nicotine, Ficine** (in which the pyrrolidine ring is attached to a flavone nucleus), **Macrostomine** (in which it is attached to a benzylisoquinoline skeleton), and **Brevicolline** (in which it is attached to a β-carboline unit). Clearly the biogenesis of these molecules requires other precursors. Macrostomine is presumably derived from tyrosine, just as Brevicolline has been shown to be derived from tryptophan.



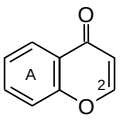
Ornithine



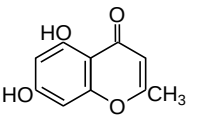
Hygrine

Chromone alkaloids (VX0340, VX0350)

A structure consisting of a pyrrolidine, piperidine or pyridine ring linked to the A ring of chromone is referred to as a chromone alkaloid. This group of compounds can be sub-divided into two types, namely those in which the chromone nucleus exists as Noreugenin (5,7-Dihydroxy-2-methylchromone) – indexed in DNP as chromone alkaloids – and those which bear a phenyl substituent at C-2 (indexed as flavonoid alkaloids). The former group is typified by **Rohitukine** and **Schumannificine**. Typical flavonoid alkaloids include **Ficine** and **Vochysine**. Compared with the noreugenin-related alkaloids, which have only been isolated from the plant families Meliaceae and Rubiaceae, the flavonoid alkaloids are more widely distributed throughout the higher plants.



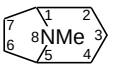
Chromone



Noreugenin

Tropane alkaloids (VX0400)

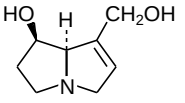
These are derived from ornithine and acetoacetate. Almost all of them are esters of mono-, di-, and tri-hydroxytropanes. They are characteristic constituents of the Solanaceae. **Atropine** and **Cocaine** are important representatives. Recent evidence suggests that for some alkaloids (Cocaine and its close relatives) malonylcoenzyme A is involved in the biosynthesis, rather than acetoacetylcoenzyme A.



Tropane
8-Methyl-8-azabicyclo[3.2.1]octane, 9CI

Pyrrolizidine alkaloids (VX0440, VX0500, VX0520, VX0540)

These occur in species of the *Senecio* genus, and elsewhere in the Compositae and Leguminosae. They have been shown to be responsible for the toxic effects, particularly liver damage, in livestock grazing on pastures infested by these species. Toxicity appears to be the result of oxidation *in vivo* to pyrrole derivatives. The majority of pyrrolizidine alkaloids are either relatively simple esters formed from a pyrrolizidine base, the necine, exemplified by Retronecine, or more complex cyclic esters formed between a necine and a necic acid (VX0500), an example being **Monocrotaline**. The necic acid units in this latter and other diester alkaloids are themselves probably derived from an aminoacid (e.g. isoleucine), rather than acetate or mevalonate.

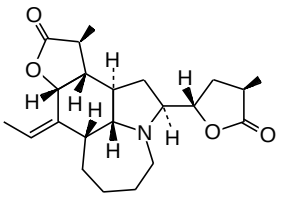


Retronecine

Many, and perhaps the majority, of pyrrolizidine alkaloids occur in the plant as *N*-oxides, the *N*-oxide function being lost during isolation.

Miscellaneous ornithine-derived alkaloids (VX1160, VX1260)

Two other small groups of alkaloids may be derived from ornithine, but this remains to be proved. The *Stemona* alkaloids, e.g. Tuberostemonine, contain a pyrrolidine ring, possibly originating from ornithine, but the structure reveals that its biosynthesis is complex. The *Elaeocarpus* alkaloids, e.g. **Elaeocarpine**, may also originate from ornithine, together with a polyketide unit. Alternatively, the whole skeleton may be polyacetate-derived. In the case of Elaeocarpidine, tryptamine is also obviously implicated.



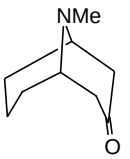
Tuberostemonine

Gellert, E. (1982) *Indolizidine Alkaloids, J. Nat. Prod.*, **45**, 50.
Gellert, E. (1987) The Phenanthroindolizidine Alkaloids, in *Alkaloids: Chemical and Biological Perspectives* (ed. S.W. Pelletier) Volume 5, Wiley-Interscience, New York.
Herbert, R.B. (1985) The Synthesis of Indolizidine and Quinolizidine Alkaloids of Tylophora, Cryptocarya, Ipomoea, Elaeocarpus, and related species, in *Alkaloids: Chemical and Biological Perspectives*, (ed. S.W. Pelletier) Volume 3, Wiley-Interscience, New York.

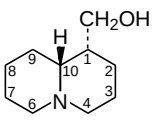
Alkaloids derived from lysine

Simple piperidine alkaloids (VX0620, VX0680)

These may be derived from lysine, acetate, acetoacetate, etc., in analogous fashion to the simple pyrrolidine alkaloids. They include Pseudopelletierine, **Anabasine** and Lupinine.

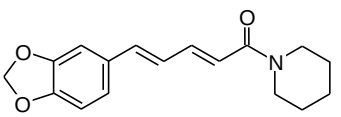


Pseudopelletierine



Lupinine

All these structural types have their analogues among the pyrrolidine alkaloids, and while it is tempting to assume biosynthesis from lysine it may not in all cases be true; **Coniine**, for example, appears to be acetate-derived.



Piperine

Relatively simple derivatives of piperidine include the alkaloids of black pepper (*Piper nigrum*), e.g. Piperine.

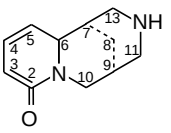
Lobelia alkaloids (VX0660)

These have no analogy among the pyrrolidine alkaloids. An example is **Lobelanine**.

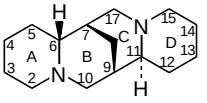
More complex lysine-derived alkaloids (VX0900, VX0920, VX0940, VX0960, VX0970, VX0980)

These include Cryptopleurine, the tri- and tetracyclic alkaloids of *Cytisus* and other species of Leguminosae, e.g. Cytisine, the poisonous principle of the

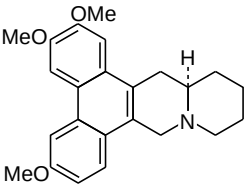
laburnum, and the closely-related sparteine group (note tricky stereochemistry owing to twofold rotation-reflection axis). Penta- and hexa-cyclic bases are found in *Ormosia* species; of these **Ormosanine** is representative.



(+)-Cytisine



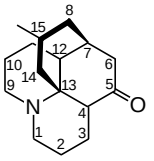
(–)-Sparteine



Cryptopleurine

Lycopodium alkaloids (VX1280)

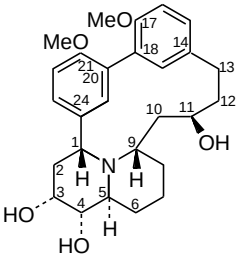
These, such as Lycopodine, are constituents of the club mosses. Whereas the earliest proposal concerning their biosynthesis implicated two C_8 units derived from acetate, it has more recently been established that two lysine units are involved. Numerous skeletal variants are known, all of which can be related to the Lycopodine skeleton; examples are **Fawcettidine** and **Serratinine**.



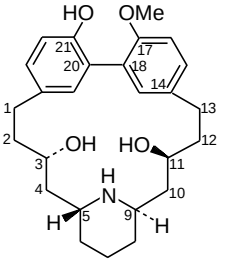
Lycopodine

Lythraceae alkaloids (VX0760)

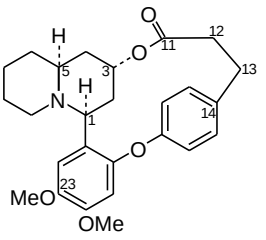
These are characterised by several unusual structural features. Lythrancine contains a quinolizidine ring system attached to a diphenyl residue, one of the benzene rings in which is derived from cinnamic acid. Other alkaloids in this group contain a diphenyl ether grouping, e.g. Decaline, and others a piperidine ring instead of a quinolizidine ring, e.g. Lythranidine. The biosynthesis of these alkaloids involves lysine as source of the quinolizidine or piperidine ring, and phenylalanine as precursor of one of the aromatic rings. The numbering system adopted throughout DNP for lactonic Lythraceae alkaloids (e.g. Decaline) is the one generally accepted. This was introduced by Horsewood *et al* (*Can. J. Chem.*, 1979, **57**, 1615) and corresponds closely to that introduced by Fujita *et al* for piperidine and quinolizidine metacyclophane alkaloids (e.g. Lythranidine, Lythrancine). The carbon atoms that correspond in biogenetic origin to the three types thus maintain corresponding numbers. Note that CA numbering is different.



Lythrancine



Lythranidine



Decaline

Blumenkopf, T.A. and Heathcock, C.H. (1985) Synthesis of *Lycopodium* Alkaloids, in *Alkaloids: Chemical and Biological Perspectives* (ed. S.W. Pelletier, Vol. 3, Wiley-Interscience, New York.

Elbein, A.E. and Molyneux, R.J. (1987) The Chemistry and Biochemistry of Recently Isolated Indolizidine Alkaloids, in Pelletier, Vol. 2.

Fodor, G. and Colasanti, B. (1985) The Pyridine and Piperidine Alkaloids: Chemistry and Pharmacology, in Pelletier, Vol. 3.

Fujita, E. and Fuji, K. (1976) The Lythraceous Alkaloids, in *MTP Series 2*, Vol. 9, *Alkaloids*, (ed. K. Wiesner), Butterworths, London.

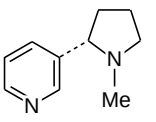
Herbert, R.B. (1985) The Synthesis of Indolizidine and Quinolizidine Alkaloids of *Tylophora*, *Cryptocarya*, *Ipomoea*, *Elaeocarpus*, and Related Species, in Pelletier, Vol. 3.

Kinghorn, A.D. and Balandrin, M.F. (1984) Quinolizidine Alkaloids of the Leguminosae: Structural Types, Analyses, Chemotaxonomy, and Biological Properties, in Pelletier, Vol. 5.

Valenta, Z. and Liu, H.J. (1976) The *Ormosia* Alkaloids, *MTP Series 2*, Vol. 9, *Alkaloids*.

Alkaloids derived from nicotinic acid (VX1020)

This is a relatively small group. Nicotine and **Anabasine** are presumably derived from nicotinic acid and ornithine or lysine, respectively. However, the piperidine ring in **Anatabine**, from *Nicotiana glutinosa*, appears not to be derived from lysine or from a polyacetate precursor; instead, both rings are derived from nicotinic acid. **Arecoline**, from betel nuts, and **Ricinine**, from the castor oil plant, are clearly derivable from nicotinic acid; in the case of Ricinine this has been established. **Dioscorine**, from *Dioscorea hispida*, provides a fascinating example of the unexpected in alkaloid biosynthesis. At first sight it seems plausible to postulate that it may be formed from lysine and a polyketide fragment. However, lysine is not a precursor, and it would appear that Dioscorine is formed from nicotinic acid and, probably, a polyacetate unit.

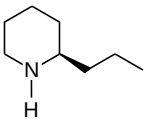


Nicotine

Fodor, G. and Colasanti, B. (1985) The Pyridine and Piperidine Alkaloids: Chemistry and Pharmacology, in *Alkaloids: Chemical and Biological Perspectives*, (ed. S.W. Pelletier) Vol. 3, Wiley-Interscience, New York.
Leete, E. (1983) Biosynthesis and Metabolism of the Tobacco Alkaloids, in Pelletier, Vol. 1.

Alkaloids of polyketide origin (VX0680, VX0700, VX1120, VX1240)

Numerous alkaloids are derived from polyacetate precursors, together with one or more amino acids. A few alkaloids, however, are almost entirely acetate-derived. These include Coniine, from hemlock; this, perhaps surprisingly, is not derived from lysine. Similarly, **Pinidine**, from *Pinus sabiniana*, is acetate-derived. Other piperidine derivatives with side-chains at position 2, which may be acetate-derived, although proof is at present lacking, include **Nigrifactine** from *Streptomyces* species, **Carpaine** and **Cassine**.



Coniine

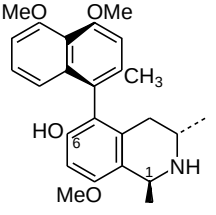
More complex examples include **Coccinellin**, the defensive agent of the common ladybird, *Coccinella septempunctata*, **Porantherine**, from the shrub *Poranthera corymbosa*, and the alkaloids of the **Ancistrocladine** group.

Still more complex are the *Galbulimima* alkaloids, e.g. **Himbacine**, which may be formed from a nonaketide unit plus acetoacetate.

Naphthalene-isoquinoline alkaloids (VX1140)

This group comprises some forty alkaloids which have been isolated mainly from the plant family Ancistrocladaceae, with some isolations from the Dionchophyllaceae. Several skeletal types are known and are based on the point of linkage between the two ring systems, e.g. 5,1′-coupled alkaloids (Ancistrocladine, **Dionchophylline C**), 5,8′-coupled (**Korupensine A**), 7,1′-coupled (**Ancistrocladisine, Dionchophylline A**), 7,2′-coupled (**Ancistrocladidine**), 7,6′-coupled (**Dionchophylline B**), etc. The recently isolated **Michellamines** are the first ‘dimeric’ alkaloids of this class to be discovered.

These alkaloids are chiral not only due to diastereoisomerism at the methyl groups but also in the biaryl linkage due to restricted rotation.

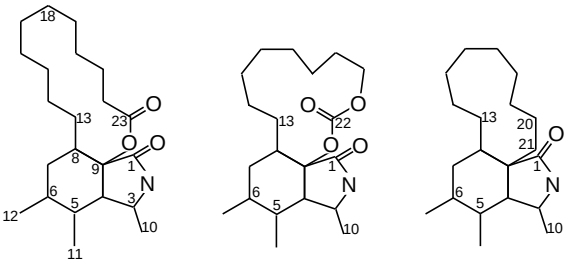


Ancistrocladine

Cytochalasan alkaloids (VX1300)

Cytochalasins are metabolites of several different and unrelated fungi. They are characterised structurally by the presence of a perhydroisoindolone system fused

to a macrocyclic ring of 11, 13 or 14 atoms. The macrocycle may be a carbocycle, a lactone or a carbonate. In addition the isoindole ring carries either a phenyl or an indolyl substituent at position 10; the latter group includes the Chaetoglobosins.



Biosynthetically, cytochalasins arise from phenylalanine or tryptophan and a polyketide derived from acetate and methionine. Cytochalasins possess a range of distinctive biological properties. These include inhibition of cytoplasmic cleavage leading to polynucleate cells, nuclear extrusion and the inhibition of cell mobility.

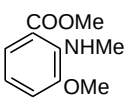
Binder, M. *et al.* (1973) *Angew. Chem. Internat. Ed. Engl.*, **12**, 370.
Cole, R.J. (1981) *Toxic Fungal Metabolites*, Academic Press, New York.
Dyke, H. *et al.* (1986) *J. Chem. Soc. Chem. Commun.*, 1447 (*synth*).
Fodor, G. and Colasanti, V. (1985) The Pyridine and Piperidine Alkaloids: Chemistry and Pharmacology, in *Alkaloids: Chemical and Biological Perspectives*, (ed. S.W. Pelletier), Vol. 3, Wiley-Interscience, New York.
Jones, R.C.F. (1984) *Nat. Prod. Rep.*, **1**, 97.
Jones, T.H. and Blum, M.S. (1983) Arthropod Alkaloids: Distribution, Functions, and Chemistry, in Pelletier, Vol. 1.
Pendse, G.S. (ed.) (1987) *Recent Advances in Cytochalasans*, Chapman & Hall, London.
Turner, W.B. (1983) *Fungal Metabolites II*, Academic Press, New York.

Alkaloids derived from anthranilic acid

A number of diverse structural groups belong in this category, the major ones being the quinolone, furanoquinoline, pyranoquinoline, acridine and quinazoline groups.

Simple anthranilic acid derivatives (VX1460)

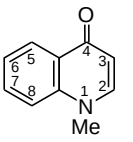
This small group includes Damascenine from *Nigella damascena*. Biosynthetically the alkaloid is derived from anthranilic acid (chorismate-derived), which is then hydroxylated and methylated.



Damascenine

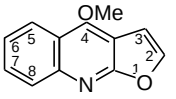
Simple quinoline alkaloids (VX1480, VX1520, VX1540, VX1560, VX1580)

These include Echinopsine and the phenethyl-quinoline, **Cusparine**.



Echinopsine

In simple quinolone derivatives C-2 and C-3 are derived from acetate; introduction of a prenyl group at C-3 followed by cyclisation then gives a furanoquinoline, e.g. **Platydesmine**, or a pyranoquinoline alkaloid, e.g. **Flindersine**. Interestingly, Dictamnine, from *Dictamnus albus*, appears to be formed by loss of acetone from an oxidation product of Platydesmine.

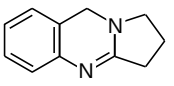


Dictamnine

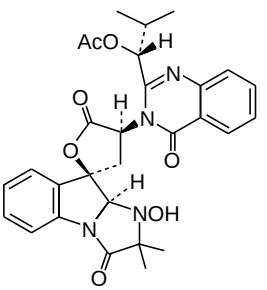
Quinazoline alkaloids (VX1600)

These include Vasicine, from *Adhatoda vasica*, and **Rutaecarpine**, from *Evodia rutaecarpa*; this latter base is clearly also derived from tryptophan. The Tryptoquivalines, which are toxic metabolites from *Aspergillus clavatus*, are also derived from anthranilic acid and tryptophan precursors, together with (presumably) valine and methylalanine-derived units.

Macrorine, from *Macrorungia longistrobus*, is obviously derived from anthranilic acid and histidine.



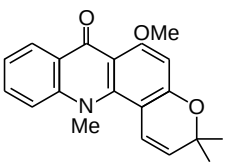
Vasicine



Tryptoquivaline

Acridone alkaloids (VX1620)

These can be broadly divided into two main subgroups. The simple acridones may be exemplified by **Melicopidine** and **Melicopicine**, and the prenylacridones in which a prenyl group introduced into the acridone nucleus is cyclised to give a pyran ring, by Acronycine.



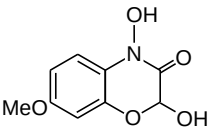
Acronycine

Acridone-coumarin alkaloid dimers (VX1690)

The isolation in 1990 of the **Acrimarines** represented the first examples of acridone-coumarin dimers from natural sources. To date, some twenty compounds have been reported from *Citrus* plant (Rutaceae) roots.

1,4-Benzoxazin-3-one alkaloids (VX1720)

2,4-Dihydroxy-7-methoxy-2*H*-1,4-benzoxazin-3(4H)-one (known as DIMBOA) is representative of this significant, yet often neglected group of alkaloids which have plant hormone significance. They usually occur in the plant as glucosides but cell injury apparently releases a glucosidase which catalyses hydrolysis to the 2-hydroxy derivatives.



DIMBOA

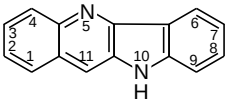
Benzodiazepine alkaloids (VX1760)

Cyclopenin and **Viridicatin**, from *Penicillium cyclopium* and *P. viridicatum*, are clearly derived from anthranilic acid and phenylalanine. Apparently no plant alkaloids with this skeleton are known to date.

Saxton, J.E. (1973) in *The Acridines*, 2nd edn (ed. R.M. Acheson), Wiley-Interscience, New York.

Cryptolepine-type alkaloids (VX1800)

Cryptolepine (5-Methyl-5*H*-quindoline) is the most prominent representative of this small but growing group of about ten related natural products isolated from *Cryptolepis* spp. The parent compound, 10*H*-Indolo[3,2-*b*]quinoline (Quindoline), has itself been isolated from *C. sanguinolenta*. More recently a spiro-alkaloid (**Cryptospirolepine**), several cryptolepine isomers (e.g. **Isocryptolepine, Neocryptolepine, Cryptosanguinolentine, Cryptotackieine), Hydroxycryptolepine, Cryptoheptine** and three dimers (**Biscryptolepine, Cryptomisrine and Cryptoquindoline**) were isolated, although the latter alkaloid was shown to be an artifact. No biosynthetic studies of this skeleton have been reported but a derivation from indole and *N*-methylanthranilic acid could be imagined for cryptolepine.



Quindoline

Alkaloids derived wholly or in part from phenylalanine or tyrosine

This extremely large and varied category consists of several different structural types, which range from simple *β*-phenylethylamine derivatives to the much more complex structures exemplified by the alkaloids of the Amaryllidaceae and the bisbenzylisoquinoline alkaloids. The isoquinoline derivatives themselves consist of a large number of structural types, which can be divided into upwards of 20 sub-groups.

Simple tyramine alkaloids (VX2000)

The simplest derivatives of phenylalanine or tyrosine are the *β*-phenylethylamines, which are presumably obtained by decarboxylation and obvious oxidative/alkylation stages.

The alkaloids in this group can be divided into four categories:

(*a*) those with a simple *β*-arylethylamine structure, e.g. **Mescaline**;

(*b*) those with the structural unit $\mathrm{ArCHCH_2N<}$ (OR), e.g. **Macromerine**;

(*c*) the Ephedra bases, e.g. **Ephedrine**;

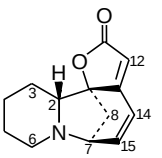
(*d*) miscellaneous alkaloids, e.g. **Aurantiamide**.

Cinnamic acid amides (VX2020)

Fagaramide, Herclavine and **Subaphylline** are simple members of this series. These amides are all derived from a trans-cinnamic acid; **Astrophylline** is an example of an *N*-cis-cinnamoyl derivative.

Securinega alkaloids (VX2100)

This small group of about 30 alkaloids occurs in the genus *Securinega* (fam. Euphorbiaceae). The biosynthesis of these alkaloids, which is not readily apparent, has been shown to involve tyrosine and lysine, in the case of Securinine. Presumably **Norsecurinine** is derived from tyrosine and ornithine.



Securinine

Phyllanthidine is presumably an oxidative transformation product of **Allosecurinine**, the diastereoisomer of Securinine, which also occurs naturally.

Betalain alkaloids (VX2140)

This is a group of some 50 alkaloidal pigments whose distribution is limited to the order Centrospermae. As a group they are zwitterionic and water soluble. **Betanidin** and **Indicaxanthin** are typical examples.

The Isoquinoline alkaloids

This extremely large and enormously varied group can be divided into approximately twenty categories.

Krane, B.D. *et al.* (1982) *J. Nat. Prod.*, **45**, 377.
Menachery, M.D. *et al.* (1986) *J. Nat. Prod.*, **49**, 745.
Phillipson, J.D., Roberts, M.F. and Zenk, M.H. (eds) (1985) *The Chemistry and Biology of Isoquinoline Alkaloids*, Springer Verlag, Berlin.
Shamma, M. (1972) *The Isoquinoline Alkaloids*, Academic Press, New York.
Shamma, M. and Moniot, J.L. (1978) *Isoquinoline Alkaloid Research*, 1972–1977, Plenum Press, New York.

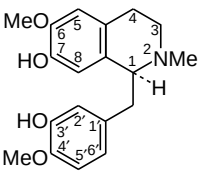
Simple isoquinoline alkaloids (VX2200, VX2300)

These can be further sub-divided into (*a*) those not bearing a carbon substituent at C-1, and which are basic, e.g. **Anhalamine**; (*b*) those with an amide carbonyl group at C-1, and therefore non-basic, e.g. **Corydaldine**; (*c*) those with a methyl group at C-1, e.g. **Salsoline**; (*d*) those with a 1,2,3,4-tetrahydroisoquinolinequinone moiety, e.g. **Mimosamycin, Renierone**, and the more complex group of **Saframycin-** and **Renieramycin-type** antibiotics; and (*e*) miscellaneous bases.

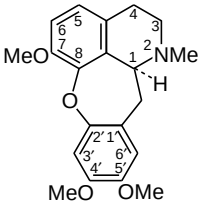
Benzylisoquinoline alkaloids (VX2320)

The simple benzylisoquinoline skeleton is derived from two molecules of tyrosine, and is the parent skeleton of a wide variety of alkaloids belonging to numerous different ring systems.

The alkaloids with the unmodified benzylisoquinoline skeleton may be divided into five subgroups; (*a*) 1,2,3,4-tetrahydrobenzylisoquinolines, e.g. Reticuline – of central importance in the elaboration of other alkaloids; (*b*) alkaloids in which all the rings are aromatic, e.g. **Papaverine**; (*c*) the cularines (VX2440), which contain an oxepine ring between C-8 and C-2′; (*d*) alkaloids with a carbon substituent at C-2′, such as **Canadaline**. These may be regarded as ring-opened berberines; (*e*) benzylisoquinolines in which a pyrrolidine ring is attached to C-4, e.g. **Macrostomine**.



Reticuline

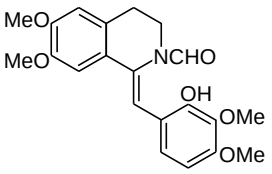


Cularine

Gözler, B. and Shamma, M. (1984) *J. Nat. Prod.*, **47**, 753.

Pseudobenzylisoquinoline alkaloids (VX2330)

The term pseudobenzylisoquinoline alkaloid is used to describe a benzylisoquinoline skeleton in which the pendant aromatic ring is oxygenated at C-2′, C-3′ and C-4′. These alkaloids are derived biogenetically from protoberberinium salts by C8-C8*a* bond scission. Polycarpine, **Taxilamine** and **Ledecorine** are typical examples.



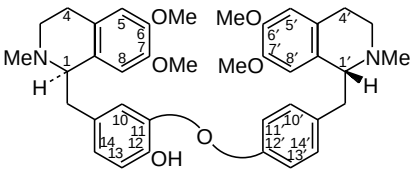
Polycarpine

Bisbenzylisoquinoline alkaloids (VX2340–VX2430)

This very large group of alkaloids is composed of two benzylisoquinoline units attached to each other by one, two, or three bonds. In most cases the units are joined *via* ether linkages, but carbon-carbon bonds between the benzyl groups are also known. The monomeric units involved are mainly hydroxylated or methoxylated benzylisoquinolines, but aporphine units occur in more than 50 alkaloids, and a few alkaloids contain a proaporphine component. The alkaloids may be subdivided into the five following major groups (the classification, proposed by Shamma, contains at least 28 subgroups, all of which are composed of unmodified benzylisoquinoline units. Dimeric alkaloids containing aporphine, proaporphine, or otherwise modified benzylisoquinoline components are not included here).

(*a*) **Alkaloids containing aryl links only**. (VX2340) The bark of *Popowia pisocarpa* has yielded a group of seven alkaloids which contain a single aromatic linkage between C-11 and C-11′. These include **Pisopowetine** and **Pisopowiaridine**

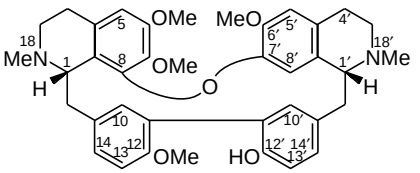
(*b*) **Alkaloids containing one ether link**. (VX2360) This is the most common structural type, the ether linkage being in most cases between C-11 and C-12′, as in Dauricine.



Dauricine

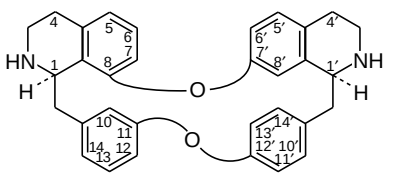
However, other attachments are known, e.g. between C-11 and C-10′, as in **Vanuatine**, between C-10 and C-7′, as in **Malekulatine** and **Ambrinine**, and between C-11 and C-7′, as in **Neferine**.

(*c*) **Alkaloids containing one aromatic link and one or two ether links**. (VX2370, VX2390) These alkaloids are mainly based on the Rodiasine and 6′,7-didemethoxy-6′,7-epoxyrodiasine skeleton, e.g. **Tiliacorine**.



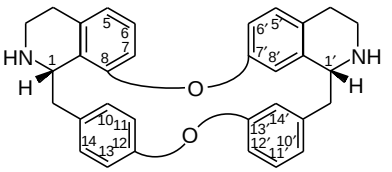
Rodiasine, 9CI

(*d*) **Alkaloids containing two ether links**. (VX2380) The largest single subgroup containing two ether linkages possesses the berbaman skeleton, exemplified by **Berbamine**, which contains ether linkages between C-8 and C-7′, and between C-11 and C-12′.



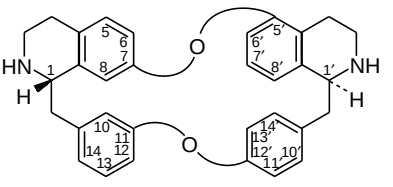
Berbaman, 9CI

Almost as large as the berbaman group is the oxyacanthan group, e.g. **Oxyacanthine**, which contains ether linkages between C-8 and C-7′, and between C-12 and C-13′. Smaller groups include the Thalicberan (C-8 to C-6′ and C-11 to C-12′), Thalidasan (C-8 to C-5′ and C-11 to C-12′), and Thalman (C-7 to C-5′ and C-11 to C-12′) types.

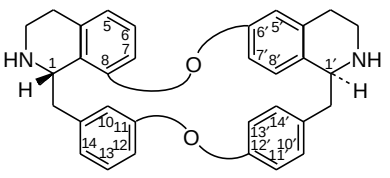


Oxyacanthan, 9CI

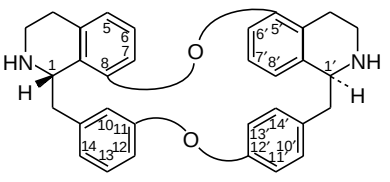
All these types contain ether linkages between the benzyl rings and between the aromatic rings of the tetrahydroisoquinoline component. The Tubocuraran sub-group contains ether linkages between the benzyl ring of one unit and the aromatic ring of the isoquinoline component of the other unit. Other linkages of this kind between benzyl and isoquinoline rings are also known.



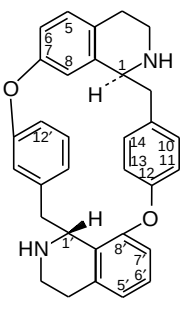
Thalman, 9CI



Thalicberan, 9CI



Thalidasan, 9CI



Tubocuraran, 9CI

(*e*) **Alkaloids with three ether links**. (VX2400) These include alkaloids with 6′,7-epoxyoxyacanthan (e.g. **Trilobine**), 7,8′-epoxyoxyacanthan, and 8,12′-epoxytubocuraran skeletons.

(*f*) **Benzylisoquinoline-Aporphine dimers**. (VX2700) Some alkaloids consist of a benzylisoquinoline unit attached to an aporphine unit, *via* a single ether linkage. Of these, **Thalicarpine** is typical.

(*g*) **Miscellaneous bisbenzylisoquinoline alkaloids** include those containing a dienone ring in one of the isoquinoline components (e.g. **Repanduline**), some with an aporphine unit attached to a pavine component, e.g. **Pennsylpavine**, those with degraded benzylisoquinoline units (e.g. **Baluchistanamine**) or with a proaporphine unit (e.g. **Epiberbivaldine**), and **Cancentrine**, which is really a combination of cularine and morphinan components.

Guha, K.P. *et al.* (1979) *J. Nat. Prod.*, **42**, 1.
Guinaudeau, H. *et al.* (1984) *J. Nat. Prod.*, **47**, 565.
Schiff, P.L. *J. Nat. Prod.*, 1983, **46**, 1; 1987, **50**, 529; 1991, **54**, 645.
Schiff, P.L. (1987) in *Alkaloids: Chemical and Biological Perspectives*, (ed. S.W. Pelletier), Vol. 5, Wiley-Interscience, New York.
Schiff, P.L. (1997) *J. Nat. Prod.*, **60**, 934.
Shamma, M. and Moniot, J.L. (1976) *Heterocycles*, **4**, 1817.

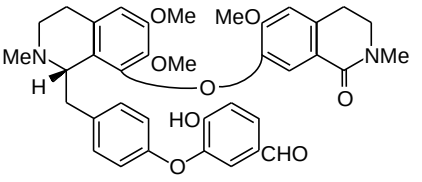
Shamma, M. and Moniot, J.L. (1978) *Isoquinoline Alkaloid Research*, 1972–1977, Plenum Press, New York.

Secobisbenzylisoquinoline alkaloids (VX2430)

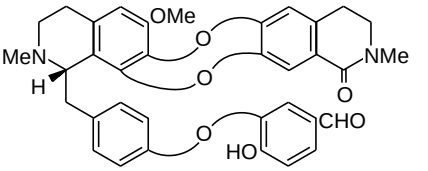
The secobisbenzylisoquinolines are alkaloids in which one of the benzylisoquinoline units is cleaved between the C-1 and the α-carbon atom. A typical example is Baluchistanamine, the apparent biogenetic precursor of which appears to be Oxyacanthine.

With regard to the oxidation state of the C-1 and C_α-atoms, secobisbenzylisoquinoline alkaloids are aldehyde lactams (e.g. Baluchistanamine, Punjabine), lactam esters (e.g. **Gilgitine, Talcamine**) or aldehyde amines (Jhelumine, **Chenabine**). Karakoramine lacks the lactam moiety but possesses a hydroxymethyl function in the C′-aromatic ring.

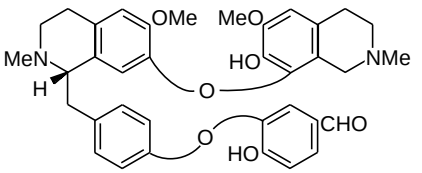
Like the bisbenzylisoquinoline precursors, these alkaloids differ in the number and position of the diphenyl ether linkages. Karakoramine has only one such bond, Baluchistanamine has two, and Punjabine has three. **Secantioquine** and Secolucidine are examples of a biphenyl system.



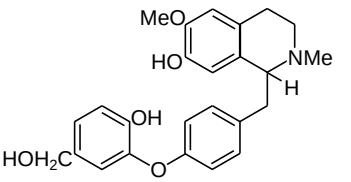
Baluchistanamine



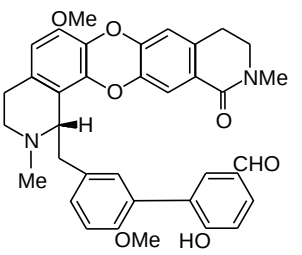
Punjabine



Jhelumine



Karakoramine



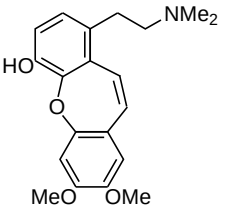
Secolucidine

Cularine group alkaloids (VX2440)

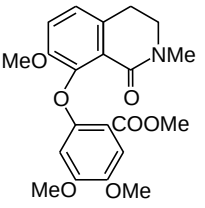
The cularines are tetracyclic isoquinoline alkaloids which contain a dihydrooxepine or oxepine ring between C-8 and C-2′. They are formed by intramolecular oxidative coupling of 7,8,3′,4′- or 7,8,4′,5′-tetraoxygenated tetrahydrobenzylisoquinolines (classical cularines and isocularines respectively) although the biogenesis of the recently isolated **Gouregine** probably proceeds via oxidation of an aporphine precursor. Structural variants in this group of alkaloids include 4-hydroxycularines (e.g. **4-Hydroxysarcocapnine**), oxocularines (e.g. **Oxocularine**), 3.4-dioxocularines (e.g. **Yagonine**) and **Aristoyagonine**, the only example of an aristocularine. Note the parallism between cularine and aporphine alkaloids.

Secocularine alkaloids (VX2450)

The secocularines can be divided into two sub-groups, namely B- and C-ring secocularines. B-Ring secocularines, exemplified by **Secocularine**, Secocularidine and **Norsecocularine**, are structurally related to phenanthrene alkaloids derived from aporphines and are probably formed *in vivo* by Hofmann degradation of cularines. C-Ring cularines, represented by Noyaine, constitute a new type of alkaloid without counterpart among aporphinoids.



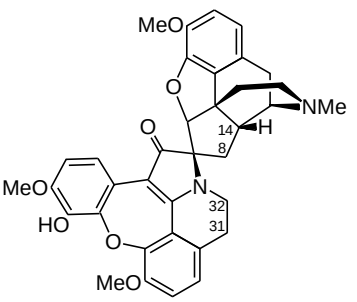
Secocularidine



Noyaine

Cancentrine-type alkaloids (VX2460)

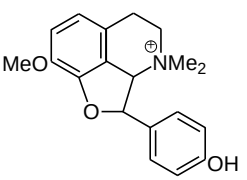
These alkaloids are dimers involving a cularine unit linked to a morphinan unit through a spiro bridge. They were found in a *Dicentra* species. Currently three examples are known: Cancentrine itself and **Dehydrocancentrines A and B**.



Cancentrine

Quettamine alkaloids (VX2470)

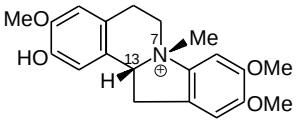
Whereas classical-type cularine alkaloids of the Fumariaceae are biogenetically derived from intramolecular oxidative coupling of tetraoxygenated tetrahydroisoquinoline precursors, the quettamines are obtained from *in vivo* intramolecular oxidation of a trioxygenated tetrahydroisoquinoline. So far only three naturally occurring quettamines are known: Quettamine, **Secoquettamine** and **Dihydrosecoquettamine**. These alkaloids, all found in *Berberis baluchistanica*, incorporate either a benzofuran or a dihydrobenzofuran ring within the molecular framework and the seco bases possess the *N*,*N*-dimethylaminoethyl substituent.



Quettamine

Dibenzopyrrocoline alkaloids (VX2480)

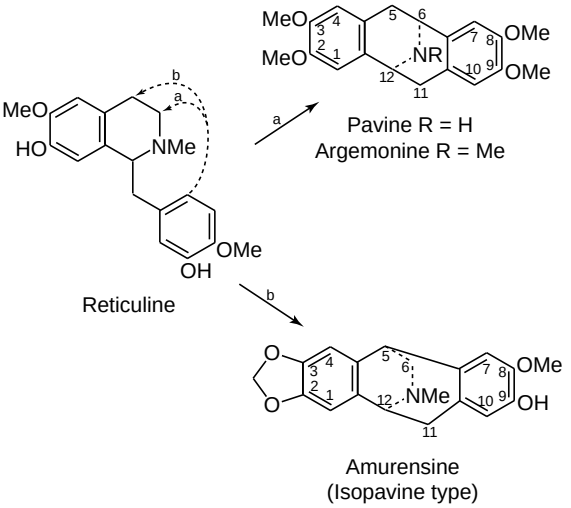
These alkaloids, only a few of which are known, are clearly derived by oxidation of a benzylisoquinoline precursor; indeed, the ring system was prepared *in vitro* by this route before it was encountered in nature.



Cryptaustoline

Pavine and isopavine alkaloids (VX2520, VX2540)

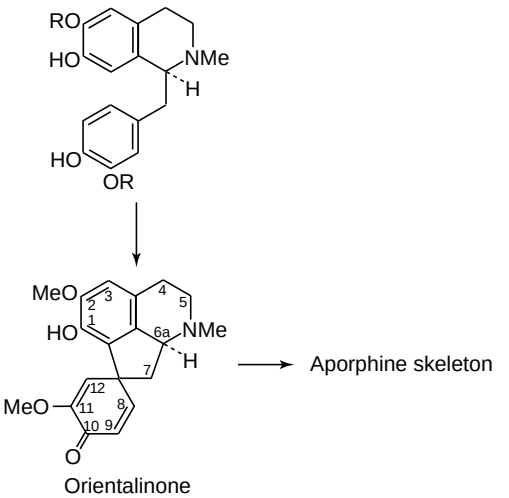
These alkaloids are formed by alternative modes of oxidative cyclisation of benzylisoquinoline precursors. In addition there are bisbenzylisoquinoline alkaloids composed of pavine and aporphine units, e.g. **Pennsylpavine**.



Gözler, B. *et al.* (1983) *J. Nat. Prod.*, **46**, 293.

Proaporphine alkaloids (VX2600)

This group of alkaloids, e.g. Orientalinone, represents an intermediate stage in the conversion of the benzylisoquinoline alkaloids by phenol oxidative coupling into the aporphines.



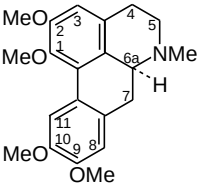
Orientalinone

The spirocyclohexenone ring may occur in various oxidation levels from cyclohexadienone to cyclohexanol.

Aporphine alkaloids (VX2610, VX2620, VX2640, VX2700, VX2750, VX2780, VX2800, VX2820, VX2840, VX6820, VX6840)

This large group of alkaloids simply contains the tetracyclic ring system formed by phenol oxidative coupling of a benzylisoquinoline precursor. The structural variations include:

(*a*) the simple aporphines, exemplified by Glaucine;



Glaucine

(*b*) dehydro derivatives of (*a*), in which the double bond is generally between C-6*a* and C-7 in *N*-methyl compounds and between C-6*a* and *N* in apo compounds;

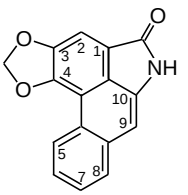
(*c*) miscellaneous oxidative derivatives of (*a*), mostly with a hydroxy or methoxy function at C-7, or two at C-4 and C-7;

(*d*) those with an aromatic isoquinoline ring and a carbonyl group at C-7, e.g. **Liriodenine**, the so-called oxoaporphines;

(*e*) miscellaneous aporphinoids. Included in this group are **Telezonine**, duguenaine-type aporphinoids, ring A quinonoid aporphinoids (e.g. **Pancoridine**), oxoisoaporphines (e.g. **Menisporphine**), azafluoranthenes (e.g. **Rufescine**), diazafluoranthenes (e.g. **Eupolauridine**), 1-azaoxoaporphinoids (e.g. **Sampangine**), azahomoaporphines (e.g. **Dragabine**), so-called catechol dioxygenase oxidized aporphinoids (e.g. **Andesine, Chiloenine, Santiagonamine**), tropoloisoquinolines (e.g. **Imerubrine**) and **Cleistopholine-** and **Onychine-** type alkaloids.

(*f*) compounds in which the heteroring has opened to give phenanthrene derivatives, with the $CH_2CH_2NR^1R^2$ chain still present, e.g. **Taspine**;

(*g*) compounds derived from (*f*) which have lost C-5, mostly containing a five-membered lactam ring (the aristolactams, e.g. Cepharanone A). The class even includes some members in which nitrogen has been oxidised to a nitro group, e.g. **Aristolochic acid A**.



Cepharanone A

Although the aristolochic acids and aristolactams are non-basic they are still classified as alkaloids since their respective skeletons bear a distinct resemblance to that of the aporphines.

(*h*) a group of dimeric aporphinoid alkaloids exemplified by the aporphine-benzylisoquinoline dimers, e.g. **Thalifaberine**, the proaporphine-benzylisoquinoline dimers (e.g. **Pakistanamine**), and the **Hernandaline**-type and **Coyhaiquine**-type alkaloids. The two latter types are, respectively, oxidation products of the aporphine-benzylisoquinolines and proaporphine-benzylisoquinolines.

A new addition to this class of compound are the proaporphine-tryptamine dimers. These heptacyclic alkaloids, found in *Roemeria hybrida* (Papaveraceae)

and *Phoebe grandis* (Lauraceae), are probably derived biogenetically by a Mannich-type condensation of a ketonic tetrahydroproaporphine with a tryptamine analogue. **Roemeridine** is a typical example.

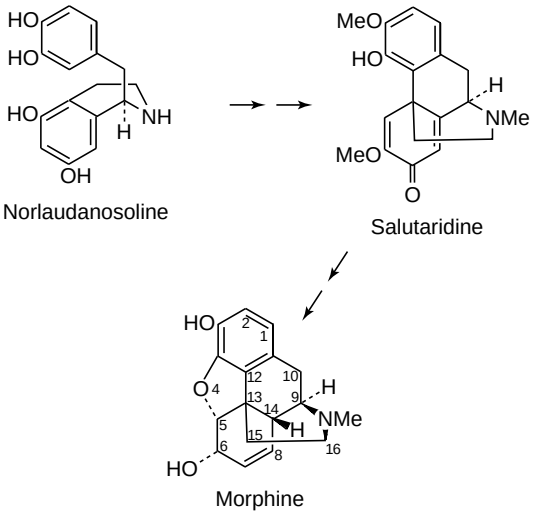
To this listing must be added a small but significant group of bisaporphines. The majority of these dimers are bonded through a carbon-to-carbon linkage at C-7 and C-7′ (e.g. **Urabaine**), although examples of C8-C8′ coupled bisaporphines [e.g. (**8, 8′-*R*)**- and (**8, 8′-*S*)-Bisisocorydine**] and oxygen-bonded dimers (e.g. **11,8′-O-Bisisocorydine, Dehatriphine**) have recently been isolated.

Cavé, A., Leboeuf, M. and Waterman, P.G. (1987) in *Alkaloids: Chemical and Biological Perspectives*, (ed. S.W. Pelletier), Vol. 5, Wiley-Interscience, New York.
Gözler, B. *et al.* (1990) *J. Nat. Prod.*, **53**, 675 (*aporphine dimers*).
Guinaudeau, H. *et al.*, *Lloydia*, (1975) **38**, 275; *J. Nat. Prod.*, (1979) **42**, 133, 325; (1983) **46**, 761; (1984) **47**, 565; (1988) **51**, 389, 1025; (1994), **57**, 1033
Jackman, L.M. *et al.* (1979) *J. Nat. Prod.*, **42**, 437.
Mix, D.B. *et al.* (1982) *J. Nat. Prod.*, **45**, 657 (*aristolochic acids and aristolactams*).
Shamma, M. and Guinaudeau, H. (1984) *Tetrahedron*, **40**, 4795.

Morphine alkaloids (VX2900)

This extremely important group of more than 30 alkaloids is formed by phenol oxidative coupling of a hydroxylated benzylisoquinoline precursor such as Norlaudanosoline, itself originating from two molecules of tyrosine.

The group may be sub-divided into bases of the Salutaridine type, those related to Morphine, which have a 4,5-oxide bridge, and those related to **Sinoacutine**, which are enantiomeric with the salutaridine group.

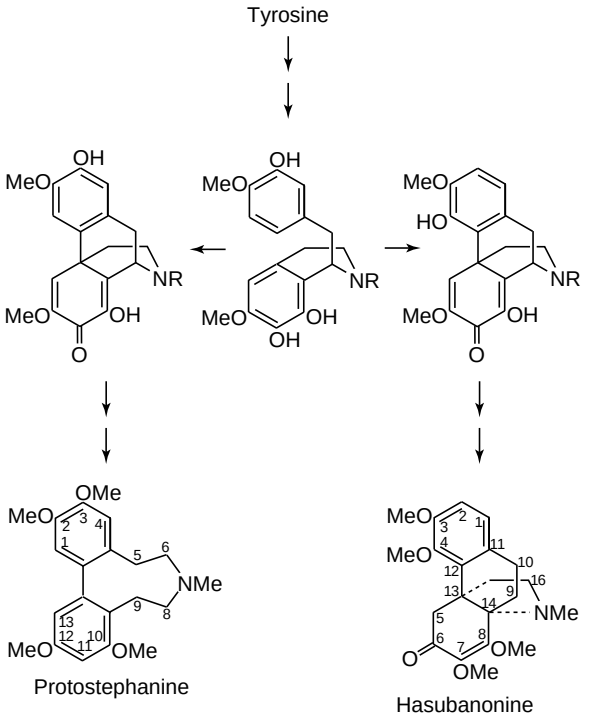


Blaskó, G. and Cordell, G.A. (1988) *Heterocycles*, **27**, 1269.

Dibenzazecine and Hasubanan alkaloids (VX2980, VX3000)

These two groups may appear at first sight to belong to quite different structural groups, but there is little doubt that biosynthetically they are not too disparate, as is evidenced by their occurrence in the same plant, *Stephania japonica*. Both groups are derived from two tyrosine units, but their biosynthesis is not simple. A central, pivotal intermediate appears to be a hydroxylated benzylisoquinoline which can cyclise, by alternative phenol oxidative coupling processes, to a hydroxysalutaridine, or isomer. This biosynthetic route is exceptional since it would appear that two hydroxy-groups need to be present in one of the aromatic rings; in all other known cases of oxidative coupling, only one hydroxy-group seems to be essential. The biosynthesis obviously has affinities with that of the

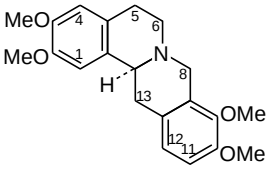
morphine alkaloids, although there is clearly an important divergence in the later stages.



Protoberberine alkaloids (VX3100, VX3240)

These tetracyclic alkaloids are derived from benzylisoquinolines by condensation with a one-carbon unit (the berberine bridge). This group of alkaloids consists of:

(*a*) the simple tetrahydroprotoberberines, e.g. Tetrahydropalmatine:

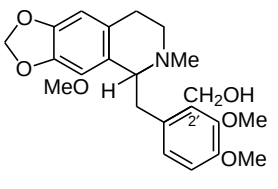


(*b*) the protoberberines, such as **Berberine**;

(*c*) 13-methyl derivatives, such as **Corydaline**;

(*d*) miscellaneous bases, e.g. **Orientalidine**, which has an extra carbon atom at C-12.

(*e*) ring-opened protoberberines (secoberberines) which can be regarded as benzylisoquinolines with a carbon substituent at C-2′. The latter may occur in different oxidation states: as an aldehyde (e.g. **Aobamine**), an alcohol (e.g. Macrantaline) or as a carboxylic acid (e.g. **Macrantoridine**).

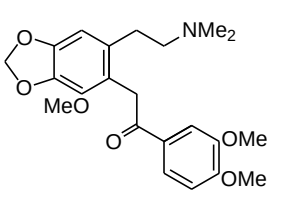


Structure determination in this series, i.e. the correct location of substituents on the protoberberine ring system, has been a matter of some difficulty and it is possible that some of the assignments currently given in DNP will prove to be incorrect.

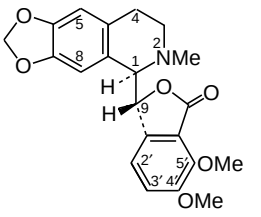
Narceine and phthalideisoquinoline alkaloids (VX3140, VX3200)

These alkaloids constitute further examples of oxidation products of protoberberines, in which the nitrogen to C-8 bond has been cleaved. The narceine group contain an ethanamine chain and, as well as relatives of narceine, include bases which contain an enol lactone or enamine lactam function, as in **Bicuculline imide**; and those with a higher (**Bicucullinine**) or lower (**Peshawarine**) oxidation level than Narceine.

The phthalideisoquinoline alkaloids contain an intact tetrahydroisoquinoline ring, but oxidation of the nitrogen to C-8 bond has been followed by γ-lactone formation; α-Hydrastine is typical of this group.



Narceine

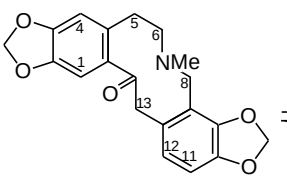


α-Hydrastine

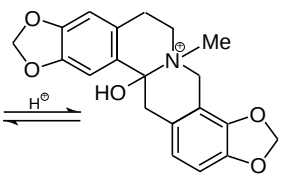
Blasko, G. *et al.* (1982) *J. Nat. Prod.*, **45**, 105.

Protopine alkaloids (VX3160)

These tricyclic bases are simply formed by oxidative ring fission of protoberberine *N*-metho salts. Two of these bases (**Corycavamine, Corycavidine**) have an additional methyl group at C-13.



Protopine

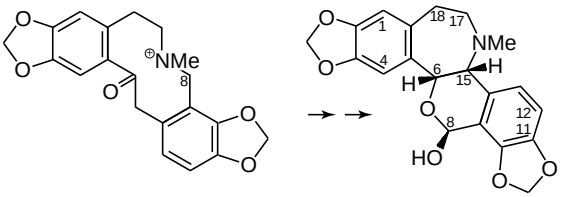


Tetracyclic protopinium cation

Guinaudeau, H. and Shamma, M. (1982) *J. Nat. Prod.*, **45**, 237.

Rhoeadine alkaloids (VX3180)

This group of alkaloids has been encountered only in the *Papaver* genus. Their biogenesis, which is not completely understood, appears to be from two tyrosine units, *via* tetrahydroberberine and protopine intermediates. Oxidative fission of the nitrogen to C-8 bond followed by oxidative cyclisation of nitrogen on to C-15 and lactol formation results in a ring system in which C-8 becomes the lactol carbon atom. (N.B. Other numberings of the ring system are in use; hence the literature can be confusing.)



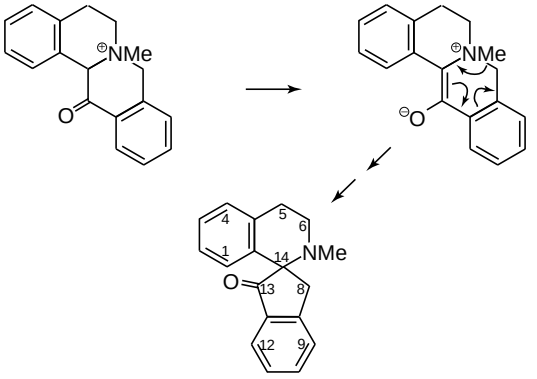
Rhoeadine
(CA numbering)

Montgomery, C.T. *et al.* (1983) *J. Nat. Prod.*, **46**, 441.

Spirobenzylisoquinoline alkaloids (VX3220)

These alkaloids are derived from protoberberines by a 1,2-shift of C-8 from nitrogen to C-14. Several mechanisms are possible, but since several of the alkaloids contain oxygen at C-13 a route *via* an enolate ylid is attractive.

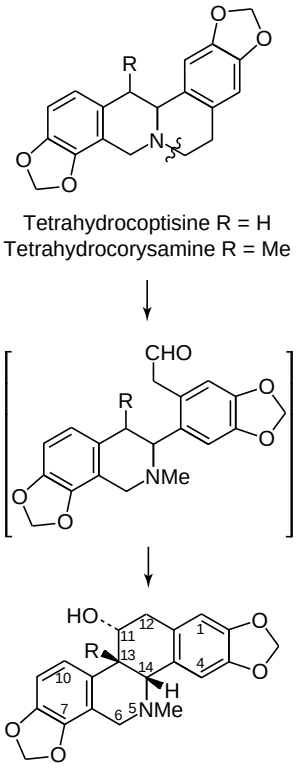
Closely related to this group is **Lahorine**, an indenobenzazepine derivative, which may be derived biogenetically from the spirobenzylisoquinolines.



Preisner, R.M. and Shamma, M. (1980) *J. Nat. Prod.*, **43**, 305.

Benzo[c]phenanthridine alkaloids (VX3300)

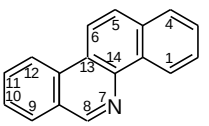
This interesting group of 100 or more alkaloids is derived from tetrahydroprotoberberine precursors by oxidation of the C-6 to nitrogen bond followed by cyclisation of C-6 on to position 13. Various oxidation stages exists, e.g. (*a*) partially reduced benzophenanthridines, as in Chelidonine; (*b*) fully aromatic systems, as in **Sanguinarine;** (*c*) tricyclic alkaloids, e.g. **Corydamine**; (*d*) tricyclic systems formed by fission of the C-6 to nitrogen bond (benzophenanthridine numbering), e.g. **Arnottianamide**; (*e*) alkaloids formed by addition of a carbon substituent to C-6, e.g. **Corynolamine**; (*f*) dimeric alkaloids, e.g. **Sanguidimerine**.



Tetrahydrocoptisine R = H
Tetrahydrocorysamine R = Me

Chelidonine R = H
Corynoline R = Me

Several numbering systems have been used for the benzophenanthridine alkaloids but the one shown below, based on biogenetic considerations, has been adopted throughout DNP (this is not the numbering scheme for Benzo[*c*]phenanthridine itself).



Krane, B.D. *et al.* (1984) *J. Nat. Prod.*, **47**, 1.

Phenethylisoquinoline alkaloids (VX3360)

This group arises from a phenethylisoquinoline precursor, which is itself generated by condensation of tyrosine with a C_6–C_3 unit derived from phenylalanine, probably *via* cinnamic acid.

In addition to the small group of simple phenethylisoquinolines, several other of the following groups are related to them by further elaboration as shown in Figure 18.

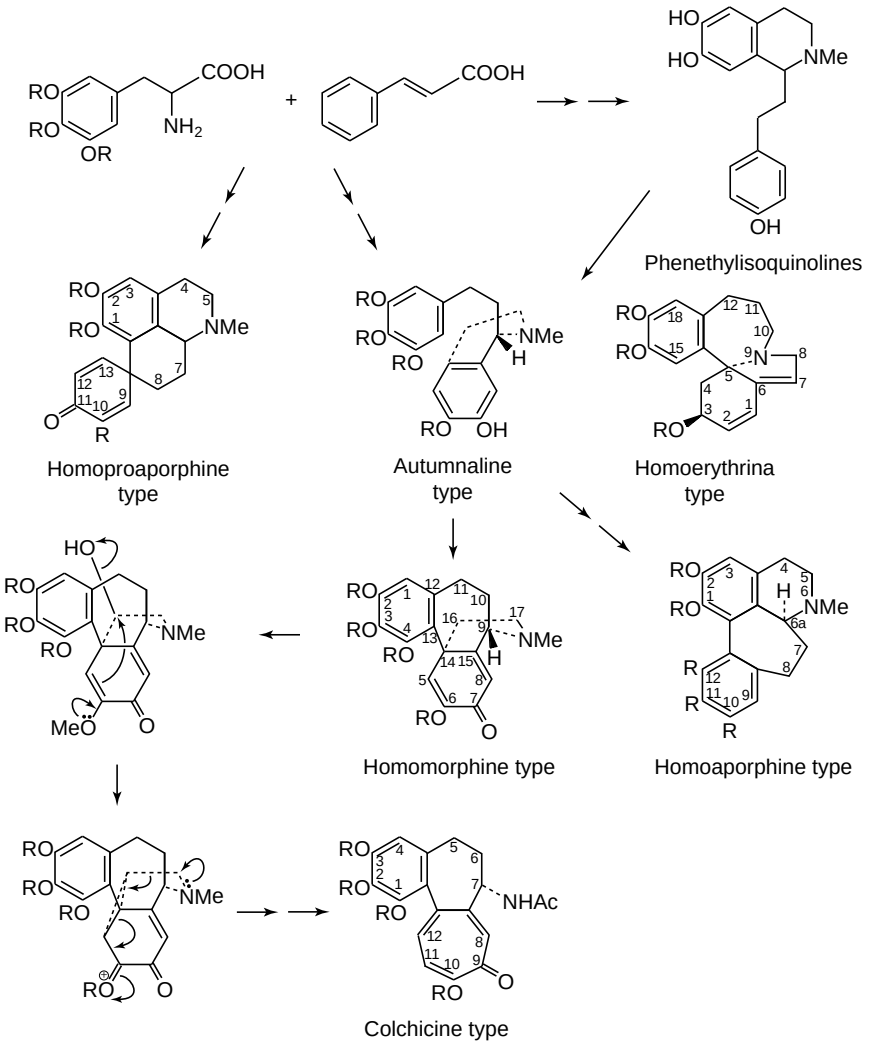


Figure 18.

Homoaporphine alkaloids (VX3380)

E.g. **Kreysigine**. The sequence from tyrosine and phenylalanine *via* a phenethylisoquinoline to homoproaporphines and homoaporphines appears superficially to be exactly analogous to the course of biosynthesis of the aporphine alkaloids. However, although Autumnaline is an efficient precursor for both Kreysiginone and Kreysigine in *Kreysigia multiflora*, dienone intermediates such as Kreysiginone are not involved in the biosynthesis of the homoaporphines, such as Kreysigine.

Tojo, E. (1989) *J. Nat. Prod.*, **52**, 909.

Homoerythrina alkaloids (VX3440)

E.g. **Schelhammeridine**. These would appear to be formed by a route analogous to that adopted in the Erythrina group.

Colchicine-like alkaloids (VX3400)

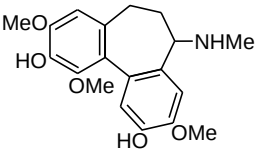
Autumnaline and **O-Methylandrocymbine** (but not **Androcymbine** itself) are efficient precursors for Colchicine. Hence the biosynthesis must involve ring enlargement of the dienone ring in the *O*-Methylandrocymbine skeleton; one attractive suggestion is that a cyclopropane intermediate may be involved.

Androcymbine arises by phenol oxidative coupling, probably of Autumnaline, by a process analogous to that involved in the biosynthesis of the morphine alkaloids.

Lumicolchicines are the product of u.v. irradiation of Colchicine, and while they have been reported to occur naturally, they could be regarded as artifacts.

Dibenzocycloheptylamine alkaloids (VX3410)

Dibenzocycloheptylamine alkaloids have recently been found in plants of the genera *Colchicum* and *Androcymbium*. To date only six naturally occurring examples are known; these include Jerusalemine and **Salimine**. Jerusalemine may be formed via decarbonylation of the tropolone ring of 2-Demethyldemecolcine by a peroxidase system present in the plant, with accompanying oxidation. Salimine, on the other hand, may arise from Colchicine by enzymatic peroxidation of ring C followed by hydroxylation and methylation.



Jerusalemine

Erythrina and cephalotaxus alkaloids (VX2940, VX3420)

This group of alkaloids is derived from two tyrosine units by oxidative coupling and intramolecular rearrangement and consists of about 100 alkaloids which may be subdivided as follows:

(*a*) those alkaloids which contain the erythrinan skeleton (e.g. Erythraline); these constitute the majority;

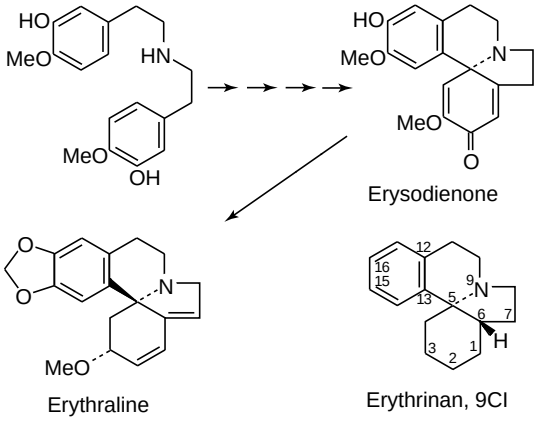
(*b*) those alkaloids in which the aromatic ring of erythraline has been replaced by an unsaturated lactone ring, e.g. *β*-**Erythroidine**;

(*c*) the cephalotaxine alkaloids, in which the hydroaromatic component has undergone a skeletal rearrangement. Some of these alkaloids occur as half-esters with dihydroxydicarboxylic acids (e.g. **Harringtonine**);

(*d*) a small group of alkaloids, known as homoerythrina alkaloids, contain an additional carbon atom in the skeleton. Such alkaloids, e.g. **Schelhammeridine**, occur in the genera *Schelhammera* and *Cephalotaxus*;

(*e*) an even smaller group of alkaloids which contain an additional nitrogen atom in ring D. These 16-azaerythrinanes, e.g. **Erymelanthine**, are possibly derived biogenetically by *in vivo* oxidation of the aromatic ring of compounds possessing the classical erythrinan skeleton followed by uptake of ammonia and recyclization;

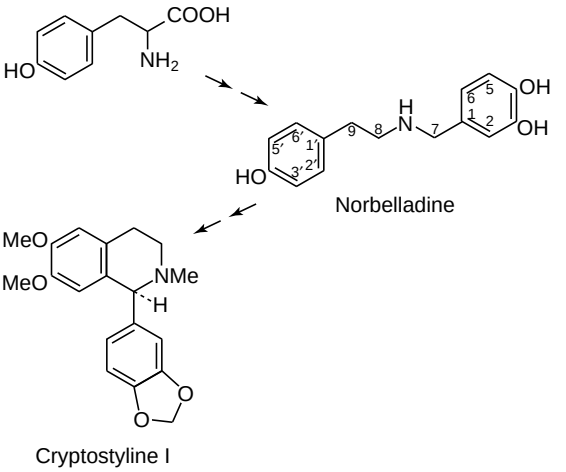
(*f*) a few so-called dimeric alkaloids that incorporate a tryptophan moiety, e.g. **Erysopinophorine**.



Amer, M.E. *et al.* (1991) *J. Nat. Prod.*, **54**, 329.
Findlay, J.A. (1976) *Cephalotaxus* Alkaloids, in *MTP Series* 2, Vol. 9, *Alkaloids* (ed. K. Wiesner), Butterworths, London.
Hudicky, T. (1987) in *Alkaloids: Chemical and Biological Perspectives*, (ed. S.W. Pelletier) Vol. 5, Wiley-Interscience, New York.
Mondon, A. (1970) *Erythrina* Alkaloids, in *Chemistry of the Alkaloids*, (ed. S.W. Pelletier), Van Nostrand Reinhold, New York.

Amaryllidaceae alkaloids (VX3500)

This group of alkaloids is also derived from two tyrosine units which combine, with loss of one carbon atom, to give a benzylphenylethylamine precursor unit, e.g. Norbelladine, which by various oxidative cyclisation processes, prominent among which are phenol oxidative coupling reactions, can give rise to the nine major skeletal groups.

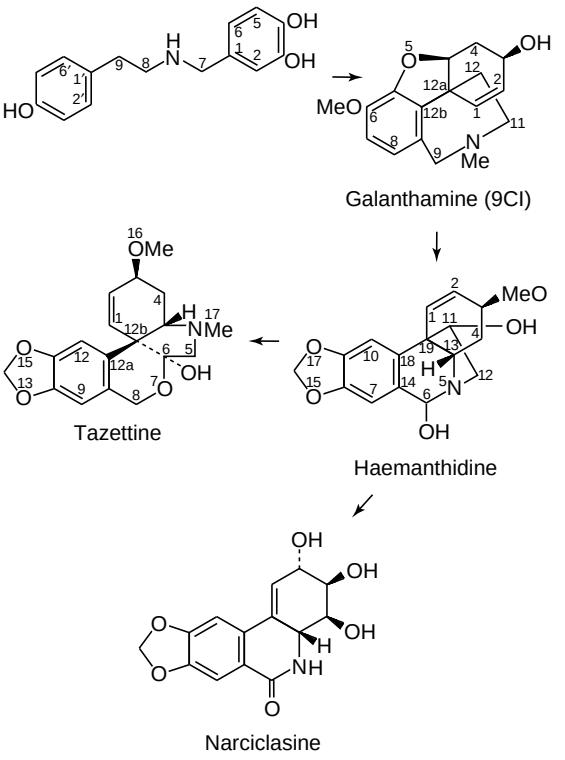


(**a**) **Cryptostyline I, cherylline, and nivalidine**.
Oxidation of a norbelladine-type precursor at C-7 followed by cyclisation at C-2′ gives the simple 1-phenyltetrahydroisoquinolines exemplified by Cryptostyline I, whereas oxidation at the alternative benzylic position (C-9) and cyclisation at C-6 gives the 4-phenyltetrahydroisoquinolines related to **Cherylline**. Oxidative coupling of positions 2 and 2′ gives the skeleton of **Nivalidine**, but this may be an artifact, derived from Galanthamine.

(**b**) **Galanthamine, haemanthidine, tazettine, and pancracine groups**.
Oxidative coupling of C-2 with C-1′ in Norbelladine gives the Galanthamine skeleton which, by an obvious cyclisation process (nitrogen to position 1), can give rise to the Haemanthidine ring system. Opening of the carbinolamine

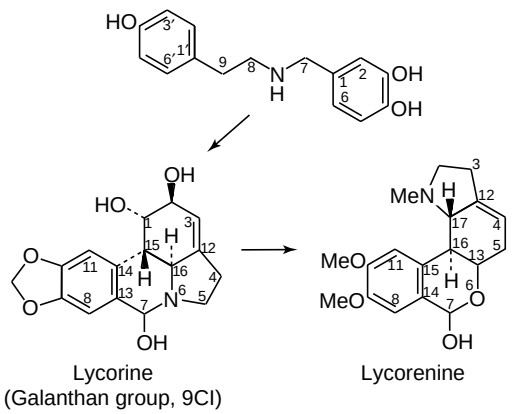
function in Haemanthidine followed by a redox reaction and cyclisation of the oxygen at C-6 on to position 11 (Haemanthidine numbering) then affords the Tazettine skeleton.

A further possibility is the migration of C-18 in the haemanthidine skeleton to position 11, which gives rise to the ring system present in **Pancracine** and **Montanine**.



(c) Lycorine and Lycorenine alkaloids

A double cyclisation of C-2 to C-3′ and nitrogen to C-2′ provides the tetracyclic skeleton characteristic of Lycorine and its analogues. Further modification of this ring system by oxidative fission of the nitrogen to C-7 bond followed by attachment of oxygen at C-7 to position 1 (galanthan numbering) then gives rise to Lycorenine.



(d) Narciclasine alkaloids

This small group, exemplified by **Narciclasine**, also stems from Norbelladine, and appears to be formed *via* Crinine (but *not* 3-Epicrinine) by loss of the two carbon bridge and appropriate oxidations.

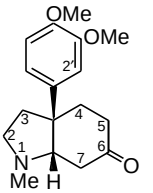
It should be noted that several of the Amaryllidaceae alkaloids occur in enantiomeric forms.

Ghosal, S. *et al.* (1985) *Phytochemistry*, **24**, 2141.

Jeffs, P.W. (1973) in *MTP Series One*, Vol. 9, *Alkaloids* (ed. K. Wiesner), Butterworths, London.

Mesembrenoid alkaloids (VX3600)

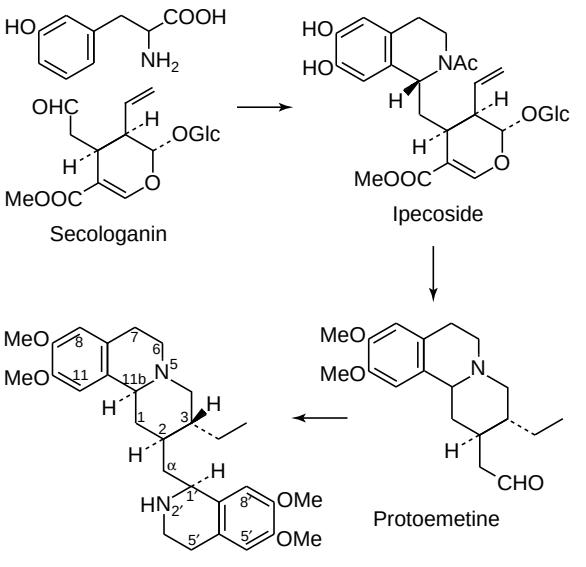
Derived from two phenylalanine units with loss of one of the ethanamine side-chains, this group of about 20 alkaloids is typified by Mesembrine; a second sub-group contains alkaloids in which the pyrrolidine ring has not been formed, as in **Joubertiamine**. A third variant contains bases in which a pyridine ring has been fused on, as in **Tortuosamine**.



Mesembrine

Emetine group alkaloids (VX3690)

The emetine group of alkaloids are unique among the isoquinoline group in that they are also derived from a monoterpenoid unit *via* Secologanin. Incorporation of one phenylalanine/tyrosine unit gives the alkaloids exemplified by Ipecoside and Protoemetine; the latter, by combination with a second amino acid unit, gives rise to the Emetine group. Alternatively, combination with a tryptamine unit gives the typical alkaloids of *Alangium lamarckii*, e.g. **Alangimarckine**.



Secologanin

Ipecoside

Protoemetine

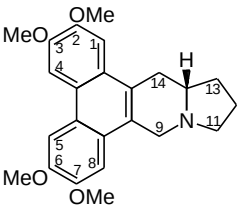
Emetine

Openshaw, H.T. (1970) in *Chemistry of the Alkaloids*, (ed. S.W. Pelletier), Van Nostrand Reinhold, New York.

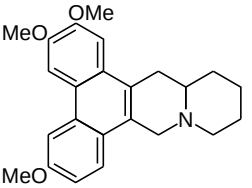
Wiegrebe, W. *et al.* (1984) *J. Nat. Prod.*, **47**, 397.

Phenanthroindolizidine and phenanthroquinolizidine alkaloids (VX3700, VX3760)

These alkaloids are derived from two molecules of phenylalanine or tyrosine, together with, presumably, ornithine (→ Tylophorine group) or lysine (→ Cryptopleurine group).



Tylophorine



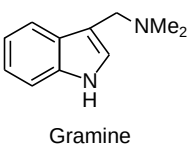
Cryptopleurine

Alkaloids derived from tryptophan

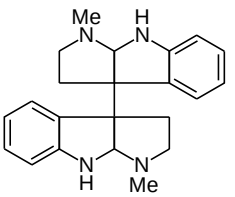
The group of alkaloids derived from tryptophan constitutes the largest, most varied and most fascinating of all alkaloid groups. The alkaloids include simple tryptamine derivatives, carbazoles (in which the ethanamine chain has been lost), a variety of alkaloids in which tryptamine has combined with one or more prenyl residues, and others in which regular monoterpenoid or diterpenoid units have been incorporated. However, the largest group, and the most extensively studied, is the alkaloids derived from tryptophan and a monoterpenoid unit based on Secologanin.

Simple tryptamine alkaloids (VX4000, VX4040, VX4140, VX4160)

The simplest indole alkaloid is Gramine. A number of simple tryptamine derivatives also occur naturally. Other relatively simple derivatives include the constituents of the Calabar bean, e.g. **Physostigmine**, several dimers, e.g. Chimonanthine, and several oligomers, e.g. **Hodgkinsine** (a trimeric species), the **Quadrigemines**, which are tetramers, and even a pentamer, **Psychotridine**.



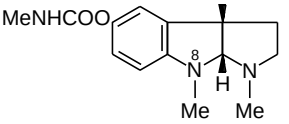
Gramine



Chimonanthine

Physostigmine-like alkaloids (VX4100)

Physostigmine, the prototype of this group of alkaloids, was first isolated from *Physostigma venenosum* and has also been produced by *Streptomyces* spp. The alkaloid is characterised by a urethane group which is readily hydrolysed with aqueous base to afford Eseroline. In addition to plant alkaloids with this skeleton, other representatives of this class have recently been isolated from the marine bryozoan *Flustra foliacea* (e.g. the **Flustramines**) and from skin extracts of the Australian frog *Pseudophryne coriacea* (e.g. **Pseudophrynamine A, Pseudophrynaminol**).



Physostigmine

Carbazole alkaloids (VX4300)

All carbazole alkaloids encountered so far contain a substituent at C-3, which seems to indicate that an isoprenoid unit, derived from mevalonate, is involved in their biosynthesis.

This group may be sub-divided into:

(*a*) simple carbazoles, e.g. **Glycozoline**, which is simply 6-Methoxy-3-methylcarbazole;

(*b*) carbazoles with an additional prenyl substituent, e.g. **Ekeberginine**;

(*c*) pyranocarbazoles, in this group a prenyl residue has cyclised on to a phenolic hydroxy-group, as in **Heptazolidine**;

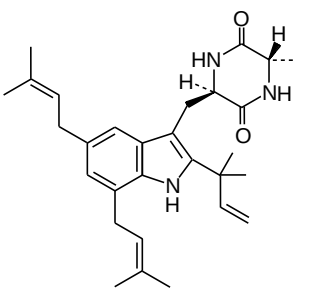
(*d*) carbazoles containing a complete monoterpene unit, e.g. **Mahanimbine**;

(*e*) bis-carbazole alkaloids, e.g. **Murrafoline C**.

Chakraborty, D.P. (1977) *Prog. Chem. Org. Nat. Prod.*, Vol. 34.

Miscellaneous tryptophan derivatives

A number of derivatives of tryptophan are known in which combination with a second amino acid affords a dioxopiperazine; introduction of one, two, or three prenyl groups is also involved in these metabolites, almost all of which occur in microorganisms, rather than higher plants.

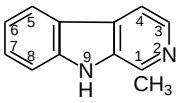


Echinulin

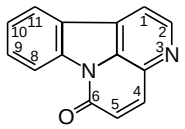
The prototype of these mould metabolites is Echinulin, from *Aspergillus echinulatus*. Others are the brevianamides, e.g. **Brevianamide A, Roquefortine** (from *Penicillium roquefortii*), and **Verruculogen** (a tremorgen from *P. verruculosum*), **Oxaline** (from *P. oxalicum*), and **Indolactam V**.

β-Carboline alkaloids (VX4240)

A large number of relatively simple *β*-carboline derivatives occur naturally. These include *β*-carbolines unsubstituted at C-1, e.g. **1,2,3,4-Tetrahydro-6-methoxy-2-methyl-*β*-carboline**, those containing a methyl group at C-1, i.e. the Harman group, and several which contain a substituent at C-1 and/or C-3. The substituents at position 1 may be acyl, carboxyl, or they may be more complex, as in **Perlolyrine**. Other bases include examples in which tryptamine has condensed with an isoprenoid unit, as in **Borrerine**, and an important group, the canthinones, in which a 3-carbon unit attached *via* the indole nitrogen and C-1 results in the introduction of another ring.

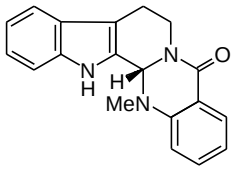


Harman



Canthin-6-one

Another group of β-carboline derivatives have been isolated from *Eudistoma olivaceum*, a Caribbean tunicate. Of these **Eudistomin C**, presumably derived from tryptophan and cysteine, is typical.



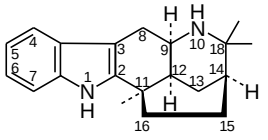
Evodiamine

The alkaloids of *Evodia rutaecarpa*, e.g. Evodiamine, which are also derived from anthranilic acid (*q.v.*), may also be included in this group.

Finally, several alkaloids from *Picrasma quassioides* are bis-β-carbolines, e.g. **Picrasidine M**.

Aristotelia alkaloids (VX4620)

This group of metabolites is notable in that the tryptamine unit is attached to an unrearranged monoterpenoid unit; Aristoteline is typical of the *Aristotelia* bases. The hapalindoles, e.g. **Hapalindole C**, are a family of metabolites which have been found in the cyanophyte, *Hapalosiphon fontinalis*.



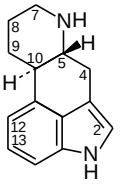
Aristoteline

Borreria alkaloids

These alkaloids also contain a regular terpenoid unit, as in **Borrecapine**, from *B. capitata*. **Borreline**, from the same plant, contains a degraded monoterpene unit. **Borreverine**, from *B. verticillata*, contains two tryptamine units and a monoterpenoid component, and may be prepared by dimerisation of Borrerine.

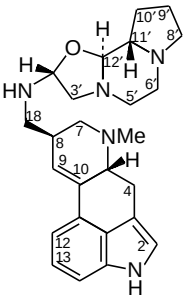
Ergot alkaloids (VX4460)

These alkaloids, which occur in the fungus *Claviceps purpurea*, are derived from 4-prenyltryptophan by cyclisation to give a tricyclic base related to **Chanoclavine I**, which is representative of the simplest subgroup. Further elaboration gives the tetracyclic ergoline skeleton, as in **Elymoclavine**, which is present in the majority of alkaloids in this group. The most important alkaloids, many of which have useful medical applications, are complex peptide alkaloids formed from lysergic acid, in which C-17 in the elymoclavine-type precursor has been oxidised to a carboxyl group, by attachment to one or more amino acids. **Ergocristine**, which is based on the ergotaman nucleus, is typical.



Ergoline, 9CI

A few other metabolites, which may also be included in this group, are the result of skeletal rearrangement. **Clavicipitic acid** is one such compound; another is α-**Cyclopiazonic acid**. However, whereas the former is definitely a product of prenyl-tryptophan metabolism the latter, from *Penicillium cyclopium* is not, since it appears to arise from reaction of a tryptophan-acetoacetate unit with prenyl pyrophosphate.



Ergotaman, 9CI

Floss, H.G. (1976) *Tetrahedron*, **32**, 873.
Horwell, D.C. (1980) *Tetrahedron*, **36**, 3123.

Monoterpenoid indole alkaloids

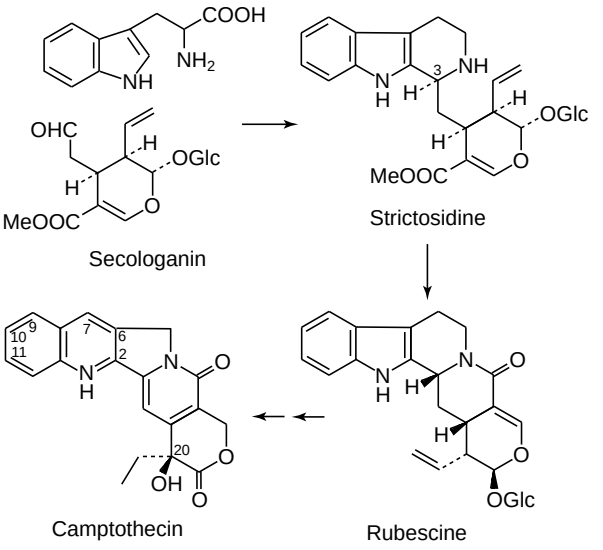
The extremely important and varied group of monoterpenoid indole alkaloids originate from the condensation of tryptophan with Secologanin to give Strictosidine, which is further elaborated to give the corynanthe alkaloids together with an impressive array of structural variants. They can be sub-divided into compounds of a dozen sub-groups.

Monoterpenoid-derived indole alkaloid glycosides (VX4640)

These are based on, e.g. Strictosidine and related compounds, mainly glycosides.

Camptothecin-like alkaloids (VX4700)

These alkaloids, which also contain the quinoline ring system, are probably derived from Strictosidine *via* an intermediate related, possibly, to Rubescine. In this case conversion of the indole group into the quinoline ring involves ring enlargement of ring B at the expense of ring C; otherwise the changes in the formation of Camptothecin from Strictosidine are trivial.



Hutchinson, C.R. (1981) *Tetrahedron*, **37**, 1047.

Indoloquinolizidine alkaloids (VX4780)

These are alkaloids in which a Strictosidine precursor has been elaborated with formation of a pyridine ring, as in **Angustine**.

The biogenesis of some of these alkaloids is not readily apparent. Since in many cases these indolopyridines occur together with related glycosidic alkaloids whose aglycones can react with ammonia to give precursors for the pyridine ring, it may be that many of these pyridinoid bases are artifacts. Corynantheine types which have lost one or more of the carbon atoms 16–19, e.g. **Deplancheine**, are also part of this group.

Kisakurek, M.V., Leeuwenberg, A.J.M. and Hesse, M. (1983) in *Alkaloids: Chemical and Biological Perspectives*, (ed. S.W. Pelletier) Vol. 1, Wiley-Interscience, New York.

Phillipson, J.D. *et al.* (1978) *J. Nat. Prod.*, **41**, 503.

Saxton, J.E. (ed.) (1983) *The Monoterpenoid Indole Alkaloids*, Wiley-Interscience, New York.

Corynanthe alkaloids (VX4800)

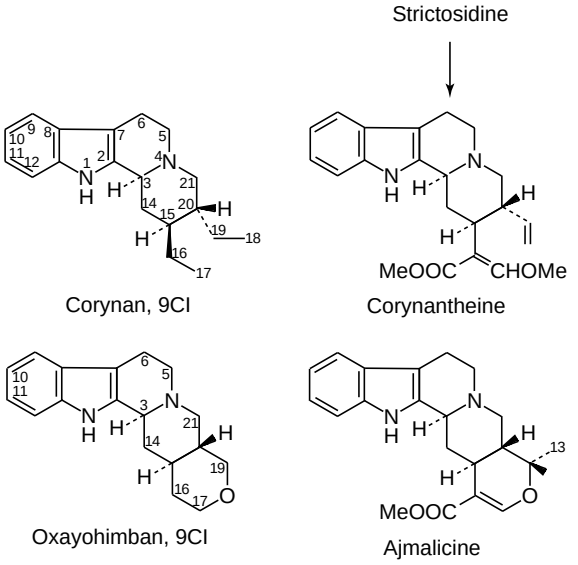
This group, based on the corynan nucleus, is exemplified by **Geissoschizine** and **Sitsirikine**.

Corynanthe tryptamine alkaloids (VX4820)

This group of about 50 alkaloids is formed from a Corynanthe moiety which is attached via C-17 to an additional tryptamine unit, as in the **Ochrolifuanines** and **Usambarines**.

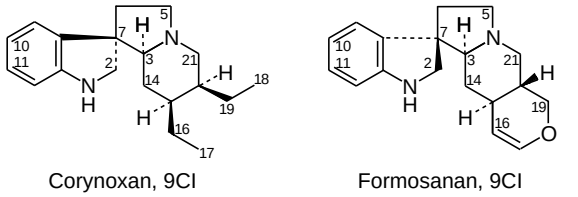
Ajmalicine-like alkaloids (VX4860)

These are based on the oxayohimban nucleus which, in common with all other ring systems in the indole alkaloid group, is numbered according to its biogenetic origin.



Oxindole alkaloids (VX4940)

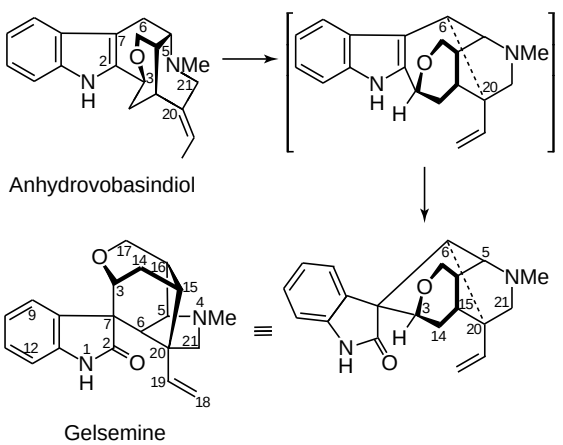
These are analogues of Corynantheine, e.g. **Rhynchophylline**, which are based on the corynoxan nucleus, or oxindole analogues of ajmalicinoid alkaloids, e.g. **Formosanine**, which are based on the formosanan nucleus.



In all the former four groups of alkaloids, stereoisomerism at positions 3, 19, and 20 is common, but C-15 has the unique configuration shown. This stereochemical constancy at C-15 is observed in virtually all known indole alkaloids.

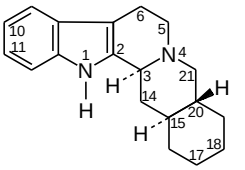
Gelsemium alkaloids (VX5000)

These alkaloids contain an oxindole function and a cage-like, hydroaromatic residue which can be imagined, in a formal sense, to arise from an intermediate related to anhydrovobasinediol by formation of a 6,20 bond and rearrangement to an oxindole. The major alkaloids in this group are related to Gelsemine; however, a smaller group, characterized by **Gelsedine**, lack the 6,20 bond, and have also lost C-21.



Yohimbinoid alkaloids (VX5040)

The Yohimbine alkaloids contain a carbocyclic ring E formed by C-17 to C-18 bond formation in a corynantheine precursor. As in the corynantheine-ajmalicine group stereoisomerism at all asymmetric centres except C-15 is known.

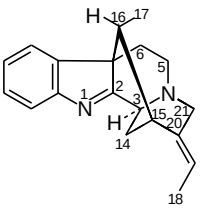


Yohimban, 9CI

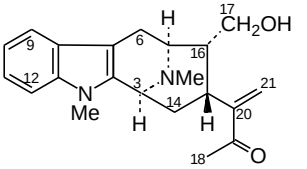
Structural variations include the presence of methoxy-groups in the aromatic ring, hydroxy- or acyloxy-groups at C-18, as in **Reserpine**, and various degrees of unsaturation in rings C-E, as in **Alstoniline**.

Akuammiline alkaloids (VX5200)

The ring system in this group is formed from a precursor of the corynantheine type by bond formation between C-16 and C-7. In addition to close relatives of **Akuammiline**, variations in this subgroup include alkaloids derived by C-3 to N-4 bond fission and C-2 to N-4 bonding, e.g. **Echitamine**; alkaloids with the Echitamine skeleton in which the C-21 to N-4 bond has been broken, e.g. **Eripine**; alkaloids in which the N-4 to C-5 bond in the Akuammiline skeleton has been severed, e.g. **Aspidodasycarpine**; a small group of alkaloids derived by fission of the C-21 to N-4 bond, as in Macroline. Most of the alkaloids have a bond between the oxygen at C-17 and C-21, as in **Alstophylline**; and finally **Nareline**, a hitherto unique alkaloid with the aspidodasycarpine carbon skeleton, and an additional bond between C-21 and C-6.



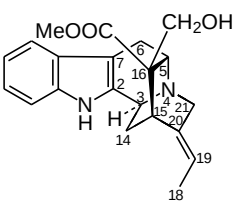
Akuammilan, 9CI



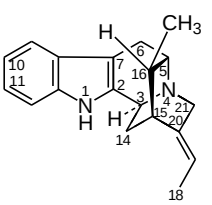
Macroline

Sarpagine alkaloids (VX5100)

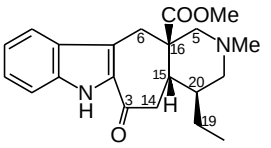
The Sarpagine (Akuammidine) group, based on the sarpagan nucleus, arises from bond formation between C-16 and C-5 of the corynantheine precursor and consists of simple Akuammidine derivatives; compounds in which the C-3 to N-4 bond has been severed, e.g. **Vobasine**; derivatives of the oxindole obtained following migration of C-3 from C-2 to C-7, e.g. **Gardneramine**; and a small group of miscellaneous bases, in which extensive rearrangement appears to have occurred. These may be exemplified by **Ervatamine, Ervitsine**, and **Koumine**. Inclusion of Ervatamine in this group receives support from the conversion of a dihydrovobasine (**Tabernaemontanine**) into Ervatamine *in vitro*.



Akuammidine



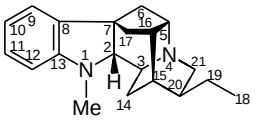
Sarpagan, 9CI



Ervatamine

Ajmaline alkaloids (VX5120)

The ajmaline group contains both 5, 16 and 7, 17 bonds. **Ajmaline** itself is the best known example. Almost all the bases in this group contain the same skeleton, but **Perakine** and **Raucaffrinoline** afford a rare structural variation in which the 21, N bond has been replaced by a 19, N bond.



Ajmalan, 9CI

Kingston, D.G.I. and Ekundago, D. (1981) *J. Nat. Prod.*, **44**, 509.

Pleiocarpamine alkaloids (VX5220)

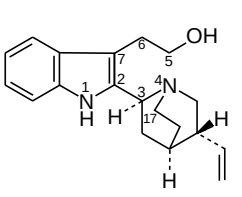
In this group a corynantheine precursor has cyclised *via* C-16 on to N-1, as in **Pleiocarpamine**.

Cinchona alkaloids (VX5240)

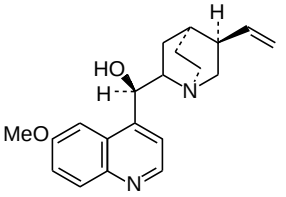
This important and well-known group, which includes the valuable antimalarial quinine, consists of two sub-groups:

(*a*) the Cinchonamine group, derived from a corynantheine-type precursor by fission of the N-4 to C-5 bond, and attachment of N-4 to C-17;

(*b*) the Quinine group, which contain a quinoline ring system generated from a precursor of the cinchonamine type by 2, 7 bond fission followed by bonding of N-1 to C-5.



Cinchonamine

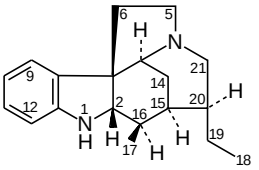


Quinine

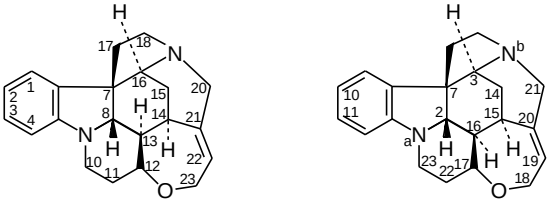
Uskokovic, M.R. and Grethe, G. (1976) in *MTP Series 2*, Vol. 9, *Alkaloids*, (ed. K. Wiesner), Butterworth, London.

Strychnos alkaloids (VX5280)

The *Strychnos* alkaloids are mainly based on the curan and strychnidine skeletons. The biogenesis presumably involves migration of C-3 in a corynanthe precursor from C-2 to C-7 followed by formation of the 2,16 bond. An early



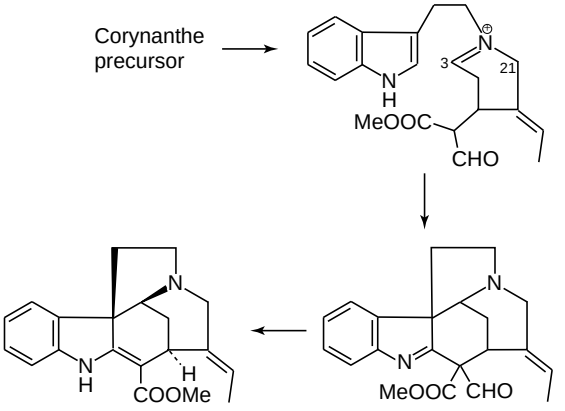
Curan (9CI)



Strychnidine, 9CI

(CA numbering) (biogenetic numbering)

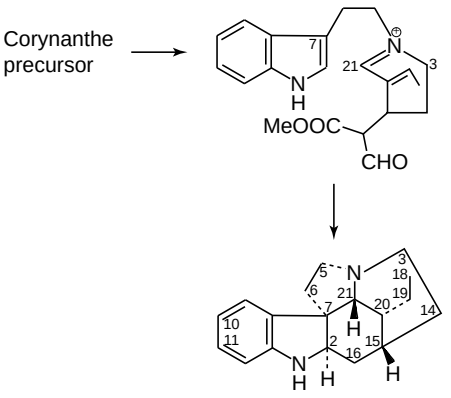
alkaloid in the curan group is therefore Preakuammicine, which loses formaldehyde to give Akuammicine. Completion of the strychnidine skeleton from the curan skeleton involves the addition of two carbon atoms, presumably from an acetate unit.



Akuammicine Preakuammicine

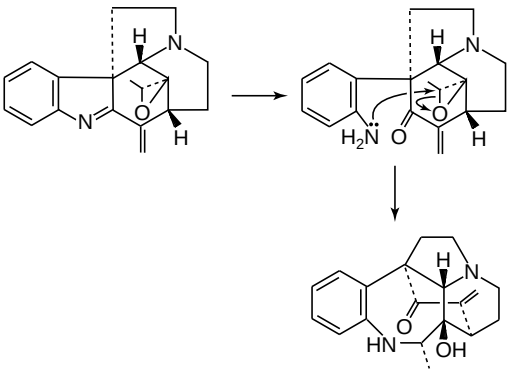
Condylocarpan alkaloids (VX5320)

These alkaloids contain a ring system similar to that of the curan group, but are formed by cyclisation of C-21 on to C-7 in a Corynanthe precursor, rather than the formation of a 3,7 bond; **Condylocarpine** is representative. Note that loss of the ethanamine carbons 5 and 6 gives the ring system of Uleine, which thus



Condylofolan, 9CI
(biogenetic numbering)

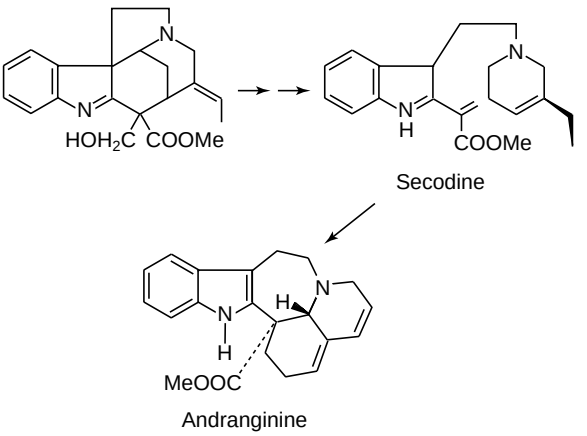
suggests an alternative biogenetic route to the one given below. Extensive modification of this skeleton appears to have occurred in the formation of Goniomine, which can be postulated to be formed by ring-opening and epoxidation of an indolenine related to Condylocarpine followed by N-1 to C-19 bonding:



Goniomine

Secodine alkaloids (VX5360, VX5380, VX4740)

This group of tricyclic alkaloids is formed by ring-opening of a precursor of the Preakuammicine type. The alkaloids occur in various stages of reduction, and in monomeric and dimeric forms. Andranginine, the product of an unusual cyclisation of a dehydrosecodine, may also be included here.



Secodine

Andranginine

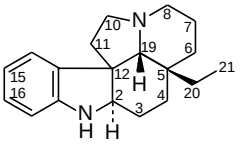
Aspidosperma alkaloids (VX5400)

The skeleton of the aspidospermidine alkaloids is formed by cyclisation of a dehydrosecodine, itself obtained from a precursor related to Preakuammicine.

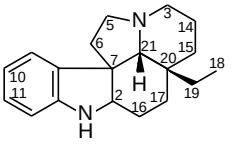
The alkaloids in this very large group are mainly based on the following structural variants:

(*a*) anilinoacrylate alkaloids, such as Tabersonine, which contain the methoxycarbonyl group at C-16. The two-carbon substituent at C-20 may be a simple ethyl group, or it may be functionalised;

(*b*) alkaloids lacking the C-16 methoxycarbonyl group, as in **Aspidospermine**. Again, C-18 and C-19 may be an ethyl group, or C-18 may be functionalised;



Aspidospermidine
(9CI numbering)



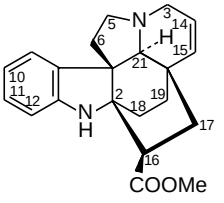
Aspidospermidine
(biogenetic numbering
used in DNP)

(*c*) alkaloids containing an ether or lactone bridge between C-18 and C-21;

(*d*) alkaloids containing an ether or lactone bridge between C-18 and C-15;

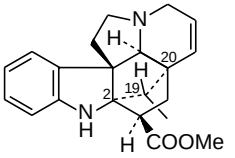
(*e*) alkaloids containing a lactone ring between C-18 and C-17, and a dihydro-1, 4-oxazine ring between N-1 and C-12, as in **Obscurinervidine**:

(*f*) alkaloids containing an additional bond between C-18 and C-2, as in Venalstonine;



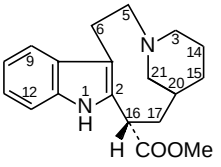
Venalstonine
(biogenetic numbering)

(*g*) alkaloids containing an additional bond between C-19 and C-2, as in Vindolinine;



Vindolinine

(*h*) the Quebrachamine group, which are derived by fission of the 7,21 bond. These may have lost the C-16 methoxycarbonyl group (e.g. **Quebrachamine**) or it may have been retained, as in Vincadine:



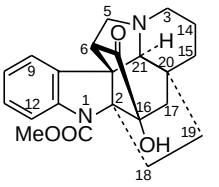
Vincadine

(*i*) miscellaneous alkaloids formed by a variety of other processes, e.g. **Aspidodispermine, Bannucine, Vincatine, Rhazinilam, Trichophylline, Meloscine, Melonine** and **Goniomitine**, which has undergone extensive rearrangement.

Overman, L.E. and Sworin, M. (1985) in *Alkaloids: Chemical and Biological Perspectives*, (ed. S.W. Pelletier), Vol. 3, Wiley-Interscience, New York.

Kopsane alkaloids (VX5560)

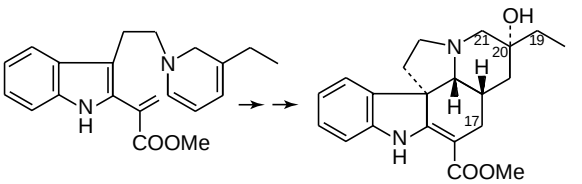
The skeleton of the kopsane group of alkaloids is simply formed by attachment of C-22 (the methoxycarbonyl carbon) of Venalstonine to C-6, as in Kopsine. Skeletal variations include the alternative acyloin structure, as in **Fruticosine**, in which C-22 is attached to C-17.



Kopsine

Quebrachamine and pandoline alkaloids (VX5500, VX5800)

The Pandoline nucleus can be imagined to be formed by cyclisation of a secodine derivative isomeric with that postulated as a precursor for Aspidospermidine:

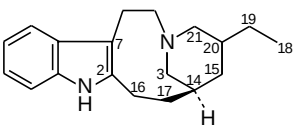


Pandoline

This group of alkaloids consists of:

(*a*) compounds containing the Pandoline (or Pseudoaspidospermidine) nucleus:

(*b*) the very small **Cleavamine** group, more often encountered as degradation products of other alkaloids, the nucleus of which may arise by fission of the 3,7 bond. Alternatively, and perhaps more likely, this ring system can be generated by fission of the 16,21 bond in an Iboga skeleton (see below).

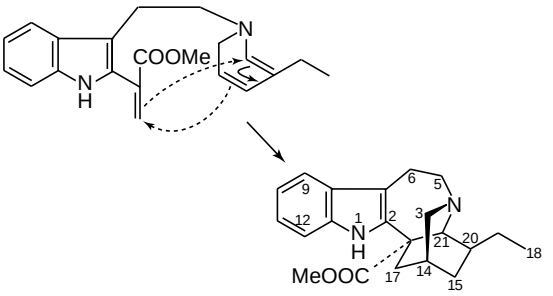


15,20-Dihydrocleavamine

(*c*) other variations in the skeleton, e.g. attachment of C-17 to C-21 (→ **Pandine**); enlargement of ring D, involving migration of C-21 from C-20 to C-19 (→ **Iboxyphylline**); contraction of ring D, involving loss of C-21 and attachment of C-20 to N-4 (→ **Ibophyllidine**).

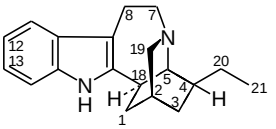
Iboga alkaloids (VX5700)

A third mode of cyclisation of a secodine-type precursor involves formation of a 16,21 bond, which gives rise to the ring system found in Ibogamine, **Catharanthine**, and numerous related alkaloids.



(–)-Coronaridine
(biogenetic numbering)

Note that the CA numbering is different.



Ibogamine, 9CI
(CA numbering)

This group of alkaloids exists in both enantiomeric series which may be defined by the chirality of C-14; thus (–)-Coronaridine is 14*R*, as shown above, and (+)-Coronaridine is 14*S*. The best-known example of the 14*S* series is probably **Catharanthine**.

Many alkaloids retain the methoxycarbonyl group, whereas others (e.g. Ibogamine) have lost it. Other variations include oxidation at C-7 to give the related hydroxyindolenines, e.g. **Ibogamine hydroxyindolenine**; oxidation followed by rearrangement to the related indoxyl, e.g. **Demethoxyiboluteine**; oxidative rearrangement to the corresponding oxindole, as in **Tabernoxidine**; and oxidation at C-19, C-3, C-5, or C-6 with, occasionally, ether formation between oxidised positions.

Pyridocarbazole alkaloids (VX5840)

This small, but pharmacologically important group is based on the 6*H*-pyrido[4,3-*b*]carbazole ring system, and is exemplified by Ellipticine and **Olivacine**.

Although these aromatic bases may superficially seem to be unrelated to the mainstream indole monoterpenoid alkaloids a possible biogenesis from Stemmadenine can be postulated, see Figure 19.

Gribble, G.W. and Saulnier, M.G. (1985) *Heterocycles*, **23**, 1277.
Kansal, V.K. and Potier, P. (1986) *Tetrahedron*, **42**, 2389.

Uleine-dasycarpidan alkaloids (VX5880)

This small group of alkaloids may well arise, like the Ellipticine group, from an oxidative fission of Stemmadenine. The genesis of the three types, e.g. Vallesamine, Uleine, and Apparicine can thus readily be explained (Figure 20).

Ngouniensine, yet another type of alkaloid that owes its origin to a stemmadenine-type precursor, may also be included in this group. The skeleton is so far unique in that it contains a 3,16 bond.

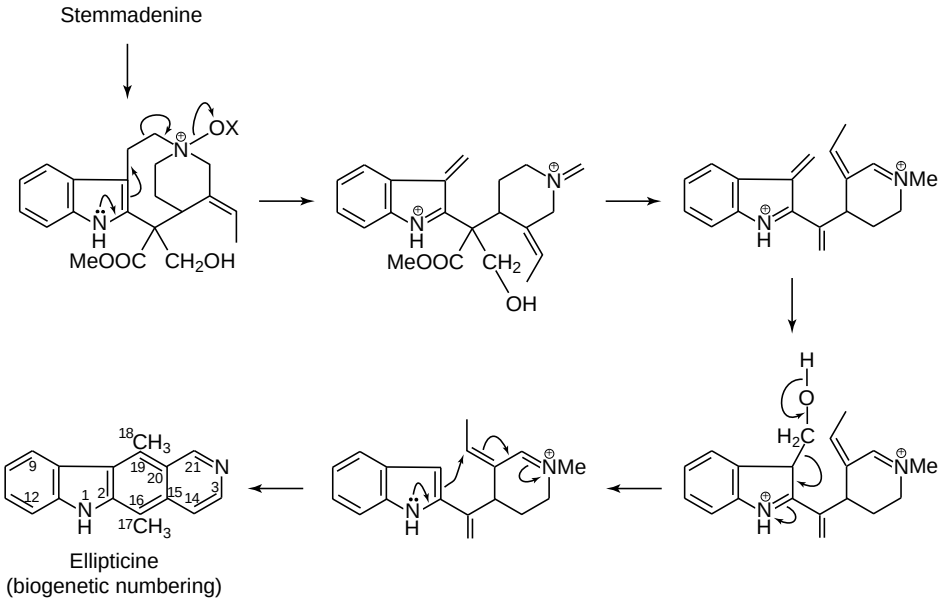


Figure 19.

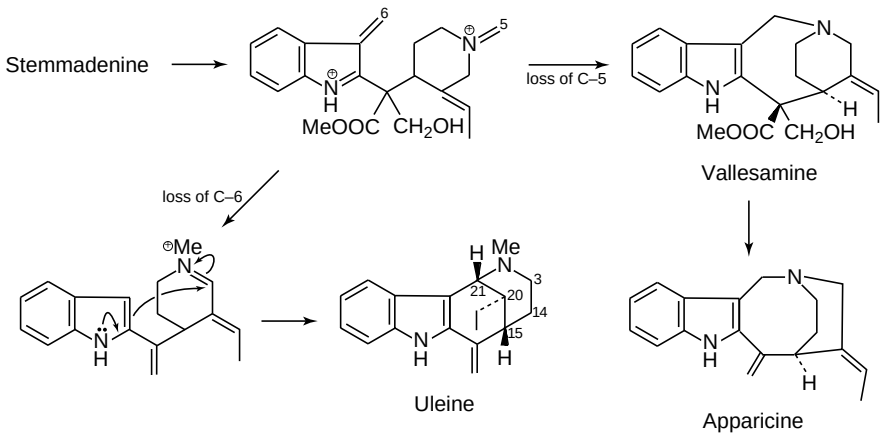


Figure 20.

Eburna alkaloids (VX5900)

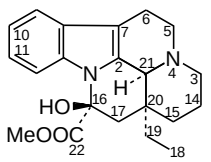
The skeleton of these alkaloids is generated by rearrangement of the aspidospermidine ring system, involving migration of C-21 from C-7 to C-2, fission of the 2, 16 bond, and attachment of C-16 to N-1. This rearrangement has been very successfully imitated *in vitro*.

The alkaloids consist of:

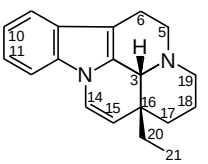
(*a*) Vincamine and its derivatives, which retain the methoxycarbonyl group;

(*b*) alkaloids such as **Eburnamine** and Eburnamenine, which have lost the C-22 ester group;

(*c*) some derivatives in which C-18 or C-19 is oxidised, as **Cuanzine**;



(+)-Vincamine
(biogenetic numbering)



Eburnamenine, 9CI
(CA numbering)

(*d*) The **Schizozygine** group, which contain an additional bond between C-2 and C-18;

(*e*) **Andrangine** and **Vallesamidine**, in which C-21 has simply migrated to C-2.

Bisindole alkaloids (VX5980)

This large group of complex alkaloids consists of a wide variety of structures, depending on the identity of the monomeric alkaloid components. Only the major sub-groups are listed here.

(*a*) alkaloids derived from a corynantheine-type unit which is attached *via* C-17 to another tryptamine unit, as in the **Ochrolifuanines**;

(*b*) alkaloids similar to those in sub-group (a) but in which further cyclisation has occurred, as in the **Roxburghines**;

(*c*) alkaloids derived from a vobasine unit, which is attached *via* C-3 to the aromatic ring of a second alkaloidal component, frequently an Iboga-type unit, or a Vobasine- or Sarpagine-type unit; **Conodurine** and **Accedinine** are examples;

(*d*) a clinically important group, in which a cleavamine-type unit is attached *via* C-16 to the aromatic ring of an Aspidosperma unit, usually Vindoline; **Vinblastine** and **Vincristine** are the best known examples;

(*e*) alkaloids derived by union of two units of the Strychnos type. Such alkaloids, which form the major constituents of calabash curare, are composed of two curan units linked *via* N-1 and C-17′, and N′-1 and C-17; ***C*-Toxiferine** is representative. In some alkaloids additional bonds are present; for example, ***C*-Curarine I** has an ether bridge between C-16 and C-16′, and ***C*-Calebassine** has an additional carbon-carbon bond between C-17 and C-17′;

(*f*) the **Vobtusine** group, which is composed of two aspidospermidine-type units linked by a spirocyclic system involving C-14 (two bonds) of one unit with C-22′ of another unit, together with an additional carbon atom attached to N-1′;

(*g*) several bases in which one component is macroline; the second component may be derived from pleiocarpamine, sarpagine, macroline, quebrachidine, or aspidospermidine;

(*h*) several bases in which one component is pleiocarpamine; the second component may be derived from vincorine, akuammiline, aspidospermidine, or tuboxenine; generally, the union of these two units involves two bonds;

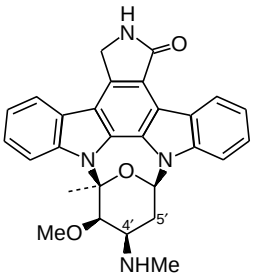
(*i*) the **Secamine** group, which is composed of two units derived from secodine;

(*j*) miscellaneous bisindole alkaloids containing one inter-unit bond;

(*k*) miscellaneous alkaloids containing two inter-unit bonds;

(*l*) a small group containing three inter-unit bonds as in **Ervafoline**.

(*m*) a small but rapidly-growing family of indolo[2,3-*a*]carbazole alkaloids and the related bis-indolylmaleimides. About sixty natural products that incorporate these ring systems are currently known. The prototype of this group is Staurosporine, originally isolated from *Streptomyces staurosporeus* AM-2282 and later found to be present in several other microorganisms. Others are the **Tjipanazoles** (from the blue-green alga *Tolypothrix tjipanasensis*) and several metabolites from slime moulds of the genus *Arcyria* (e.g. the **Arcyriarubins** and **Arcyriaflavins**).



Staurosporine

Gribble, G.W. and Berthel, S.J. (1993) *Stud. Nat. Prod. Chem.*, **12**, 365 (*indolocarbazoles*).
Lounasmaa, M. and Nemes, A. (1982) *Tetrahedron*, **38**, 223.

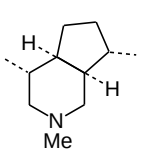
Terpenoid alkaloids

(excluding those involving tyrosine or tryptophan)

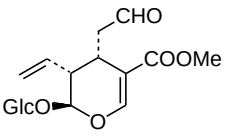
Monoterpenoid alkaloids (VX6240, VX6260)

These form a small but varied class; most of them are derived from iridoid precursors (see Terpenoid section) and may contain a pyridine or piperidine ring. The carbon skeleton is mostly C_{10}, but in many it is C_9 and in some it is C_{11}. There are two major groups:

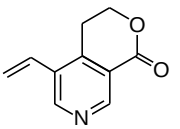
(*a*) those derived from iridodial-like precursors, e.g. α-Skytanthine;
(*b*) a diverse group derived from Secologanin, typified by Gentianine, **Bakankoside**, and **Gentioflavin**.



α-Skytanthine



Secologanin

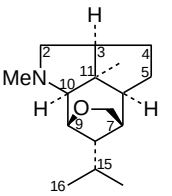


Gentianine

Dendrobium alkaloids (VX6340)

These alkaloids fall biogenetically into two quite distinct groups:

(*a*) a group of sesquiterpene alkaloids typified by **Dendrobine**, from *Dendrobium nobile*, with variants involving oxygenation at C-2 or C-6, and fission of the nitrogen to C-2 bond;



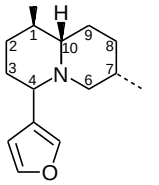
Dendrobane, 9Cl

(*b*) a group of indolizidine bases exemplified by **Crepidamine** and **Crepidine**, from *D. crepidatum*. These are probably not terpenoid in origin, and may be derived from shikimic acid, acetate, and ornithine.

Nuphar alkaloids (VX6360)

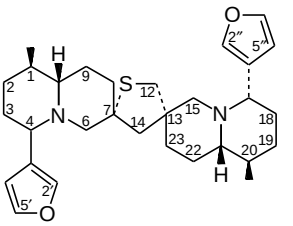
The *Nuphar* (water-lily) alkaloids contain a normal sesquiterpene carbon skeleton, and can be divided into three main sub-groups:

(*a*) the furylpiperidine derivatives, e.g. **Nuphamine**;
(*b*) the furylquinolizidine derivatives, e.g. Deoxynupharidine;



Deoxynupharidine

(*c*) a group of dimeric, sulfur-containing furylquinolizidines, e.g. Neothiobinupharidine.

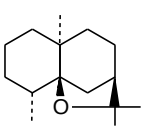


Neothiobinupharidine, 9CI

There are also a few miscellaneous bases which are based on variants of the above major sub-groups.

Macrocyclic sesquiterpene alkaloids (VX6320)

This group contains the ring system of dihydroagarofuran, a sesquiterpene of the eudesmane group (see Terpenoid section), esterified with nicotinic acid or with any of several dicarboxylic acids, e.g. **Evoninic acid**. Most of the alkaloids, which occur in *Euonymus* and *Maytenus* species, among others, contain a medium ring dilactone involving one of the dicarboxylic acids; **Evonine** is typical. Some alkaloids with two dilactone medium rings have also been isolated.

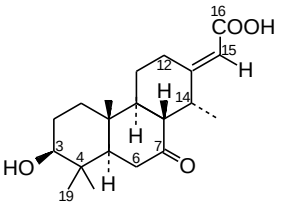


Dihydroagarofuran

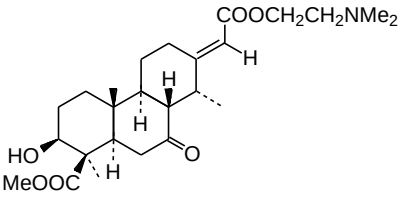
Erythrophleum alkaloids (VX6460)

The alkaloids of *Erythrophleum* species are based on the diterpene skeleton related to Cassaic acid. The oxidation pattern is relatively simple, involving only C-3, C-6, C-7 and C-19. In most alkaloids C-19 is at the carboxylic acid oxidation level. All the alkaloids are esters or amides of a C-16 carboxylic acid with *N*,*N*-dimethylethanolamine or *N*-methylethanolamine. Erythrophlamine is a typical example.

There have been some confusing structure revisions in this series.



Cassaic acid



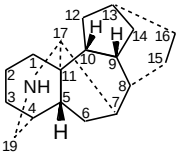
Erythrophlamine

C_{19} and C_{20} Diterpenoid alkaloids and 4-nor analogues (VX6400, VX6420)

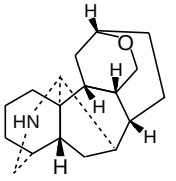
The alkaloids of this group may be divided into three major structural types, which can be further subdivided into twelve sub-groups. Although they are obviously diterpenoid in origin few biogenetic studies have been reported, apart from relatively early reports of the incorporation of acetate and mevalonate into **Browniine** and **Lycoctonine**, and of mevalonate and glycine into **Delcosine**.

The structural types are as follows:

(*a*) Alkaloids based on the C_{19} aconitane ring system. This accounts for the majority of the alkaloids of this group, which differ only in the pattern of substitution by hydroxy, methoxy, acetyloxy, benzoyloxy and, occasionally, other acyloxy groups in the ring system.



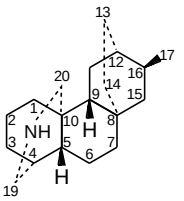
Aconitane, 9CI



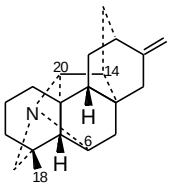
Heteratisane, 9CI

(*b*) A few alkaloids belong to the heteratisane group, formed from the aconitane framework by oxidative fission of the 13,14 bond.

(*c*) The second major group are the C_{20} alkaloids, based on atidane. Few alkaloids, as it happens, are based on the parent ring system, since many skeletal variations are known;



Atidane, 9CI



Hetisan, 9CI

(*d*) The hetisane group, in which additional rings are introduced into the atidane ring system by formation of 14,20 and N,6 bonds.

(*e*) A small group of atidane 7,20 cyclic ethers, as in **Ajaconine**.

(*f*) A small group of bases in which an additional carbocyclic ring is introduced by attachment of C-7 to C-20, as in **Denudatine**.

(*g*) Complex hetisane derivatives, e.g. **Delnudine**, in which further modification of the ring system has occurred. In the case of Delnudine this has involved the contraction of ring C.

(*h*) A group of dimeric atisines, exemplified by **Staphisine**.

(*i*) The third major group of alkaloids is based on the C_{20} veatchine skeleton, as in **Cuauchichicine**.

(*j*) 7,20-Cycloveatchine bases, e.g. **Lucidusculine**.

(*k*) 14,20-Cycloveatchine bases, e.g. **Anopterine**.

(*l*) Miscellaneous bases.

Benn, M.H. and Jacyno, J.M. (1983) in *Alkaloids: Chemical and Biological Perspectives*, (ed. S.W. Pelletier), Vol. 1, Wiley-Interscience.
Pelletier, S.W. and Page, S.W. (1976) in *MTP Series 2*, Vol. 9, *Alkaloids*, (ed. K. Wiesner), Butterworths, London.
Pelletier, S.W., Mody, N.V., Joshi, B.S. and Schramm, L.C. (1984) in *Pelletier*, Vol. 2.
Wiesner, K. (1985) *Tetrahedron*, **41**, 497.

Miscellaneous diterpenoid alkaloids (VX6480)

This category contains diterpenes linked by an ester function to a non-terpenoid nitrogen-containing unit. Examples are **Ryanodine** and **Taxine I**. Also included in this group are the indoloditerpenes. Some indole-diterpene metabolites have lost one carbon atom from the diterpene skeleton. These include **Paspalinine**, a potent tremorgen from *Claviceps paspali*. The penitrems, e.g. **Penitrem A**, which are mycotoxins from *Penicillium crustosum*, are yet more complex metabolites which have an affinity with Paspalinine but have an additional terpene unit attached to the aromatic ring.

Miller, R.W. (1980) *J. Nat. Prod.*, **43**, 425.

Olivoretin group

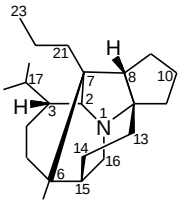
These, such as **Olivoretin D** (Teleocidin B), are metabolites of *Streptoverticillium olivoreticuli*, and show pronounced vesicant activity. **Teleocidin A1** is clearly terpenoid in origin, and so presumably are the other teleocidins.

Daphniphylline alkaloids (VX6500)

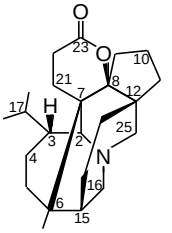
The alkaloids of *Daphniphyllum* species constitute a unique group of complex bases derived from squalene. They can be divided into six sub-groups which differ skeletally:

(*a*) the **Daphniphylline** group;
(*b*) the **Secodaphniphylline** group;
(*c*) the Daphnane (9CI) group, e.g. Daphnilactone A (note that the CA numbering differs from that most often used);
(*d*) the **Daphnilactone B** group;
(*e*) the **Yuzurimine** group;
(*f*) the **Yuzurine** group.

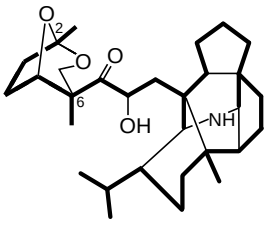
The secodaphniphylline group contain the carbon framework closest to that of squalene, one carbon atom having been transferred from C-2 to C-6. Subgroups (*a*) and (*b*) have a C_{30} skeleton, whereas those in (*c*)–(*f*) have lost a C_8 unit, and the Yuzurine skeleton has lost an additional carbon atom from the terminal isopropyl group.



Daphnane, 9CI



Daphnilactone A



Secodaphniphylline (heavy lines trace the precursor squalene)

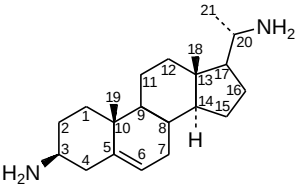
Yamamura, S. and Hirata, Y. (1976) in *MTP Series 2*, Vol. 9, *Alkaloids*, (ed. K. Wiesner), Butterworths, London.

Steroidal alkaloids

This very large group may be divided into nine subgroups. For further information on steroid structure and biosynthesis, see the Steroid section above.

Steroidal alkaloids (pregnane type) (VX6780)

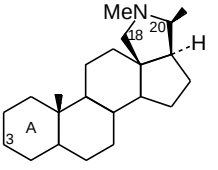
Pregnane steroids containing one or more amino groups at C-3 and/or C-20, such as Irehdiamine A, or with an amino group at C-18 as in the **Batrachotoxins**.



Irehdiamine A

Steroidal alkaloids (conanine type) (VX6700)

Alkaloids containing the conanine skeleton. Nearly all these bases contain an amino or an oxygen function at C-3.

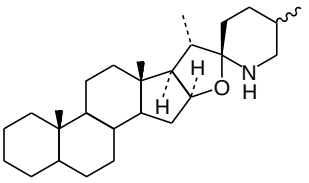


Conanine, 9CI

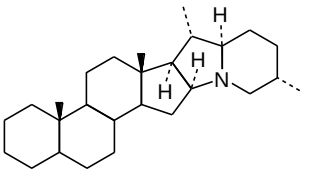
Steroidal alkaloids (spirosolane and solanidine type) (VX6660, VX6680, VX6720, VX6740)

Alkaloids in which a cholestane side-chain has been converted into:

(*a*) a piperidine ring, to give the secosolanidane skeleton;
(*b*) a bicyclic system containing a piperidine and a tetrahydrofuran ring to give the spirosolane skeleton;

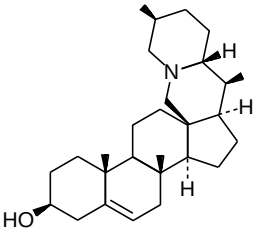


Spirosolane, 9CI



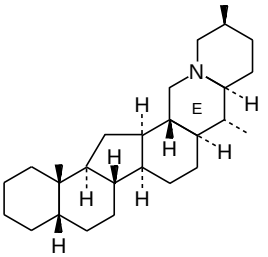
Solanidane, 9CI

(*c*) a bicyclic system to give the solanidane skeleton;
(*d*) a bicyclic system in which the piperidine nitrogen has been linked to the C-18 methyl group. This has been found only in Procevine so far, and is of special interest because it can be regarded as a precursor of the rearranged cevane skeleton.

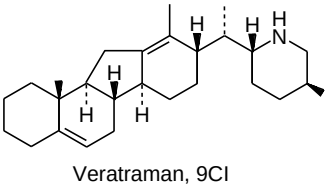


Procevine

(*e*) a pyrrolidine ring, as in **Tomatillidine**.
In addition, two sub-groups of alkaloids with rearranged skeletons are known. These are;
(*f*) the cevane group, in which ring D in a Procevine-type precursor has been enlarged at the expense of ring C;



Cevane, 9CI



Veratraman, 9CI

(*g*) the veratraman group, in which ring E of Cevane has been opened.

Steroidal alkaloids (buxus type) (VX6760)

The *Buxus* alkaloids are a large group of bases, the great majority of which fall into three sub-groups:

(*a*) those containing the pentacyclic 4,4,14-trimethyl-9,19-cyclopregnane skeleton. The majority of the *Buxus* alkaloids belong to this category.
(*b*) those containing a tetracyclic system in which 9,19 bond fission has occurred to give a seven-membered ring B.
(*c*) those alkaloids in which one or both of the carbon atoms attached to C-4 have been lost.

All the alkaloids have a nitrogen function at C-3 and/or C-20, which may be unmethylated, partially methylated, or fully methylated.

The suffix letters used in the nomenclature of this group indicate the degree of methylation of the nitrogen atoms:

	Substitution at N-3		Substitution at N-20	
	R^1	R^2	R^3	R^4
A	Me	Me	Me	Me
B	Me	Me	Me	H
C	H	Me	Me	Me
D	H	Me	Me	H
E	Me	Me	H	H
F	H	H	Me	Me
G	Me	H	H	H
H	H	H	H	Me
I	H	H	H	H
K	Me	Me	–	–
L	–	–	Me	Me
M	Me	H	–	–
N	–	–	Me	H
O	H	H	–	–
P	–	–	H	H

In DNP the entries for these alkaloids are organised under the (usually unknown) unsubstituted parents of the I, O, or P type.

Steroidal alkaloids (salamandra type) (VX6640)

In many of these alkaloids ring A has been enlarged, with incorporation of nitrogen, as in **Samandarine**.

Miscellaneous steroidal alkaloids (VX6790)

Non-nitrogenous steroids linked by an ester or acetal bond to a nitrogen-containing unit, as in **Bufotoxin**.

Imidazole alkaloids (VX6920)

This group, obviously derived from histidine, consists of:

(*a*) the *Pilocarpus* alkaloids, of which **Pilocarpine** is typical;

(*b*) miscellaneous bases obviously containing a histamine moiety, e.g. **Casimiroedine**;

(*c*) miscellaneous bases, e.g. **Isolongistrobine**.

Oxazole alkaloids (VX6930)

Upwards of thirty naturally occurring oxazoles are currently known. They have been isolated from various sources – plants of the Gramineae (e.g. **Annuloline**) and Rutaceae (e.g. **Halfordinol**), nudibranch egg masses (**Ulapualides**) and microorganisms. The latter have furnished the majority of the compounds, ranging from the simple indolyl alkaloids **Pimprinine, Pimprinethine** and **Pimprinaphine**, to complex peptide antibiotics such as the **Mikamycin/Streptogramin/Virginiamycin** family. The marine and bacterial oxazoles appear to have been formed from peptides of aliphatic amino acids

while the oxazoles of the Gramineae and Rutaceae arise from the chorismic acid-phenylalanine pathway.

Thiazole alkaloids (VX6935, VX6937)

More than 100 naturally occurring compounds that incorporate the thiazole moiety have been isolated to date. These alkaloids are a heterogeneous group ranging in complexity from **Aeruginoic acid** and the simple peptide **Herbamide A** to antineoplastic cyclopeptides such as **Ulicyclamide, Ulithiacyclamide, Patellamides** and **Dolastatins**.

Pyrazine and quinoxaline alkaloids (VX6940)

Pyrazines have been isolated from widely differing sources: from microorganisms, plants, mushrooms, animals, insects (especially ants, where they are considered to function as alarm pheromones) and more recently from marine organisms, where they are the actual light emitters in bioluminescence processes. A series of tetrahydroquinoxalines has been isolated from the scent gland of the Canadian beaver, *Castor fiber*.

Pyrazines also contribute to the aroma of various foodstuffs, including coffee, cocoa, tea and cooked meats, but from these sources they are generated by pyrolytic processes.

Pyrrole alkaloids (VX7010)

The pyrrole alkaloids are a heterogeneous group ranging in complexity from the very simple brominated pyrroles (e.g. **2,3-Dibromo-1*H*-pyrrole**), simple amino acids (**Kainic acid**) and peptides to the lipophylic **Malyngamides**, porphorins and other tetrapyrrole pigments (see following section). Compounds that incorporate the pyrrole moiety have been isolated primarily from marine sources (sponges, bacteria and algae) and microorganisms.

Putrescine alkaloids (VX7020)

These alkaloids can be subdivided into

(*a*) simple derivatives of putrescine with one or two cinnamic acid amide linkages, e.g. **4-Coumaroylputrescine, Feruloylputrescine (Subaphylline), Dicaffeoylputrescine**

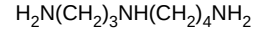
(*b*) derivatives of 2-hydroxyputrescine, e.g. ***N*-(4-Coumaroyl)-** and **N-Feruloyl-2-hydroxyputrescine**

(*c*) agmatine derivatives, e.g. **4-Coumaroylagmatine, Hordatine A**

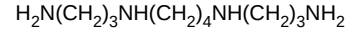
(*d*) miscellaneous, e.g. **Aerothionin, *N*-Carbamoylputrescine**.

Spermine and spermidine alkaloids (VX7030, VX7040, VX7050, VX7060, VX7070)

A number of alkaloids are derived from Spermine or Spermidine, themselves derived from ornithine *via* Putrescine.



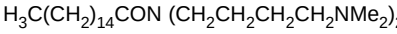
Spermidine



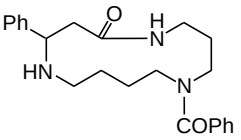
Spermine

Condensation of either Spermidine or Spermine with one or two cinnamic acid units, or with an unbranched carbon chain, gives the skeleton of these alkaloids. Aside from these aliphatic amines, therefore, phenylalanine or tyrosine, and long-chain fatty acids are involved. The biosynthesis clearly also involves phenol coupling processes in certain cases, e.g. **Codonocarpine**.

The spermidine alkaloids can be subdivided into four sub-groups: (*a*) simple diamides, e.g. **Maytenine**; (*b*) medium-ring compounds involving one cinnamic acid unit in the ring, e.g. Celabenzine; (*c*) medium-ring compounds involving two cinnamic acid units in the ring, e.g. **Codonocarpine**, **Lunaridine**; (*d*) medium ring compounds containing a C_{10} to C_{16} unbranched carbon chain in the ring, e.g. **Cannabisativine**; (*e*) derivatives of spermidine lengthened by one methylene group. This small group of amides of Solamine, so-called homospermidine alkaloids, was isolated from plants of the family Solanaceae. Solapalmitine is representative of this class.



Solapalmitine



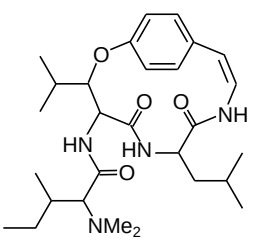
Celabenzine

The spermine-derived alkaloids may be divided into three sub-groups: (*a*) those in which the two terminal putrescine chains form an eight-membered ring with a cinnamic or a C_8 or C_{10} unbranched carbon chain, e.g. **Homaline**; (*b*) **Pithecolobine** in which the one large ring involves a C_{12} chain; (*c*) two alkaloids in which two medium rings are formed with two cinnamic acid units.

Hesse, M. and Schmid, H. (1976) Macrocyclic Spermidine and Spermine Alkaloids in *MTP Series 2*, Vol. 9, *Alkaloids* (ed. K. Wiesner), Butterworths, London.

Peptide alkaloids (VX7100)

There are now over 250 cyclopeptide alkaloids, which by definition are composed of a number of amino acids, among which phenylalanine or tyrosine are frequently found. Almost all of these alkaloids contain a medium ring (13–15 membered) incorporating a *β*-aminostyryl component. Examples are Frangulanine and **Zizyphine A**. As far as is known the component aminoacids have the L-configuration, with a few exceptions, e.g. D-phenylserine in **Lasiodine A** and D-*threo*-*β*-phenylserine and D-*erythro*-*β*-hydroxyleucine in **Scutianine E**.



Frangulanine

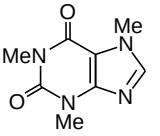
Amanita alkaloids (VX7120)

The toxins of the European death cap mushroom *Amanita phalloides* and other *A.* spp. constitute an even more complex group of macrocyclic peptides, mostly containing sulphur. These include the amatoxins (e.g. **α-Amanitin, β-Amanitin**), the phallotoxins (e.g. **Phalloidin**) and the virotoxins (e.g. **Viroidin**). The fly agaric (*A. muscaria*) also contains the low molecular weight compound **Muscarine**.

Purines (VX7300)

Purines are involved along with pyrimidines as bases in DNA and RNA. These and other purines may be divided into;

(*a*) the ubiquitous, well-known oxypurines, exemplified by Caffeine;
(*b*) derivatives of adenine, e.g. the plant hormone, **Zeatin**;
(*c*) miscellaneous.



Caffeine

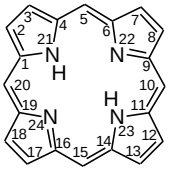
Pteridines and analogues (VX7350)

Pteridines are a widely distributed class of naturally occurring compounds. They owe their exceptional position in the field of heterocyclic chemistry mainly to their unusual chemical properties, their conspicuous fluorescence and their importance in metabolism, and partly to their discovery as pigments in butterfly wings. Three of the most common butterfly pigments are **Leucopterin, Xanthopterin and Isoxanthopterin**. The red pigments in the eye of the fruitfly *Drosophila melanogaster*, e.g. **Drosopterin, Isodrosopterin and Neodrosopterin** are complex pteridine derivatives. **Folic acid**, a water-soluble growth factor in bacteria and an anti-pernicious anaemia factor in animals is also a pterin derivative with a *p*-aminobenzoylglutamic acid sidechain at the 6-position. It occurs naturally as the dihydro derivative. Marine pteridines are represented by **Leucettidine** and **Urochordamines A and B**.

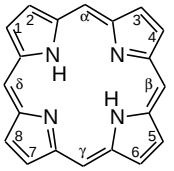
Polypyrroles (VY)

The polypyrroles (tetrapyrroles) are a numerically limited class of natural products that are mostly strictly functional. The main examples are haems, chlorophylls, bilins and **Vitamin B_{12}**. All types of organism use tetrapyrroles of one or more of these classes and all the functional tetrapyrroles derive from one common tetrapyrrolic intermediate, Uroporphyrinogen III (Uro'gen III). Uro'gen III is derived entirely from eight molecules of 5-Aminolaevulinic acid (ALA) by the action of three enzymes, *via* Porphobilinogen (PBG) and Hydroxymethylbilane (HMB) as intermediates. A particularly important feature in Uro'gen III is the fact that ring D has been inverted and so the acetate and propionate side-chains are not in the same order as on the other three pyrrolic rings, A to C. This feature can be found in virtually all naturally occurring tetrapyrroles. Some organisms, however, have a low activity of the enzyme uro'gen III synthase (as occurs in the human disease, congenital erythropoietic porphyria). In these cases non-enzymic cyclisation of HMB occurs to give Uro'gen I, which has the regular alternating pattern of the acetate and propionate side-chains, and a number of derived type I porphyrins can be isolated from these organisms.

The main system of nomenclature used in DNP is that recommended by the IUPAC-IUB Joint Commission on Biochemical Nomenclature. For the cyclic tetrapyrroles this is based on the porphyrin with the carbon atoms numbered 1 to 20 and the nitrogen atoms numbered 21 to 24. This has superseded the older 'Fischer' numbering which numbered only the eight β-positions of the five-membered pyrrole rings and labelled the four bridging *meso*-carbons α, β, γ and δ.

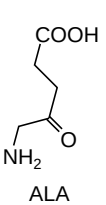


Porphyrin; IUPAC-IUB numbering

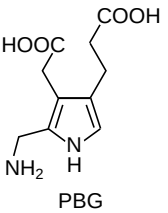


Porphyrin; Fischer numbering

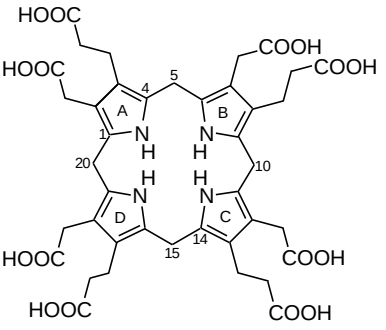
For natural porphyrins the IUPAC-IUB numbering starts on ring A and continues to rings B, C, and D, as shown below for Uro'gen III (ring D is always the inverted ring, see above). *Chemical Abstracts* on the other hand, though it uses the same 1 to 20 numbering for the carbon atoms, starts the numbering at such a position and in such a direction that the propionate side-chains get the lowest possible locants (thus for Uro'gen III the numbering would start at the position shown as 14 and proceed anticlockwise).



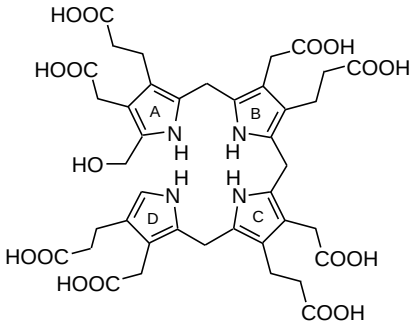
ALA



PBG

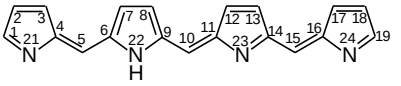


Uro'gen III



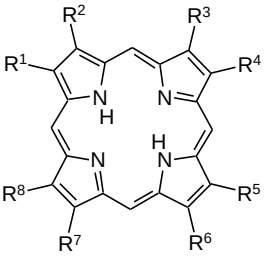
HMB

Certain reduced porphyrins have recognised names: **Chlorin** is 2,3-dihydroporphyrin, **Bacteriochlorin** is 7,8,17,18-tetrahydroporphyrin, **Isobacteriochlorin** is 2,3,7,8-tetrahydroporphyrin and **Porphyrinogen** is 5,10,15,20,22,24-hexahydroporphyrin.



Bilin; IUPAC-IUB numbering

Although several of the naturally occurring intermediates in tetrapyrrole biosynthesis are at the porphyrinogen oxidation level (e.g. Uro'gen III), these compounds are generally readily oxidised in air to the corresponding aromatic porphyrins. Thus it is the porphyrins that are isolated. In addition to the naturally occurring types I and III porphyrins (as explained above), DNP includes other isomers for comparison purposes in many cases. For **Uroporphyrin**, for example, assuming each ring has one acetate and one propionate side-chain, there are four possible isomers or types and these are given in the table below. In the **Protoporphyrin** series, however, there is a further degree of isomerism because two of the rings have a methyl and a propionate sidechain whereas two have a methyl and a vinyl sidechain. This results in 15 different types, numbered by Fischer I to XV as shown in the second table. Protoporphyrins I and II are related to Uroporphyrin I, III to V are related to Uroporphyrin II, VI to XI are related to Uroporphyrin III and XII to XV are related to Uroporphyrin IV. The naturally occurring Protoporphyrin is type IX. If the Roman numeral is omitted IX is assumed.



Type	R^1	R^2	R^3	R^4	R^5	R^6	R^7	R^8
I	A	P	A	P	A	P	A	P
II	A	P	P	A	A	P	P	A
III	A	P	A	P	A	P	P	A
IV	A	P	P	A	P	A	A	P

The substitution patterns for uroporphyrins I to IV (A = CH_2COOH, P = CH_2CH_2COOH)

Type	R^1	R^2	R^3	R^4	R^5	R^6	R^7	R^8
I	Me	V	Me	V	Me	P	Me	P
II	Me	P	Me	V	Me	P	Me	V
III	Me	V	V	Me	Me	P	P	Me
IV	Me	P	V	Me	Me	V	P	Me
V	Me	P	V	Me	Me	P	V	Me
VI	Me	P	Me	P	Me	V	V	Me
VII	Me	P	Me	V	Me	P	V	Me
VIII	Me	V	Me	P	Me	P	V	Me
IX	Me	V	Me	V	Me	P	P	Me
X	Me	V	Me	P	Me	V	P	Me
XI	Me	P	Me	V	Me	V	P	Me
XII	Me	V	V	Me	P	Me	Me	P
XIII	V	Me	Me	V	Me	P	P	Me
XIV	Me	P	V	Me	V	Me	Me	P
XV	Me	P	V	Me	P	Me	Me	V

The substitution patterns for protoporphyrins I to XV (V = $CH{=}CH_2$, P = CH_2CH_2COOH)

Porphyrins and porphyrinogens (VY0905)

The main biosynthetic pathway from Uro'gen III starts with the stepwise decarboxylation of each of the four acetate side-chains to give **Coproporphyrinogen III**, then oxidative decarboxylation of two of the propionate side-chains to give **Protoporphyrinogen**. These porphyrinogens and the partly decarboxylated intermediates are always isolated after aerial oxidation to give the corresponding porphyrin which is much more stable. Enzymic oxidation of Protoporphyrinogen gives **Protoporphyrin**, which is the branch point in the pathways to the haems and bilins and to the chlorophylls. Other porphyrins that can be found, in faeces for example, are mostly bacterial degradation products of Protoporphyrin with modification of the vinyl groups, e.g. **Mesoporphyrin, Deuteroporphyrin** and **Haematoporphyrin**.

Haems and metal-free haems (VY0910)

Insertion of Fe^{2+} into the centre of Protoporphyrin gives the uncharged Haem (American spelling, heme), also known as Haem b or Protohaem. This is the oxygen-carrying pigment of haemoglobin and myoglobin and the prosthetic group of Cytochrome b. Other cytochromes have closely related haems, e.g. haems a, c, d, and o. The Fe^{3+} form is called **Haemin**; it is positively charged and requires a counterion, as in Haemin chloride. **Haematin** is specifically the hydroxide salt. Many other metals can be inserted into porphyrins synthetically but apart. from magnesium (see Chlorophylls), zinc is the only metal that is sometimes found naturally as a result of disorders in iron metabolism.

Bile pigments (bilins) (VY0915)

In animals the degradative pathway for haem is *via* an oxidative ring cleavage to give **Biliverdin** (Biliverdin IXα). This is then reduced to **Bilirubin** (10,23-dihydrobiliverdin) and excreted, as a bis-glucuronide ester, through the bile duct into the gut, where further reduction of double bonds by bacteria occurs. The same oxidative cleavage of haem can be effected non-enzymically by the coupled action of oxygen and a reducing agent such as ascorbic acid. In this reaction, cleavage can occur at any one of the four *meso* positions (C-5, 10, 15 and 20) and thus four isomeric Biliverdins (IXα, β, γ and δ) are produced.

In plants the same oxidative cleavage of haem leads to the photoresponsive pigment **Phytochromobilin** and, in algae, to the light-harvesting pigments such as **Phycocyanobilin**. Both of these are found *in vivo* covalently attached to proteins by thioether links.

Chlorophylls and derivatives (VY0920)

Insertion of Mg^{2+} instead of Fe^{2+} into protoporphyrin is the start of the pathway that leads to the chlorophylls. A key intermediate in this pathway is **Protochlorophyllide**, in which the carbocyclic ring E, found in all chlorophylls and bacteriochlorophylls, has been formed in an oxidative cyclisation reaction. The chlorophyll c family, found in phytoplankton, have a porphyrin skeleton derived from Protochlorophyllide by insertion of a double bond into the propionate side-chain but the plant chlorophylls are all chlorins, having the C-17/18 double bond reduced in a photochemical, NADPH-dependent reduction of Protochlorophyllide giving **Chlorophyllide a**. Esterification with phytol gives **Chlorophyll a; Chlorophyll b** has, in addition, the 7-methyl oxidised to a formyl group.

There are a number of compounds in DNP which are the result of chemical degradation of chlorophylls and were used in the classical proof of its structure. For example **Phaeophytin a** is the magnesium-free derivative and **Phaeophorbide** has, in addition, the phytyl ester hydrolysed. Under more vigorous conditions further degradation occurs, especially of the sensitive β-ketoester functionality in ring E.

Bacteriochlorophylls and derivatives (VY0925)

Photosynthetic bacteria rely on a slightly more diverse range of tetrapyrrole pigments. Purple photosynthetic bacteria contain **Bacteriochlorophyll a**, which is a bacteriochlorin, having two opposite pyrrole rings reduced. Green sulfur bacteria on the other hand contain **Bacteriochlorophylls c, d** and **e**, which are in fact chlorins not bacteriochlorins and are each a family of pigments with varying numbers of extra methyl groups introduced onto the C-8 and C-12 side-chains. Other less common pigments have been named **Bacteriochlorophyll b** and **g** but **Bacteriochlorophyll f**, proposed to be the 20-desmethyl derivative of **Bacteriochlorophyll e**, has not yet been discovered.

Bacteriochlorophylls also have a wider range of esterifying groups than do the chlorophylls. Thus, whereas Bacteriochlorophyll a usually has the normal phytyl ester, Bacteriochlorophylls c, d and e commonly have a farnesyl group. Geranylgeranyl and straight-chain hydrocarbon esters are also found in some organisms.

Vitamin B_{12}, precursors and variants (VY0930, VY0935)

Another pathway of tetrapyrrole biosynthesis, found in bacteria, leads to **Vitamin B_{12}** and related compounds. Methylation of the macrocycle gives **Precorrin 2**, which is also thought to be the precursor of Sirohaem (prosthetic group of sulfite and nitrite reductases), **Factor F430** (cofactor involved in methanogenesis) and **Haem d_1,** (prosthetic group of Cytochrome cd_1, a dissimilatory nitrite reductase). Further intermediates on the route to Vitamin B_{12} are **Precorrin 6x** and either **Hydrogenobyrinic acid** or its cobalt derivative **Cobyrinic acid**.

Vitamin B_{12} can exist with the cobalt in oxidation state (I), (II) or (III), though (III) is the most stable. In this oxidation state the metal requires an upper axial ligand (the nucleotide loop provides the lower ligand). The normal

isolation procedure introduces a cyanide ligand (**Cyanocobalamin**) but water, hydroxide, or nitrite are other possibilities. The two biologically active forms of Vitamin B_{12} both have Co-C bonds: **Coenzyme B_{12}** has an axial adenosyl group and is the cofactor for a number of enzymic rearrangement reactions and **Methylcobalamin** has a methyl as the axial group and is involved in enzymic methyl-transfer reactions.

A number of close relatives of vitamin B_{12} have been found in various anaerobic bacteria. These have the dimethylbenzimidazoyl group of the nucleotide loop replaced by other substituted benzimidazoyl groups or by purines or even by simple phenoxy groups. The latter type of group cannot provide the lower ligand to the cobalt ion.

Geoporphyrins (VY0940)

A wide range of tetrapyrrole derivatives have been found in sedimentary deposits derived from organic matter, such as crude oil, oil shales and lignite. These porphyrins have undergone various degradative reactions: generally the macrocycle is at the stable porphyrin oxidation level but the side chains are fully reduced and very often decarboxylated; they very often occur as nickel or vanadyl complexes. Thus one of the most common geoporphyrins is **Deoxophylloerythroetioporporphyrin (DPEP)**, which could well be derived from degradation of Chlorophyll a. The arrangement of the substituents on the geoporphyrins gives a clue to their origin and thus the type of organic material which gave rise to the deposit. For example, geoporphyrins that lack a substituent at C-7 are assumed to derive from Chlorophyll b or related compounds, whereas those in which a cyclisation onto the C-17 side-chain has occurred (such as examples with a methylated five-membered ring) may derive from the Chlorophyll c family. Other geoporphyrins have been isolated which have the additional carbon atoms on the C-8 and C-12 side-chains characteristic of Bacteriochlorophylls c, d and e. The types of reaction undergone by the porphyrins also give useful information about the conditions that they have experienced over the history of the deposit.

Miscellaneous polypyrroles (VY0945)

Although the majority of tetrapyrroles found in organisms are the ones described above, widely-distributed and having specific well understood catalytic functions, there are a few that are found in individual organisms and have more obscure or unusual functions. Some are present purely as pigments, for example in certain birds' eggs and butterfly wings. Among the unusual functions are those of **Bonellin**, produced as a hormone by the females of a certain marine worm to ensure her offspring remain male, and Substance F, which is responsible for the bioluminescence of krill. Other examples, such as **Corallistin A** from a sponge or **Chlorophyllone a** from a type of clam, do not have any recognised function.

Blanche F. *et al*. (1995) *Angew. Chem., Int. Ed. Engl.*, **34**, 383 (*biosynth*).
Dolphin, D. (ed.) (1978–79) *The Porphyrins*, Vols 1–7, Academic Press, New York.
Dolphin, D. (ed.) (1982) *B_{12}*, Vols 1–2, Wiley, New York.
Jordan, P.M. (ed.) (1991) *Biosynthesis of Tetrapyrroles*, Elsevier, Amsterdam.
Leeper, F.J. (1985) *Nat. Prod. Rep.*, **2**, 19, 561; (1987) **4**, 441; (1989) **6**, 171 (*biosynth*).
Scheer, H. (ed.) (1991) *Chlorophylls*, CRC Press, Boca Raton.
Smith, K.M. (ed.) (1975) *Porphyrins and Metalloporphyrins*, Elsevier, Amsterdam.
Zagalak, B. and Friedrich, W. (eds) (1979) *Vitamin B_{12}*, de Gruyter, Berlin.
The Biosynthesis of the Tetrapyrrole Pigments *(1994) Ciba Foundation Symposium* , **180,** Wiley, Chichester.

Pure Appl. Chem., Vol. 71, No. 4, pp. 587–643, 1999.
Printed in Great Britain.

INTERNATIONAL UNION OF PURE AND APPLIED CHEMISTRY
ORGANIC CHEMISTRY DIVISION
COMMISSION ON NOMENCLATURE OF ORGANIC CHEMISTRY (III.1)

REVISED SECTION F: NATURAL PRODUCTS AND RELATED COMPOUNDS

(IUPAC Recommendations 1999)

Prepared for publication by

P. M. GILES, Jr
Chemical Abstracts Service, Columbus, OH 43210, USA

Revised section F: natural products and related compounds (IUPAC Recommendations 1999)

CONTENTS

Preamble 588
Introduction 589
Rules 590
RF-1. Biologically based trivial names 590
RF-2. Semisystematic nomenclature for natural products 591
RF-3. Fundamental parent structures 592
3.1 General guidelines for choice 592
3.2 Structural features allowed 592
3.3 Semisystematic names 592
3.4 Indicated hydrogen 593
3.5 Numbering 593
3.6 Identification of individual rings 593
3.7 Stereochemistry 594
RF-4. Skeletal modification of fundamental parent structures 594
4.1 Removal of skeletal atoms 594
4.2 Addition of skeletal atoms 594
4.3 Bond formation 598
4.4 Bond cleavage 598
4.5 Bond migration 601
4.6 Terminal ring removal 602
4.7 Combinations of operations 602
RF-5. Replacement of skeletal atoms 605
RF-6. Fusion of additional rings 606
RF-7. Bridged parent structures 610
RF-8. Bond order modifications 610
RF-9. Derivatives of parent structures 611
RF-10. Stereochemical modifications 614
References 618

Abstract: The nomenclature of natural products has suffered from much confusion, mostly for historical reasons. The isolation of a new substance, in the early days of the science, generally preceded its characterization by a lengthy period. Thus, these compounds were often assigned trivial names that gave no indication of the structure of the molecule and were often found afterwards to be misleading. Even when the original names were later revised (for example: glycerin to glycerol) the new names often expressed the structure imperfectly and were thus unsuitable for the nomenclatural manipulation that is required to name derivatives or stereoisomers. The result was a proliferation of trivial names that taxed the memory of chemists and obscured important structural relationships.

The resultant disorder in the literature led to the creation of committees of specialists with the task of codifying the naming of compounds in various connected areas of natural-product

chemistry, such as steroids, lipids, and carbohydrates. As far as their recommendations have been followed, their efforts have been successful in eliminating confusing or duplicate nomenclature.

It is the aim of the IUPAC Commission on Nomenclature of Organic Chemistry to unite as far as possible all the specialist reports into a single set of recommendations that can be applied in most areas of natural-product chemistry. Accordingly, provisional recommendations were prepared and published as Section F of the IUPAC Organic Nomenclature Rules, first in 1976 [1], and then in the 1979 edition of the Rules [2].

INTRODUCTION

Section F of these IUPAC Organic Nomenclature Rules is intended to help those working with natural products and related compounds to overcome two difficulties that are encountered frequently in their work:

a. A new compound has been isolated from a natural source, but its structure may be imperfectly known. To allow ready reference to such a compound, a *trivial* name may be coined for it according to the provisions of Rule RF-1, below.

b. A new compound has been isolated and its structure determined; its *systematic* name can thus be generated. However, this name may be too cumbersome to be continually inserted into the text of a scientific paper. To overcome this difficulty and show the close similarity to related compounds, a *semisystematic* name can be created in accordance with Rules RF-2 through RF-10.

From the above explanation, three definitions follow:

i. A *trivial* name is the name given for convenience to a new compound of yet uncertain structure. Such a name carries no (or minimal) structural information and is generally derived from the biological origin of the material. Trivial names should be formed according to the principles of RF-1, which also offers advice on how the change from a trivial to a semisystematic name can be accomplished in such a manner that the literature can still be retrieved without difficulty.

ii. A *systematic* name is one based on Sections A–E of the IUPAC Organic Nomenclature Rules [1] and on the more recent Guide to IUPAC Organic Nomenclature [3]. It will generally comprise stereo descriptors, followed by prefixes, the 'parent', and finally a suffix.

iii. A *semisystematic* name created according to Rules RF-2 through RF-10, provides a simplified alternative to the systematic name. This is normally achieved by the creation of a 'semi systematic parent.' There are two general types of semisystematic parents used for naming natural products and related compounds:

(a) *Parent hydrides*, which are structures that do not have terminal hetero atoms or functional groups and therefore consist only of skeletal atoms and hydrogen, for example, in steroid [4], terpene, carotene [5], and alkaloid nomenclature; this type of semisystematic parent can be treated according to the rules in Sections A–D of the IUPAC Organic Nomenclature Rules [2] and the Guide to IUPAC Organic Nomenclature [3], i.e. a suffix and prefixes indicating substituents and modifications to the skeletal structure may be added.

(b) *Functional parents*, which are structures that have certain terminal heteroatoms or groups, such as are found in carbohydrates, amino acids, and nucleosides; suffixes are usually not added to this type of semisystematic parent.

Those who wish to name a new compound will obtain the greatest benefit from the present rules if they bear the following in mind:

i. If the structure is unknown, a trivial name may be formed according to Rule RF-1.

ii. If the structure is known and not unduly complex, systematic nomenclature should be used [2,3].

iii. If the structure is known and complex, there are three possibilities:

(a). A semisystematic parent name describing most of the skeleton of the compound is already in existence in the literature. This should be used and the new name derived from it by the operations of substitutive nomenclature.

(b). The literature does not contain a semisystematic parent name that describes most of the skeleton of the compound, or a structure closely resembling that skeleton. In that case, a new parent name should be created according to Rules RF-2 through RF-7. If necessary, the degree of saturation or unsaturation is indicated according to Rule RF-8; prefixes and suffixes according to Rule RF-9 and stereo descriptors according to Rule RF-10 are added to arrive at the name of the new compound.

(c). The literature contains a semisystematic name describing a structure not identical with, but closely resembling the skeleton in question. In that case, a number of 'structure-modifying prefixes' (such as homo, seco) are available under Rule RF-4. Different skeletal atoms, additional rings and bridges can be described according to RF-5 through RF-7. Placed in front of the existing parent name, such prefixes modify the meaning so that it describes the skeletal structure of the new compound. Rules RF-8 through RF-10 are then brought into play to generate the name of the compound itself.

The appendix to this Section lists illustrative semisystematic parent names and the structures they define.

RULES

Rule RF-1. Biologically based trivial names

RF-1.1. When a compound is isolated from a natural source and a trivial name is required, the name should be based whenever possible on the family or genus or species name of the biological material from which the compound has been isolated. (Names based on a presumed metabolic activity of the isolated substance should not be generated.) As far as possible, the name should be selected to reflect the known or the likely distribution of the natural product. For example, a hypothetical species *Paradigma exemplare* (family Beispieliae) might yield products named on the basis of beispieliae, paradigma, or exemplare. If appropriate, the class or order might also be used as the basis for the name of a compound that occurs in a number of related families.

RF-1.2. The trivial name should not give a false implication as to structure or identity of principal groups present.

RF-1.3. The following groups of letters have significance as terminations in organic chemical nomenclature, and therefore should not be used as terminations to trivial names coined for natural products of unknown structure (except for the special case discussed in Rule RF-1.5 below).

—	—	al	—	am	an	ane	—	—	ate	—
–	—	—	—	—	en	ene	et	—	ete	—
ic	ide	—	ile	—	in	ine	—	—	—	ium
—	—	ol	ole	—	—	one	—	ose	—	—
		olide						oside		
—	yde	yl	—	—	—	yne	—	—	—	—

RF-1.4. As the vowel 'u' does not occur as the first letter in the above list, the ending -une, or where euphony so dictates -iune, has been chosen to indicate that the trivial name it terminates describes a compound of unknown structure (-une can be taken to symbolize *unknown* or the German *unbekannt*).

Such a name implies all substituent groups, including a principal characteristic group; hence, no prefixes or suffixes indicating substituent groups can be used. Thus, products derived from the hypothetical species *Paradigma exemplare* (family Beispieliae) noted above, might be named paradigmune, exemplarune, or beispiefiune.

> Note: In the past, the ending -ine has been commonly used for compounds of unknown structure; however, this ending is not recommended under Rule RF-1.3 because of its significance in Hantzsch-Widman ring nomenclature (R-2.3.3).[3]

RF-1.5. Two or more compounds isolated from the same source or obtained by subsequent separation from a substance previously considered as homogeneous may be differentiated by adding a capital letter, e.g. A, B, etc., to a name formed according to RF-1.1 through RF-1.4. Further refinements can be distinguished by the addition of subscript numbers to these letters.

RF-1.6. Names ending in -une or -iune are only temporary in that some unknown function and/or skeleton exists. As soon as the terminal heteroatoms and groups attached to a parent skeletal structure are identified, the ending is changed to -ane, or other suitable ending (see RF-3.3), and the terminal heteroatoms, groups, and other features should be expressed as a suffix and prefixes in the usual manner. For example, the compound paradigmune is found to be fully saturated and have an oxo and two hydroxy substituents; the name then becomes dihydroxyparadigmanone.

RF-1.7. As soon as the structure of a new natural product has been fully determined, if it is relatively simple the trivial name established above in RF-1.4 through 1.6 should be abandoned in favor of a systematic one based on the established principles of organic nomenclature [2,3]. For a more complicated structure, if a previously known parent structure can be easily used to derive a semisystematic name for the new natural product, the trivial name should be abandoned in favor of such a semisystematic name based on the name for the established parent structure. If a previously known parent structure cannot be found, the stem name assigned above becomes that of a new parent which is numbered according to the rules given below.

RF-1.8. If it is subsequently found that the compound is identical with a known natural product structure, the trivial or semisystematic name formed later should be abandoned in favor of the previously recorded one, or a name derived systematically from it. In the latter case, the new name should be as close as possible to the one recorded earlier.

Rule RF-2. Semisysternatic nomenclature for natural products

Many naturally occurring compounds belong to well-defined structural classes, each of which can be characterized by a set of parent structures that are closely related structurally, that is, each can be derived from a fundamental structure by one or more well-defined operations.

RF-2.1. A semisystematic name for a naturally occurring compound or a synthetic derivative should be based on the name of an appropriate fundamental parent structure as described in RF-3.

> [Examples of fundamental parent structures and names for a variety of natural compound classes are given in the Appendix.]

RF-2.2. To the name of a fundamental parent structure are added affixes denoting: (1) modifications to the skeletal structure (RF-4, 6, and 7); (2) replacement of skeletal atoms (RF-5); (3) changes in the state of hydrogenation implied by the name of the parent structure (RF-8); (4) atoms or groups substituting hydrogen atoms of the parent structure (RF-9); (5) configurations not already implied by the name of the parent structure, or changed from that implied (RF-10). Unless specified otherwise in the rules that follow, methods of construction of the name and the principles of organic nomenclature as given in the IUPAC Organic Nomenclature Rules [2,3] are followed.

Rule RF-3. Fundamental parent structures

RF-3.1. General guidelines for choosing a fundamental parent structure.

RF-3.1.1. A fundamental parent structure should reflect the basic skeleton (including nonterminal hetero atoms and hetero groups) that is common to most compounds in that class.

RF-3.1.2. Fundamental parent structures should be chosen so that as many related natural products as possible can be derived from each by well-defined operations and principles of organic nomenclature.

RF-3.1.3. A fundamental parent structure should include as much stereochemistry as possible that is common to the relevant class of natural products.

RF-3.2. Structural features allowed for fundamental parent structures:

RF-3.2.1. A fundamental parent structure should only exceptionally include rings that are part of a characteristic group, such as a lactone or cyclic acetal. However, there are a number of widely used names that do include cyclic characteristic groups, for example, spirostan and cardenolide and these are allowed by these rules.

RF-3.2.2. A fundamental parent structure should not contain terminal hetero atoms or groups, except as provided by Rule RF-3.2.5.

RF-3.2.3. A fundamental parent structure should contain acyclic hydrocarbon groups that occur in most of the compounds in the natural product class.

RF-3.2.4. A fundamental parent cyclic structure should be as nearly fully saturated, or fully unsaturated in terms of the maximum number of noncumulative double bonds, as possible, while still representing the level of saturation (or unsaturation) of as many related compounds as possible. This principle should not be applied so rigorously that a large number of unsaturated bonds must be expressed by subtractive suffixes or that a large number of unsaturated bonds must be removed by hydro prefixes.

RF-3.2.5. Certain large classes of natural products, such as carbohydrates, nucleosides and peptides, are characterized by the occurrence of relatively simple but highly functionalized units and their oligomers and polymers. For these classes of compounds, called fundamental functional parents (see Introduction), the provisions of RF-3.2.2 do not apply in that fundamental parent structures may also contain functional groups.

RF-3.3. A semisystematic name for a fundamental parent structure should be derived as far as possible from a trivial name formed according to Rule RF-1. The endings to be used in place of '-une' or '-iune' should follow the following guidelines:

(a). '-an(e)', if the entire parent structure is saturated;

(b). '-en(e)', if the cyclic or the main chain of the acyclic part, or both, of the parent structure contain the maximum number of noncumulative double bonds;

(c). '-aran(e)', if, in an otherwise fully saturated parent structure, there occur one or more individual rings that contain the maximum number of noncumulative double bonds. Most examples of such substances already have names ending in '-an' (see Note 2, below) or are alkaloids whose names end in '-ine', e.g. ergoline, aspidospermidine, and strychnidine.

Note 1: In these rules the final 'e' will be used. The omission ofthis final 'e', or the use of a final ''a', may occur in languages other than English.

Note 2: The ending '-an' has been used for names of some heterocyclic parent structures having partial unsaturation, e.g. morphinan, rheadan, yohimban. Hence, even though the ending '-aran(e)' has been suggested above for structures of these types, no change to '-aran(e)' is required for names already established. However, for such compounds, the '-an(e)' ending cannot be used to indicate fully saturated analogs, which must be described using hydro prefixes (see RF-6).

RF-3.4. Indicated hydrogen, as described in Rule A-21.6, may be used to describe isomers of fundamental parent structures that have saturated skeletal atoms within a ring system or portion of a ring system having the maximum number of noncumulative double bonds.

RF-3.5. Numbering of fundamental parent structures.

RF-3.5.1. A numbering pattern established among a group of structurally related natural products is used for numbering the skeletal atoms of the fundamental parent structure, providing all skeletal atoms have been included in the numbering system.

RF-3.5.2. If no numbering pattern has become established among the members of a group of structurally related natural products, the fundamental parent structure is numbered according to the following guidelines:

(a). The skeleton is examined to identify a preferred ring system. This will be the ring system defined as 'senior' according to the criteria of Rule C-0.14 [2]. The locant '1' is assigned to the atom of the preferred ring system whose locant would be '1' according to systematic numbering for that particular ring system.

(b). All skeletal atoms of the preferred ring system are assigned consecutive Arabic numbers, including atoms at fusion positions in fused ring systems, beginning with the locant '1', and following the path prescribed for systematic numbering of that particular type of ring system.

(c). Acyclic substituents to skeletal atoms of ring components or connecting acyclic structures are numbered each in its entirety, including branches, in order of the increasing value of the locant of the skeletal atom to which each is attached.

(d). Skeletal atoms of acyclic connections to other rings or ring systems, if any, are numbered consecutively beginning with the atom next to the preferred ring system, followed by the skeletal atoms of the other rings or ring systems as prescribed by (b) above; if two or more acyclic connections to other rings or ring systems are present, the one attached to the preferred ring system at the lowest numbered position is numbered first, then the ring attached to it, followed by the acyclic connector at the next lower numbered position of the preferred ring or ring system, etc.

(e). Between two groups at a gem-disubstituted position, the larger group, in terms of the number of skeletal atoms, is numbered first; if there is still a choice, the principles of Rule C-15.11(c)–(e) [2] are followed. If the two groups are then identical and attached to a cyclic structure properly drawn (see Appendix) the group stereochemically α is numbered first (see RF-4.5.1); if the two groups are identical and attached to an acyclic structure, the group *trans* to the main chain is numbered first as described in the carotenoid recommendations Rule 12.4

RF-3.6. Identification of individual rings. Certain modifying prefixes for names of fundamental parent structures (see Rule RF-4) have used a ring identifier rather than specific locants of skeletal atoms. Accordingly, the identification of individual rings for some of the more common natural products has become well established. However, since in these recommendations locants of skeletal atoms are used to describe structural modifications instead of letters, except for the rather special case of removal of a terminal ring (see RF-4.6), no attempt has been made here to codify a system for lettering rings. Nevertheless, to provide continuity with the use of this system, names using letters to identify rings are given where appropriate.

RF-3.7. Stereochemical configuration of fundamental parent structures. The name of the fundamental parent structure implies, without further specification, the absolute configuration at all chiral centers and the configuration at double bonds corresponding to the group of natural products from which the parent name was derived, except as specified in these rules, or in rules for specific kinds of natural product compounds. Such stereochemistry for many fundamental parent structures is shown by the drawings in the Appendix. In some instances the configuration at a center is not implied by the parent name and must always be specified.

Rule RF-4. Skeletal modifications of fundamental parent structures.

RF-4.1. Removal of skeletal atoms without affecting the number of rings.

RF-4.1.1. The removal of an unsubstituted skeletal atom, saturated or unsaturated, from a ring or an unsubstituted skeletal atom from an acyclic portion of a fundamental parent structure with its attached hydrogen atom(s) is described by the prefix 'nor-'; the loss of two or more skeletal atoms is indicated by combining an appropriate numerical prefix with 'nor-', e.g. 'dinor .... trinor-', etc.

Note 1: The provisional Section F Rules [1,2] require the skeletal atoms removed to be saturated carbon atoms by using the prefix 'nor' to indicate the removal of methylene groups. The carotenoid recommendations (5(a)) [5] provide that 'nor' be used to indicate the removal of CH groups as well. These revised recommendations are more precise by permitting removal of CH groups only in a ring having the maximum number of noncumulative double bonds; they are also more general by allowing 'nor' to indicate the removal of hetero atoms.

Note 2: The special use of the prefix 'nor-' without multiplying prefixes or locants to indicate the replacement by hydrogen atoms of all methyl groups attached to the ring is discouraged.

The position of the skeletal atom that is removed is denoted in all cases by its locant in the numbering of the parent structure.

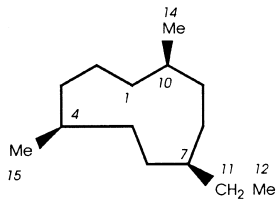
Note: A capital letter, associated with the locant of a skeletal atom where needed, has been used with prefixes such as 'nor-' and 'dinor-' to indicate removal of methylene groups from a particular ring. This system is used in *Chemical Abstracts* index nomenclature, but is not included in Section F because it is not as general as the locant system recommended here.

Although, because the locant of each skeletal atom removed is cited, an unambiguous name can be generated by the removal of any skeletal atom, it is traditional to remove skeletal atoms with the highest possible locant in an atomic connector in a cyclic portion of the skeletal structure. An atomic connector is a chain of homogeneous skeletal atoms of the same element connecting any combination of bridgehead or ring junction atoms, rings or ring systems (i.e. ring assemblies), substituted skeletal atoms in the parent structure, or heteroatoms. In an acyclic portion of a skeletal structure, the skeletal atom removed preferably is the one of an acyclic atomic connector or a terminal segment nearest to the free end of the acyclic part of the structure. (This is done in order to maintain as far as possible traditional numbering of structural features of the compound and of compounds derived from it.) A terminal segment of a skeletal structure is an acyclic segment of homogeneous skeletal atoms connected at only one end by the features of structure that terminate atomic connectors (see above) (Scheme 1, 2).

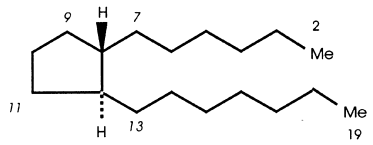
RF-4.1.2 When the removal of an unsaturated skeletal atom from a ring containing the maximum number of noncumulative double bonds in the fundamental parent structure results in the creation of a saturated ring position, this position is described by indicated hydrogen symbolism. When there is a choice, the indicated hydrogen symbol (*H*) is assigned to the lowest numbered nonangular position in the unsaturated portion of the structure (e.g. Scheme 3).

RF-4.2. Addition of skeletal atoms without affecting the number of rings.

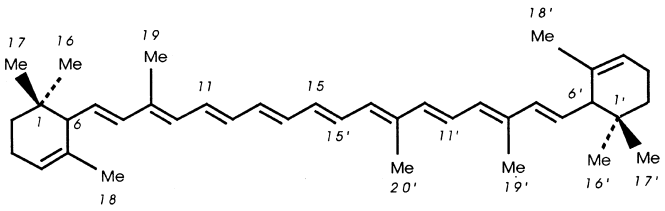
RF-4.2.1. The addition of a methylene (-CH_2-) group between two skeletal atoms of a fundamental parent structure is described by the prefix 'homo-'; the addition of two or more methylene groups is indicated by combining an appropriate numerical prefix with 'homo-', e.g. 'dihomo', 'trihomo-', etc.



13-Norgermacrane

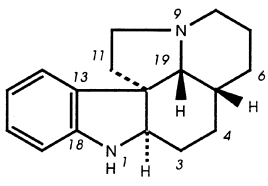


1,20-Dinorprostane

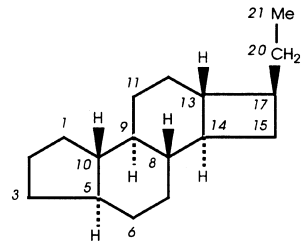


20-Nor-ε,ε-carotene

Scheme 1

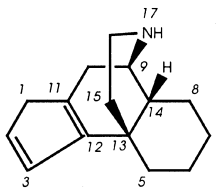


20,21-Dinoraspidospermidine



4,16,18,19-Tetranor-5α-pregnane
(has been called *A,D*(15),18,19-Tetranor-5α-pregnane)

Scheme 2



1*H*-4-Normorphinan

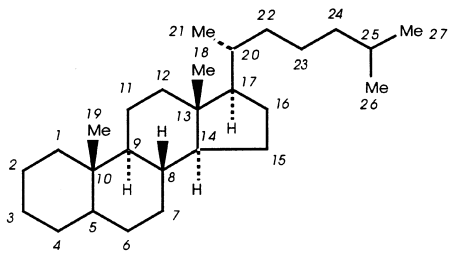
Scheme 3

Positions of inserted methylene groups in the modified fundamental parent structure are indicated by the locants of the added methylene group which are cited in front of the prefixes, 'homo', 'dihomo-', etc.

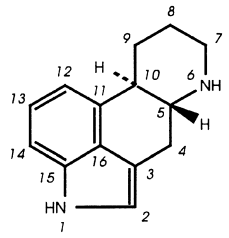
Note: Capital letters, associated with locants of added methylene groups, where needed, have been used with prefixes such as 'homo-' and 'dihomo-' to indicate insertion of methylene groups into particular rings. This system is used in Chemical Abstracts index nomenclature, but is not included in these Section F recommendations because it is not as general as the locant system recommended here.

The assignment of the locant to an added methylene group depends on whether it is considered to be

inserted into an atomic connector or terminal acyclic segment (see also RF-4.1.1) or into a bond connector. A bond connector is a connection between any combination of bridgehead or ring junction atoms, rings or ring systems (i.e. ring assemblies), substituted skeletal atoms, or heteroatoms. The structures below illustrate atomic connectors, bond connectors, and terminal segments (Scheme 4).



Cholestane



Ergoline

Scheme 4

Atomic connectors:

In cholestane: 1–4, 6–7, 11–12, 15–16 and 22–24.

In ergoline: 2, 4, 7–9 and 12–14.

Terminal segments:

In cholestane: 18, 19, 21, 26 and 27.

In ergoline: None.

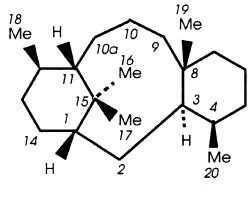
Bond connectors:

In cholestane: 5–10, 8–9, 8–14, 9–10, 13–14, 13–17 and 17–20.

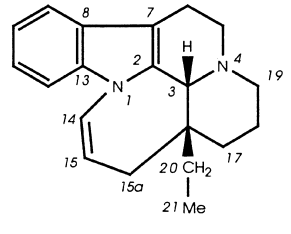
In ergoline: 1–15, 3–16, 5–6, 5–10, 10–11 11–16 and 15–16.

RF-4.2.2. Numbering of additional skeletal atoms.

RF-4.2.2.1. Methylene groups inserted into an atomic connector or into a terminal segment are identified by adding a letter ‘a’, ‘b’, etc., to the locant of the highest numbered skeletal atom of the atomic connector or terminal segment consistent with the location of double bonds remaining in the structure (compare example 2 below and example 3 under RF-4.2.2.2). If there are equivalent atomic connectors, the highest numbered atomic connector is chosen, and the methylene group is inserted after the highest numbered skeletal atom in that connector (e.g. Schemes 5, 6).



10a-Homotaxane

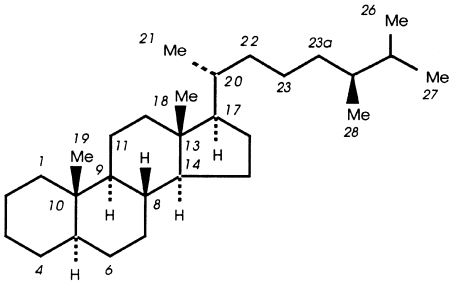


15a-Homoeburnamenine

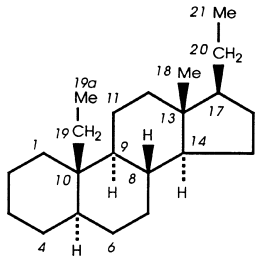
Scheme 5

Note: Addition of acyclic side chains or extension of terminal segments of a side chain already attached to the basic skeleton of a fundamental parent structure may also be done by principles of substitutive nomenclature.

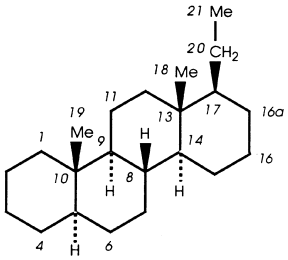
RF-4.2.2.2. Methylene groups inserted into a bond connector are identified by citing both locants of



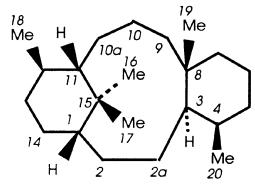
23a-Homo-5α-ergostane



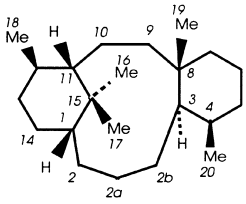
19a-Homo-5α-pregnane



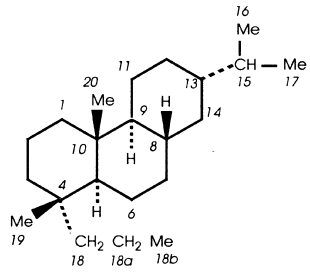
16a-Homo-5α-pregnane



2a,10a-Dihomotaxane



2a,2b-Dihomotaxane



18a,18b-Dihomoabietane

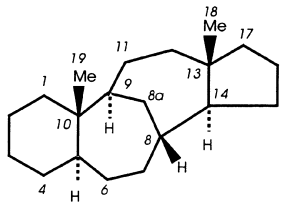
Scheme 6

the skeletal atoms terminating the bond connector enclosing the second (higher) number in parentheses, followed by a letter ‘a’, ‘b’, etc. according to the number of methylene groups.

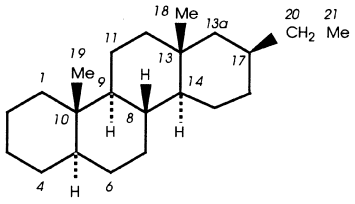
Note: The insertion of a methylene group into a bond connector has been described by combining the capital letter(s) of the expanded ring(s) with the locant of the inserted atom derived by adding a letter ‘a’, ‘b’, etc., to one of the locants for the skeletal atoms terminating the ring bond connector (Rule 3S-7.3 [4]) (e.g. Scheme 7).

RF-4.2.3. The insertion of a methylene group into a ring or a ring system of a fundamental parent structure that contains the maximum number of noncumulative double bonds or into a cyclic system of conjugated double bonds may create a saturated ring position that is indicated by ‘indicated hydrogen’. The position of the inserted methylene group is prescribed by Rule RF-4.2.2, even though the saturated ring position may be elsewhere in the unsaturated ring system as denoted by the appropriate locant for the indicated hydrogen (e.g Scheme 8).

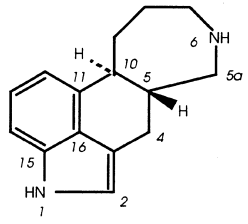
Note 1: Although it would be quite easy to justify the omission of indicated hydrogen when the saturated position is the same as the inserted methylene group, this should not be encouraged.



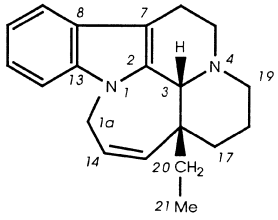
8(9)a-Homo-5α-androstane



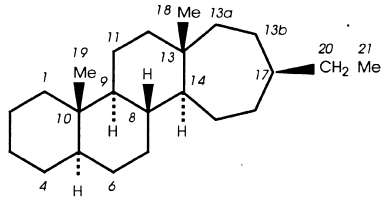
13(17)a-Homo-5α-pregnane



5(6)a-Homoergoline

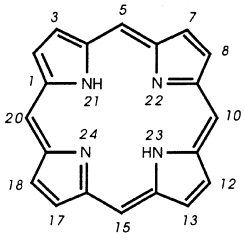


1(14)a-Homoeburnamenine

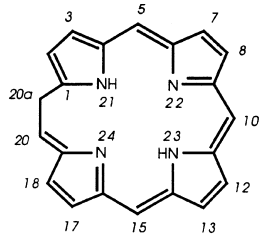


13(17)a,13(17)b-Dihomo-5α-pregnane
(has also been called
D(17a,17b)-Dihomo-5α-pregnane)

Scheme 7



Porphyrin (fundamental parent structure)



20a*H*-20a-Homoporphyrin

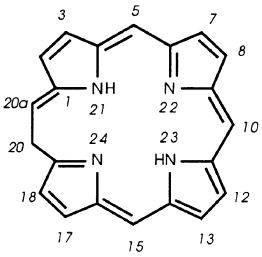
Scheme 8

Note 2: The name 20*H*-20a-Homoporphyrin describes an alternative tautomeric form shown in Scheme 9.

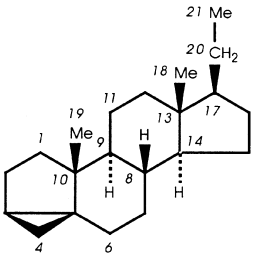
RF-4.3. Bond formation. The creation of an additional ring by means of a direct link between any two atoms of a fundamental parent structure is described by the prefix cyclo preceded by the locants of the skeletal atoms so connected. Where necessary, the stereochemical configurations created by the new bond are denoted by α, β, or ξ as described under RF-10 (e.g. Scheme 10, 11).

RF-4.4. Bond cleavage

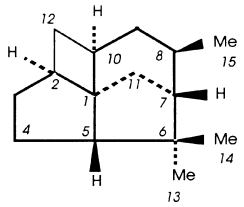
RF-4.4.1. Cleavage of a ring bond (saturated or unsaturated) with addition of the appropriate number



Scheme 9

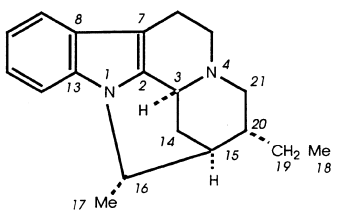


3α,5-Cyclo-5α-pregnane
(note that the cyclo bond in this drawing is α relative to the other three bonds)

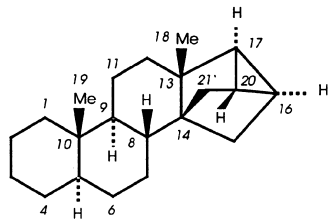


10β,12-Cyclocedrane
(note that the locants for the methyl groups at position 6 were reversed in the provisional Section F Rules[1,2])

Scheme 10



16β*H*-1,16-Cyclocorynan
(see RF-4.5.1 for the use of β*H*)



(20*S*)-14,21:16β,20-Dicyclo-5α,14β-pregnane (see RF-10.2.1 for the use of β at position 14)

Scheme 11

of hydrogen atoms at each new terminal group thus created, is indicated by the prefix 'seco-' and the locants of the cleaved bond. The original numbering is retained (e.g. Schemes 12, 14).

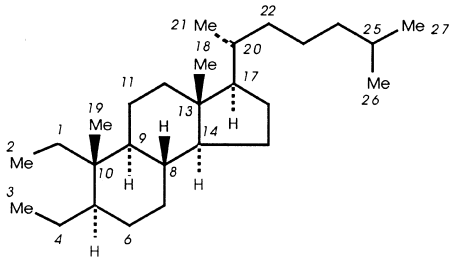
Note: The configuration at positions 15 and 20 of the secoyohimban in Scheme 12 relates to the stereochemistry of the parent structure only if drawn in the same orientation.

Although having the same absolute configuration, the stereochemistry denoted by α- and β- at these positions could be reversed in an alternative orientation as shown in Scheme 13.

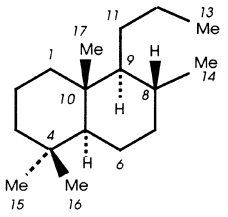
This situation has been recognized in the recommendations for naming Vitamin D compounds (9, 10-seco steroids [6]) where sequence rule descriptors (*R/S*) are recommended for describing all configurations in ring A, since these compounds are often drawn in an alternative orientation.

RF-4.4.2. The unitalicized prefix 'apo-' preceded by a locant is used to indicate removal of all of a side chain of a fundamental parent structure beyond the skeletal atom corresponding to that locant. Removal of two or more side chains is indicated by the prefixes 'diapo', 'triapo-', etc., preceded by appropriate locants. Numbering of the skeletal atoms in the parent structure is retained in the resulting fragment.

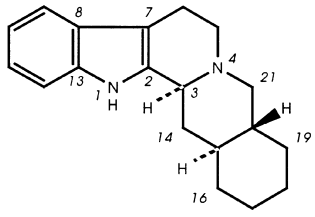
Note: This procedure has been used only in carotenoid nomenclature (Scheme 15) [5].



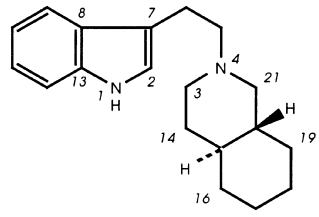
2,3-Seco-5α-cholestane



13,14-Secopodocarpane
(also may be named 8β-14,15,16-Trinorlabdane by RF-4.1)

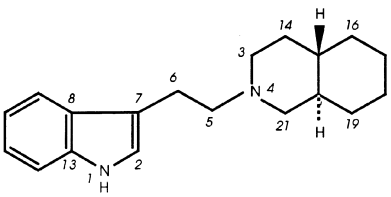


Yohimban (fundamental parent structure)

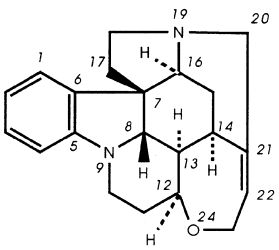


2,3-Secoyohimban

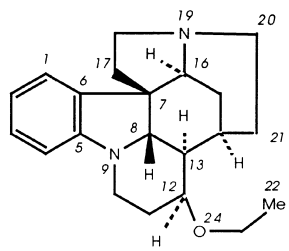
Scheme 12



Scheme 13

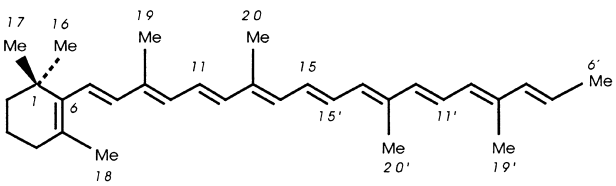


Strychnidine (fundamental parent structure)



21,22-Secostrychnidine

Scheme 14



6′-Apo-β-carotene

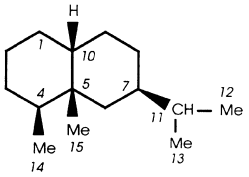
Scheme 15

RF-4.5. Bond migration. Parent structures that are not simple derivatives of accepted fundamental parents, but may be considered to arise from such parents by migration of one or more bonds, may be named by the following methods.

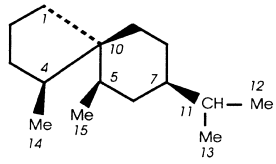
RF-4.5.1. The prefix X(Y→ Z)-abeo- designates the migration of one end of a single bond from its original position in a fundamental parent structure to another position. In this prefix, X is the locant of the stationary, i.e. unchanged, end of the migrating bond; Y is the locant of the position of the moving end of the migrating bond in the fundamental parent structure; and Z is the locant of the position of the moving end of the migrating bond in the final structure. The numbering of the fundamental parent structure is retained in the new structure.

Stereochemical configurations of the fundamental parent structure are retained. New stereochemistry of ring atoms having one hydrogen atom still present is indicated by the α/β system, or, if necessary by the Sequence Rule method (R/S). The projection of the hydrogen atom below (α) or above (β) the reference plane of the ring system is indicated by the appropriate symbol and a capital italic letter *H* following the locant of the ring atom in the structure, all enclosed in parentheses, and cited before the 'abeo-' prefix described above.

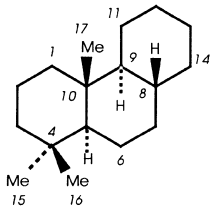
Other new stereochemistry is described by the Sequence Rule System (Scheme 16).



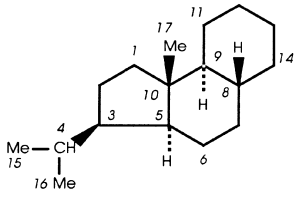
Eremophilane (fundamental parent structure)



(10*R*)-4(5→10)-Abeoeremophilane



Podocarpane (fundamental parent structure)



(3α*H*)-5(4→3)-Abeopodocarpane

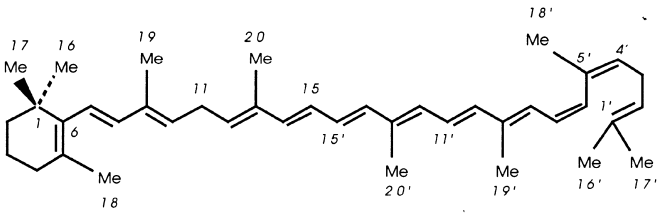
Scheme 16

Note: The prefix '*A*-Neo-' has been used to describe the bond migration shown in this last example. However, the 'abeo' operation is preferred because it is general.

RF-4.5.2. The italic prefix '*retro*-', preceded by a pair of locants is used to indicate a shift, by one position, of all single and double bonds of a conjugated polyene system delineated by the pair of locants that are not part of a system of maximum number of noncumulative double bonds in a ring or ring system. The first locant is the skeletal atom that has lost a hydrogen atom and the second the one that has gained a hydrogen atom. 'Retro' has been used in this manner in the nomenclature of carotenoids (e.g. Scheme 17).

Note 1: The prefix '*retro*-' has been used in steroid nomenclature to indicate a 9β,10α-configuration instead of the implied configuration, but is not recommended.

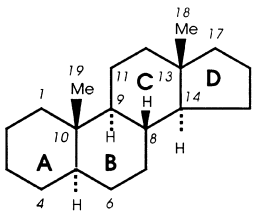
Note 2: The prefix '*retro*-' is also used to designate a peptide in which the amino acid sequence is the reverse of the sequence of the naturally occurring peptide.



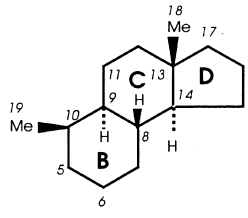
4′,11-*retro*-β,ψ-Carotene

Scheme 17

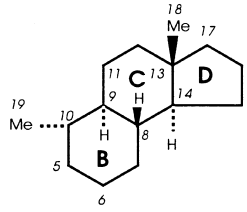
RF-4.6. Removal of a terminal ring. The removal of a terminal ring from a fundamental parent structure of a natural product with the addition of an appropriate number of hydrogen atoms at each junction with the adjacent ring is indicated by the prefix 'des-' followed by the capital italic letter of the ring removed (see RF-3.6). Stereochemistry implied by the name of the fundamental parent structure remains the same unless otherwise specified. Numbering of skeletal atoms in the parent structure is retained in the modified structure (e.g. Scheme 18).



5α-Androstane (fundamental parent structure)



Des-*A*-androstane



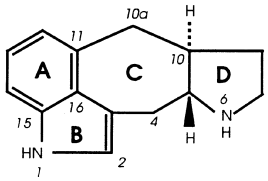
Des-*A*-10α-androstane (see RF-10.2.1 for use of α at position 10)

Scheme 18

RF-4.7. Combinations of prefixes for modifying a fundamental parent structure. The modifications to a fundamental parent structure prescribed by the prefixes in the preceding recommendations (RF-4.1 through RF-4.6) may be combined to effect even more drastic changes in structure. The operation indicated by each prefix is applied to the fundamental parent structure sequentially as one 'advances backwards', i.e. moves from right to left, from the name of the fundamental parent structure. The following recommendations are not rigorous rules for choosing a unique name, but are intended to be guidelines for choosing combinations of prefixes and for the order of citation in generating an unambiguous name.

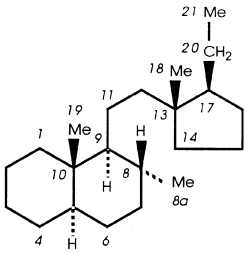
RF-4.7.1. When different combinations of prefixes can be used to effect the same transformation in structure, the combination of choice should express the fewest number of operations. When the number of operations is the same, the combination of homo/nor is preferred to cyclo/seco; choice between other combinations expressing the same number of operations is based on earliest alphabetic order of the prefixes (Scheme 19).

RF-4.7.2. The order of citation of combinations of structure modifying prefixes must avoid improper use of the prefixes as defined above or impossible situations when the corresponding operations are carried out in the manner prescribed above (Scheme 20).



10(11)a-Homo-9-norergoline
(not 5,9-Cyclo-5,10-secoergoline;
also has been called
C(10a)-Homo-*D*-norergoline)

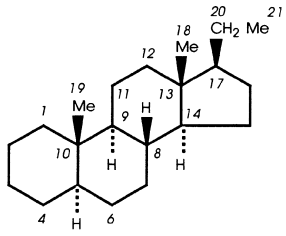
Scheme 19



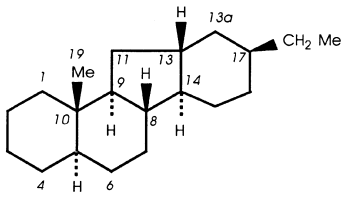
8a,14-Seco-8(14)a-homo-5α-pregnane
(the ring expansion is followed by
the bond cleavage; a methyl-seco
name might also be used)

Scheme 20

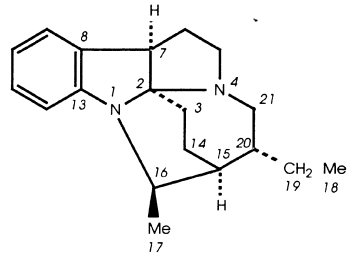
RF-4.7.3. (Alternative to RF-4.7.4). After satisfying RF-4.7.1. and RF-4.7.2., combinations of structure modifying prefixes are cited in alphabetic order from left to right proceeding towards the name of the fundamental structure. Numerical prefixes denoting multiple operations of the same kind do not affect the order (Scheme 21).



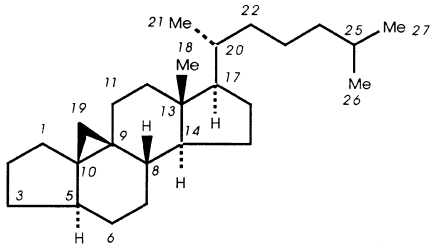
5α-Pregnane (fundamental parent structure)



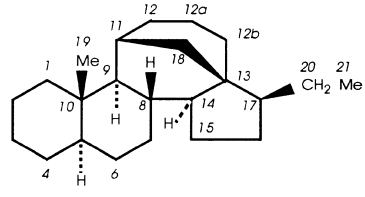
13(17)a-Homo-12,18-dinor-5α-pregnane



(2*S*,7α,16α*H*)-7*H*-1,16:2,4-Dicyclo-
3,4-secocorynan



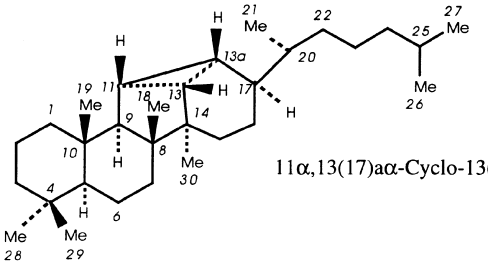
9β,19-Cyclo-4-nor-5α,9β-cholestane



11β,18-Cyclo-12a,12b-dihomo-
5α-pregnane

Scheme 21

Note: The cyclodihomopregnane example above can be named in at least two other ways: (1) 11α,18b-Cyclo-18a,18b-dihomo-5α,13α-pregnane; (2) 11α,13-Propano-18-nor-5α,13α-pregnane. The first name uses the same number of operations, but extends a side chain rather than enlarging a ring (the latter operation seems more usual) resulting in higher locant numbers. The second name uses only one operation of the type discussed in this section, but requires the use of a bridge (see RF-6), which may or may not be considered preferable.



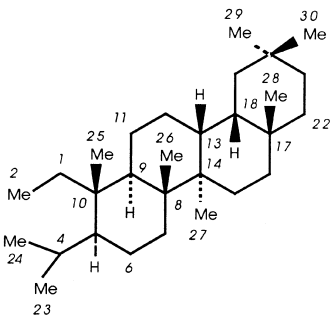
11α,13(17)aα-Cyclo-13(17a)-homo-12-nordammarane

Scheme 22

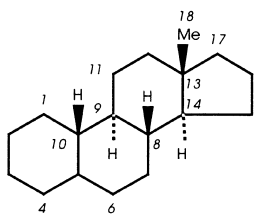
RF-4.7.4. (Alternative to RF-4.7.3). After satisfying RF-4.7.1 and RF-4.7.2., the prefixes that indicate atomic replacement (oxa, aza, etc.) are cited, followed by those showing bond rearrangements (abeo, cyclo, retro, seco, etc.), followed by those that indicate addition or removal of skeletal atoms (apo, de, des, nor). If more than one of any of these operations is needed, they are cited in alphabetic order before the name of the parent structure. Schematically, this order is as follows:

replacement	bond rearrangements	removal/addition of skeletal atoms	parent structure
aza, oxa, etc.	abeo, cyclo, retro, seco	apo, de, des, homo, nor	

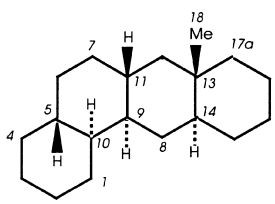
(Examples in Scheme 23)



2,4-Seco-3-noroleanane
(may be named 3-Nor-2,3-secooleanane by RF-4.7.3, above, which is preferable to 24-Nor-2,3-secooleanane or 3-Nor-3,4-secooleanane on the basis of lower locants)



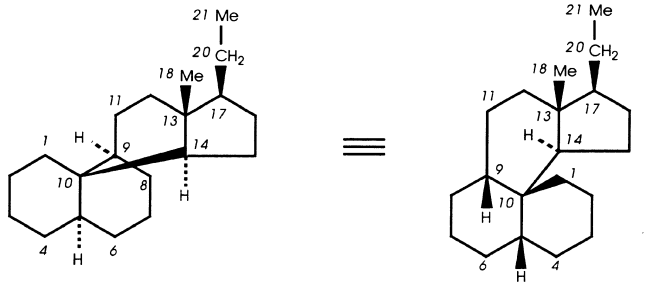
Estrane (fundamental parent structure)



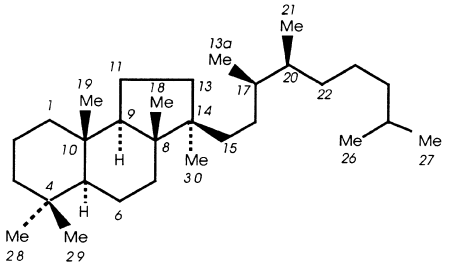
7,11α-Cyclo-7,8-seco-17a-homo-5α-estrane
(7,11α-Cyclo-17a-homo-7,8-seco-5α-estrane by RF-4.7.3)

Scheme 23

Note. This example can also be named 11β*H*-7(8→11)-abeo-17a-homo-5β,10α-estrane, which uses only two operations; however, it requires the use of modified stereochemistry at two implied centers, which may be less preferable.



10,14β-Cyclo-8,14-seco-19-nor-5α-pregnane (may also be named 10β,14-Cyclo-19-nor-8,14-seco-5α-pregnane by RF-4.7.3; note the apparent change in configuration in the alternative drawing -- see the note following the third example under RF-4.4.1)



(17α*H*,20*S*)-13,13(17)a-Seco-13(17)a-homo- 12-nordammarane

Scheme 24

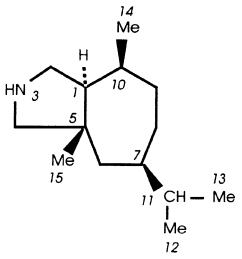
RF-5. Replacement of skeletal atoms in parent structures whose names are formed according to RF-3 and RF-4 above is indicated by extending the principles of organic replacement nomenclature (see rules B-4 [2] and C-0.6 [2]).

RF-5.1. The replacement of a carbon atom in the skeletal system of a parent structure by a heteroatom is described by appropriate replacement 'a' prefixes (see R-9.3 [3]). This procedure is used even though the parent structure may already be a heterocycle. The numbering of the parent structure is retained (e.g Scheme 25).

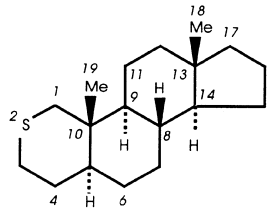
RF-5.2. The replacement of a heteroatom in a parent structure of a natural product by a carbon atom is described by the replacement prefix 'carba'. The original numbering is retained. If the heteroatom in the parent structure is unnumbered, the replacing carbon atom is numbered by affixing the letter 'a' to the locant of the immediately adjacent lower numbered skeletal atom. If the immediately adjacent lower numbered skeletal atom is a 'homo' atom, the letter 'b', 'c', etc., as appropriate is used. Stereochemical configuration at the new carbon skeletal atom is described by methods for specifying additional stereochemistry (see RF-10.2) (e.g. Scheme 26).

RF-5.3. Replacement of a heteroatom in a parent structure by another heteroatom is denoted by the appropriate replacement ('a') prefix (e.g. Scheme 27).

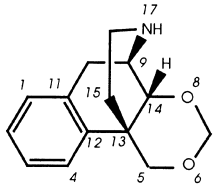
RF-5.4. Indicated Hydrogen. When the replacement of a skeletal atom in a portion of the structure of a fundamental parent structure that contains the maximum number of noncumulative double bonds or an extended conjugated system of double bonds results in the creation of a saturated skeletal position, that position is indicated by the symbolism of indicated hydrogen (e.g. Scheme 28).



3-Azaambrosane

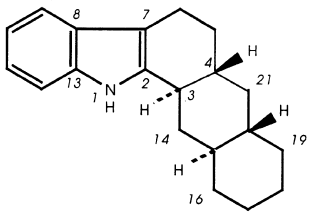


2-Thia-5α-androstane

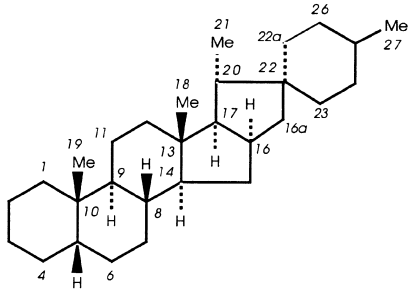


6,8-Dioxamorphinan

Scheme 25

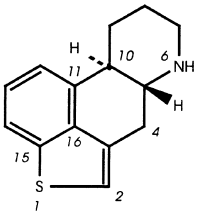


(4β)-1*H*-4-Carbayohimban



16a,22a-Dicarba-5β-spirostan

Scheme 26

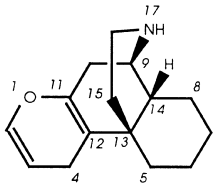


1-Thiaergoline

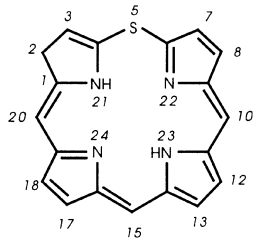
Scheme 27

RF-6. Fusion of additional rings. It is often convenient to retain the advantages of a sernisystematic natural product name, particularly with regard to stereochemistry, for naming structures having rings or ring systems fused to a fundamental parent structure of a natural product.

Since most natural products with cyclic structures have fundamental parent structures that are fully saturated and fusion nomenclature principles are based on the concept of the presence of the maximum number of noncumulative double bonds, adaptations of general fusion nomenclature are necessary. These



4*H*-1-Oxamorphinan

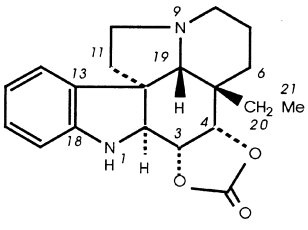


2*H*-5-Thiaporphyrin (the name porphyrin implies two saturated positions normally shown at 21 and 23)

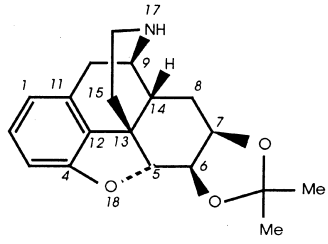
Scheme 28

adaptations are described in the recommendations that follow. No attempt is made to legislate rigidly or to cover every case. The decision between a fusion-natural product name and a systematic fusion name is left to choice based on particular circumstances in each case.

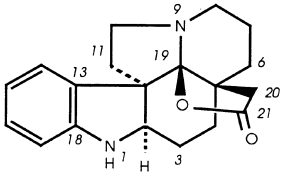
RF-6.1. Rings or ring systems fused to a fundamental parent structure of a natural product when derived from functional groups, such as lactones or cyclic acetals, are preferably named by the usual methods of organic nomenclature for such cyclic functional groups (see also RF-6.2.2) (e.g Scheme 29).



Aspidospermidine-3α,4α-diyl carbonate



Acetone 4,5α-epoxymorphinan-6β,7β-diyl acetal (for use of epoxy as a bridge, see RF-7)



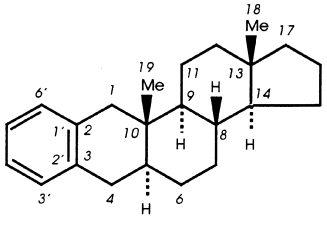
21-Noraspidospermidine-20,19-carbolactone (also named 19-Hydroxyaspidospermidino-21,19-lactone)

Scheme 29

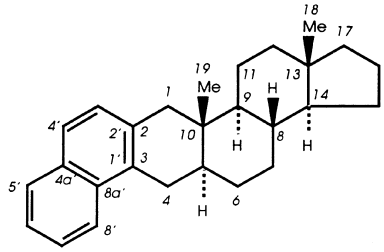
RF-6.2. The fundamental parent structure of the natural product as a component is used in its normal state of saturation or unsaturation. Accordingly, a double bond is not cited in the natural product parent at the fusion site just because the other component contains the maximum number of noncumulative double bonds. Hence, the bonding at a fusion site is governed usually by the non-natural product component; fusion resulting in a quaternary carbon atom at the fusion site also affects the bonding (see RF-6.2.3, below).

RF-6.2.1. A ring or ring system of systematic organic nomenclature [2] (carbocyclic or heterocyclic) fused to a fundamental parent structure is described by its fusion prefix name (see A-21.4 and B-3) [2] prefixed to the name of the fundamental parent structure. The position of the fusion is indicated by sets of locants, as needed, separated by a colon, enclosed in square brackets, and inserted into the name between the components. The skeletal atoms of the natural product are identified by plain

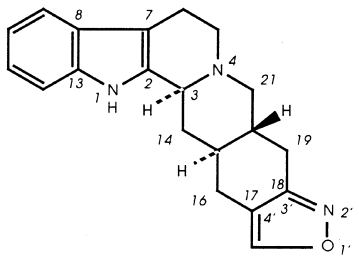
(unprimed) locant numbers and the 'systematic' component by primed locant numbers. Where there is a choice, the locant numbers for the systematic component are as low as possible and are cited in the same direction of numbering as for the natural product component. Terminal vowels of the names of systematic components are not elided when followed by a vowel (this is consistent with the recommendation contained in the fused ring nomenclature report and is a change from previous recommendations) [7] (e.g. Scheme 30).



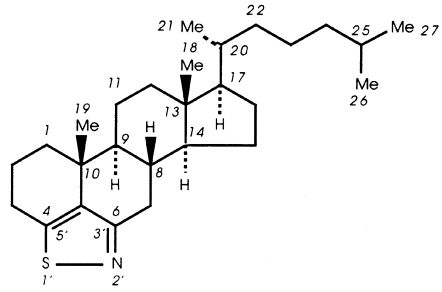
Benzo[2,3]-5α-androstane
(note that only one set of locants, that of the androstane component, is needed here; the CA Index name is Benz[2,3]androst-2-ene, (5α)-)



Naphtho[2′,1′:2,3]-5α-androstane



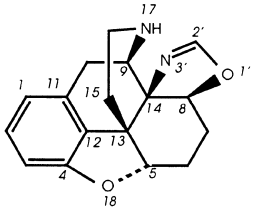
1*H*-Isoxazolo[4′,3′:17,18]yohimban



Isothiazolo[5′,4′,3′:4,5,6]cholestane

Scheme 30

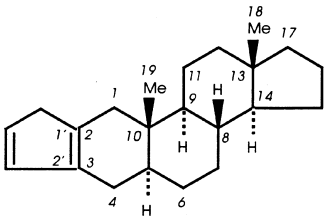
RF-6.2.2. The systematic organic component fused to the fundamental parent component of a natural product structure contains the maximum number of noncumulative double bonds consistent with the bonding requirements at the fusion site. Saturated positions on the systematic component, including the fusion sites, that have at least one hydrogen atom are designated by the indicated hydrogen symbolism (see A-21.6) [2]. Locants of the systematic component are used to identify the position of the indicated hydrogen, where there is a choice (e.g. Scheme 31,32).



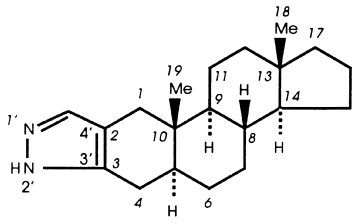
8α*H*-4,5α-Epoxyoxazolo[5′,4′:8,14]morphinan

Scheme 31

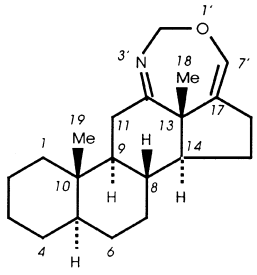
RF-6.2.3. Saturated, or partially saturated, carbocyclic and heterocyclic ring components fused to the fundamental parent structure of a natural product are named using hydro prefixes. Where there is a choice, the unprimed locant numbers of the systematic component are used (e.g. Scheme 33).



5′*H*-Cyclopenta[2,3]-5α-androstane

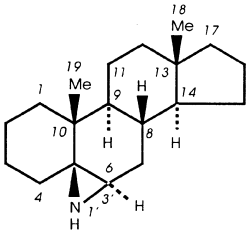


2′*H*-Pyrazolo[4′,3′:2,3]-5α-androstane

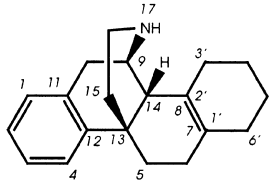


2′*H*-[1,3]Oxazepino[4′,5′,6′:12,13,17]-5α-androstane

Scheme 32



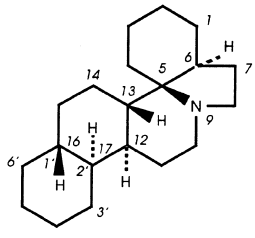
1′,6α-Dihydroazirino[2′,3′:5,6]-5β-androstane



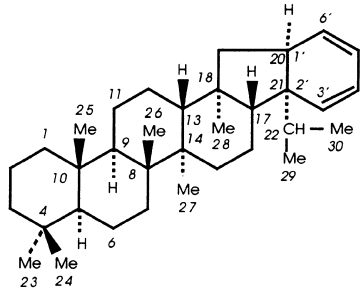
3′,4′,5′,6′-Tetrahydrobenzo[7,8]morphinan

Scheme 33

RF-6.2.4. Stereochemistry in fused rings or hydrogenated derivatives is cited by the α/β symbolism in accordance with its use in the steroid nomenclature recommendations (see 3S-1.4 [4] and RF-10) or by the Sequence Rule method (*R/S*) (see also RF-10) (e.g. Scheme 34).



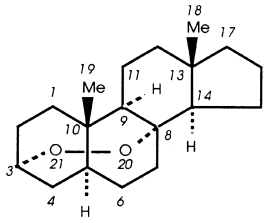
3′,4′,5′,6′,12α,13β,14,15,16β,17α-Decahydrobenzo[16,17]erythrinan



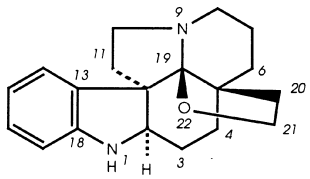
20α*H*-Benzo[20,21]-(21β*H*)17(22→21)-abeogammacerane (also has been named 1′*H*-Benzo[20,21]-*A*′-neogammacerane)

Scheme 34

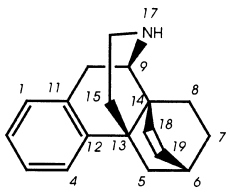
RF-7. Bridged fundamental parent structures. Atomic bridges added to fundamental parent structures of natural products may be described by the methods used in systematic organic nomenclature (A-34 [2], B-15.1 [2] and R-9.2 [3]). This method is often used with hetero atom bridges. In fact, this method is often more useful than fusion procedures for describing certain types of heterocyclic rings fused to a fundamental parent structure, for instance, oxireno (epoxy) and thiireno (epithio). Bridge prefixes are always nondetachable (e.g. Schemes 35, 36).



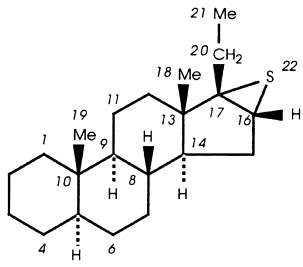
3α,8-Epidioxy-5α,8α-androstane



19,21-Epoxyaspidospermidine

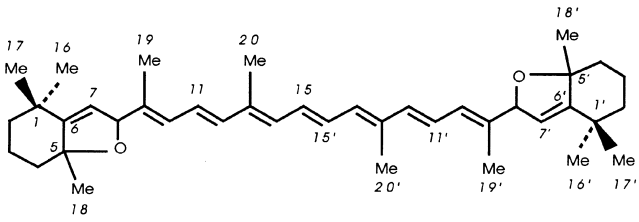


6β,14-Ethenomorphinan (may also be named 6α,14-Ethano-14α-morphin-7-ene)



16α,17-Epithio-5α-pregnane (may also be named 16β*H*-Thiireno[16,17]-5α-pregnane by RF-6)

Scheme 35

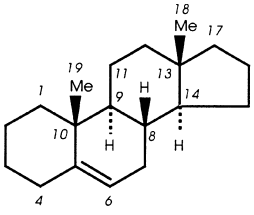


5,8:5′,8′-Diepoxy-5,8,5′,8′-tetrahydro-β,β-carotene (note that in the cartenoid recommendations[2] the bridge prefix "epoxy" is detachable, resulting in the name shown.)

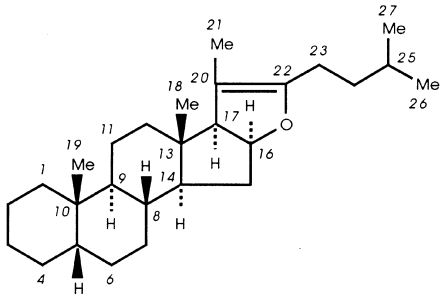
Scheme 36

RF-8. Bond order modification

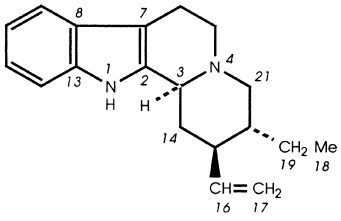
RF-8.1. Unsaturation in a compound whose parent structure (see RF-3 through RF-7) is fully saturated or in the portion of a parent structure that is otherwise fully saturated and whose name ends in 'an', 'ane', or 'anine' (see 3S-2.5 in the steroid nomenclature recommendations [4]) is indicated by changing -an or -ane to -ene, -adiene, -yne, etc. or -an to -en-, -adien-, -yn-, etc. (Rule RF-3.1.1). Locants (numerals and/or letters) are placed immediately before the part of the name to which they relate (see R-0.1.2 [3]) (Scheme 37, 38).



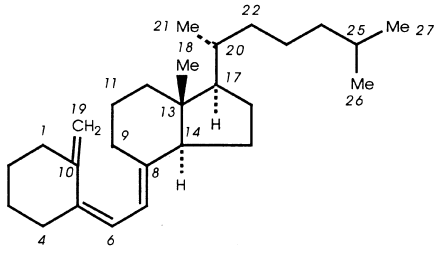
Androst-5-ene



5β-Furost-20(22)-ene

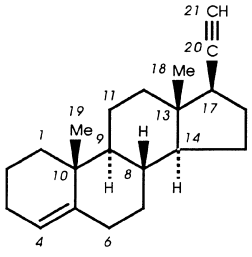


Coryn-16-ene

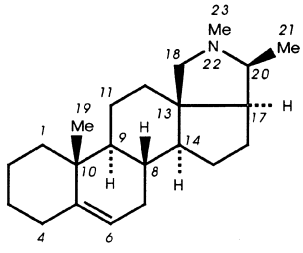


(5*Z*,7*E*)-9,10-Secocholesta-5,7,10(19)-triene

Scheme 37



Pregn-4-en-20-yne



Con-5-enine

Scheme 38

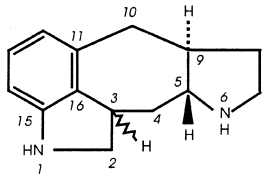
RF-8.2. Saturation of double bonds in a parent structure (see RF-3 through RF-7) whose name implies the presence of isolated double bonds and/or a system of conjugated double bonds is described by the prefix 'hydro-', itself prefixed by an appropriate numerical term (e.g. Scheme 39).

RF-8.3. The introduction of unsaturation additional to any implied in a parent structure (see RF-3 through RF-7) whose name does not end in 'an', 'ane' or 'anine'; the conversion of an implied double bond to a triple bond; and the introduction of an additional double bond with rearrangement of an implied double bond, are denoted by the prefix 'dehydro-', itself prefixed by a numerical term equal to the number of hydrogen atoms removed and the appropriate locants (e.g. Schemes 40, 41).

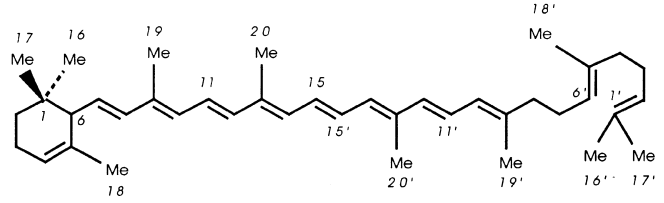
RF-8.4. Rearrangement of a double bond may be indicated by a combination of hydro and dehydro prefixes (e.g. Scheme 42).

RF-9. Derivatives of parent structures are named according to the usual methods of systematic organic nomenclature [2,3] as far as possible.

RF-9.1. The prefixes and suffixes of organic nomenclature are used in the prescribed manner to name atoms and groups that are considered to substitute for hydrogen atoms of parent structures. In naming

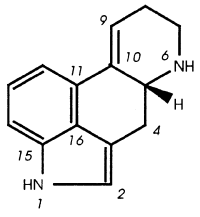


2,3ξ-Dihydro-9α*H*-5(10→9)-abeoergoline



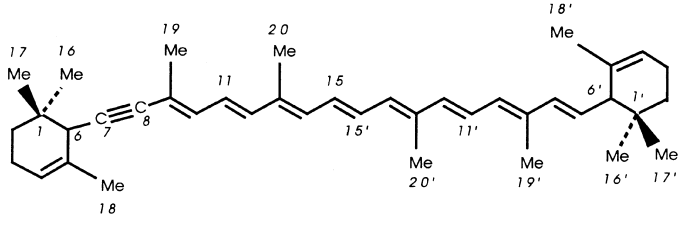
7′,8′-Dihydro-ε,ψ-carotene

Scheme 39

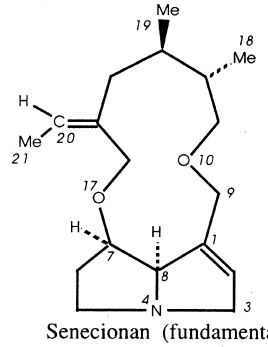


9,10-Didehydroergoline

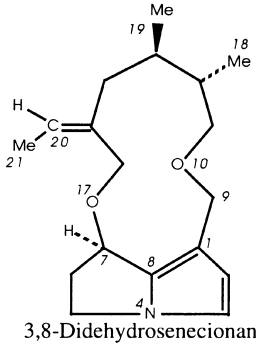
Scheme 40



7,8-Didehydro-ε,ε-carotene

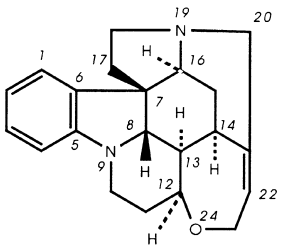


Senecionan (fundamental parent structure)

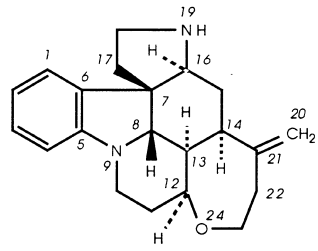


3,8-Didehydrosenecionan

Scheme 41



Strychnidine (fundamental parent structure)

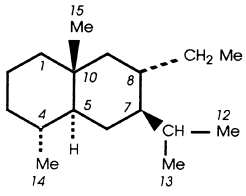


20,21-Didehydro-21,22-dihydro-= 19,20-secostrychnidine

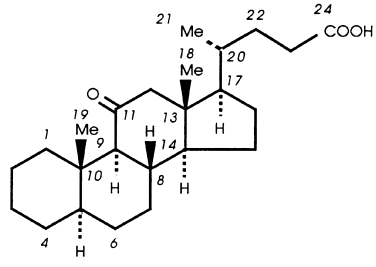
Scheme 42

derived acids and related characteristic groups, unmodified parent structures are used as far as possible (e.g. Scheme 43).

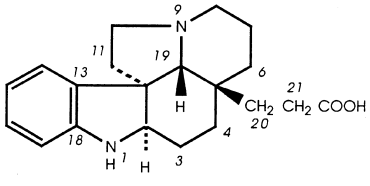
RF-9.2. Modifications to principal characteristic groups, such as esters, acetals, etc., are named by the usual methods of organic nomenclature. Cyclic modifications, such as lactones, cyclic acetals, etc., are named preferably as such rather than as fused ring, bridged, or spiro modified parent structures (see also RF-6.1) (e.g. Scheme 44).



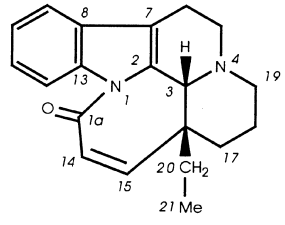
8α-Ethyleudesmane



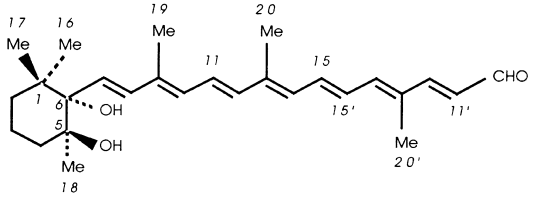
11-Oxo-5α-cholan-24-oic acid



Aspidospermidine-21-carboxylic acid

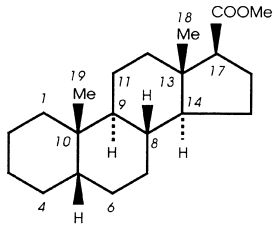


1(14)a-Homoeburnamenin-1(14)a-one

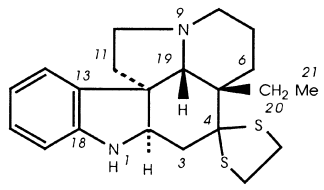


(5*S*,6*S*)-5,6-Dihydroxy-5,6-dihydro-10′-apo-β-caroten-10′-al

Scheme 43



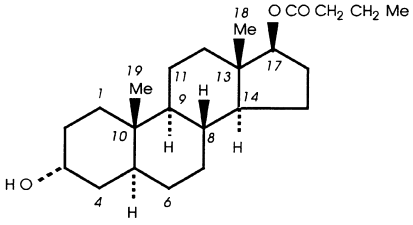
Methyl 5β-androstane-17β-carboxylate



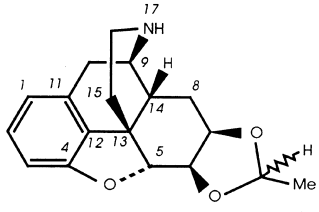
Aspidospermidin-4-one ethylene dithioacetal

Scheme 44

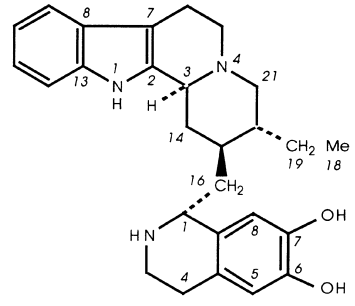
RF-9.3. Substituent prefix names for natural product parent structures may be formed in the usual way by adding a suffix, such as '-yl', '-diyl', '-ylidene', to the name of the parent structure with elision of the final 'e', if any, of the parent structure name before 'y' (e.g. Scheme 45).



3α-Hydroxy-5α-androstan-17β-yl butyrate



Acetaldehyde 4,5α-epoxy morphinan-6β,7β-diyl acetal



1,2,3,4-Tetrahydro-1α-(17-nor corynan-16-yl)isoquinoline-6,7-diol

Scheme 45

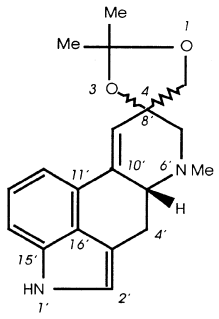
RF-9.4. Ring assemblies and spiro derivatives of parent structures are named in the usual manner [2,3] (e.g. Scheme 46).

Note: In the ring assembly name (dione) in Scheme 46 the hydroxy group is created by the seco operation. To provide for the explicit citation of the hydroxy group as a substituent, the oxygen atom could be removed first using the nor operation; this would be followed by the seco operation and the addition of the prefix 'hydroxy-'.

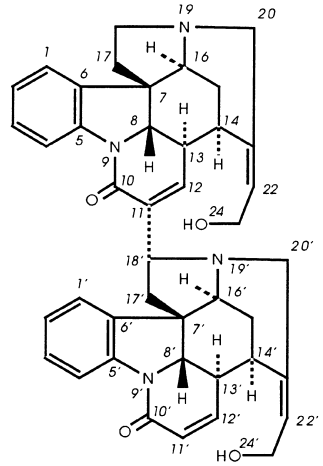
RF-10. Stereochemical configuration

RF-10.1. Names of the fundamental and modified parent structures (see RF-3 through RF-7) imply, without further specification, absolute configuration at all chiral elements as depicted in these recommendations and the following Appendix.

When a planar or quasi-planar system of rings is denoted as a projection on paper, as in these recommendations, an atom or group attached to the ring is called α if it lies below or β if it lies above the plane of the paper. Use of this system requires the orientations of structure as given herein. In the example in Scheme 47,



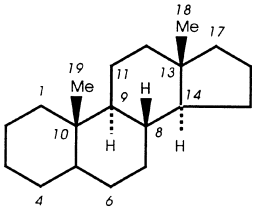
(4ξ)-9′,10′-Didehydro-2,2,6′-trimethyl spiro[1,3-dioxolane-4,8′-ergoline]



11,11′,12,12′-Tetradehydro[11,18′α-bi-12,24-secostrychnidine]-10,10′-dione

Scheme 46

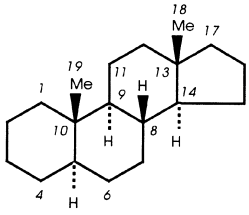
the implied configuration shown defines the attached hydrogen atoms or methyl groups at positions 8, 10 and 13, as β-, and at positions 9 and 14 as α-; here, the hydrogen atom at the chiral position 5 is not known and thus the orientation is ξ (xi). In the case of a racemic compound, that enantiomeric structure drawn should be the one that shows the lowest numbered chiral center in the α- configuration (see also RF-10.4). This may differ from the usual practice, which is to draw the enantiomeric structure having the same absolute configuration as the naturally occurring substance.



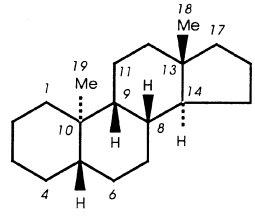
Scheme 47

RF-10.2. Stereochemical configurations that are different from those in the parent structure or that have been generated by substitution, etc.

RF-10.2.1. At chiral centers, the α/β system as described above or by 3S-1.4 in the IUPAC-IUB recommendations for the nomenclature of steroids [4] and Rule E-4.11 of the IUPAC Nomenclature of Organic Chemistry [2] is used (see also Section 2 in the recommendations for the nomenclature of Vitamin D [6]) (Schemes 48, 49).

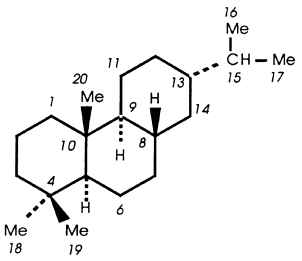


5α-Androstane (fundamental parent structure)

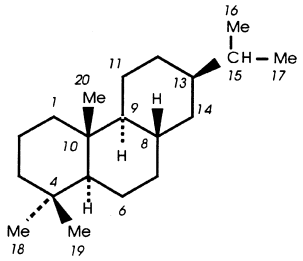


5β,9β,10α-Androstane

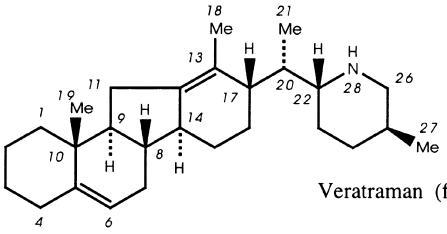
Scheme 48



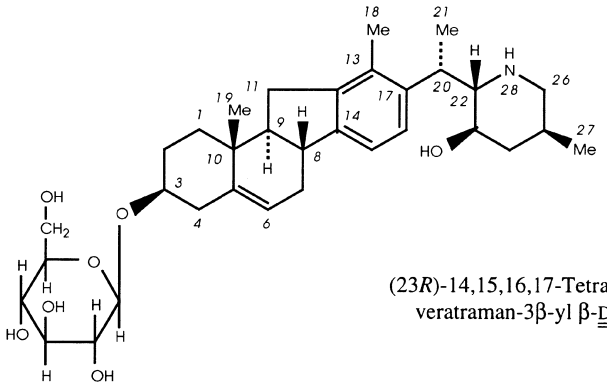
Abietane (fundamental parent structure)



13α*H*-Abietane



Veratraman (fundamental parent structure)

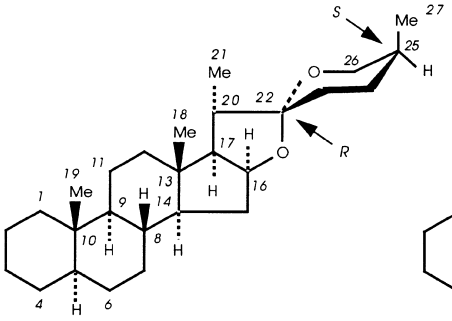


(23*R*)-14,15,16,17-Tetradehydro-23-hydroxy veratraman-3β-yl β-D-glucopyranoside

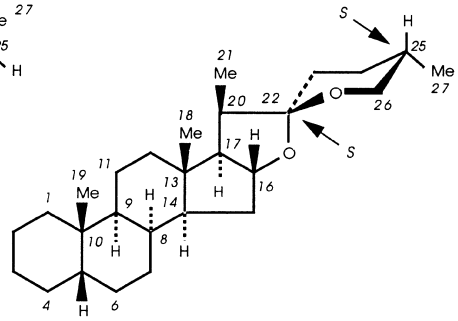
Scheme 49

RF-10.2.2. When the α/β method is not applicable or is not acceptable for the specific natural product class, the *R/S* symbolism of the Sequence Rule System is used (e.g. Schemes 50, 51).

RF-10.2.3. The descriptors *cis/trans* or *E/Z* are used to describe modified or additional stereochemical configurations for double bonds (e.g. Scheme 52).

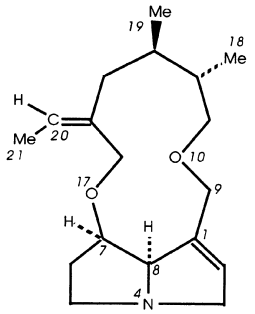


(25*S*)-5α-Spirostan (fundamental parent structure; configurations at positions 5 and 25 are not implied)

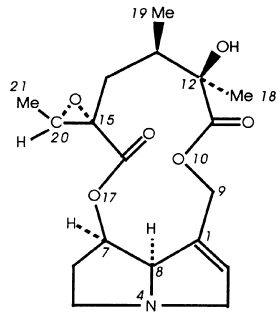


(22*S*,25*S*)-5β,8α,16β*H*,20*R*-Spirostan

Scheme 50

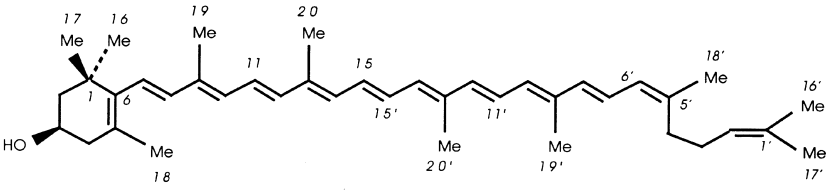


Senecionan (fundamental parent structure)



(20*R*)-12-Hydroxy-15α,20-epoxy senecionan-11,16-dione

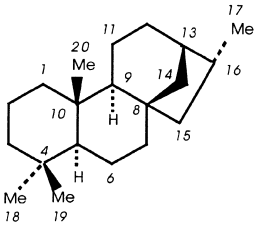
Scheme 51



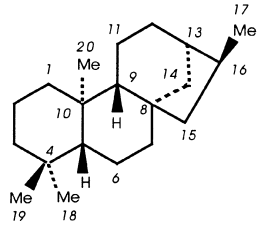
(3*R*)-5′-*cis*-β,ψ-caroten-3-ol (all *trans*- double bonds are assumed in carotenoid names; the *E/Z* method is not normally used with carotenoid double bonds)

Scheme 52

RF-10.3. Configurational inversion at all asymmetric centers whose configurations are implied by or stated in the name for the fundamental parent structure is indicated by the italicized prefix *ent*- (a contracted form of *enantio*-) placed in front of the name of the parent structure. (This is a change from that given by 3S-5.1 in the steroid recommendations [3] and by Rule F-6.4 in the provisional Section F recommendations [1,2]) (e.g. Scheme 53, 54).



Kaurane (fundamental parent structure)



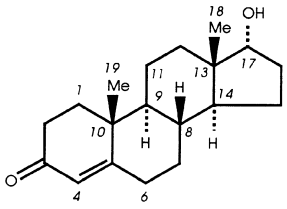
ent-Kaurane

Scheme 53

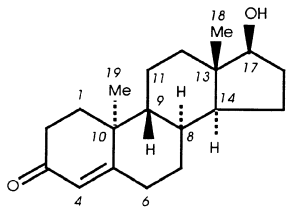
Note: There is confusion in the literature here; some authors (and *Chemical Abstracts*) use Kaurane for the enantiomer called *ent*-Kaurane above.

RF-10.4. Configurational inversion at one asymmetric center whose configuration is implied or stated in the name for the fundamental parent structure can be indicated by the italicized prefix *epi*- (derived from word '*epimer*') placed in front of the name of the parent structure and prefixed by the locant of the affected atom (e.g. Scheme 55).

RF-10.5. Racemates are named by citing the italicized prefix *rac*- (an abbreviation for *racemo*-) in front of the whole name of the compound including the prefix *ent*-, if present. The enantiomer for naming is chosen in accordance with RF-10.1.

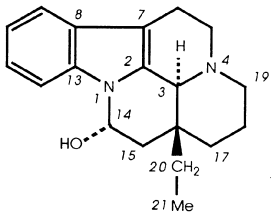


17α-Hydroxyandrost-4-en-3-one

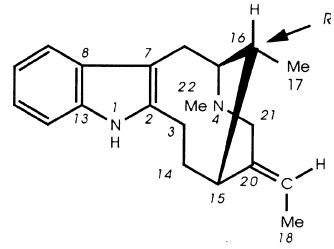


17β-Hydroxy-*ent*-13α,14β-androst-4-en-3-one or 17β-Hydroxy-8α,9β,10α-androst-4-en-3-one (see RF-10.2.1) (but not *ent*-17α-Hydroxy-13α,14β-androst-4-en-3-one)

Scheme 54



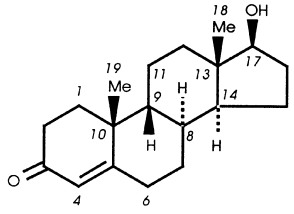
3-*epi*-Eburnamin-14α-ol (the name Eburnamin implies a 3β configuration)



16-*epi*-Vobasan (Vobasan implies a 16*S* configuration)

Scheme 55

RF-10.6. When the relative, but not the absolute, configurational relationships among asymmetric centers are known, the symbols *R** and/or *S** are used in accordance with Rule E-4.10 [2]. Alternatively, enantiomers of known relative, but unknown absolute configuration may be distinguished by a prefix (+)-*rel*-or (−)-*rel*-, where the plus and minus sign refer to the direction of rotation of polarized light at the sodium-D line. Hence, the dextrorotatory form of the following structure would be named: (+)-*rel*-17β-Hydroxy-8α,9β-androst-4-en-3-one (Scheme 56).



Scheme 56

REFERENCES

1 International Union of Pure and Applied Chemistry. Nomenclature of Organic Chemistry. Section F—Natural Products and Related Compounds, Recommendations 1976. *1UPAC Information Bulletin Appendices on Tentative Nomenclature, Symbols, Units and Standards*, no. 53, December (1976). [also in: *Eur. J. Biochem*. **86**, 1–8 (1978)].

2 International Union of Pure and Applied Chemistry. *Nomenclature of Organic Chemistry, Sections A, B, C, D, E, F and H*. 1979 edn. Pergamon Press, Oxford (1979).

3 International Union of Pure and Applied Chemistry. *A Guide to IUPAC Nomenclature of Organic Compounds*. Blackwell Scientific Publications, Oxford (1993).

4 International Union of Pure and Applied Chemistry Joint Commission on Biochemical Nomenclature.

Nomenclature of steroids. *Pure Appl. Chem.* **61**, 1783–1822 (1989). [also in: *Eur. J. Biochein*. **186**, 429–458 (1989), and in *Dictionary of Steroids* (R. A. Hill, D. X. Kirk, H. L. J. Makin, G. M. Murphy, eds), pp. xxx–lix. Chapman & Hall, London (1991)].

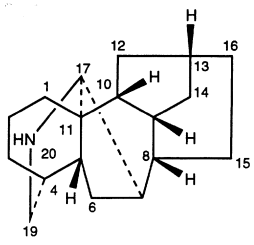
5 International Union of Biochemistry, International Union of Pure and Applied Chemistry, Joint Commission on Biochemical Nomenclature. Nomenclature of carotenoids. *Pure Appl. Chem*. **41**, 405–431 (1975).

6 International Union of Biochemistry, International Union of Pure and Applied Chemistry, Joint Commission on Biochemical Nomenclature. Nomenclature of vitamin D. *Pure Appl. Chem*. **54**, 1511–1516 (1982). [also in: *Arch. Biochem. Biophys*. **218**, 342–346 (1982); *Endokrinol. Inf*. **2**, 53–64 (1982); *Eur. J. Biochem*. **124**, 223–227 (1982), and *Mol. Cell. Biochem*. **49**, 177–181 (1982)].

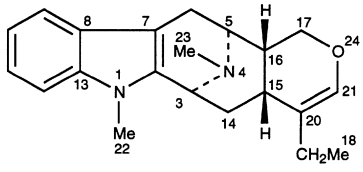
7 International Union of Pure and Applied Chemistry. Nomenclature of fused and bridged fused ring systems. *Pure Appl. Chem*. **70**, 143–216 (1998).

APPENDIX

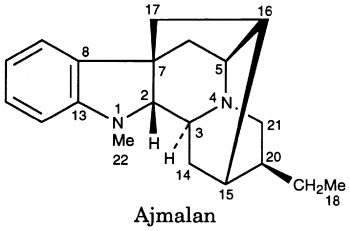
Alkaloids



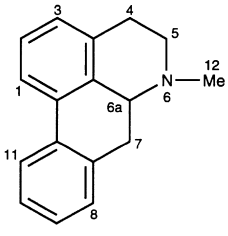
Aconitane



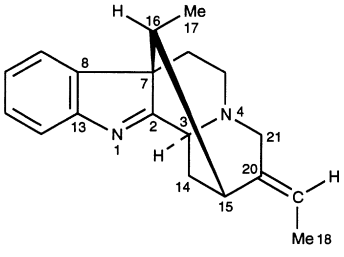
Alstophyllan



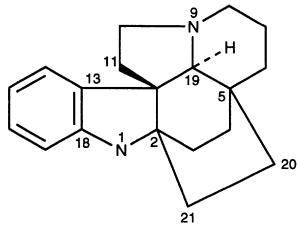
Ajmalan



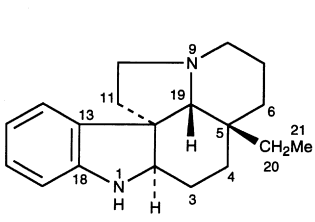
Aporphine



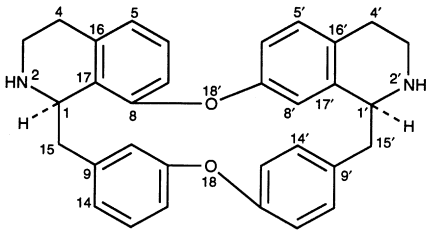
Akuammilan



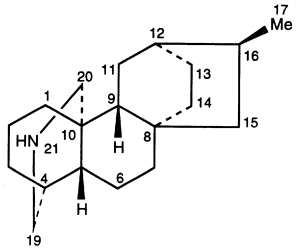
Aspidofractinine



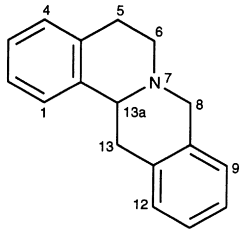
Aspidospermidine



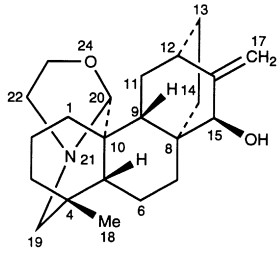
Berbaman



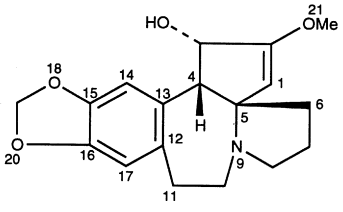
Atidane



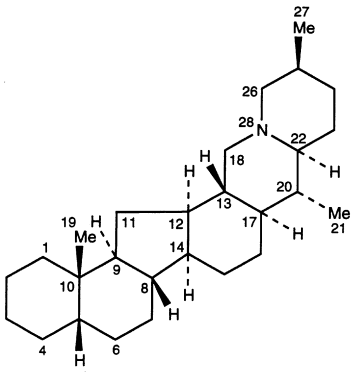
Berbine



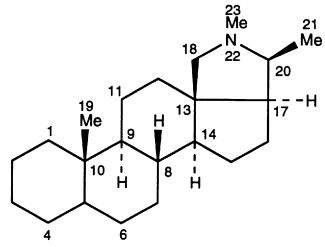
Atisine



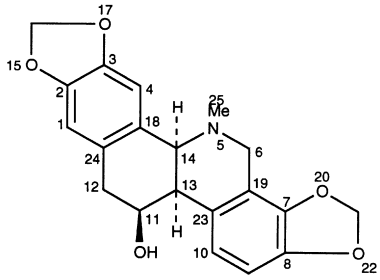
Cephalotaxine



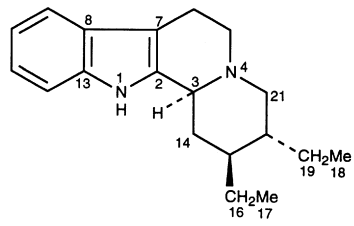
Cevane



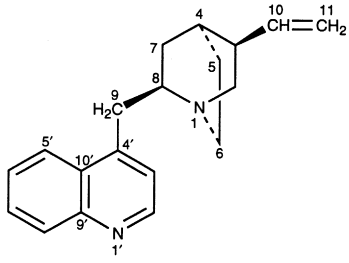
Conanine



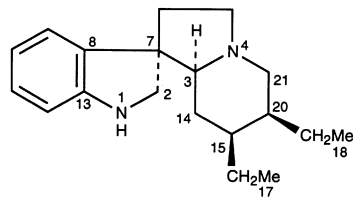
Chelidonine



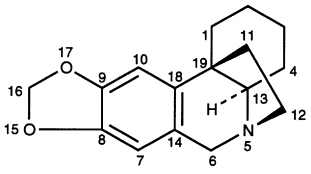
Corynan



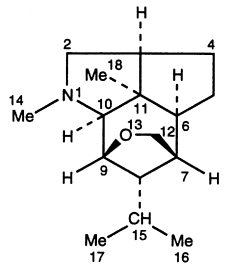
Cinchonan



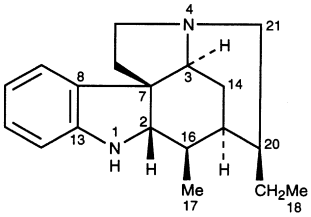
Corynoxan



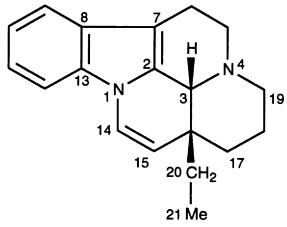
Crinan



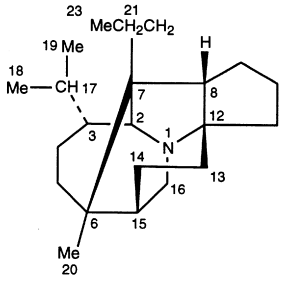
Dendrobane



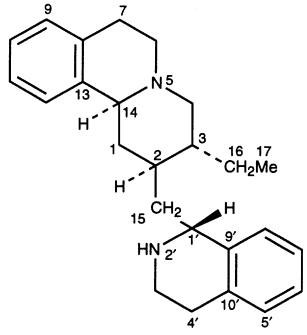
Curan



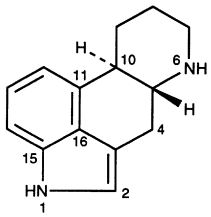
Eburnamenine



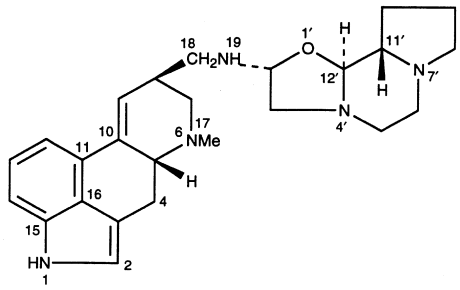
Daphnane



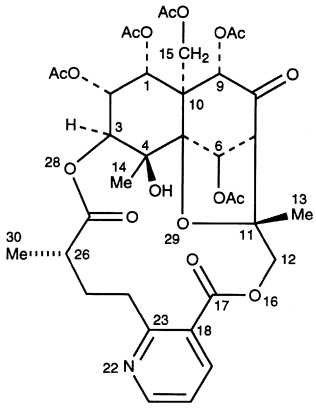
Emetan



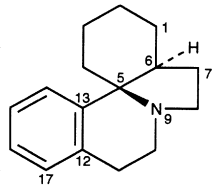
Ergoline



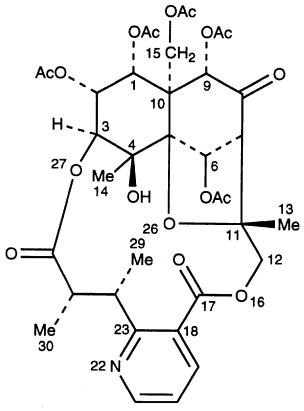
Ergotaman



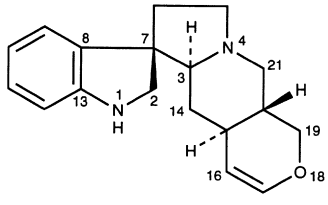
Evonimine



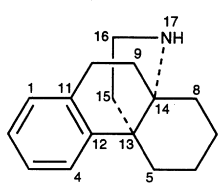
Erythrinan



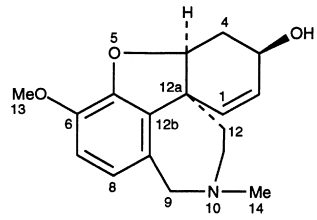
Evonine



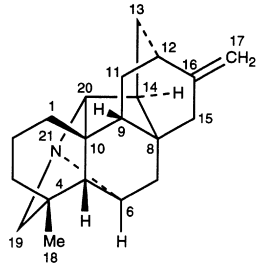
Formosanan



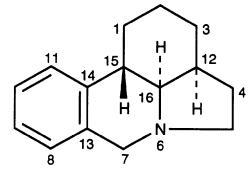
Hasubanan



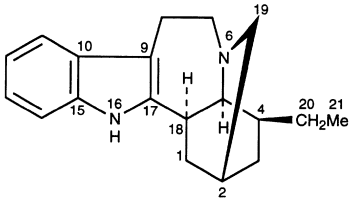
Galanthamine



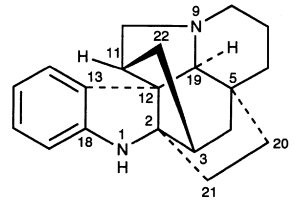
Hetisan



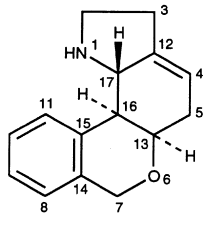
Galanthan



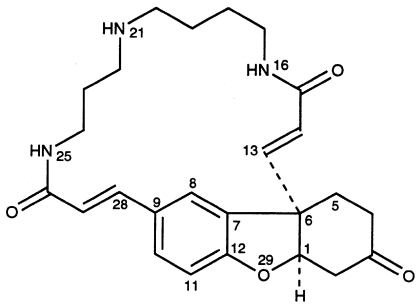
Ibogamine



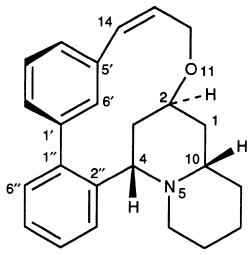
Kopsan



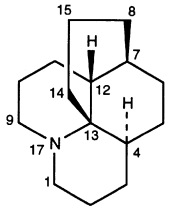
Lycorenan



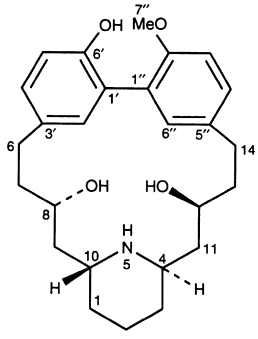
Lunarine



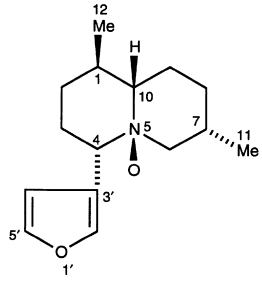
Lythran



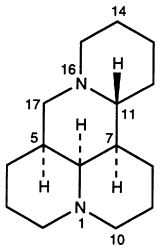
Lycopodane



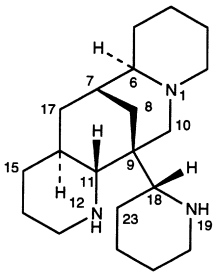
Lythranidine



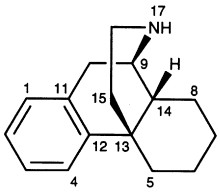
Nupharidine



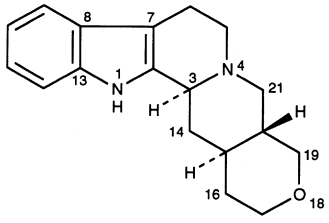
Matridine



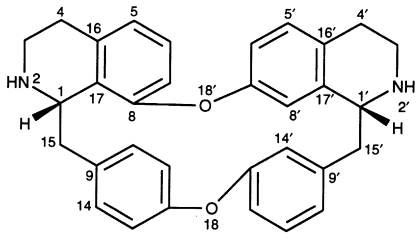
Ormosanine



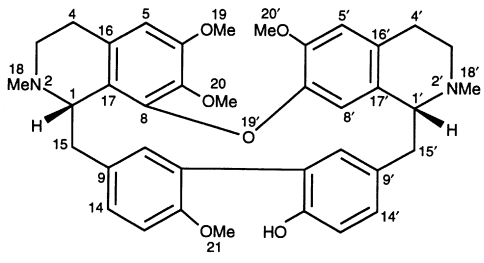
Morphinan



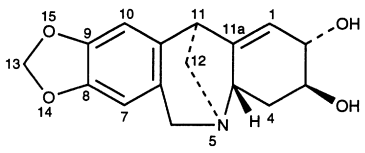
Oxayohimban



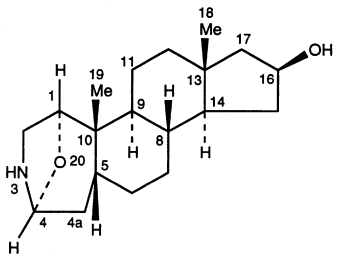
Oxyacanthan



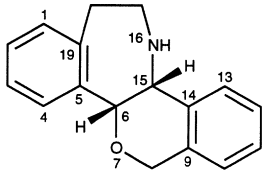
Rodiasine



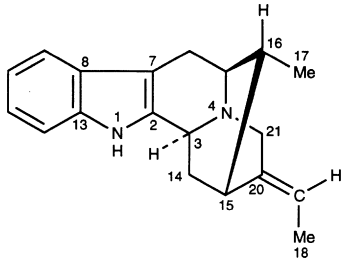
Pancracine



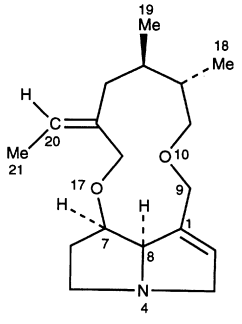
Samandarine



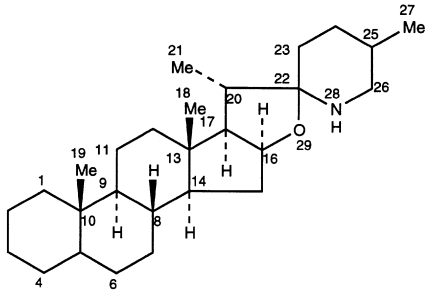
Rheadan



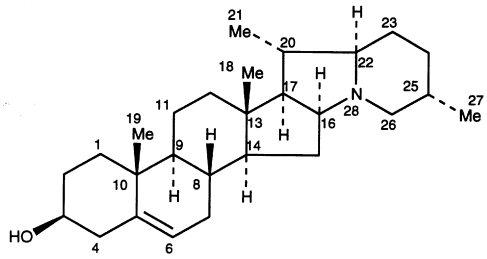
Sarpagan



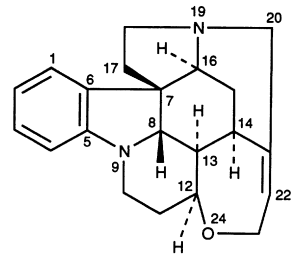
Senecionan



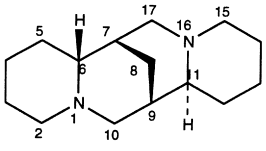
Spirosolane



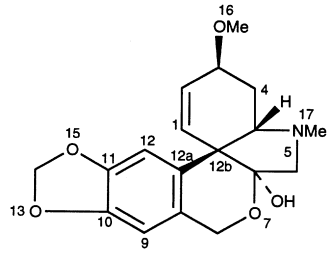
Solanidine



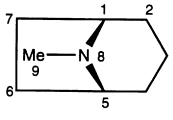
Strychnidine



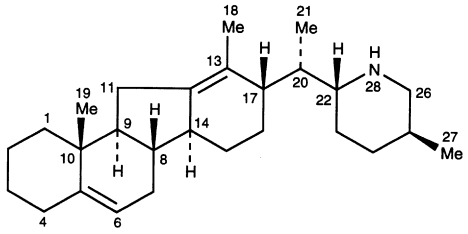
Sparteine



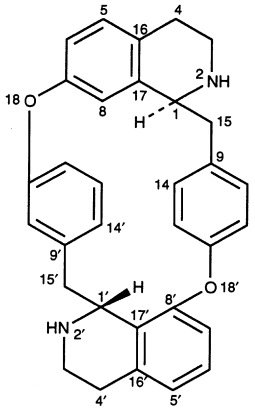
Tazettine



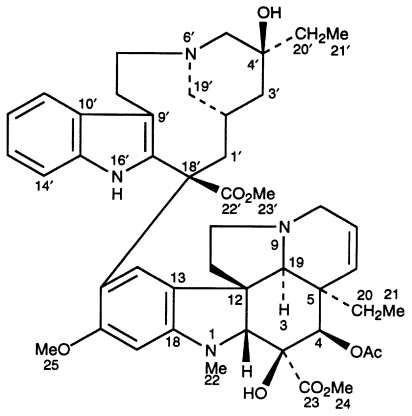
Tropane



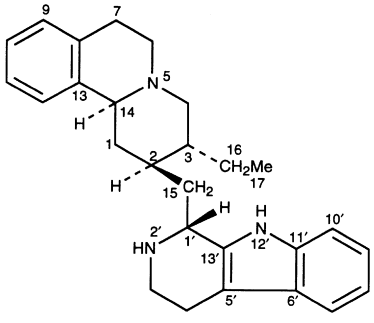
Veratraman



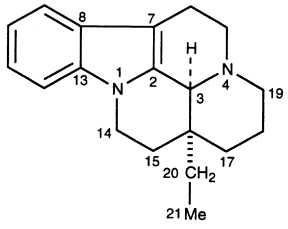
Tubocuraran



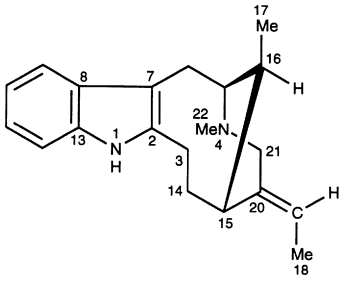
Vincaleukoblastine



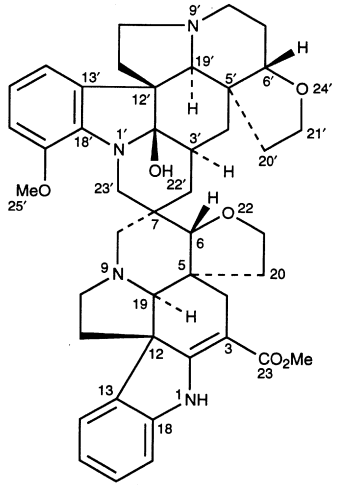
Tubulosan



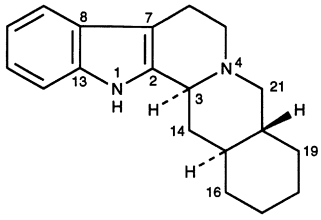
Vincane



Vobasan

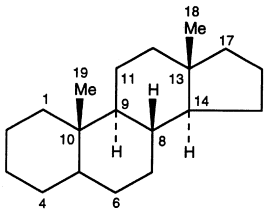


Vobtusine

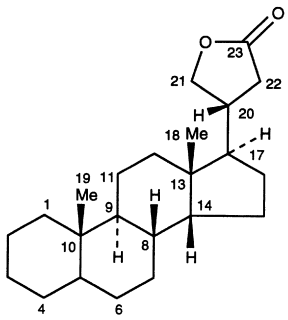


Yohimban

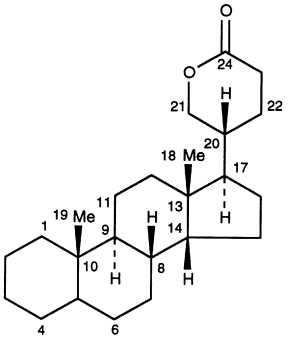
Steroids



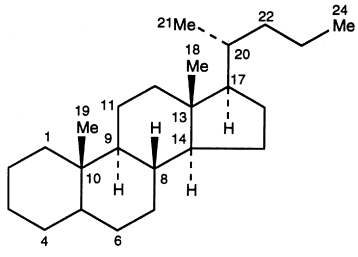
Androstane



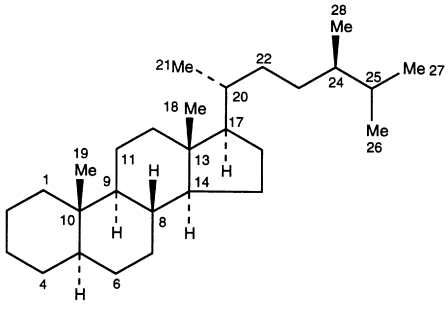
Cardanolide



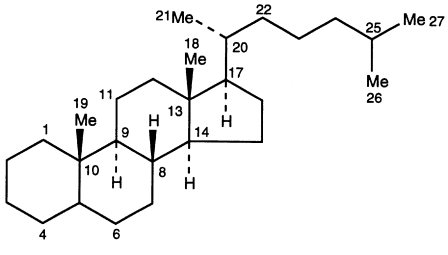
Bufanolide



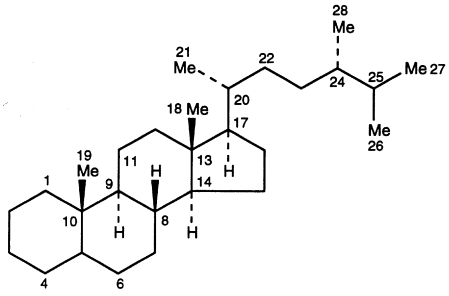
Cholane



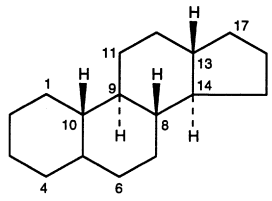
Campestane



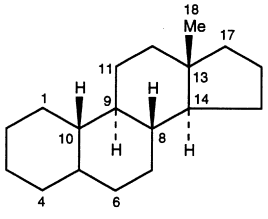
Cholestane



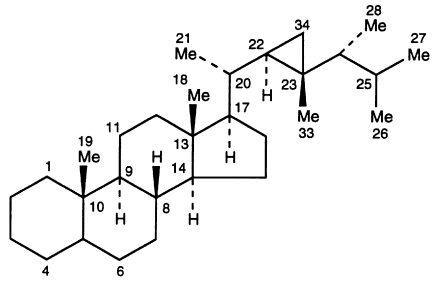
Ergostane



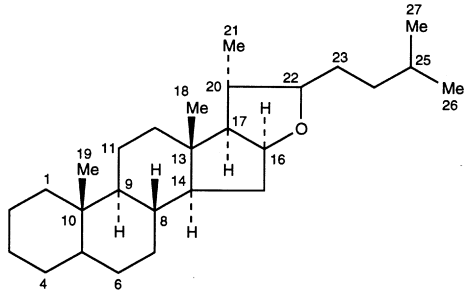
Gonane



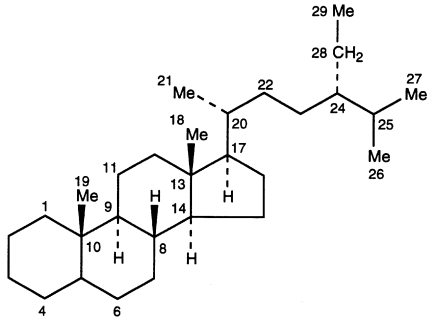
Estrane



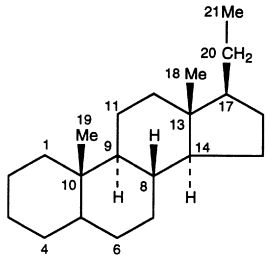
Gorgostane



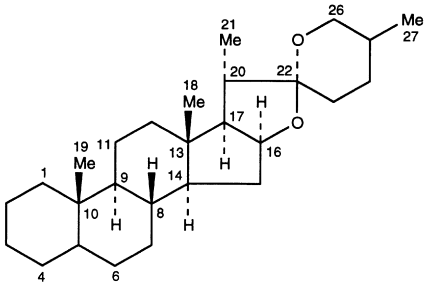
Furostan



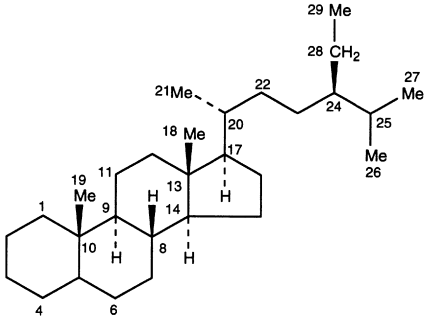
Poriferastane



Pregnane

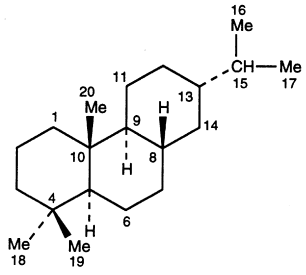


Spirostan

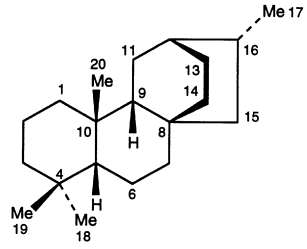


Stigmastane

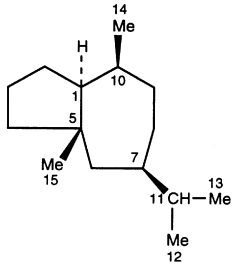
Terpenoids



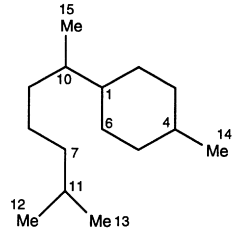
Abietane



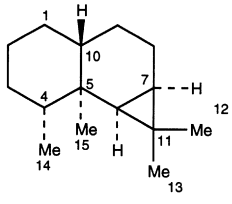
Atisane



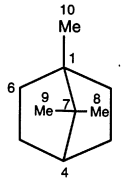
Ambrosane



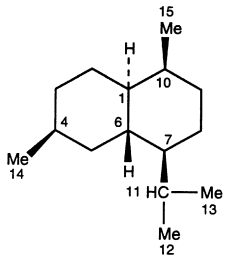
Bisabolane



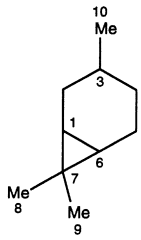
Aristolane



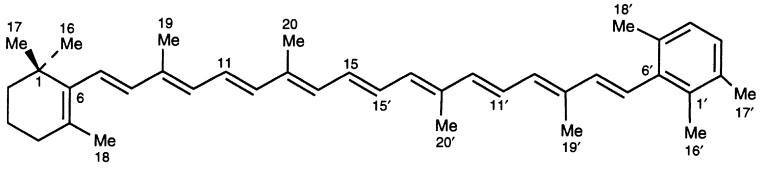
Bornane



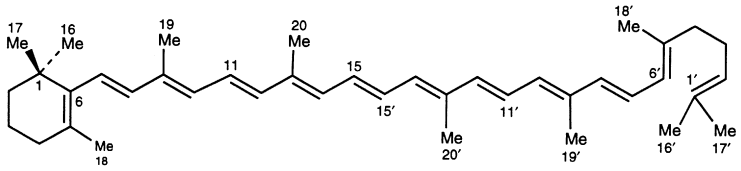
Cadinane



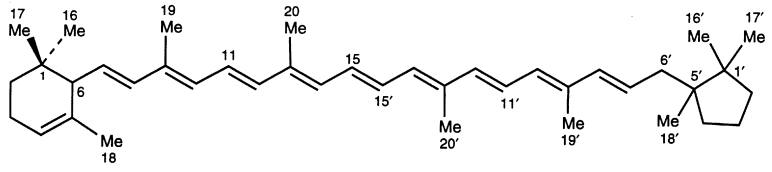
Carane



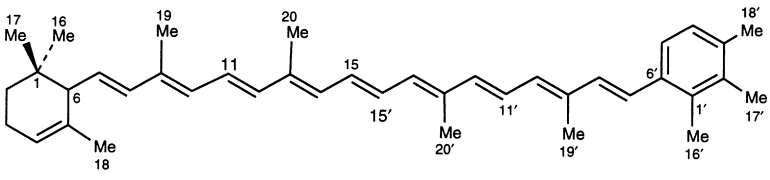
β,φ–Carotene



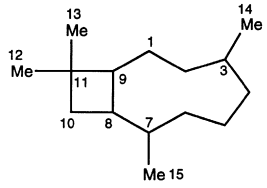
β,ψ–Carotene



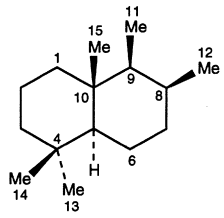
ε,κ–Carotene



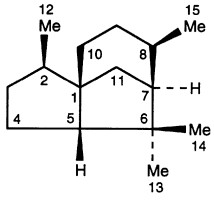
ε,χ–Carotene



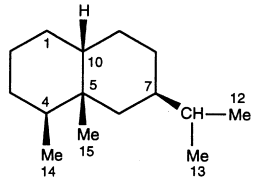
Caryophyllane



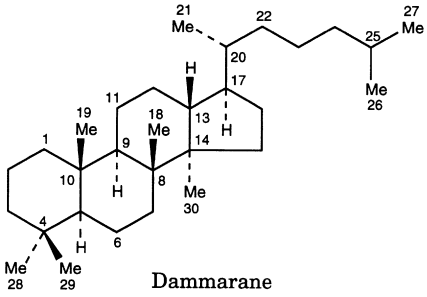
Drimane



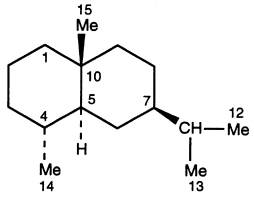
Cedrane



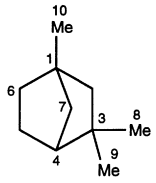
Eremophilane



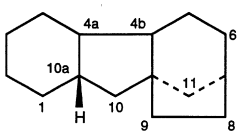
Dammarane



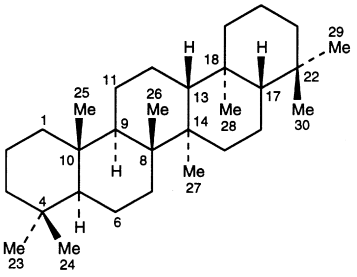
Eudesmane



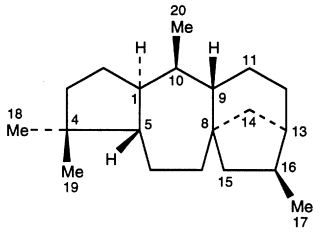
Fenchane



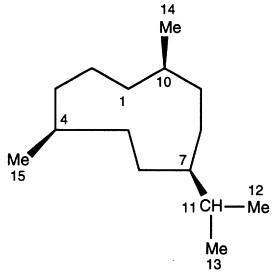
Gibbane



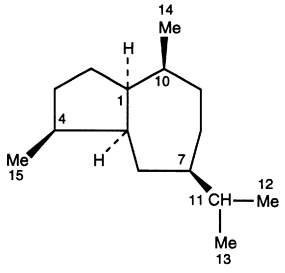
Gammacerane



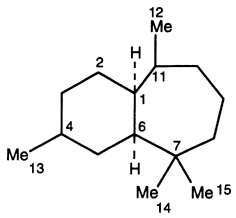
Grayanotoxane



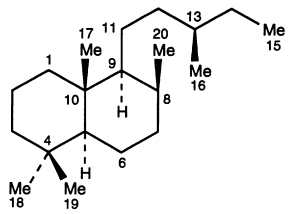
Germacrane



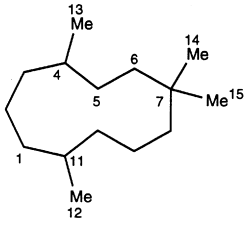
Guaiane



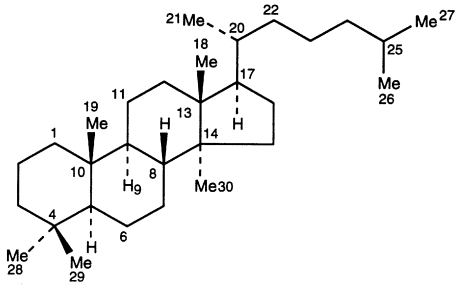
Himachalane



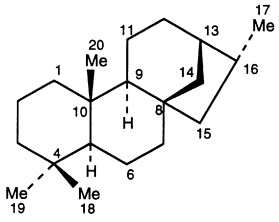
Labdane



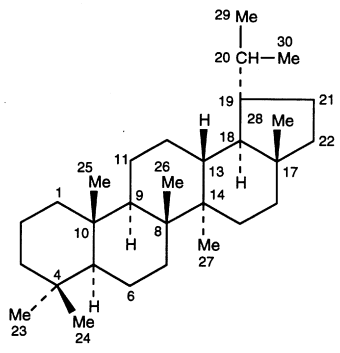
Humulane



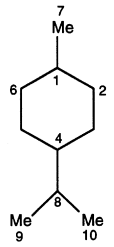
Lanostane



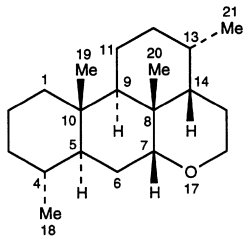
Kaurane



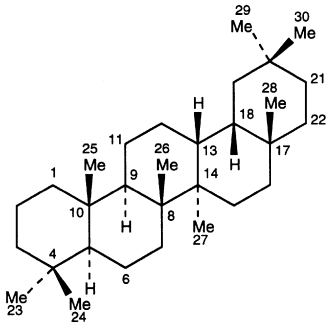
Lupane



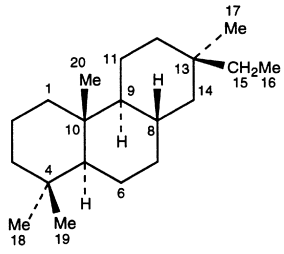
p-Menthane



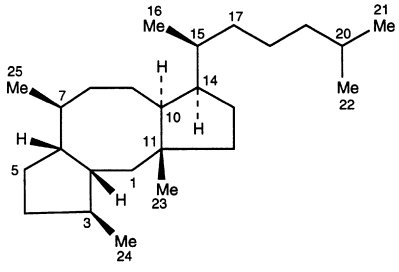
Picrasane



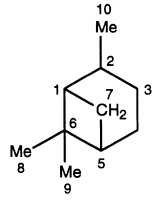
Oleanane



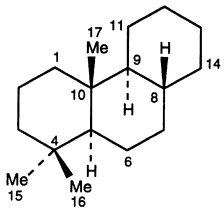
Pimarane



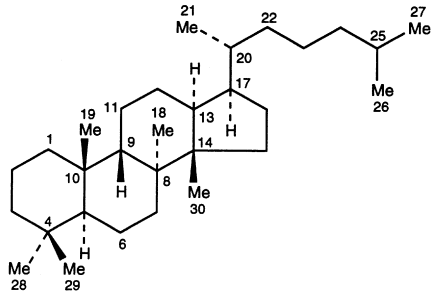
Ophiobolane



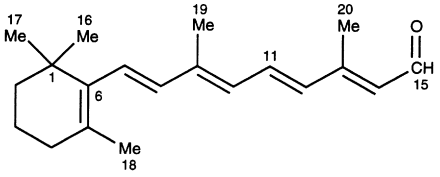
Pinane



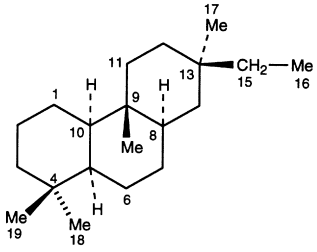
Podocarpane



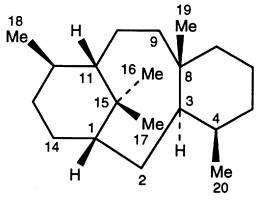
Protostane



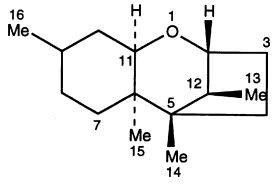
Retinal



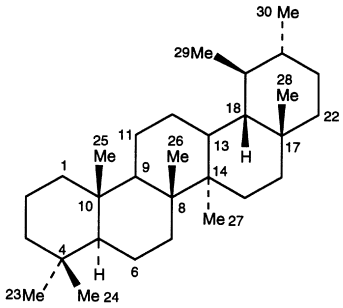
Rosane



Taxane

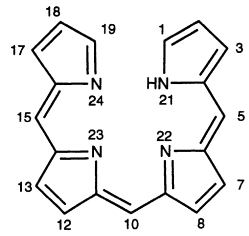


Trichothecane

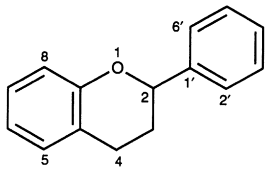

30 Me
29Me
28
Me
18
22
25
Me
11
26
Me
13
H
17
1
9
14
10
8
Me 27
4
H
6
23Me
Me 24

Ursane

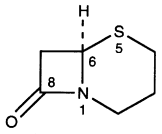
Miscellaneous



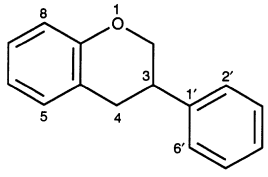
Biline



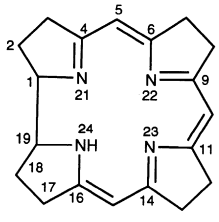
Flavan



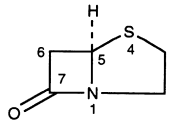
Cepham



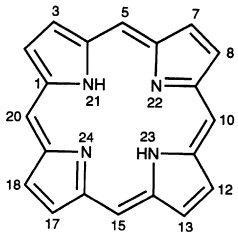
Isoflavan



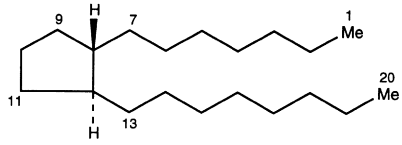
Corrin



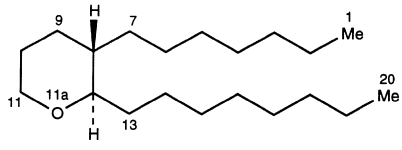
Penam



Porphyrin



Prostane



Thromboxane