

# The Organic Chemist's Desk Reference

A COMPANION VOLUME TO THE  
*DICTIONARY OF ORGANIC COMPOUNDS*, Sixth Edition

P.H. Rhodes

*With contributions from other DOC editors and contributors  
Hazard and toxicity information by R. Purchase*

## Contents

# The Organic Chemist's Desk Reference

A COMPANION VOLUME TO THE  
DICTIONARY OF ORGANIC COMPOUNDS, Sixth Edition

P.H. Rhodes

*With contributions from other DOC editors and contributors*  
*Hazard and toxicity information by R. Purchase*



**CHAPMAN & HALL**

Electronic Publishing Division

London · Glasgow · Weinheim · New York · Tokyo · Melbourne · Madras

**Published by**  
**Chapman & Hall, 2-6 Boundary Row, London SE1 8HN, UK**

---

Chapman & Hall, 2-6 Boundary Row, London SE1 8HN, UK

Blackie Academic & Professional, Wester Cleddens Road, Bishopbriggs, Glasgow G64 2NZ, UK

Chapman & Hall GmbH, Pappelallee 3, 69469 Weinheim, Germany

Chapman & Hall USA, 115 Fifth Avenue, New York, 10003, USA

Chapman & Hall Japan, ITP-Japan, Kyowa Building, 3F, 2-2-1 Hirakawacho, Chiyoda-ku,  
Tokyo 102, Japan

Chapman & Hall Australia, 102 Dodds Street, South Melbourne, Victoria 3205, Australia

Chapman & Hall India, R. Seshadri, 32 Second Main Road, CIT East, Madras 600 035, India

---

First edition 1995

© 1995 Chapman & Hall

Typeset by EJS Chemical Composition, Midsomer Norton, Bath, Avon

Printed in Great Britain at the University Press, Cambridge

ISBN 0 412 64410 X (HB) 0 412 54100 9 (PB)

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the UK Copyright Designs and Patents Act, 1988, this publication may not be reproduced, stored, or transmitted, in any form or by any means, without the prior permission in writing of the publishers, or in the case of reprographic reproduction only in accordance with the terms of licences issued by the appropriate Reproduction Rights Organization outside the UK. Enquiries concerning reproduction outside the terms stated here should be sent to the publishers at the London address printed on this page.

The publisher makes no representation, express or implied, with regard to the accuracy of the information contained in this book and cannot accept any legal responsibility or liability for any errors or omissions that may be made.

A catalogue record for this book is available from the British Library

Library of Congress Catalog Card Number: 94-69993

# The Organic Chemist's Desk Reference

A COMPANION VOLUME TO THE  
FIFTH EDITION OF ORGANIC CHEMISTRY, SIXTH EDITION

PH. RE. J. ...  
... 20% errors and corrections  
... information by the publisher

JOHN WILEY & SONS  
New York, London, Sydney, Melbourne, India, Japan  
... ..

# Acknowledgements

---

I would like to thank the following people for their contributions, suggestions and additions:

H.M. Bradley

J. Buckingham

R. Edmundson

H. Hailes

R.A. Hill

S.M. Jeffery

F.M. Macdonald

G. Pattenden

R. Purchase

R.A. Raphael

C.W. Rees

M. Reynolds

# Preface

---

Organic chemistry is literally a vital science, in two respects. First, it controls the functioning of living organisms through the operation of the complex series of reactions that power them. The details of these reactions are studied in tandem with the sister discipline, biochemistry. Secondly, an enormous range of organic reactions are carried out daily throughout the world, some on a very large scale, producing the fuels, materials, medicines, insecticides and other essential organic compounds in daily use.

Organic chemists are sometimes accused of being insular, but success in organic chemistry needs a lively appreciation of the other sciences – and not just other branches of chemistry. An assertion perhaps closer to the truth is that in order to make a success of organic chemistry, the practitioner needs to know so many specialist facts and methods that the subject can be daunting to a non-specialist.

Anything that helps the organic chemist, or any other scientist using organic chemical information, to access rapidly as much of this knowledge as possible is to be welcomed. There are many review series and reference works in daily use throughout the world, and one of these is the *Dictionary of Organic Compounds* (DOC), with which some of us have been associated for many years. As part of the major revision of this work that has taken place for the preparation of the new Sixth Edition, the idea arose of a slender companion volume that would include a wide variety of the miscellaneous facts that organic chemists constantly need but which would be out of place in DOC 6 itself.

We make no apology for the fact that one of the functions of this volume is to publicise the new DOC 6 and to help users to get the most out of it. But in addition they will find collected here a mass of useful up-to-date information, which is often irritatingly scattered and difficult to locate, including details of the content of most of the other major reference sources.

This companion volume is published both in hard covers as part of the DOC 6 set, and separately as a modestly priced paperback, which is available in bulk to teachers and research directors. We are confident that it will be widely used, and expect that further editions will be called for: any suggestions concerning data for future inclusions will be welcomed and should be addressed to the DOC editors at Chapman & Hall.

The Editorial Board  
J.I.G. Cadogan, S.V. Ley  
G. Pattenden, R.A. Raphael  
C.W. Rees



# Contents

<b>Acknowledgements</b>	<b>page vii</b>
<b>Preface</b>	<b>ix</b>
<b>1 History of the Dictionary of Organic Compounds (DOC)</b>	<b>1</b>
1.1 The Chapman & Hall Chemical Databank and the Sixth Edition of DOC	1
1.2 The content of DOC 6	2
<b>2 User guide to DOC</b>	<b>3</b>
2.1 Using DOC 6	3
2.2 Compound selection policy	3
2.3 Literature coverage	4
2.4 Organisation of entries	4
2.5 Abbreviations used in DOC 6	11
<b>3 Organic chemistry journals</b>	<b>13</b>
<b>4 Useful reference works in organic chemistry</b>	<b>22</b>
<b>5 Abstracting and indexing publications</b>	<b>25</b>
5.1 <i>Chemical Abstracts</i>	25
5.2 <i>Chemisches Zentralblatt</i>	26
5.3 <i>Beilstein</i>	26
5.4 Other publications	27
5.5 Electronic publications	28
<b>6 Patent literature</b>	<b>30</b>
<b>7 Nomenclature in <i>Chemical Abstracts</i></b>	<b>32</b>
7.1 Introduction	32
7.2 CA Index Names	32
7.3 Older names encountered in CA	37
<b>8 Glossary of terms used in describing organic structures</b>	<b>42</b>
8.1 Introduction	42
8.2 Glossary	43
8.3 Acronyms in organic chemistry	90
<b>9 Molecular formulae</b>	<b>106</b>
9.1 The Hill system	106
9.2 <i>Chemical Abstracts</i> conventions	106
9.3 Checking molecular formulae	106
<b>10 CAS Registry Numbers</b>	<b>108</b>
10.1 Introduction	108
10.2 The Registry Number	108
10.3 Specificity	108
10.4 Duplicate Registry Numbers	108
10.5 Registry Numbers without CA abstract entries	109
10.6 Registry Numbers with asterisks	109
10.7 Registry Numbers and Collective Indexes	109

## Contents

---

<b>11</b>	<b>Linear notations</b>	<b>110</b>
11.1	WLN (Wiswesser Line Notation)	110
11.2	SMILES (Simplified Molecular Input Line Entry System)	110
11.3	Connection tables	111
<b>12</b>	<b>Hazard information</b>	<b>112</b>
12.1	Risk and hazard assessment	112
12.2	Physical properties	112
12.3	Occupational exposure limits	112
12.4	Reactive hazards	113
12.5	Toxicology	113
12.6	Health effects of chemicals	114
12.7	Storage and handling of chemicals	115
12.8	Hazardous reaction mixtures	115
12.9	Disposal of chemicals	115
12.10	Peroxide-forming chemicals	123
12.11	Conclusion	124
12.12	Further reading	124
<b>13</b>	<b>Languages</b>	<b>127</b>
13.1	A German–English dictionary	127
13.2	Russian–English transliteration	133
13.3	The Greek alphabet in organic chemistry	133
13.4	Multiplicative prefixes from Greek and Latin	137
<b>14</b>	<b>SI units</b>	<b>138</b>
14.1	Basic SI units	138
14.2	Common derived SI units	138
14.3	Prefixes used with SI units	138
14.4	Conversion factors	139
<b>15</b>	<b>Commonly used solvents</b>	<b>140</b>
15.1	Boiling points and fire hazards	140
15.2	Handling solvents	140
<b>16</b>	<b>Miscellaneous</b>	<b>144</b>
16.1	Buffer solutions	144
16.2	Resolving agents	144
16.3	Solvents for recrystallisation	147
16.4	Materials used for heating baths	147
16.5	Freezing mixtures	147
16.6	Solvents for extraction of aqueous solutions	147
16.7	Drying agents	149
16.8	Pressure–temperature nomograph	149
16.9	Acid and base dissociation constants	150
<b>17</b>	<b>Spectroscopy</b>	<b>153</b>
17.1	Infrared spectroscopy	153
17.2	Ultraviolet spectroscopy	153
17.3	Nuclear magnetic resonance	158
17.4	Mass spectroscopy	161
	<b>Index</b>	<b>167</b>



# 1 History of the *Dictionary of Organic Compounds* (DOC)

---

The First Edition of DOC, edited by Sir Ian Heilbron and H.M. Bunbury, was published in 1934 in three volumes. DOC was periodically updated, leading to the Fourth Edition (1964), which was updated by annual supplements in the years 1965–79. For the Fifth Edition (1982), DOC was completely revised and transferred onto a fielded database. This database became the Chapman & Hall Chemical Databank (CHCD), which is now the basis for the publication of not only DOC, but also other chemical dictionaries such as the *Dictionary of Natural Products* (DNP) and the *Dictionary of Organometallic Compounds*, which have resulted from the subsequent considerable enhancement of the database. A CD-ROM version of DNP was released in 1992 and of DOC in 1993.

**Sir Ian Morris Heilbron** (1886–1959) was an organic chemist distinguished for his work on natural products. Born in Glasgow, he studied chemistry there at the Royal Technical College (now the University of Strathclyde) before taking his Ph.D. degree at Leipzig, where he studied under Hantzsch. In 1909 he returned to the Royal Technical College, where he was a lecturer until the outbreak of World War I. He served with distinction as an officer during the war, being awarded the DSO and promoted to Lieutenant-Colonel, finishing as Assistant Director of Supplies at GHQ, Salonika. After the war he worked for a short time at the British Dyestuffs Corporation in Manchester, which became part of ICI, before returning to the Royal Technical College as Professor of Chemistry. He later held Chairs at Liverpool (1920–33), Manchester (1933–38) and Imperial College (1938–49). During World War II he acted as a scientific adviser to the Ministry of Production, and he played a leading part in the introduction of DDT. He retired from academic life in 1949 and was appointed the first Director of the Brewing Industry Research Foundation, where he remained until his death in 1959.

**Hugh Mills Bunbury** was born on 21 September 1889. During World War I he served in the City of London Regiment (Royal Fusiliers) and was attached

to the Ministry of Munitions. In 1919 he joined the British Dyestuffs Corporation in Manchester, which became part of ICI, where he became a colleague of Heilbron's. He later trained as a barrister, retiring from ICI in 1967. He wrote *The Industrial Applications of Coal Tar Products* with A. Davidson (1925), *The Destructive Distillation of Wood* (1923) and, much later, edited *Chemists and the Law*, published by E & FN Spon in 1967. He died on 17 May 1972.

The original proposal for a *Dictionary of Organic Compounds* was sent to the publishers by Bunbury in about 1931, and the First Edition was put together largely by Heilbron's research groups in Liverpool and Manchester.

## 1.1 The Chapman & Hall Chemical Databank and the Sixth Edition of DOC

The Chapman & Hall Chemical Databank, from which DOC is produced, is a selected and carefully edited resource containing chemical and physical data together with key literature references and structural information for approximately 300 000 chemical substances. The database was originally set up to produce the Fifth Edition of DOC, which was published in 1982. It has subsequently been considerably expanded to produce various further dictionaries, notably:

- *Dictionary of Organometallic Compounds* (First Edition 1984; Second Edition 1994)
- *Dictionary of Organophosphorus Compounds* (1988)
- *Dictionary of Drugs* (1990)
- *Dictionary of Inorganic Compounds* (1992)
- *Dictionary of Analytical Reagents* (1993)
- *Dictionary of Natural Products* (1994)

The new, Sixth Edition of DOC (1995) is the second to be produced from the database. Each of the main

## History of the dictionary of organic compounds (DOC)

---

dictionaries is continually updated. In print format, an annual Supplement is produced for each Dictionary.

Since 1993, these dictionaries have also been available in CD-ROM format, which features text searching and display by Headfast software provided by Head Software International, and substructure searching using PsiBase from Hampden Data Services. The two packages function on the same disk and are seamlessly linked to provide very flexible searching of the database by text, structure or a combination of both. The three main CD-ROM dictionaries are available on subscription, which provides a complete update every six months; the *Dictionary of Analytical Reagents* and the *Dictionary of Drugs* are available as one-off purchases.

### 1.2 The content of DOC 6

Considerable expansion of the databank that supports DOC has taken place since the Fifth Edition Main Work was published in 1982. Some reorganisation of the content of DOC has taken place in the new Sixth Edition, most notably concerning the coverage of natural products.

#### *Natural products*

In the late 1980s it was decided to publish a separate *Dictionary of Natural Products*, which would have the aim of being a comprehensive record of all known natural substances. DOC 6 therefore contains fewer natural product entries than DOC 5 but with a more rational focus, giving (sometimes abbreviated) entries only for the most widespread, practically important and structurally typical natural products. The *Dictionary of Natural Products* (published 1994) and its supplements (like DOC, available also in

substructure-searchable CD-ROM format) gives a comprehensive coverage of the nearly 100 000 natural products now known.

#### *Drugs*

Similarly, DOC 6 focuses on a limited coverage of the most important drugs in common use: in 1996, DOC will be joined by a *Dictionary of Pharmacological Agents*, giving a much wider coverage, including substances of pharmaceutical interest that are not yet marketed.

#### *Organometallics*

DOC 5 contained a very limited number (about 500) of entries concerning the most fundamental organometallic structures, e.g. ferrocene. This coverage was discontinued with the compilation of the much more comprehensive *Dictionary of Organometallic Compounds* (First Edition 1984; Second Edition 1994). In this edition of DOC there are no entries for true organometallic compounds, but a good coverage of the organic compounds of boron and silicon, especially those with synthetic applications.

#### *Increased coverage of core compounds*

DOC 5 contained information on approximately 90 000 compounds, including numerous natural products, whereas DOC 6 documents 160 000 substances, with only restricted natural product coverage. Therefore, it will be clear to users that there has been a substantial increase, about 250%, in the fundamental starting materials, reagents and interesting target molecules that are the categories of compound of most interest to mainstream organic chemists, particularly those involved in synthesis. In order to ensure that this coverage corresponds as closely as possible to users' needs, the literature has been carefully reviewed to mid-1994.

## 2 User Guide to DOC

---

### 2.1 Using DOC 6

As with previous editions, the arrangement of entries is alphabetically by DOC Name. Thus, in cases where there is no possible ambiguity about the name, the compound can be located immediately without using the index. However, the majority of even quite simple organic compounds have at least two completely valid names (e.g. methylbenzene/toluene), and therefore, if a compound cannot be located immediately, the Name Index should be used. This includes all DOC Names and alternative names given throughout the DOC, including those applicable to derivatives. A Molecular Formula Index and CAS Registry Number Index are also provided.

Each entry is numbered to assist its ready location. The DOC Number consists of a letter of the alphabet, a supplement number, followed by a five-digit number. In the 'Main Work' volumes the first digit is invariably 0; entries in the First Supplement will carry a first digit 1; entries in the Second Supplement will carry the first digit 2; and so on. All index entries refer to the DOC Number. Each index is described in detail in the appropriate volume.

Many compounds appear as derivatives of parent compounds; compounds such as ethers, esters and *N*-methyl derivatives, if they cannot be located immediately, should be looked for (a) under the parent, (b) using the Name Index (in which all names of derivatives appear) or (c) using the Formula Index (which includes molecular formulae for all derivatives except those of a characterisation nature such as 2,4-dinitrophenylhydrazones).

### 2.2 Compound selection policy

#### 2.2.1 General

Many thousands of new entries have been added to the new edition, and extensive reviews have been carried out of suitable sources to ensure that as far as possible the needs of all potential users of DOC have

been taken into account. In general, DOC 6 covers the following classes of compounds:

- The basic fundamental organic compounds of simple structure that are frequently required as starting materials, and which have usually been the subject of extensive physicochemical study.
- The most important and widespread natural products. (The *Dictionary of Natural Products* provides comprehensive coverage of natural products, including more extensive entries for many of those which have DOC 6 entries.)
- Compounds with a well established use, e.g. pesticides and drugs in current use.
- Laboratory reagents and solvents.
- Other compounds with interesting chemical, structural or biological properties, including 'intriguing' molecules that have been specially synthesised in order to investigate their chemical and physical properties.

This coverage has been especially reviewed and enhanced for the new edition, and users should find that, for the majority of queries, it corresponds to their needs. The extensive inclusion of further orientating information about the compounds covered, such as CAS Names and CAS Registry Numbers, makes it easier to carry out a deeper search if more extensive information is needed about a particular compound.

The Editors are always pleased to receive comments on the selection policy, and in particular to receive specific suggestions for compounds or groups of compounds for inclusion.

#### 2.2.2 Derivatives

Some types of compound are usually treated as derivatives of parent compounds, for example:

- Hydrates, complexes (e.g. picrates).
- Salts and quaternary salts, e.g. hydrochlorides and methobromides.
- Classical organic derivatives; for example, entries for esters, acid chlorides, amides and nitriles will be found under the parent acid unless they are

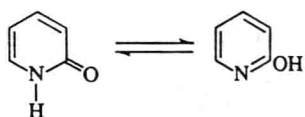
particularly important compounds in their own right, justifying their own entry.

All important derivatives occur in the Formula Index and can be traced in that manner.

### 2.2.3 Tautomerism in DOC 6

For the Sixth Edition of DOC the presentation of information on tautomeric compounds has been improved and standardised.

The commonest type of tautomerism encountered is the  $\text{NH} \rightleftharpoons \text{OH}$  equilibrium shown by many heterocyclic compounds, e.g. 2(1*H*)-pyridinone  $\rightleftharpoons$  2-pyridinol (2-hydroxypyridine):



The position of equilibrium for a particular compound depends on various factors, but in general the NH form predominates for most heterocyclic compounds under most conditions. Simple derivatives may be either: (a) derivatives that are themselves capable of tautomerism, e.g. hydrochloride, *N*-oxide; (b) derivatives of the NH form, e.g. *N*-methyl; or (c) derivatives of the OH form, e.g. methyl ether. The placement of the derivatives within the DOC entry reflects these possibilities.

The policy for naming such compounds in DOC 6 is as follows:

1. All possible synonyms relating to both the NH and OH forms are given and can readily be found in the Name Index regardless of which tautomer is looked up.
2. The entry name for very simple compounds such as that shown above, for which the tautomerism is well studied, is the predominant tautomer, i.e. 2(1*H*)-Pyridinone.
3. In more complex cases where the tautomerism may or may not have been carefully investigated for individual members of the series, the hydroxy-heteroaryl name is normally used as the entry name, with a note where appropriate indicating that the structure represented by the entry name may not be the predominant tautomer. This greatly simplifies the presentation of series of compounds (e.g. chlorodihydroxypyridines) where some are capable of NH tautomerism and some are not.

The same general principle is used for more complex cases where several different tautomers are present and where even 'blocked' derivatives such as *N*-methyl may still be capable of restricted tautomerism. For an example in DOC 6, see the series of naphthyridinediols.

The presentation of other types of tautomerism in DOC 6 follows the same general policy. In a few cases involving ring  $\rightleftharpoons$  chain tautomerism, the two tautomers have separate entries with cross-references between them.

### 2.2.3 Anions and cations

For ionic substances such as quaternary ammonium salts, the entry refers to the anion or cation and the molecular formula and molecular weight given are those of the ion. The various salts (e.g. chloride, nitrate) are treated as derivatives, each with its own molecular formula. Where a substance such as a dye is normally prepared and handled as, for example, a sodium salt, the entry usually refers to the parent acid.

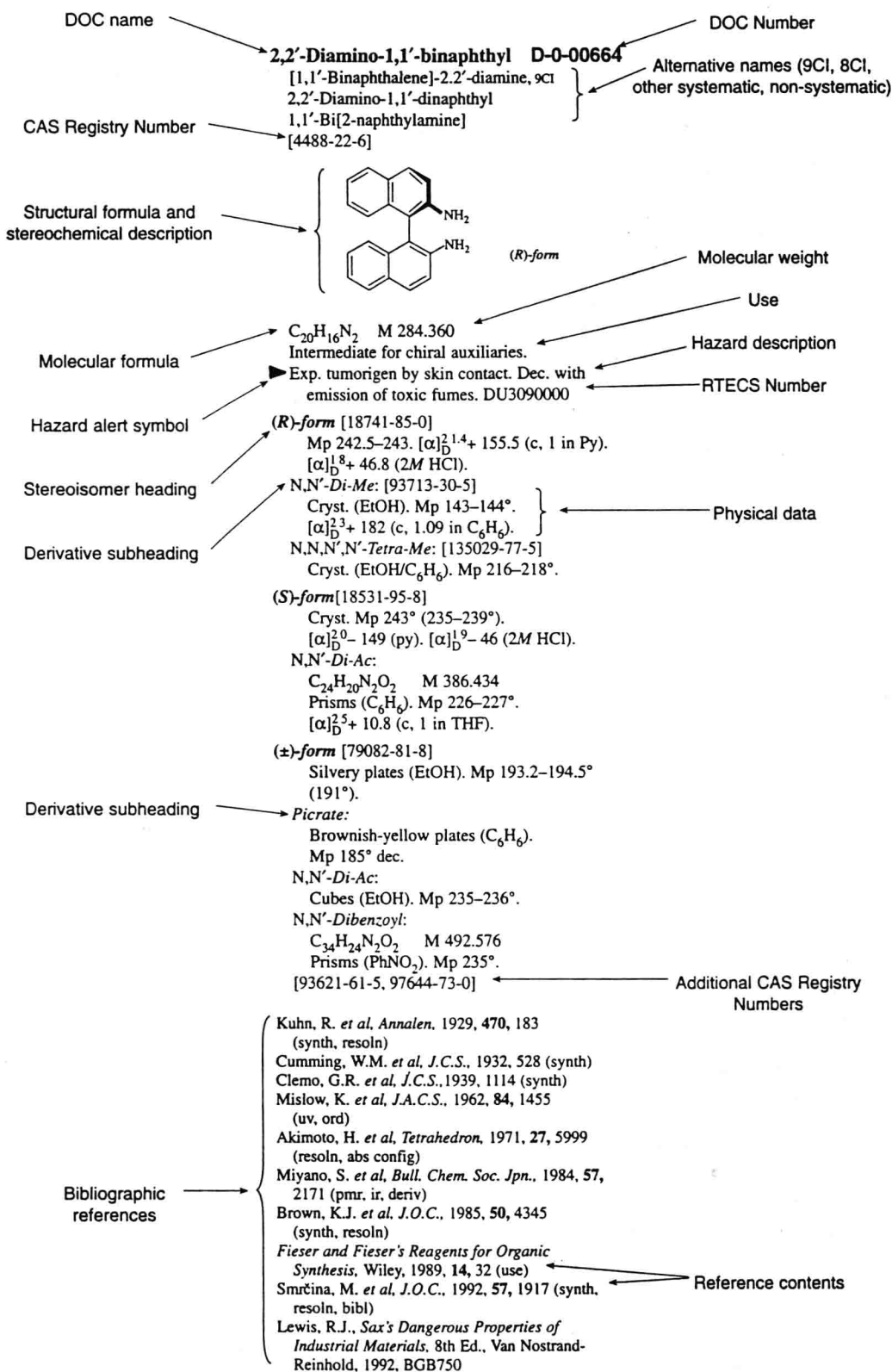
## 2.3 Literature coverage

In compiling this edition the primary literature has been surveyed to mid-1994, and outstandingly important information from the second half of 1994 has been incorporated.

The first annual Supplement, which will appear in 1996, will survey the literature to mid-1995; thereafter annual Supplements will appear in the middle of each year and will be based on the literature to the middle of the previous calendar year. An even more topical coverage of the literature can be obtained from the CD-ROM version, which is republished complete on one disk every six months.

## 2.4 Organisation of entries

In general, the format of individual entries remains similar to that of previous editions. Figure 2.1 illustrates the format of a typical entry within which the individual types of data have been labelled. The range of information included within the entries is described below.



**Figure 2.1** The format of a typical entry in DOC 6, showing the individual types of data that may be included.

The *Dictionary of Organic Compounds* is also available as a database on CD-ROM. Entries may be displayed on screen or output in a format similar to that of the printed work. Powerful text retrieval and structure search software allows searches to be carried out not only on chemical names, molecular formulae and CAS Registry Number but also text words, hazard and toxicity information, labelled references, physical description and the physical properties using browsable indexes.

Structure and substructure searching using PsiBase for Windows™ rapidly allows the user to find all entries with particular structural features, and searches may be readily combined or refined.

The CD-ROM version of DOC 6 is available on subscription, which provides bi-annual updates in the form of a replacement CD-ROM containing updates and important new compounds from the chemical literature.

### 2.4.1 Chemical names and synonyms

#### (a) DOC Names

The DOC Name is that chosen to head each entry and is that which, in the opinion of the Editors, is most likely to be known by, and of use to, most readers. Systematic names following IUPAC conventions are used wherever feasible, but trivial names are used for most natural products, where systematic names are too cumbersome for convenient use. In cases where no one name stands out as being clearly more familiar or convenient than others, the *Chemical Abstracts* name is usually given precedence. In this edition, there has been some further revision of DOC nomenclature towards greater standardisation with CAS, especially for example in the choice of CAS Names such as '1,2-benzenediamine' as the entry names for aromatic and heterocyclic compounds in preference to the formerly preferred '1,2-diaminobenzene', etc.

The legislating body for chemical nomenclature is the International Union of Pure and Applied Chemistry (IUPAC). It is important to recognise that, while IUPAC lays down rules and principles for good nomenclature, application of those rules will not necessarily lead to a unique name for each compound. In other words, a compound may have more than one valid name conforming to IUPAC principles. Developments of IUPAC nomenclature, especially

CAS nomenclature, introduce more rigid principles so as to arrive at a unique name for every possible compound.

It is important to emphasise this point, because it is a common misconception that the correct application of IUPAC rules will produce one, and one only, name for each compound, and therefore if a publication such as DOC presents two or more names, one of them must be 'wrong'. The situation is described in more detail in the preamble to the publication *A Guide to IUPAC Nomenclature of Organic Compounds, Recommendations 1993*, eds R. Panico *et al.* (Blackwell):

It is important to recognize that the rules of systematic nomenclature need not necessarily lead to a unique name for each compound, but must always lead to an unambiguous one. Lucidity in communication often requires that the rules be applied with different priorities. A comparative discussion of the compounds  $\text{CH}_3\text{-CH=CH}_2$ ,  $\text{Cl-CH}_2\text{-CH=CH}_2$ ,  $\text{C}_6\text{H}_5\text{-CH}_2\text{-CH=CH}_2$ ,  $\text{H}_2\text{N-CH}_2\text{-CH=CH}_2$ , and  $\text{HO-CH}_2\text{-CH=CH}_2$ , might be easier to follow if they were all named as propenes, even though with the last three, the benzene ring, amino, and hydroxy groups may have seniority over the double bond for citation as a parent or as a suffix. In other cases, a set of rules that generates clear and efficient names for some compounds can lead to clumsy and nearly unrecognizable names for others, even closely related ones. To force the naming of all compounds into the Procrustean bed of one set of rules would not serve the needs of general communication, and the Commission believes that the majority of organic chemists would not accept such a policy for general communication. This situation can be illustrated by a compound that most chemists would probably name as 'pentaphenylethane', instinctively, whereas the application of a principle favouring rings over chains leads to a name such as 'ethanepentaylpentabenzene'. The first name is certainly more easily recognized than the second ....

In view of the foregoing considerations, this *Guide to IUPAC Nomenclature of Organic Compounds* often presents alternative sets of rules, equally systematic, wherever available and justifiable, to enable a user to fit the name to a particular need.

Lastly, the Commission recognizes that for certain types of compounds, there is significant disagreement among chemists in different fields as to what should be the preferred nomenclature. This situation leads to an apparent lack of decisiveness in some of the recommendations in this document. This is unavoidable, because long experience has taught that formulating rules not having general support is a futile exercise; such rules will be widely ignored. Therefore, the Commission's policy is to offer critically examined alternatives, some of which may be new proposals, and to observe how they are accepted and used. If one of the alter-



natives subsequently becomes preferred to an overwhelming extent by the community of chemists, a future edition of recommendations can reflect that fact.

In this Guide, some practices of the Chemical Abstracts Service and/or of the Beilstein Institute have been mentioned. This is done only for informational purposes, and such instances are not necessarily recommendations of the Commission. The Commission recognized that there are circumstances that require a preferred (i.e. unique) name. These include comprehensive indexing (such as for the volume indexes to *Chemical Abstracts*) in order to avoid an intolerable amount of cross-indexing and multiple entries. This need is being met in a particular way by Chemical Abstracts Service as in-house procedures designed to place compounds with the same parent skeleton together while at the same time minimizing the number of rules. The *Chemical Abstracts Index Guide* treats the majority of compounds, but is not complete. There are a number of other in-house procedures applied elsewhere, such as in *Beilstein* (not yet explicitly published).

### (b) Synonyms

An important function of DOC is to present a wide range of synonyms. In general the selection is made as useful as possible, but no attempt is made to provide exhaustive lists of proprietary names for pharmaceuticals and other commercially available compounds.

Editorial policy on minor points of nomenclature is as follows:

- Propanoic vs propionic and butanoic vs butyric; the former are preferred, particularly in entry names.
- Sulfur vs sulphur; the former is used.
- *Chemical Abstracts* practice regarding the position of locant numerals is generally followed; e.g. 2-hepten-1-ol is preferred to hept-2-en-1-ol.
- Eicosa vs icosia and oestr vs estr; the former are given priority but the alternative spellings are given as synonyms and thus appear in the Name Index.

The editorial generation of new synonyms that are not in the literature has been kept to a minimum.

### CA names

Names corresponding to those used by Chemical Abstracts Service during the Eighth and Ninth Collective Index periods (1967–71 and 1972–76 respectively) are labelled with the suffixes 8CI and 9CI respectively. It should be noted that 9CI nomen-

clature is defined as that brought into use by CAS at the beginning of the Ninth Collective Index period, and that for organic compounds has been carried over essentially unchanged into subsequent Collective Index periods. Therefore the suffix '9CI' does *not* mean that the compound can necessarily be found in the Ninth Collective Index, as it may have been indexed only since 1976.

### Recommended names

Names recommended by various regulatory bodies and standards associations are given suffixes to denote their origin:

ANSI	American National Standards Institute
BAN	British Approved Name
BSI	British Standards Institution
INN	International Nonproprietary Name
ISO	International Standards Organisation
JAN	Japanese Accepted Name
USAN	United States Adopted Name
WSSA	Weed Science Society of America

Note that no distinction is made between proposed and recommended INNs.

### 2.4.2 CAS Registry Numbers

CAS Registry Numbers have been included for as many compounds as possible. For a description of CAS Registry Numbers, see Chapter 10. While every attempt has been made to achieve *accuracy* of the reported Registry Numbers, no guarantee can be provided as to the *comprehensiveness* of the range of numbers presented. Thus, the absence of a Registry Number does not necessarily imply that one has not been allocated.

Registry Numbers that clearly belong to an entry but which cannot be unequivocally matched up to any of the individual compounds covered by that entry are given at the end of the entry. These additional Registry Numbers fall into one of the following categories:

- Duplicate Registry Numbers.
- Registry Numbers assigned to trivial variants of a compound, e.g. hydrates.
- Registry Numbers of stereoisomers or derivatives for which no physical data can readily be found.
- Registry Numbers referring to entities that are not specifically treated by the DOC entry.

## User guide to DOC

- Registry Numbers referring to non-specific isomers.

For further information about CAS Registry Numbers, see Chapter 10.

### 2.4.3 Structural formulae

The structures in DOC are drawn as accurately as possible according to best current 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 orientated in the standard Haworth manner so that structural comparisons can quickly be made.

In a series of closely related compounds, e.g. a series of aromatic or heteroaromatic isomers, the structural formula is given only for the first member.

### 2.4.4 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.

Various methods of describing stereochemistry are used (these are described elsewhere in this volume):

- The (*R,R*)-system for chiral molecules (see Section 8.2, under *R*-).
- The *D,L*-system for sugars and amino acids (see Section 8.2, under *D*-).
- The  $\alpha,\beta$ -system for complex natural products such as steroids, for other cyclic molecules and for anomers of sugars (see Section 8.2, under *steroids*).
- The (*E,Z*)-system for specifying configurations at double bonds (see Section 8.2, under *E*-).
- The *ent*-convention where there is configurational inversion at all of the chiral centres whose configuration is implied in a name (see Section 8.2, under *ent*-).
- The *sn*-convention for glycerides (see Section 8.2, under *sn*-).

### 2.4.5 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 computer calculated from the most current IUPAC table of atomic weights and rounded to three decimal places. 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.

### 2.4.6 Importance/use

Care has been taken in DOC to make the information given on the importance and uses of chemical substances as accurate as possible. Wherever possible, information on a particular use has been checked against a critical source, such as the *Kirk-Othmer Encyclopedia of Chemical Technology* or *Ullmann's Encyclopedia of Industrial Chemistry*.

### 2.4.7 Physical data

#### (a) Melting points and boiling points

Melting and boiling points are reported in degrees Celsius. The policy followed in the case of conflicting data is as follows:

- Where the literature melting points are closely similar, only one figure (the highest or most probable) is quoted.
- 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° (135–136°).
- Where quoted figures differ widely and some other explanation such as polymorphism or incorrect identity seems the most likely explanation, both figures are quoted without parentheses; thus 142°, 205–206°.
- Known cases of polymorphism or double melting point are noted.
- Boiling points are given at atmospheric pressure unless otherwise indicated. The pressure in mmHg (if not atmospheric) is given as a subscript, e.g.

Bp<sub>100</sub> 85°. Some boiling points are now quoted in the literature with the pressure given using the SI unit kPa (kilopascals); for the time being, DOC retains mmHg. The conversion factor is: 1 mmHg = 0.133 222 kPa; 1 kPa = 7.500 64 mmHg.

**(b) Optical rotations**

Optical rotations are given wherever possible. They are expressed in the form:  $[\alpha]_D^{20} +30.6$  (c, 1.2 in CHCl<sub>3</sub>). This denotes a temperature of 20 °C, wavelength at the sodium D line (589 nm) and a concentration of 1.2 g/100 ml in chloroform solution. In many cases, an indication of optical purity (op) or enantiomeric excess (ee) is reported after the value, if it is given in the literature. The degree sign formerly given following optical rotations, and which is still extensively found in the primary literature, has been dropped as it is dimensionally incorrect.

**(c) Densities and refractive indices**

Densities and refractive indices are now of less importance for the identification of liquids than has been the case in the past, but are still quoted for relatively common substances such as solvents. Many literature values for refractive indices are based on relatively impure samples obtained before modern purification methods were available. This applies particularly to natural products such as monoterpenes. Densities and refractive indices are not quoted where the determination appears to refer to undefined mixtures of stereoisomers.

**(d) Solubilities**

Solubilities are given only where the solubility is unusual for an organic compound. Most organic compounds not containing polar groups are soluble in typical organic solvents, e.g. ether, and insoluble in water.

**(e) pK<sub>a</sub> values**

pK<sub>a</sub> values are given for both acids and bases. The pK<sub>a</sub> of a base can be obtained by subtracting its pK<sub>b</sub> from 14.17 (at 20 °C) or from 14.00 (at 25 °C).

**(f) Spectroscopic data**

Spectroscopic data, such as uv wavelengths and extinction coefficients, are given 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 located through the references quoted. Spectroscopic data (initially uv and cmr) are now being introduced into the CD-ROM versions in a form that makes the data both numerically searchable and displayable on-screen.

## 2.4.8 Toxicity and hazard information

**(a) General**

An important function of a DOC 6 entry is to alert the user to potential hazards associated with the use of the compound. This information is highlighted by the sign ▷ (which also appears in the indexes). For this edition of DOC, all hazard and toxicity data have been carefully and critically assessed, and re-edited by a specialist Editor. Brief summaries of hazard and toxicity information have been included with the data for many chemical substances, particularly where there have been reports of adverse effects in people or where an incident in a laboratory has revealed the reactive nature of a chemical. As with the RTECS Accession Numbers, the sign ► is used to highlight summaries of hazard and toxicity data. Any references to this information in the primary literature or the standard monographs that have been quoted in this edition will usually carry the tags (*haz*) or (*tox*). A more detailed explanation of the choice of hazard and toxicity information for DOC may be found in Chapter 12.

Although much care has been taken to ensure the accuracy and completeness of reported data, *DOC must not be considered a comprehensive source on hazard data*. The function of hazard data in DOC 6 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.

Further details of DOC coverage of toxicity and hazard information can be found in Chapter 12.

**(b) RTECS Accession Numbers**

Many entries in DOC 6 contain one or more RTECS Accession Numbers. These numbers refer to toxicity information on the relevant compounds from the