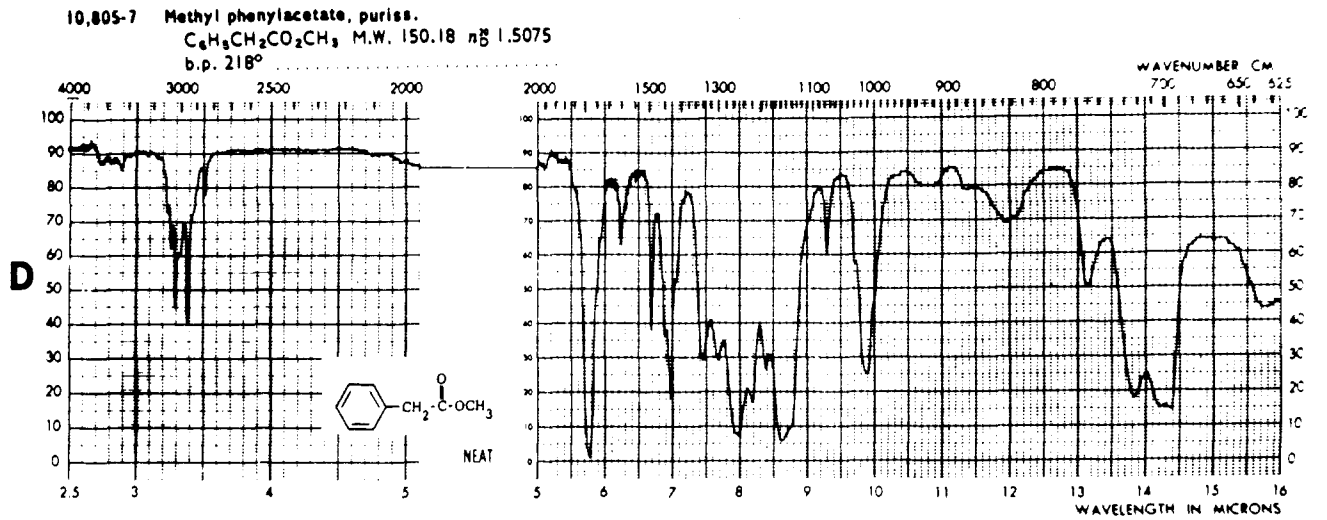

Figure 9. The spectrum of the compound in Figures 7 and 8 as given in the Aldrich catalog.[9]

ROG, Chemical and Cosmetical works, Ljubljana, and the Research Community of Slovenia.

REFERENCES AND NOTES

(1) Zupan, J. *Anal. Chim. Acta* **1978**, *103*, 273.
(2) Clerc, J. T.; Zupan, J. *Pure Appl. Chem.* **1977**, *49*, 1827.
(3) Zupan, J.; Penca, M.; Razinger, M.; Barlič, B.; Hadži, D. *Anal. Lett.* **1979**, *12*, 109.
(4) Zupan, J.; Hadži, D.; Penca, M.; Razinger, M.; Barlič, B. *Anal. Chim. Acta* **1980**, *122*, 103.
(5) Penca, M.; Zupan, J.; Hadži, D. *Kem. Ind.* **1978**, *27*, 137.
(6) Razinger, M.; Zupan, J.; Penca, M.; Barlič, B.; Hadži, D. Proceedings of the World Spectroscopy Conference, Lisbon, 1979, in press.
(7) Smith, E. G.; Baker, P. A. "The Wiswesser Line-Formula Chemical Notation", 3rd ed.; CIMI: Cherry Hill, N.J., 1975.
(8) Aldermaston Mass Spectrometry Data Collection, Reading, U.K.
(9) The Aldrich Library of Infrared Spectra, Pouchert, C. J., Ed.; Aldrich Chemical Inc.: Milwaukee, Wis., 1970.
(10) Wyandotte-ASTM Infrared Data File; ASTM: Philadelphia, Pa., 1964.
(11) CNMR Data Base, NIC, Delft, P.O.B. 36, 2600AA Delft, The Netherlands.

Parent Compound Handbook—Successor to The Ring Index†

J. E. BLAKE,* S. M. BROWN, T. EBE, A. L. GOODSON, J. H. SKEVINGTON, and C. E. WATSON

Chemical Abstracts Service, P.O. Box 3012, Columbus, Ohio 43210

Received March 10, 1980

This paper outlines the development of the PCH, reviews its growth in the three years since its introduction, briefly discusses its production, and illustrates some of the ways it can be useful to the chemical community.

INTRODUCTION

Chemical Abstracts Service (CAS) first published the "Parent Compound Handbook" (PCH) in March 1977 as the successor to "The Ring Index".[1] Since that time it has been updated regularly. The PCH is both a current-awareness service and a major reference work on ring systems and natural products for those who use *Chemical Abstracts* (CA) and its associated indexes. At present it contains more than 50 000 entries. These include chemical structure diagrams, CA index names, Wiswesser Line Notations (WLN),[2,3] CAS Registry Numbers, molecular formulas, systematic names for natural products, and ring data for cyclic parent compounds. The more recent entries also include CA references indicating the first occurrence of the ring systems in CAS processing. Access to the information that the PCH contains is provided by means of a set of six indexes, which allow searching by ring analysis, ring substructure, name, WLN, molecular formula, and CAS Registry Number.

BACKGROUND AND HISTORY

From the initial publication of *Chemical Abstracts* in 1907, chemical structure diagrams have been used in CA issues to illustrate specific chemical substances. However, they were not used for the first ten years in the annual indexes because the alphabetical index entries for related substances (such as bromophenanthrene, chlorophenanthrene, and nitrophenanthrene) were widely scattered, depending on the names of the substituents. Only with the adoption of inverted index nomenclature in Volume 10 and in the First Decennial Index in 1916 did the inclusion of chemical structure diagrams in the indexes become feasible. Index entries for closely related substances were no longer scattered throughout the index at several headings depending on the substituent name, but were

† Presented in part to the Division of Chemical Information, 173rd National Meeting of the American Chemical Society, New Orleans, La., March 22, 1977.

0095-2338/80/1620-0162$01.00/0

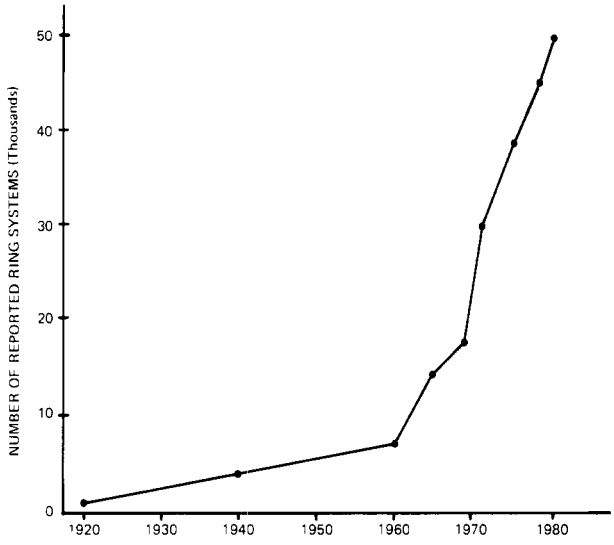

Figure 1. Growth of reported ring systems.

collected together at a common index heading such as phenanthrene. It thus became possible to include a diagram showing the structure and nomenclature locants of some of the index heading compounds.

In 1921, it was suggested that a catalog of the more than 800 ring systems identified during CA processing would be a very useful publication. The diversity of numbering systems and the absence of a uniformly accepted system of nomenclature for rings forestalled publication of a ring catalog at that time. In the ensuing years, strong efforts were directed toward developing an acceptable set of uniform rules for numbering ring systems. Finally, in 1938, a set of ring-numbering rules was adopted by the Commission on Nomenclature of Organic Chemistry of the International Union of Chemistry. In 1940, Patterson and Capell published as an ACS monograph a ring catalog, "The Ring Index", containing 3978 ring systems.

In the period between 1940 and 1960, the number of ring systems nearly doubled and the rules for naming and numbering rings were further refined. These events led to the publication of the second edition of "The Ring Index" by Patterson, Capell, and Walker in 1960, containing information on 7272 rings. Supplements to this edition of "The Ring Index" were published in 1963, 1964, and 1965, by which time the total number of rings covered was 14 265. The third supplement extended the literature coverage through the end of 1963. Although no further supplements were published, "The Ring Index" continued to be used, but a more up-to-date ring collection was desired.

The next collection of ring systems, although incomplete, was published in 1969 in the first issue of the CA Index Guide. In addition to other information, it included structure diagrams for almost 18 000 rings. By the time the Eighth Collective Index Guide was published in 1972, the Index Guide included over 30 000 structure diagrams. This continuing growth of the number of organic ring systems is illustrated in Figure 1. CAS currently processes 230 new rings every month.

At the beginning of the Ninth Collective Period (1972–1976), ring parent and stereoparent structure diagrams were removed from the Index Guide so that they could be incorporated into a new publication. As the successor to "The Ring Index", this new publication would serve as a reference support for the users of CA. The main goals for the new publication were that it should be a concise reference work for ring parents, natural product parents (stereoparents), and boron parents (cage parents); an up-to-date reflection of new parents as they were encountered in routine CA processing; and a useful reference with the data contained therein accessible via several different approaches.

The "Parent Compound Handbook", first published in March 1977, was developed with these goals in mind. The PCH has grown from 40 466 entries in August 1976[6] to 49 906 entries in November 1979. It has acquired new ring parents, stereoparents, and cage parents at an annual rate of 7.2%. The number of new ring parents added each year has remained fairly constant over the lifetime of the PCH.

PARENT COMPOUNDS

CAS defines the term *Parent Compound* as a member of the general class of substances which comprise the fundamental building blocks of CA index nomenclature. Parent compounds are specific molecules upon which derivation of CA index substance names is based. They include classes of CA index heading parents such as chains, rings, peptides, metallocenes, carbohydrates, steroids, alkaloids, boron cages, and so forth. For most parent compounds, an illustrative structure diagram is created for the CA Chemical Substance Index to show the structure, nomenclature locants, and stereochemistry (if any) for an entire set of index heading parents.[4] It is not necessary for the particular parent compounds themselves to have been described as such in the literature, since a bond variation or a derivative may actually have been reported instead. The subset of parent compounds contained in the PCH includes four types: ring parents such as naphthalene, cyclic stereoparents such as pregnane, acyclic stereoparents such as L-alanine, and cage parents such as 1,2-dicarbadodecaborane-(12).

THE PARENT COMPOUND FILE

The "Parent Compound Handbook" consists of two major divisions. The first division, known as the Parent Compound File (PCF), contains structure diagrams and associated information on the parent compounds. The PCF is subdivided into four sections: Cage Parents, Acyclic Stereo Parents, Cyclic Stereo Parents, and Ring Parents. Each section is ordered by a five-letter alphabetic code known as the Parent Compound Identifier, which conveys no information other than the parent compound's location in the PCF.

Each of the four sections of the Parent Compound File is assigned a range of alphabetic identifiers as follows:

- Cage Parents (containing boron parents and metallocenes): BBBBB–BPZZY;
- Acyclic Stereo Parents (including amino acids, carbohydrates, and other stereoparents which contain no parent rings): BQBBR–BZZZP;
- Cyclic Stereo Parents (including alkaloids, steroids, carbohydrates, cyclic amino acids, and other stereoparents which contain one or more parent rings): CBBBC–DZZZR; and
- Ring Parents (containing all the parent rings in the CAS Chemical Registry System): FBBBF–ZZZZH.

These alphabetic identifiers simply serve as the link between the indexes and the PCF.

Each unique parent compound is assigned its identifier sequentially just prior to publication. It is then automatically entered into the proper section of the file according to the identifier that has been assigned it. The ranges contain enough identifiers to allow for the future growth of each section of the PCF.

Examples of entries from the Parent Compound File are shown in Figures 2 through 5. Although there is some variation from section to section, the information provided for each entry includes the CA index name (both current and Eighth Collective Period), the CAS Registry Number, the molecular formula, the Wiswesser Line Notation, the ring analysis data,

BBDSV 41655-26-9
5,6-Dicarbadecaborane(12)
$C_2H_{12}B_8$

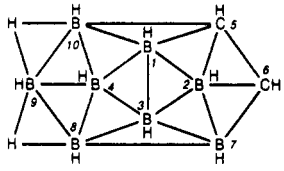

Figure 2. Cage Parent entry.

BQCHK 327-57-1
L-Norleucine ((*S*)-2-*aminohexanoic acid*)
8CI Norleucine, L-
$C_6H_{13}NO_2$
ZY4&VQ -B
Text descriptor 5:L

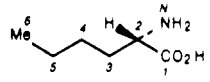

Figure 3. Acyclic Stereo Parent entry.

CGXLE 64785-10-0
Androst-6-eno[6,7-*d*][1,2,3]oxadiazole (*[7a*R*-(7aα,7bβ,9aα,12aβ,12bα)]-3b,5,6,7,7a,7b,8,9,9a,10,11,12,12a,12b-tetradecahydro-7a,9a-dimethyl-4H-cyclopenta[1,2]phenanthro[9,10-d]-1,2,3-oxadiazole*)
$C_{19}H_{28}N_2O$
T C5 H6 B665 DONN&TTTTJ A1 Q1
Text descriptor 4:.ANDROST..5
5-Ring System
5,5,6,6,6
C_2N_2O-C_5-C_6-C_6-C_6
CA 87:184766e

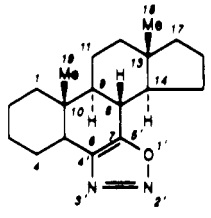

Figure 4. Cyclic Stereo Parent entry.

GPRQD 24592-58-3
Spiro[furo[2',3':4,5]cyclopenta[1,2-*b*]naphtho[2,1-*d*]pyran-8(1*H*),2'-[2*H*]pyran]
$C_{22}H_{14}O_3$
T F5 E5 B666 DO HXO QHJ H-& BT6 OXJ
6-Ring System
5,5,6,6,6,6
C_4O-C_5-C_5O-C_5O-C_6-C_6

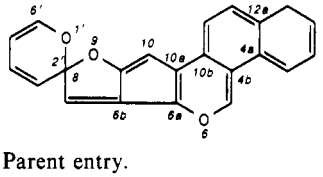

Figure 5. Ring Parent entry.

the CA reference (limited to the more recent entries only), and a structure diagram which shows the nomenclature-locant numbering system for the CA index name. In addition, each stereoparent entry includes a systematic name and the CAS Chemical Registry System Text Descriptor which describes the stereochemistry of the entry.

THE INDEX OF PARENT COMPOUNDS

The other major division of the PCH is known as the Index of Parent Compounds (IPC). The IPC includes six different indexes or access points: CA Index Name, WLN, CAS Registry Number, molecular formula, ring analysis, and ring substructure. These provide the user with a variety of ways to locate the entries in the Parent Compound File. As illustrated in Figure 6, the key to the PCF in each case is the Parent Compound Identifier assigned to each parent compound.

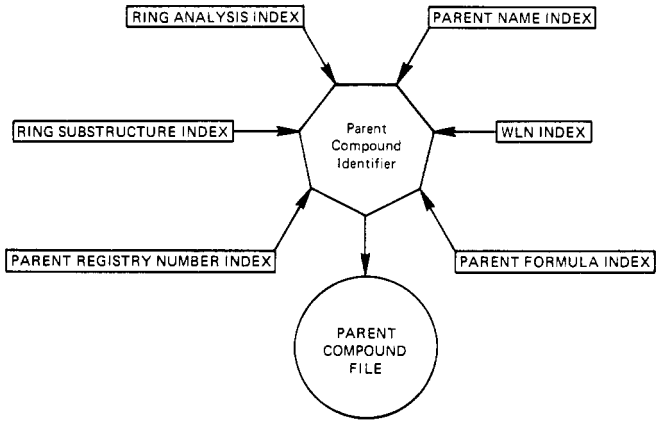

Figure 6. Access points for the PCH.

GPRTB 7247-00-9
● Spiro[furo[3',2':1,2]indeno[5,4-*g*]-3-benzazepine-9(1*H*),2'-[2*H*]pyran]
$C_{23}H_{15}NO_2$
T F5 E5 B667 HXO SN RHJ H-& BT6 OXJ
6-Ring System
5,5,6,6,6,7
C_4O-C_5-C_5O-C_6-C_6-C_6N
●

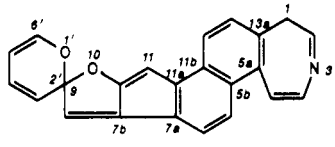

Figure 7. A replacement page entry.

The Ring Analysis Index, like "The Ring Index",[1] is arranged according to ascending order of ring complexity. In addition to parent compounds, it contains names of stereoparent ring modifications as well as CA references for cyclic natural products whose structures are not yet completely elucidated.

The Ring Substructure Index lists in ascending order all cyclic stereoparents and ring parents by the individual component rings of their ring analyses. Entries for component rings up to ring size 9 are subdivided by ring line formulas if necessary. The index is also permuted so that ring components containing common heteroatoms are grouped for easy user access.

The Parent Name Index is an alphabetical listing of the current CA index names of all the parent compounds, as well as undefined or partially defined natural products. The names are permuted to group identical name roots.

The Wiswesser Line Notation Index is a useful reference tool for checking the notations of parent compounds. Entries are listed in a modified alphanumeric order. It is the only index that groups parents having the same ring skeleton.

The Parent Formula Index, which includes parent names, is ordered by the ascending molecular formula of parent compounds.

The Parent Registry Number Index is a collection of all the CAS Registry Numbers of parent compounds and undefined or partially defined natural products. Its entries are listed in ascending numerical order.

UPDATING THE PARENT COMPOUND HANDBOOK

The PCH is updated every two months by a set of supplements. Each supplement, whose cover date indicates the time at which the data was obtained from the CAS Information Base, consists of three parts. The first part contains a set of

new pages for the sections of the Parent Compound File (PCF). The second part supplies replacement pages for the PCF. These pages contain previously published updates or changes to entries. Each change is flagged by a large black dot. Figure 7 illustrates an entry in which both the name and the structure diagram were changed.

The third part of the supplement, which consists of an Index of Parent Compounds Cumulative Supplement, contains updates for all six indexes for both new and changed entries. The supplement indexes are cumulative from the date of issue of the full index set. A new, updated IPC is issued every two years. Thus, to locate the identifier of a parent compound, the user must look in only two places: the full IPC and the latest supplement. By means of these supplements, the PCH remains current and meaningful to the user as an up-to-date reference tool. In addition, each supplement also serves as a current-awareness service for new rings and other new parent compounds.

PUBLICATION SYSTEM

The PCH is entirely computer-produced. Its production cycle includes building the information base, selecting the parent compounds from that information base, editing the selected data, generating the index entries, and formatting and packaging the various PCH sections, indexes, and supplements.

The data for each parent compound, including the date of entry and the parent type, is keyed into and stored by the CAS Information Base. Online graphics terminals are used to enter the structure image for each parent compound into the system.[5] The date on which the last PCH supplement was published is retained by the PCH system, and this data is used in determining which parent compounds will be selected from the information base. Any parent compound that has an entry date or a change date later than the date of publication of the previous supplement, and that has been approved for publication, is selected from the information base and is entered into the PCH publication system.

Before being accepted by the system, however, the parent compound data must first pass a number of system edits which ensure that specific required items are present for each parent compound, and that these items are internally consistent. In the ring analysis for a cyclic parent, for example, the number of component rings, the size of each component ring, and the elemental analysis of each component ring must be consistent. If there are four component rings, then there must also be four components for the ring size and four components for the elemental analysis. Also, if the first entry in the ring size list indicates a component ring size of six, then there must be exactly six atoms in the formula for the first ring component of the elemental analysis.

Once it has passed the system edits, each parent compound is then assigned its unique alphabetic identifier by the PCH publication system. The PCH system maintains a master file, in ascending alphabetic identifier order, of all parent compounds that have been published. When a previously published parent compound is selected from the information base for inclusion in the current supplement, it is compared against its corresponding entry on the master file. In this manner the data items that have changed for the selected parent compound are identified, replaced on the master file, and highlighted by a black dot in the supplement.

Since two complete pages of a PCF section must be reprinted when a change to a parent compound on one of the pages is identified, the system must maintain a file in page number order. The PCH data sections are arranged in ascending alphabetic identifier order so that the system can list the range of alphabetic identifiers printed on each page. By combining this file with the master file, the system is able to select from the master file all of the parent compounds that appear on a given page of the PCF. This capability is utilized in producing each supplement.

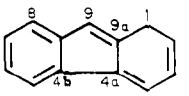

Figure 8. Simple ring substance.

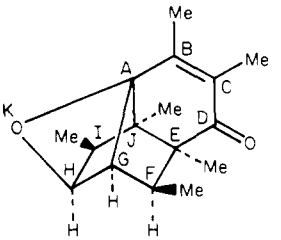

Figure 9. Complex ring substance.

The PCH publication system then formats and photocomposes each page (text data and structure diagrams) of the data sections and indexes. To facilitate the addition of new and replacement pages, the PCF is contained in screw-post binders.

USING THE PARENT COMPOUND HANDBOOK

The Parent Compound File and the individual indexes contained in the Index of Parent Compounds each include explanatory introductions to facilitate use of the "Parent Compound Handbook". In addition to the material in these introductions, the following examples illustrate some of the ways that the PCH can be used by the chemical community.

The user of the PCH can search for acyclic natural products, boron compounds, or metallocenes simply by scanning the Acyclic Stereo Parents or Cage Parents sections of the PCF. Both sections are small enough (4 and 8 pp, respectively) to allow direct manual searching with little expenditure of time.

In the case of compounds containing rings, the PCF sections are much larger and therefore the user must generally refer to the indexes. To use the indexes the user must first perform a ring analysis for the compound. Ring analysis is a method of describing a cyclic system at three levels: the number of rings in the system, the size of those component rings (together giving the smallest set of smallest rings), and the elemental analysis of each of the individual component rings in the ring system. The use of ring analysis in locating a PCH entry is illustrated first with the simple chemical structure shown in Figure 8, and then with the complex structure shown in Figure 9.

To determine the CA name for the ring system shown in Figure 8, the user first determines that the compound is a three-ring, all-carbon system containing one five-membered and two six-membered rings. Its entry in the Ring Analysis Index is therefore found at 3-RING SYSTEMS; 5,6,6; C_5–C_6–C_6. There are five fused-ring systems at that analysis in the index; by checking each entry, the user determines that the ring is listed under the identifier FWZLE, and that it is named 1*H*-Fluorene.

The existence of bridges in a ring system makes the ring analysis a little more difficult. In a search of this type, the user may want to identify the name, the ring-numbering scheme, and the WLN of the substance shown, and may then want to determine whether it is a parent ring of any stereoparents or ring-modified stereoparents.[7]

To perform a ring analysis on a substance of this type, the user must ignore all the substituents and consider only the ring system itself.[8] First the user must determine the number of component rings in the system. This number is equal to the minimum number of bond scissions required to convert the

KTGCN 58004-30-1
2,3a-Epoxy-3,7-methano-3a*H*-indene
$C_{10}H_6O$
T565 C4/EK D 3AEE K DOX JHJ
4-Ring System
4,5,5,6
C_3O-C_4O-C_5-C_6
CA 83:192184w

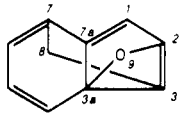

Figure 10. PCF entry for compound in Figure 9.

ring system to a noncyclic entity.[9] In Figure 9, scissions of bonds A–B, A–K, E–J, and G–H are sufficient. Thus, the number of component rings contained in the example is four.

The user must then identify the smallest set (i.e., four) of smallest component rings that include all of the atoms in the ring system. Several sets of four rings are apparent. The AGHK ring has four atoms and is the smallest ring in the system, so any smallest set must include it. The bicycloheptane bridge system (EFGHIJA) can be described as containing either two five-membered rings or a five- and six-membered ring. Because five–five is a smaller set, it is preferred to five–six. The remaining ring (ABCDEJ) has six atoms. The entire ring system is therefore listed under "4,5,5,6" in the Ring Analysis Index.

The next step is to identify the elements in each of the four component rings. The elemental analysis of the four-membered ring is C_3O; i.e., it contains three carbons and one oxygen. Component rings containing heteroatoms are preferred to all-carbon rings of the same size, so the ring component AKHIJ (C_4O) is preferred to AGHIJ (C_5). The second component is therefore C_4O, and the two remaining ring components are C_5 and C_6. Thus the total ring analysis is 4-RING SYSTEMS; 4,5,5,6; C_3O–C_4O–C_5–C_6.

The Ring Analysis Index and its latest Cumulative Supplement show nine entries at this ring analysis. Four of these are spiro ring systems (which our example is not) and they are therefore eliminated. The remaining five entries are

1*H*-Cyclopent[*f*]oxeto[2,3-*g*]benzofuran [HLKBX]
2,3a-Epoxy-3,7-methano-3a*H*-indene [KTGCN]
1,3,5-Methenoisobenzofuran [HDJDM]
1*H*-Oxeto[3′,2′:1,5]cyclopenta[1,2-*b*]benzofuran [GJKNQ]
2*H*-Oxeto[2′,3′:2,3]cyclopenta[1,2-*b*]benzofuran [GJKPZ]

A user unfamiliar with chemical nomenclature might well need to examine each of these entries in the Ring Parents section to match the structure in Figure 9. The second name, KTGCN, is the desired structure. The absence of any stereoparent names indented under the ring parent name in the index indicates that the ring is not the basis for any stereoparents.

The PCF entry at KTGCN is shown in Figure 10. The structure diagram confirms that this is indeed the parent compound being sought. By using the CA index name, the structure locants, and the WLN for the ring parent, the user can derive the name and WLN for the chemical substance in Figure 9: (1α,2α,3β,3aα,7β,7aβ,8*R**)-2,3,7,7a-Tetrahydro-1,4,5,7,7a,8-hexamethyl-2,3a-epoxy-3,7-methano-3a*H*-inden-6(1*H*)-one and T565 C4/EK D 3AEE K DOX HV FUTJ A1 B1 F1 G1 I1 J1, respectively. This information provides a starting point for searching the literature for this compound and others related to it. The CA reference provided with the PCH entry denotes the first appearance in the CAS Information Base of any derivatives of this ring parent.

A user interested in this compound might well be interested in related compounds. An examination of the WLN index at the entry for the parent compound shows no other parent compounds with the same basic ring skeleton. What, then,

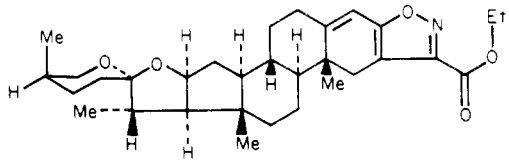

Figure 11. A natural product.

KZRML 60527-35-7
Spiro[8*H*-furo[2″,3″:4′,5′]cyclopenta[1′,2′:7,8]phenanthro[3,2-*d*]isoxazole-8,2′-[2*H*]pyran]
$C_{24}H_{13}NO_3$
T H5 F6 E6 B655 JNO UOXJ V-& BT6= OXJ
7-Ring System
5,5,5,6,6,6,6
C_3NO-C_4O-C_5-C_5O-C_6-C_6-C_6
CA 85:78257p

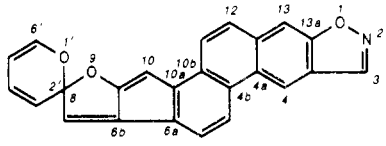

Figure 12. PCF entry KZRML.

about similar compounds with the epoxide ring opened? If the CA index name of the ring parent is used with the epoxide portion omitted, the name is 3,7-methano-3a*H*-indene. The Parent Name Index has no entries at that name, but there are 11 entries for various other methanoindenes. Examination of each of these entries in the Ring Parents section of the PCF reveals that GCWLN (1,4-Methano-1*H*-indene) is a related parent compound. A check of the Ring Analysis Index shows that no stereoparents have this ring as a basic structural component. The WLN index shows nine related heterocyclic ring parents with the same molecular skeleton as 1,4-methano-1*H*-indene. Using the original epoxy compound and these ten related methano compounds, the user can now search the CA indexes for other substances related to the original substance.

If the user wanted to determine the CA name and ring-numbering scheme for the natural product in Figure 11, several approaches could be used to identify this substance and to determine whether it is named by CAS as a stereoparent or ring system. The ring parent can be identified by a different method from that used in the first two examples. Removal of all the acylic substituents identifies the ring skeleton as a spiro system. Ignoring the hydrogens, the formula for the skeleton is $C_{24}NO_3$. The Parent Formula Index and its most recent supplement include only one entry which is at C_{24}-$H_{13}NO_3$ in the Supplement and no entry at all in the main index. The identifier listed at that formula is KZRML. Although the parent name is given, the Parent Compound File is checked under the identifier KZRML to see if this is the parent compound in question.

The PCF entry at KZRML, shown in Figure 12, confirms that this is the ring parent sought. The user then consults the Ring Analysis Index, using the ring analysis shown in Figure 12, to identify stereoparents based on this ring system. At this ring analysis there is but one entry, the ring parent name of the example, but indented under the name is a stereoparent name with its identifier, CGNMP. The user can find that stereoparent, shown in Figure 13, in the Cyclic Stereo Parent section of the PCF. By comparing the substance shown in Figure 11 with the stereoparent shown in Figure 13, the user can determine that they are closely related. Naming the substance in Figure 11 requires the addition of the ethoxycarbonyl substituent at position 3′, the double bond between positions 4 and 5, and the stereochemistry of the methyl group at position 25. The name is therefore ethyl (25*R*)-spirosta-2,4-dieno[2,3-*d*]isoxazole-3′-carboxylate; this ester would

CGNMP 60439-60-3
Spirost-2-eno[2,3-d]isoxazole ([4aS-(4aα,4bβ,6aα,6bβ,7β,8β,9aβ,10aβ,10bα)]-3',4,4',4a,4b,5,5',6,6',6a,6b,7,9a,10,10a,10b,11,12,12a,13-eicosahydro-4a,5',6a,7-tetramethylspiro[8H-furo[2'',3'':4',5']cyclopenta[1',2':7,8]phenanthro[3,2-d]isoxazole-8,2'-[2H]pyran])
$C_{28}H_{41}NO_3$
T H5 F6 E6 B655 JNO UOX&TTTTTJ
B1 F1 W1 V-& BT6OXTJ E1
Text descriptor 4:.SPIROST..7
7 Ring System
5,5,5,6,6,6,6
C_3NO–C_4O–C_5–C_5O–C_6–C_6–C_6
CA 85:78257p

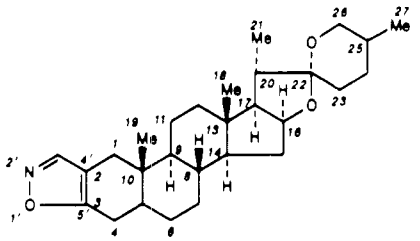

Figure 13. Stereo Parent entry CGNMP.

appear in CA indexes under the name of the acid.

SUMMARY

The PCH has been in use now for more than three years. It contains more than 50 000 entries and serves as a valuable reference tool for those engaged in searching the chemical literature. It is the only complete listing of organic ring systems available. With its periodic updates, it continues to be a complete reference tool. It is especially useful as a current-awareness service for identifying new ring systems, boron cages, metallocenes, and stereoparents.

REFERENCES AND NOTES

(1) A. M. Patterson, L. T. Capell, and D. F. Walker, "The Ring Index", 2nd ed., American Chemical Society, Washington, D.C., 1960, and its three supplements published in 1963, 1964, and 1965.
(2) E. G. Smith and P. A. Baker, "The Wiswesser Line-Formula Chemical Notation (WLN)", CIMI, Cherry Hill, N.J., 1976.
(3) T. Ebe and A. Zamora, "Wiswesser Line Notation Processing at Chemical Abstracts Service", *J. Chem. Inf. Comput. Sci.*, **16**, 33–35 (1976).
(4) The CA Index Guide, American Chemical Society, Columbus, OH, 1977, pp 761–771.
(5) J. E. Blake, N. A. Farmer, and R. C. Haines, "An Interactive Computer Graphics System for Processing Chemical Structure Diagrams", *J. Chem. Inf. Comput. Sci.*, **17**, 223–228 (1977).
(6) Although not published until March 1977, the data for the initial PCH was extracted from the information base in Aug 1976. This date is carried at the head of each page. Similarly, each supplement carries the date its data was extracted.
(7) See the Introduction to the Ring Analysis Index.
(8) Cyclic substituents are considered a separate ring system and would be treated separately.
(9) Alternatively, the number of rings may be calculated from the formula: no. of rings = no. of edges − no. of atoms + 1 (where the number of edges is the number of distinct atom-to-atom connections). For the ring system of interest, the formula is: 14 edges − 11 atoms + 1 = 4 rings.

Graph-Based Chemical Nomenclature. 1. Historical Background and Discussion

ALAN L. GOODSON
Chemical Abstracts Service, P.O. Box 3012, Columbus, Ohio 43210

Received March 31, 1980

Events leading to development of chemical nomenclature based on graph theory are described. Chemical nomenclature is defined and distinguished from chemical line notations and connection tables, with which it is sometimes confused. Development of chemical nomenclature from the era of ancient civilizations to the present time is reviewed and the recognized limitations of current practices are described. The case for consideration of a new approach to chemical nomenclature is presented, and the requirements for such a nomenclature to be acceptable to the chemical community are discussed. The development of graph theory and its application to chemistry are reviewed; its advantages as the basis of a single, comprehensive chemical nomenclature are presented.

1. INTRODUCTION

Nomenclature is defined[1a] as: "a system or set of names or designations used in a particular science, discipline, or art and formally adopted or sanctioned by the usage of its practitioners." Chemical nomenclature is more specifically defined as: "a set of chemical names that may be systematic...or not and that aims to tell the composition and often the structure of a given compound by naming the elements, groups, radicals, or ions present and employing suffixes denoting function..., prefixes denoting composition..., configuration prefixes..., operational prefixes..., arabic numbers or Greek letters for indicating structure (as positions of substituents), or Roman numerals for indicating oxidation state." Chemical nomenclature can be illustrated with reference to Figure 1. The substance depicted is named (systematically) 1,4-benzenediol in the *Chemical Abstracts* (CA) Ninth Collective Chemical Substance Index but is more commonly known by its trivial (i.e., nonsystematic) name: hydroquinone.

Nomenclature should not be confused with chemical line notations or connection tables, as sometimes occurs. A notation is defined[1b] as: "a system of characters, symbols, or abbreviated expressions used in...science to express technical facts, quantities or other data". To illustrate a notation, the Wiswesser Line Notation (WLN)[2] for 1,4-benzenediol (hydroquinone) is QR DR. A connection table[3] is "a uniquely ordered list of the node symbols of the structure (or graph) in which the value (atomic symbol) of each node and its attachment (bonding) to the other nodes of the total structure are described." The Chemical Abstracts Service (CAS) Registry III connection table for 1,4-benzenediol is illustrated in Figure 2. Thus, nomenclature, notations, and connection tables are quite distinct from each other and need not be confused.

Nomenclature, notations, and connection tables are of value in different ways. Connection tables are useful as detailed, atom-by-atom computer records of chemical structures, as in the CAS Registry File. Chemical line notations, such as WLN's, are convenient tools for manual or machine registration and substructure searching of files of moderate size (50 000–100 000 structures), although they are used for some

0095-2338/80/1620-0167$01.00/0