are zero: eq 26 reduces to about one-third of its original length.

II. COMPUTER SIMULATION OF STEADY-STATE KINETICS

Three lessons in Computer-Aided-Instruction are offered to the students. The first lesson introduces the concept of basic steady-state kinetics and shows the student the steady-state treatment of Michaelis and Menten as well as the various forms of the equations. A typical output from this lesson is given in Figure 3. The second "lesson" concentrates on inhibition of steady-state enzyme kinetics and attempts to bring the basic concepts to the full attention of the student. Various plots are produced, allowing student's inference. One such plot is presented in Figure 4.

The use of the bicyclic scheme allows not only the use of a single cycle, but provides also rather extensive freedom in the choice of values for the constants and analytical concentrations. The user may also choose various values for the rate constants, which generate the product. Two types of curves are available to the user, presented in Figures 5 and 6.

ACKNOWLEDGMENT

The technical assistance of Byron Long and Julian Aguda is gratefully acknowledged.

REFERENCES AND NOTES

(1) L. Michaelis and M. Menten, *Biochem. Z.*, **49**, 333 (1913).
(2) H. E. Briggs and J. B. S. Haldane, *Biochem. J.*, **19**, 338 (1925).
(3) V. Henri, *Arch. Fisiol.*, **1**, 299 (1904).
(4) J. M. Reiner, "Behavior of Enzyme Systems", 2nd ed, Van Nostrand-Reinhold, New York, N.Y., 1969, 345 pp.
(5) G. H. Czerlinski, "Chemical Relaxation", Marcel Dekker New York, N.Y., 1966, p 33. Preequilibrium in the above case means FAST in relation to the time range of the steady state, which is not yet "turned on" at the time range, the above equations apply to.
(6) H. Lineweaver and D. Burk, *J. Am. Chem. Soc.*, **56**, 658 (1934).

Wiswesser Line Notation Processing at Chemical Abstracts Service†

TOMMY EBE* and ANTONIO ZAMORA

Chemical Abstracts Service, The Ohio State University, Columbus, Ohio 43210

Received October 16, 1975

A WLN data base has been created for 40,000 ring systems, representing 18,000 unique ring graphs identified by the CAS Chemical Registry System. The manual encoding effort is supported by a canonical WLN generation program for complex ring graphs. WLNs are automatically machine edited through syntax checks and by comparison of molecular formulas derived from the WLN with those in Registry Files.

INTRODUCTION

In 1968, Chemical Abstracts Service (CAS) began associating Wiswesser Line Notations[1] (WLNs) with ring systems by publishing a list[2] of WLNs for structures in "The Ring Index".[3] These notations were obtained from files made available by Elbert G. Smith, Professor of Chemistry, Mills College, and from tapes provided by the Dow Chemical Company. Since this initial effort, CAS has continued to generate WLNs for new ring systems being added to the ring file. The ring file has grown since 1968 to include almost 40,000 ring systems, representing 18,000 unique ring graphs identified by the CAS Chemical Registry System.[4] In order to generate WLNs for these ring systems efficiently and accurately, the manual encoding and editing effort is supported by the current version of the Registry System, Registry III, and by two computer programs. One program generates canonical WLNs for complex polycyclic structures, and an editing program automatically checks all WLNs added to the data base. As illustrated by Figure 1, this paper discusses the manual encoding procedure and the relevant computer support for processing WLNs at CAS.

IDENTIFICATION OF RING SYSTEMS

Registry III identifies ring systems within chemical substances to assist in the generation of substance names and to support structure display programs. A unique ring system identifier (RID), which is contained within the structure record for a substance, is assigned to each ring system. The RID is actually a composite identifier, as illustrated by Figure 2. Each unique ring graph (cyclic skeleton) is assigned a basic identifier, to which is appended an identifier for the ring atom variation, and also an identifier for the bond variation. The RID, therefore, uniquely identifies a ring system and can be used to associate related ring systems.

During substance registration, whenever Registry III identifies a ring system with a new ring graph, it assigns a new RID, adds the ring system to the file, and generates a New Ring Graph Alert message. If the ring system has a ring graph that is already in the file, but has a new heteroatom variation, then an RID is constructed using the previously assigned ring graph identifier with a new heteroatom identifier, the ring system is added to the file, and a New Ring System Alert message is generated. These new ring systems, identified by either message, are encoded into WLN by one of the following procedures.

RING SYSTEMS WITH COMPLEX RING GRAPHS

The canonical notation for a ring system is generated by describing the structural features of the ring graph according to the hierarchical requirements of WLN Rule 30.[5] The function of Rule 30 is to determine the preferred locant path that unambiguously describes the ring graph. The most difficult and time-consuming task in generating a WLN for a complex polycyclic structure is finding all the possible locant paths in order to choose the preferred path. Since this task is ideally suited for a computer, and since the number of complex ring graphs in the CAS ring file is large, it was necessary to develop computer support for this manual encoding effort.

Based on a program[6] written at the Dow Chemical

† Presented to the Division of Chemical Information, 170th National Meeting of the American Chemical Society, Chicago, Ill., Aug 25, 1975.

* To whom correspondence should be addressed.

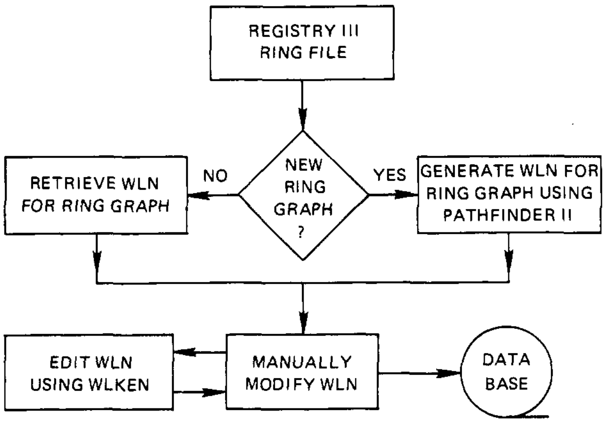
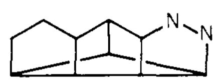
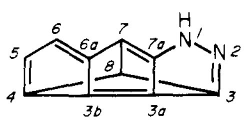


Figure 1. WLN processing scheme at CAS.

159868



159868.26

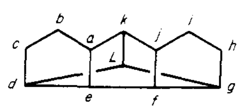


159868.26.14

3,4,7-Metheno-1*H*-pentaleno[2,1-*c*]pyrazole

[CAS Registry Number 54532-20-6]

Figure 2. Ring identifier (RID): graph, heteroatom, and bond variation.



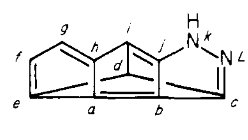
ae fj ak dl gl

(nonconsecutive locant pairs)

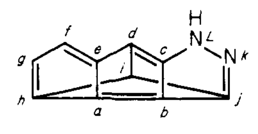
Figure 3. PATHFINDER II input data.

NCLPS:	ae ah di bj cl
WLN:	L5555 B5/DI/CL 5AABEH LTJ
MESSAGES:	Preferred path starts from input locant e.
	Input locant e starts 2 equivalent preferred paths.
	Input locant e is equivalent to input locant f.
TIME:	0.44 seconds

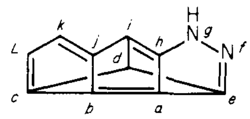
Figure 4. Partial PATHFINDER II ring graph analysis report.



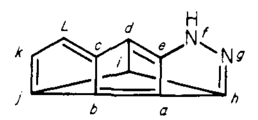
A



B



C



D

	A:	T5555 B5/DI/CL 5AABEH L KMN DHJ
	B:	T5555 B5/DI/CL 5AABEH L KNM IHJ
	C:	T5555 B5/DI/CL 5AABEH L FNM DHJ
Preferred	D:	T5555 B5/DI/CL 5AABEH L FMN IHJ

Figure 5. Equivalent paths found by PATHFINDER II.

Company, which subsequently became commonly known as PATHFINDER, PATHFINDER II was developed to meet the need for computer assistance for WLN generation. More fully described elsewhere,[7] PATHFINDER II reliably generates canonical WLNs for the most complex polycyclic structures at much less overall cost than manual generation.

A typical example of a ring system with a new ring graph is shown in Figure 3. The encoder assigns an arbitrary locant path to the structure, describing it as a set of nonconsecutive locant pairs (NCLPs) in the order in which the rings are closed by the locant path. The NCLPs are keypunched and input to PATHFINDER II.

A partial description of the program output, the PATHFINDER II Ring Graph Analysis Report, is illustrated in Figure 4. As shown by the second and third messages, PATHFINDER II also determined for this ring graph that input locants *e* and *f* were equivalent, and that each one started two equivalent preferred paths.

This report enables the encoder to transfer the preferred path locant data to the ring system in Figure 5, and to modify the canonical notation to describe the heteroatoms present and the degree of saturation for the ring system. The equivalent path messages alert the encoder that the preferred path may be completed in four different ways, as shown by structures **A**, **B**, **C**, and **D**. The encoder can readily determine by Rules 30h and 30i that structure **D** produces the correct notation.

After the notation is completed for the ring system, the ring graph analysis report is added to the PATHFINDER II report file. This hardcopy file, accessible by RID, is a permanent record of the program output for every ring graph processed, and it serves to prevent the processing of any ring graph more than once.

When a WLN must be generated for a new ring system with a previously identified ring graph, i.e., the ring graph is already in the file but the heteroatom combination is new, the encoder first checks the PATHFINDER II report file using the RID. The canonical notation is then readily modified to describe the new ring system as outlined above.

RID-WLN LISTING

The first portion of the WLN for a ring system is a description of the ring graph in terms of an unambiguous locant path unique to that ring graph. The first part of the RID, i.e., the graph identifier, is also unique to a particular ring graph. Therefore, the computer listing of all the WLNs in the ring file, printed in ascending order of RID, is a very useful encoding tool, since it groups together all the WLNs for ring systems having the same ring graph.

Whenever a WLN must be generated for a new ring system that is not particularly complex, it is very convenient for the encoder to access this listing by RID. After finding a WLN for a different heteroatom variation for the same ring graph, i.e., same graph identifier, the encoder then easily modifies the notation to describe the heteroatoms present in the new ring system.

When the RID-WLN listing was first created in 1974, one immediate and expected benefit was the easy visual identification of different notations for ring systems with the same ring graph. More than 500 such errors were found in the 40,000 ring notations at that time. PATHFINDER II was used to generate the correct notations for those ring graphs.

EDITING WLNS

Once the complete WLN for a ring system is generated, it is added to the machine file without any manual review. The notation is, however, automatically checked by a WLN editing program[8] (WLKEN). This program first decodes the WLN and generates a diagnostic error message if it finds a notation inconsistency. If none is found, the program calculates a molecular formula, compares it with one obtained by Registry III, and generates an error message if the formulas are not identical.

If the formulas do match, then WLKEN tries to find a better notation following WLN Rule 30. Although it does this only for ortho-fused ring systems, i.e., ring systems whose notations do not cite bridges, multicyclic points, or branch locants, WLKEN does go beyond the ring graph level (Rule 30a to 30g) by checking the order of locants for cited ring segment symbols (Rule 30h), the order of cited ring segment symbols (Rule 30i), etc.

If a better notation is found, WLKEN prints the input locants and the corresponding locants for the better notation. The encoder can then write the complete correct notation. This editing capability gives added confidence to the RID-WLN listing for manually encoding large ortho-fused ring systems. There is no overlap in the functions performed by WLKEN and PATHFINDER II since the latter program is not used for ortho-fused ring systems.

Whenever an error is discovered by any means, the correct notation is keyboarded and replaces the notation in the data base. These WLN corrections are also automatically edited by WLKEN. As used at CAS, therefore, WLKEN keeps a close check on encoding, transcribing, and keyboarding errors. In addition, since the data base was constructed from WLNs that were obtained from several sources and were not previously machine edited at CAS, the entire WLN file was edited when WLKEN was installed. Several thousand relatively simple errors were thereby discovered and corrected.

CONCLUSION

The CAS involvement with WLNs stems from a desire to provide a bridge between the CAS ring system data base and the WLN files of other organizations. This CAS data base will be published in "The Parent Compound Handbook" (PCH), which will replace "The Ring Index". Accordingly, a great effort was made to generate the data accurately and efficiently.

PATHFINDER II was developed to encode the many hundreds of complex ring graphs, with much more reliability and at much less overall cost than would be possible by unassisted manual effort. Once all the complex graphs have been processed, use of PATHFINDER II will be sharply reduced and limited only to new ring graphs. The RID-WLN listing, resulting from the ring graph identification feature of Registry III, would then be the basic manual WLN encoding tool.

WLKEN was installed as part of a computer system designed to replace almost all of the time-consuming manual review and proofing of WLNs. The system automatically monitors all input data and reduces manual effort by requiring review only of error messages. Since its installation late in 1974, this system has both reduced the overall cost of editing notations, and also given the WLN data base a high degree of accuracy. Subscribers to the PCH should therefore be able to confidently use these WLNs in their own encoding efforts.

REFERENCES AND NOTES

(1) E. G. Smith, "The Wiswesser Line-Formula Chemical Notation", McGraw-Hill, New York, N.Y., 1968.

(2) "Wiswesser Line Notations Corresponding to Ring Index Structures", PB 180901, American Chemical Society, Washington, D.C., 1968.

(3) A. M. Patterson, L. T. Capell, and D. F. Walker, "The Ring Index", 2nd ed, American Chemical Society, Washington, D.C., 1960; Suppl. I, 1963; Suppl. II, 1964; Suppl. III, 1965.

(4) P. G. Dittmar, R. E. Stobaugh, and C. E. Watson, "The Chemical Abstracts Service Chemical Registry System. I. General Design", presented to the Division of Chemical Literature, 169th National Meeting of the American Chemical Society, Philadelphia, Pa., April 7, 1975.

(5) Reference 1, p 168.

(6) C. M. Bowman, F. A. Landee, N. W. Lee, and M. H. Reslock, "A Chemically Oriented Information Storage and Retrieval System. II. Computer Generation of the Wiswesser Notations of Complex Polycyclic Structures", *J. Chem. Doc.*, **8**, 133–8 (1968).

(7) A. Zamora, and T. Ebe, "PATHFINDER II. A Computer Program that Generates Wiswesser Line Notations for Complex Polycyclic Structures", *J. Chem. Inf. Comput. Sci.*, following paper in this issue.

(8) A. Leo, D. Elkins, and C. Hansch, "Computerized Management of Structure-Activity Data. III. Computerized Decoding and Manipulation of Ring Structures Coded in WLN", *J. Chem. Doc.*, **14**, 65–69 (1974).