(7) N. Carpenter, "Syntax-Directed Translation of Organic Chemical Formulas into Their Two-Dimensional Representation", *Comput. Chem.*, **1**, 25-28 (1976).
(8) P. L. Fehder and M. P. Barnett, "Syntactic Scanning of Chemical Information", *J. Chem. Doc.*, **5**, 8-13 (1965).
(9) P. G. Barker, "Syntactic Definition and Parsing of Molecular Formulas", *Comput. J.*, **18**, 355-359 (1975); **21**, 224-233 (1978).
(10) K. Rankin and S. J. Tauber, "Linguistics as a Basis for Analyzing Chemical Structure Diagrams", *J. Chem. Doc.*, **11**, 139-141 (1971).
(11) S. J. Tauber and K. Rankin, "Valid Structure Diagrams and Chemical Gibberish", *J. Chem. Doc.*, **12**, 30-34 (1972).
(12) W. E. Underwood and L. N. Kanal, "Structural Descriptions, Transformation Rules, and Pattern Analysis", Proceedings of the First International Joint Conference on Pattern Recognition, Washington DC, 1973, pp 434-444.
(13) J. L. Pfaltz and A. Rosenfeld, "Web Grammars", Proceedings of the International Joint Conference on Artificial Intelligence, Washington, DC, 1969, pp 609-619.
(14) C. M. Bowman, L. C. Davison, and P. F. Roush, "On-line Storage and Retrieval of Chemical Information. I. Structure Entry", *J. Chem. Inf. Comput. Sci.*, **19**, 228-230 (1979).
(15) W. T. Wipke, G. I. Ouchi, and S. Krishnan, "Simulation and Evaluation of Chemical Synthesis—SECS. An Application of Artificial Intelligence Techniques", *Artif. Intelligence*, **11**, 173-193 (1978).
(16) H. L. Gelernter, A. F. Sanders, D. L. Larsen, K. K. Agarwal, R. H. Boivie, G. A. Spritzer, and J. E. Searleman, "Empirical Explorations of SYNCHEM", *Science (Washington, DC)*, **197**, 1041-1049 (1977).
(17) H. W. Whitlock, "An Organic Chemist's View of Formal Languages", *ACS Symp. Ser., No. 61*, 1977.
(18) H. W. Whitlock, "A Heuristic Solution to the Functional Group Switching Problem in Organic Synthesis", *J. Am. Chem. Soc.*, **98**, 3225-3233 (1976).
(19) P. E. Blower and H. W. Whitlock, "An Application of Artificial Intelligence to Organic Synthesis", *J. Am. Chem. Soc.*, **98**, 1499-1510 (1976).
(20) E. V. Krishnamurthy and M. F. Lynch, "Coding and Analysis of Generic Chemical Formulae in Chemical Patents", *J. Inf. Sci.*, in press.
(21) M. F. Lynch, J. M. Barnard, and S. M. Welford, "Computer Storage and Retrieval of Generic Chemical Structures in Patents. 1. Introduction and General Strategy", *J. Chem. Inf. Comput. Sci.*, paper 1 in this series.
(22) J. M. Barnard, M. F. Lynch, and S. M. Welford, "Computer Storage and Retrieval of Generic Chemical Structures in Patents. 2. GENSAL, a Formal Language for the Description of Generic Chemical Structures", *J. Chem. Inf. Comput. Sci.*, preceding paper in this issue.
(23) P. Naur, Ed., "Report on the Algorithmic Language ALGOL 60", *Commun. ACM*, **3**, 299-314 (1960).
(24) A. V. Aho and J. D. Ullman, "The Theory of Parsing, Translating and Compiling", Prentice-Hall, Englewood Cliffs, NJ, 1972.
(25) S. A. Greibach, "Inverses of Phrase Structure Generators", Ph.D. Thesis, Harvard University, Cambridge, MA, 1963.
(26) J. E. Ash, "Connection Tables and Their Role in a System" in "Chemical Information Systems", J. E. Ash and E. Hyde, Eds., Ellis Horwood, London, 1975.
(27) H. Hiz, "A Linearization of Chemical Graphs", *J. Chem. Doc.*, **4**, 173-180 (1964).
(28) S. H. Eisman, "A Polish-type Notation for Chemical Structures", *J. Chem. Doc.*, **4**, 186-190 (1964).
(29) A corollary of this rule is that terminal atoms which are connected directly to a branching atom are specified as late as possible, consistent with the requirements of a Polish notation.
(30) H. I. Scoins, "Placing Trees in Lexicographic Order", *Mach. Intelligence*, **3**, 43-60 (1968).
(31) D. J. Cohen and C. C. Gotlieb, "A List Structure Form of Grammars for Syntactic Analysis", *Computing Surveys*, **2**, 65-82 (1970).
(32) J. Lederberg, G. L. Sutherland, B. G. Buchanan, E. A. Feigenbaum, A. V. Robertson, A. M. Duffield, and C. Djerassi, "Applications of Artificial Intelligence for Chemical Inference. I. The Number of Possible Organic Compounds. Acyclic Structures Containing C, H, O and N", *J. Am. Chem. Soc.*, **91**, 2973-2976 (1969), and subsequent publications.
(33) L. Masinter, N. S. Sridharan, J. Lederberg, and D. H. Smith, "Applications of Artificial Intelligence for Chemical Inference. XII. Exhaustive Generation of Cyclic and Acyclic Isomers", *J. Am. Chem. Soc.*, **96**, 7702-7714 (1974).
(34) Y. Kudo and S.-I. Sasaki, "Principle for Exhaustive Enumeration of Unique Structures Consistent with Structural Information", *J. Chem. Inf. Comput. Sci.*, **16**, 43-49 (1976).
(35) C. A. Shelley, T. R. Hays, M. E. Munk, and R. V. Roman, "An Approach to Automated Partial Structure Expansion", *Anal. Chim. Acta*, **103**, 121-132 (1978).
(36) J. Earley, "An Efficient Context-Free Parsing Algorithm", *Commun. ACM*, **13**, 94-102 (1970).
(37) The acyclic grammar is described as deterministic. This is a consequence, in part, of the fact that each terminal symbol appears in only a single replacement rule in the grammar. The appropriate replacement rule is therefore readily located.
(38) G. W. Adamson, J. Cowell, M. F. Lynch, A. H. W. McLure, W. G. Town, and A. M. Yapp, "Strategic Considerations in the Design of a Screening System for Substructure Searches of Chemical Structure Files", *J. Chem. Doc.*, **13**, 153-157 (1973).
(39) G. W. Adamson, J. A. Bush, A. H. W. McLure, and M. F. Lynch, "An Evaluation of a Substructure Search System Based on Bond-Centered Fragments", *J. Chem. Doc.*, **14**, 44-48 (1974).

Specification and Unconstrained Enumeration of Conformations of Chemical Structures for Computer-Assisted Structure Elucidation

JAMES G. NOURSE

Department of Chemistry, Stanford University, Stanford, California 94305

Received January 22, 1981

A symmetry group called the conformation symmetry group (CFSG) which provides a method of uniquely specifying molecular conformation based on an appropriate discrete bond property (such as rotameric state) is formulated. The method is applicable to entire chemical structures as well as substructures. The CFSG can be used to build a simple "acyclic" conformation generator and leads to a solution for a heretofore unsolved problem in conformation enumeration.

For a number of purposes, including computer representation, the specification of an organic chemical structure starting from its molecular formula can be done in several stages. In the CONGEN program (CONstrained GENeration of isomers) for computer-assisted structure elucidation, we have chosen to do this chemical structure specification in three stages. Starting from the molecular formula, the first stage is to specify the constitution or atom-to-atom connectivity of the structure. The CONGEN program is capable of generating all the possible constitutional isomers for any molecular formula.[1] Furthermore this generation can be constrained to just those possibilities consistent with partial substuctural information.[2]

The second stage is to specify the configuration of stereocenters (chiral centers and double bonds) in the structure to give the configurational stereoisomer. The CONGEN program is capable of generating all the possible configurational stereoisomers for any molecular formula[3] and this generation can also be constrained to just those possibilities consistent with partial stereochemical information.[4] The third stage is to specify the conformation (i.e., position of torsional rotation about single bonds) of the structure, and the approach to this

0095-2338/81/1621-0168$01.25/0

problem is the subject of this paper. This sequence can be extended to include specification of refined atomic coordinates, energetics, etc.

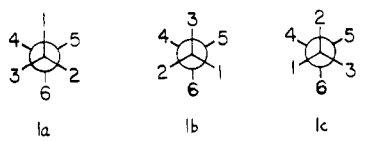
For the present purposes the problem of conformation specification starts with a structure (which can also be a substructure) of definite constitution (atom-to-atom connectivity) and configuration (or chiral centers and double bonds). This structure is represented by a graph augmented with a twofold property on some nodes (stereocenters) which includes the configurations on those centers.[3,5] This is not an explicit geometric representation although geometry is implied by the standard bond lengths and atomic valence angles. The problem is to add to this representation information about the dihedral angles around rotatable single bonds. A byproduct of our specification of constitution and configuration is the symmetry group of the structure,[3,5] and it is assumed that this is available for conformation specification.

The approach taken to conformation specification is in analogy to that taken previously for configuration specification.[5] This approach requires the use of two key properties, one local and one global to the chemical structure. The local property is the torsional position about the single bond, and this will be discussed in the first section. The global property is the symmetry group of the structure properly represented, and this will be discussed in the second section. This symmetry group is used to solve three problems central to the use of conformations in computer-assisted structure elucidation. The first is the unconstrained, nonredundant generation of possible conformations for a structure of given constitution and configuration. This is called an "acyclic" generation and is discussed in section III. The second is the specification and canonicalization of the structure's conformation represented this way. This result can be applied to both complete structures and substructures and is described in section IV. The third is the enumeration of conformations (i.e., independent of generation) and represents a heretofore unsolved problem for chemical structures of arbitrary branching and symmetry. This result is discussed in section V.

I. REPRESENTATION OF SINGLE-BOND TORSIONAL POSITIONS

Unlike the specification of constitution and configuration which can be done with discrete variables, the specification of conformation involves a continuous variable—the torsional angle around a single bond. For present purposes this must be "converted" into a discrete variable. Methods for doing this conversion are discussed below.

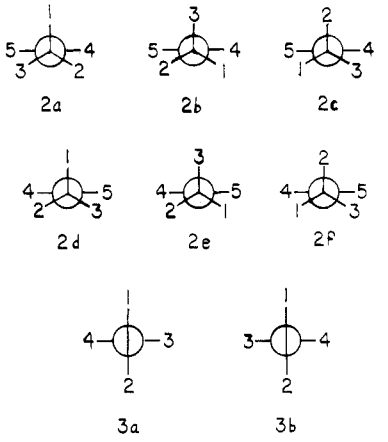
In graph-theoretic terms the present problem of conformation specification requires the assignment of a discrete property to some of the edges (bonds) of the graph (chemical structure). This graph will have all the nodes (atoms) numbered uniquely and will have configurations assigned to all stereocenters based on this numbering.[3,5] This bond property should be one which can be assigned independently to the various bonds. Any inconsistencies in this independent assignment are constraints which will not be discussed in the present paper. Furthermore, the edge property should be based on the atom numbering and stereocenter configurations already present in the chemical structure representation. An example of such a discrete bond property for an ethane molecule would be the choice of the three possible staggered torsional positions for a numbered molecule, **1a–c**. Following a suggestion of



1a 1b 1c

Wipke and Dyott,[6] these can be described in terms of the atom numbers by choosing the two smallest numbered atoms connected to the two atoms of the central bond. The three conformations can then be named 1,4(−), 1,4(trans), and 1,4(+) (respectively, **1a–c**) in which the sign refers to the sign of the torsional angle [clockwise is (+)]. This is a three-valued property, and specification of a conformation would require the assignment of one of the three values to each bond in a structure in which the atoms are suitably substituted. Alternatively, one could define a six-valued variable (three staggered and three eclipsed conformations) or a finer grid of smaller torsional increments. These torsional positions could be interpreted as ranges of the torsional variable rather than just discrete positions. The point here is that some choice of a discrete variable for the various bonds must be made depending on the problem at hand. (Staggered rotamers may suffice for problems in which only local bond torsional energy potentials are important, while a finer grid may be desired for problems in which many local and global energy potentials are important.) The methods described below can be used for any such choice so long as they can be independently assigned per bond and can be represented by using atom numbers and stereocenter configurations.

Other kinds of bonds will require different variables. For example, the bond between an sp^3 and sp^2 atom has six different "staggered" positions, **2a–f**, while sp^2–sp^2 bonds have two staggered positions, **3a,b**. While these discrete variables (staggered, eclipsed, etc.) are the ones most commonly used, one cannot exclude other types of possibilities in all cases.[7] It is therefore important to have a method which can make use of any bond property satisfying the above conditions.



2a 2b 2c

2d 2e 2f

3a 3b

II. CONFORMATION SYMMETRY GROUP (CFSG)

Once the choice of the discrete variables (usually torsional positions or ranges) for the various bonds has been made, the next step is to make proper use of any symmetry in the structure. The resulting symmetry group representation called the conformation symmetry group (CFSG) is the key structure which is used in the generation, specification, and enumeration of conformations.

The formulation of the conformation symmetry group (CFSG) will be presented by means of an example. The CFSG for 2,4-dimethylpentane (**4**), as numbered in Figure 1, will be explicitly constructed. The three possible staggered rotamers for the two central bonds will be used as the discrete variable and are also shown in Figure 1. The symmetry group for 2,4-dimethylpentane (**4**) is given in Table I by atomic permutations, spatial operations, and central bond permutations (first three columns, respectively). The method of finding this symmetry group differentiates between operations which are spatial rotations and those which are spatial reflections, since the latter will invert the configurations of all stereocenters.[3,5] The CFSG is constructed by determining the effect

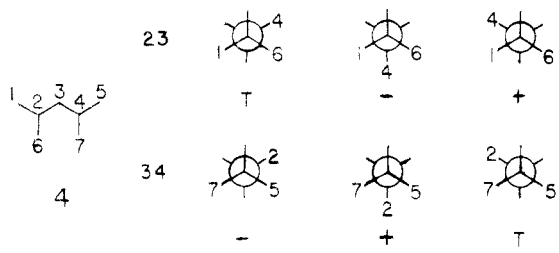


Figure 1. Three staggered rotameric positions about the two central bonds of 2,4-dimethylpentane (**4**). Top row shows the rotamers about the bond between atoms 2 and 3, the bottom row for 3 and 4. The rotamers are named with reference to the smallest numbered substituents on either side of the bond.

Table I. Symmetry Group for 2,4-Dimethylpentane (4)

atomic permutation	point group operation	central bond permutation	bond and rotamer permutation	bond and overall rotamer permutation
E	E	(23)(34)	(23)(34)	(23)(34)
(17)(24)(56)	C_2	(23 34)	(23(T+−)-34(T−+))	(23 34)
(15)(24)(67)	σ_v	(23 34)	(23(+−)-34(+−))	(23 34)
(16)(57)	σ_v	(23)(34)	(23(T+))-(34(T−))	(23(T+))-(34(T−))

of these symmetry operations on the possible rotamers for the two central bonds (shown in Figure 1).

The four symmetry operations in Table I will be discussed in sequence. The identity operation leaves all atoms, bonds, and rotamers fixed and does not illustrate any nontrivial properties of the CFSG. The C_2 operation exchanges the two central bonds but also exchanges the possible rotamers as follows,

23 ↔ 34 (23(T+−)34(T−+))
T ↔ −
− ↔ +
+ ↔ T

as can be verified by following the effect of atom permutations on the rotamers in Figure 1. A shorthand representation of this permutation is given above which indicates the bonds exchanged (23 and 34) and the fate of the three rotamers for each bond. The fate of the rotamers around the 23 bond is indicated after the 34 bond. The first σ_v operation exchanges the two central bonds and the rotamers as follows:

23 ↔ 34 (23(+−)34(+−))
T ↔ T
+ ↔ −
− ↔ +

Since this is a reflective operation, the dihedral angles all reverse sign. These results can be verified by determining the effect of the permutations on the rotamers in Figure 1. The final reflective operation fixes both bonds but permutes the rotamers as follows:

23 ↔ 23 34 ↔ 34 (23(T+))(34(T−))
T ↔ + T ↔ −
+ ↔ T − ↔ T
− ↔ − + ↔ +

The fifth column in Table I gives the CFSG in the compact notation which will be used in the applications of the CFSG in sections III and IV.

The CFSG is a subgroup of a particular type of permutation group called an exponentiation group.[8] With each cycle of a permutation of the bonds for which rotamers (torsional positions) are designated, there is associated some permutation of those rotamers. This permutation of rotamers is computed by multiplying the permutations of the rotamers within each cycle of the bond permutations as follows:

(23(T+−)34(T−+)) = (23 34 (T)(+)(−)) = (23 34)

In this case the identity permutation of rotamers was obtained. This has been done for all four permutations in the CFSG of the 2,4-dimethylpentane (**4**) example, and these are given in the fifth column of Table I. This multiplication is done to put the permutation into its most compact form. This representation of the permutations of the CFSG is most useful in enumeration and will be used in section V.

III. ACYCLIC CONFORMATION GENERATOR

The CFSG can be used in an algorithm which completely and nonredundantly generates the possible conformations for a chemical structure with given constitution and stereocenter configurations. The algorithm discussed here is termed "acyclic" since constraints which result from ring closure and excluded volume (steric interactions of overlapping chains) are not considered. This is a first step toward the construction of an algorithm and computer program which performs conformation generation for computer-assisted structure elucidation. An analogy can be found to the sequence of steps followed in the development of an algorithm and program for generation of all possible constitutional isomers.[1,9] The acyclic (structures with no double bonds or rings) generator was developed first. The novel result here for conformation generation is that the generation using such an algorithm will be nonredundant for structures of arbitrary symmetry.

The generation of conformations will be described by using the 2,4-dimethylpentane (**4**) example (Figure 1). A conformation of this structure is represented as an ordered pair of rotamers for the two central bonds. Thus, [−T] represents the conformation with the 23 bond in the − conformation and the 34 bond in the T conformation (Figure 1). Distinct conformations correspond to equivalence classes of these ordered pairs based on the CFSG. Using the entire CFSG will result in conformations which are mirror images of each other being collected into the same equivalence class. Using only the rotational subgroup (the top two permutations in Table I) of the CFSG will give separate equivalence classes for mirror image conformations.

A generation algorithm would be the following: (1) Take one of the ordered pairs and form its equivalence class by means of the CFSG operations. (2) Save one of the members of this equivalence class. (3) Choose another ordered pair not previously found in an equivalence class and repeat (1)–(3) until no pairs remain. An alternative algorithm would take each pair in sequence and check if the operations of the CFSG gave a "lower" pair based on some predefined ordering of the rotamers and bonds. Only those which gave no lower pair would be retained. This latter method would probably require less storage but would also probably take longer.

For any such algorithm a method of forming equivalence classes (distinct conformations) based on the CFSG is required. To do this the operation of the permutations in the CFSG (Table I, column 5) must be defined. Consider the operation of the C_2 rotation represented as (23(T+−)34(T−+)).

(23(T+−)34(T−+))
[T+] → [−−]

This operation takes the starting 23(T) to 34(−) and the starting 34(+) to 23(−). These are symmetry equivalent

Table II. Generation of Conformations for 2,4-Dimethylpentane (4)

(23)(34)	(23(T+−)-34(T−+))	(23)(34)	(23(T+−)-34(T−+))
[−−]	[T+]	[T−]	[T−]
[−T]	[++]	[TT]	[+−]
[−+]	[−+]	[+T]	[+T]

Table III. Generation of *dl* Pairs and Achiral Conformations for 2,4-Dimethylpentane (4)

(23)(34)	(23(T+−)-34(T−+))	(23(+−)-34(+−))	(23(T+))-(34(T−))	point group
[−−]	[T+]	[++]	[−T]	C_1
[−+]	[−+]	[−+]	[−+]	C_{2v}
[T−]	[T−]	[+T]	[+T]	C_2
[TT]	[+−]	[TT]	[+−]	C_s

conformations ([T+] and [−−]) as can be verified by reference to the rotamers in Figure 1. The operation of this CFSG permutation on all nine possible pairs is summarized in Table II. This gives six equivalence classes which correspond to the six distinct conformations possible. Conformations which are left unchanged by this operation (there are three, [−+], [T−], [+T]) have twofold symmetry. It is in this way that the symmetry groups of the conformations are computed by the above algorithms. These three conformations would have statistical weight 1 while the other three would have statistical weight 2 (apart from energetics). This information is also retreivable by using the above algorithms.

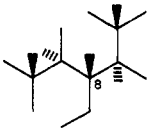
The rest of the CFSG operations collect the conformations which are mirror images into equivalence classes. This result is summarized in Table III. There are four equivalence classes which correspond to the two chiral conformations (*dl* pairs) and two achiral conformations. These four cases are identified by their point group symmetries in Table III. The permutations in the symmetry group of each conformation are also derived by this method.

IV. CONFORMATION SPECIFICATION

A unique specification of the conformation of an organic structure of definite constitution and configuration is provided by establishing the equivalences class of the CFSG to which the conformation belongs. This assumes some choice of the possible torsional positions for each rotatable bond (section II). A unique (canonical) name for a conformation can be obtained by choosing the "lowest" member of the CFSG equivalence class based on some orderings of the torsional positions and rotatable bonds. Thus for the example of 2,4-dimethylpentane (4) used above, if the orderings (− < T < +) and (23 < 34) were chosen for the torsional positions and rotatable bonds, respectively, the lowest or canonical member of each equivalence class of the CFSG would be the leftmost as shown in Tables II and III.

An advantage of this method of conformation specification is that a compact nongeometric representation is obtained which can be turned into a geometric (i.e., atomic coordinates) representation if desired. A structure of given constitution, configuration, and conformation would be represented as a graph (for the bond connectivity) with additional node properties (for the stereocenter configurations) and edge properties (bond torsional positions). This is just an augmented graphical representation; however, by use of standard bond lengths and angles along with the designated torsional angles, a geometric representation could be derived. Another advantage of this method is that the specification simultaneously considers a local (bond torsional position) and global property (overall symmetry) of the structure. The importance of considering these two types of properties (global and local) for stereoisomer specification has been discussed previously.[3,5] The final conformation specification consists of just labeling the augmented graph properly by use of the symmetry group. This final specification does not include the symmetry group nor is the symmetry group necessary to translate it into another representation (such as a geometric representation).

While this method of conformation specification can be used for structures of arbitrary symmetry, it is relatively rare to find chemical structures with any symmetry, particularly natural products and molecules of biomedical interest. However, this method can also be used to specify (i.e., uniquely name) the conformation of substructures (i.e., partial chemical structures) with definite constitution and configuration. There will often be substructures with symmetry contained in complete structures with no overall symmetry. For example, the substructure **5** is contained in the steroid nucleus. In order



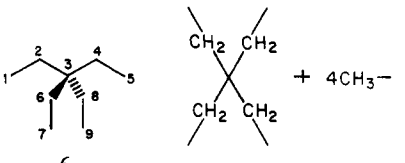
5

to properly specify the conformation of arbitrary structures and substructures, it is particularly important that any symmetry be accomodated.

V. CONFORMATION ENUMERATION

The present method yields a single formula for the unconstrained enumeration of the conformations (based on a discrete number of torsional positions possible for each rotatable bond) of a chemical structure of arbitrary branching and symmetry, a heretofore unsolved problem. The key components of this formula are the conformation symmetry group (CFSG) derived above (section II) and the recently derived exponentiation group enumeration formula.[8b] The derivation of this enumeration formula for conformations parallels the derivation of a similar enumeration formula for configurational stereoisomers.[5,10]

The objective here is to give an enumeration formula which requires only the symmetry group and possible bond torsional positions for the chemical structure under study. This type of enumeration problem has received considerable attention.[11] A common approach has been to consider such problems as substitutional isomerism problems and to use the well-known formulas for enumeration of substitutional isomers.[11c,d] Thus to compute the number of staggered rotamers possible for 3,3-diethylpentane (**6**), one might factor the problem into a staggered neopentane skeleton (**7**) and four methyl ligands which must be substituted onto this skeleton in all possible consistent ways.



6 7

For present purposes this type of approach is unacceptable. First, the chemical structure must be divided or "factored" into two parts: a skeleton (the neopentane skeleton) and ligands (methyl groups). This distinction is arbitrary and will lead to a nondeterministic step in any algorithm which must make use of this method for arbitrary chemical structures. The problem of ligand–skeleton factoring has been discussed in more detail previously.[5] Secondly, the symmetry group of the resulting skeleton will generally be larger and often much

Table IV. Enumeration of Conformations of 2,4-Dimethylpentane (4) Using (1)

symmetry	E	C_2	σ_v	σ_v
permutation	(23)(34)	(23 34)	(23 34)	(23(T+))(34(T−))
n_{ijk}	3 × 3	3	3	1 × 1

larger than that of the original structure.[11a,c,d] Since the enumeration formulas compute terms based on this symmetry group, the number of terms which must be computed in such cases can be excessive. This can lead to a time-consuming step in both an enumeration algorithm and any generation algorithm based on such substitutional methods. Since most chemical structures have little symmetry, a method which requires only this symmetry group and no larger group will be advantageous for these purposes.

The counting formula for the number of conformations, N, can be given by (1) where g is the size (number of elements

$$N = \frac{1}{g}\sum_{i=1}^{c} h_i \prod_{j=1}^{p} \prod_{k=0}^{n_j} n_{ijk} \qquad (1)$$

in) of the CFSG, c is the number of conjugacy classes in the CFSG, h_i is the size of the ith conjugacy class, p is the number of bonds for which rotamers (torsional positions) are designated, n_j is the number of permutation cycles of length j, and n_{ijk} is the number of rotamers fixed in the kth cycle of length j; n_{ij0} is defined to be 1.[12]

To illustrate the use of (1) consider the 2,4-dimethylpentane (**4**) example above. Table IV summarizes the computation for the number of achiral conformations and *dl* pairs. The n_{ijk} terms are easily obtained by using the permutation representation derived above (section II and column 5 of Table I). The first three permutations fix all rotamers in all cycles (there are no permutations of rotamers indicated in the bond permutation cycles, see section II); so each cycle contributes 3 to the total as indicated in Table IV. The final permutation, (23(T+))(34(T−)), exchanges two rotamers in each cycle and hence fixes only one. Thus each cycle contributes 1 to the total as shown in Table IV. The overall total is therefore (1/4)(9 + 3 + 3 + 1) = 4 which is correct as shown earlier (Table III). The total number of conformations considering mirror images separately is obtained by using only the first two terms. The total is (1/2)(9 + 3) = 6 which is correct as indicated in Table II.

As a second example, consider the enumeration of the conformations of 3,3-diethylpentane (**6**). Enumeration of conformations of such a symmetrical structure has proved difficult.[13] The summary of the computation for the numbered structure **6** is given in Table V. Each column has one representative permutation of each conjugacy class. The number of conformations (*dl* pairs and achiral) = (1/24)(81 + 27 + 18 + 18) = (1/24)(144) = 6. The number of conformations (chiral and achiral) is (1/12)(81 + 27) = (1/12)(108) = 9. These results have been given by Prelog[14] and structure drawings have been by Balaban[11b] (although the structure designated as having C_{2v} symmetry actually had D_{2d} symmetry).

Table V. Enumeration of Conformations of 3,3-Diethylpentane (6) Using (1)

E	$3C_2$	$8C_3$	$6\sigma_v$	$6S_4$
(23)(34) (36)(38)	(23 34)-(36 38)	(23 34 38)-(36(T−+))	(23 34)-(36(T+))-(38(T−))	(23 38 34 36)
3 × 3 × 3 × 3	3 × 3 × 3	8 × 3 × 0	6 × 3 × 1 × 1	6 × 3
81	27	0	18	18

Proof of this formula (1) is analogous to that given for a configurational stereoisomer enumeration formula[10] and follows from the results of Kerber for the exponentiation group enumeration formula.[8b] The results are invariant to renumbering or relabeling since these operations have the effect of conjugation, and all the relevant group-theoretical properties of the individual permutations are dependent only on the conjugacy class to which the permutation belongs.[12] The numerical contributions of the various counting terms can be visualized in some cases. For example, the fact that the C_3 permutations in Table V contribute 0 to the total is related to the fact that no conformations of 3,3-diethylpentane (**6**) can have C_3 symmetry.

ACKNOWLEDGMENT

We thank the National Institutes of Health for their generous financial support (RR-0612) and their support of the SUMEX computer resource on which computer programs are developed and made available to outside collaborators (RR-00785).

REFERENCES AND NOTES

(1) Masinter, L. M.; Sridharan, N. S.; Lederberg, J.; Smith, D. J. *J. Am. Chem. Soc.* **1974**, *96*, 7702.
(2) Carhart, R. E.; Smith, D. H.; Brown, H.; Djerassi, C. *J. Am. Chem. Soc.* **1975**, *97*, 5755.
(3) Nourse, J. G.; Carhart, R. E.; Smith, D. H.; Djerassi, C. *J. Am. Chem. Soc.* **1979**, *101*, 1216.
(4) Nourse, J. G.; Smith, D. H.; Carhart, R. E.; Djerassi, C. *J. Am. Chem. Soc.* **1980**, *102*, 6289.
(5) Nourse, J. G. *J. Am. Chem. Soc.* **1979**, *101*, 1210.
(6) Wipke, W. T.; Dyott, T. M. *J. Am. Chem. Soc.* **1974**, *96*, 4834.
(7) Hounshell, W. D.; Dougherty, D. A.; Mislow, K. *J. Am. Chem. Soc.* **1978**, *100*, 3149.
(8) (a) Harary, F. *Am. Math. Monthly* **1959**, *66*, 572. (b) Kerber, A. "Representations of Permutation Groups"; Springer-Verlag: New York, 1975; Vol. II.
(9) Lederberg, J.; Sutherland, G. L.; Buchanan, B. G.; Feigenbaum, E. A.; Roberson, A. V.; Duffield, A. M.; Djerassi, C. *J. Am. Chem. Soc.* **1969**, *91*, 2973.
(10) Nourse, J. G. in "The Permutation Group in Physics and Chemistry". *Lect. Notes Chem.* **1979**, *12*, 19.
(11) (a) Davidson, R. A. Ph.D. Thesis. The Pennsylvania State University, 1976. (b) Balaban, A. T. *Rev. Roum. Chim.* **1976**, *7*, 1049. (c) Frei, H.; Bauder, A.; Gunthard, H. H. *Top. Curr. Chem.* **1979**, *81*, 1. (d) Balasubramanian, K. *Theor. Chim. Acta* **1979**, *53*, 129.
(12) Concise definitions of conjugacy classes, permutation cycles, and other relevant group theoretical concepts have been given previously.[3,5] More complete definitions can be found in any text on Group Theory. For example: Hall, M., Jr. "The Theory of Groups"; MacMillan: New York, 1959.
(13) See ref 11a, p 224 and ref 11b, p 1068.
(14) Cahn, R. S.; Ingold, C. K.; Prelog, V. *Angew. Chem., Int. Ed. Engl.* **1966**, *57*, 385.