244

Tetrahedron. Vol. 32, pp. 2513 to 2519. Pergamon Press 1976. Printed in Great Britain

STRUCTURAL ISOMERISM OF MONO- AND SESQUITERPENOID SKELETONS[1,2]

DENNIS H. SMITH* and RAYMOND E. CARHART
Departments of Chemistry and Computer Science, Stanford University, Stanford, CA 94305, U.S.A.

(*Received in USA* 13 *February* 1976; *Received in UK for publication* 17 *May* 1976)

Abstract—A generator of chemical structures (CONGEN) has been utilized to investigate two aspects of the structural isomerism of mono- and sesquiterpenoid skeletons: (1) the scope of possible isomers under various structural constraints; and (2) the scope of possible isomers based on a mechanistic model which allows interactive exploration of reactions of formation and interconversion. The possibilities, even under severe constraints, are many more than the structural types commonly encountered in nature. These results indicate the potential danger of structural assignment based in part on biogenetic grounds.

INTRODUCTION

Many years of isolation and characterization of terpenoid natural products, together with recognition of the isoprene unit as a common structural building block[3] have provided a historical record which is used frequently in the characterization of new, unknown terpenoids. When spectroscopic and chemical data by themselves are insufficient for unambiguous assignment of structure, candidate structures are frequently evaluated using reasoning loosely termed "biogenetic considerations". These considerations take several forms, including: (a) reasoning by analogy with skeleton types characterized previously in some closely related study; (b) correlation with co-occurring, known seletons; (c) direct invocation of the isoprene rule;[3] (d) selection from among known skeletons those which best fit available data. A brief, and far from comprehensive, examination of recent issues of some journals where new terpenoid skeletons are frequently reported reveals many instances of structural assignment based on one or more of (a)–(d) above.[4]

We stress at the outset that we have no data to contradict the conclusions made previously in references cited[4] and similar studies. The structural assignments seem plausible and may well be correct. However, recent systematic investigations of molecular structure problems utilizing the CONGEN program for computer-assisted structure elucidation[5] have indicated the large number of potential solutions when constraints are insufficient. Intuitively, we expect that there are many other plausible solutions to the cited problems. The problem of Stoessl *et al.*,[4a] has 123 possible solutions;[5] the problem of Tada *et al.*,[4c] 40 solutions; the problem of Takahashi *et al.*,[4b] 876 from spectroscopic data alone, reduced to 20 after consideration of their evidence from chemical degradation work. Of these 20, only the proposed structure possesses a known skeleton; this structure and four others satisfy the isoprene rule with three units linked head-tail, head-tail.

We recognize that reasoning by analogy on biogenetic considerations is a powerful method for focussing on the most plausible structures(s) in such large problems. However, this reasoning carries with it the danger of overlooking novel skeletal types; the extent of the danger depending on the strength or weakness of the biogenetic arguments relating to the formation and interconversion of the various terpenoid skeletons. Although mechanistic formation of the larger terpenoids has been explored with considerable success,[6,7] formation of the various mono- and sequiterpenoids are largely open questions. Complex mechanistic schemes have been proposed[3,8] based on co-occurrence[9] and structural and stereochemical control in laboratory studies. These schemes are recognized to be simplifications.[3,8,9] They have been used to rationalize, retrospectively, apparent interrelationships among members of several classes of structurally related terpenoids.

We undertook this investigation because, to our knowledge, no one previously has explored the potential scope of structural isomerism in the mono- and sesquiterpenoids. Some definition of the realm of possibilities would give a measure of how well the known systems represent all possible systems and, indirectly, how safe one is using known systems for structural assignment. Also, arguments based on mechanistic schemes are invoked in structure elucidation work, even though the prospective, predictive power of the schemes to our knowledge has never been examined. In subsequent sections we discuss results obtained (using CONGEN[5] and extensions) in preliminary investigations of the potential structural isomerism of mono- and sesquiterpenoid skeletons under under various structural and mechanistic constraints. Those interested in pursuing any aspect of this problem in greater depth are referred to the Experimental.

Method

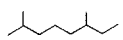
We have investigated two approaches to generation of possible structural isomers of mono- and sesquiterpane skeletons. Method I includes various ways of structure generation under constraints, except those constraints imposed by mechanistic considerations. Method II is based on such mechanistic considerations. These methods construct, or generate, structural isomers.[5,10] We presently take no account of the potential variety of stereoisomers for a given molecular skeleton. Thus, no constraints which speak of stereoisomeric properties can be directly used to restrict structural possibilities; we indicate subsequently the effects of this limitation. With that exception, however, the methods are exhaustive. The structures produced in a given problem should be viewed as all those which are possible under the given constraints. The task of sorting out those possibilities and applying further constraints can also be done with the aid of the program.[5]

Method I

Isomer construction under constraints. We used two approaches to construct isoprenoid skeletons by solving constrained problems.

Method I.A

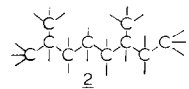
Exhaustive generation with constraints. This approach generates complete sets of structural isomers, under optional constraints, based only on the empirical formula of the structures. To focus initially on skeletons devoid of multiple bonds, each problem was done with the initial constraint that no multiple bonds be formed. For example, for C_{10} isomers, acyclic skeletons were generated from $C_{10}H_{22}$, monocyclics from $C_{10}H_{20}$ and bicyclics from $C_{10}H_{18}$, the last two cases including the constraint of no multiple bonds. During or subsequent to the generation, constraints can be applied to "prune",[5] or reduce, the set. In this example, we could apply the isoprene rule to require a head-to-tail linkage (or any other linkage).[5] This would yield the complete set of structures which could arise from a C_{10}, head-to-tail isoprenoid skeleton (1) disregarding any mechanistic considerations.



1

Method I.B

Formation of rings among atoms of a given skeleton. This approach assumes a given configuration of atoms, for example, the head-to-tail isoprenoid skeleton (1), and assumes initially that all remaining valences of C atoms are potential points of formation of new rings. This is done by allotting "free valences" (bonds with an unspecified terminus)[5,10] to all possible positions of substitution, i.e. yielding 2, for a head-to-tail linked-monoterpane skeleton.



2

CONGEN has the capability of forming a given number of new bonds in a structural fragment such as 2 which possesses some number of free valences (a "superatom").[5] Again, constraints can be used during this procedure, for example, to forbid formation of multiple bonds. One new bond, with this constraint, yields monocyclics, formation of two new bonds yields bicyclics, and so forth. Of course, the results of Methods I.A and I.B must agree for structures based on the same skeleton.[11] Method I.B is more efficient if one desires to examine structures based on a single skeleton.

Method II

A mechanistic model. Facilities have been added to CONGEN to permit modelling of certain cyclization and rearrangement processes.[12] Given a starting structure, or structures, these processes can be exercised stepwise by the chemist, with the capability of examining intermediates and implementing constraints. The available processes, structural requirements and some important constraints which can be used if desired are summarized in Table 1.

The processes in Table 1 are carried out by the chemist interactively with the program. They may be exercised in any meaningful order at the discretion of the chemist. For example, formation of the monoterpanes can be simulated

Table 1. Cyclization and rearrangement processes, structural requirements and available constraints under the mechanistic model, Method II

Process	Structural Requirements	Constraints[a]
Cyclization	Initiated by ⊕ Must cyclize with C=C	Markovnikov and/or Anti-Markovnikov Allylic rearrangement[b]
Hydride Shift	Initiated by ⊕	Path length(s)[c] Form vinylic carbonium ion ? Degree of atom from which shift occurs
Alkyl Shift	Initiated by ⊕	Bond type[d] Path length(s)[c] Form vinylic carbonium ion ? Degree of atom from which shift occurs
Quench[e]	Must be a ⊕	-
Saturate[f]	Must be C=C	-

[a] The constraints mentioned are selected by the user as "switches" to select an option, either as a yes/no question (e.g., "Allow $C{=}C{-}C^{+} \rightarrow {}^{+}C{-}C{=}C$?") or specification of one or more items from a list of possibilities (e.g., various path lengths for hydride or alkyl shifts).

[b] The allylic rearrangement $C{=}C{-}C^{+} \rightarrow {}^{+}C{-}C{=}C$ is viewed formally as a cyclication with formation of a small (2-membered) ring.

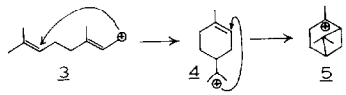
[c] Path lengths refer to the conventional chemical specification of shifts as 1,2, 1,3 and so forth.

[d] May be a ring bond, a chain bond, or either, which is moved from one atom to another, thus effectively migrating an atom and its attached components.

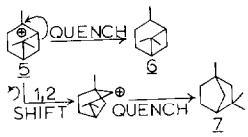
[e] Removes the ⊕ charge by replacing with H^{-}.

[f] Saturates multiple bonds to yield saturated skeletons.

by beginning with structure **3**, which represents geranyl (or nerol) pyrophosphate. The cyclize command (Table 1) would carry out a single cyclization, one of which is indicated as **3** → **4**. A further cyclization of **4** would yield **5**. The cyclizations can be preceeded or followed by any specified sequence of hydride or alkyl shifts.



In general, each step begins with a number of structures and a new group of structures results. The full capabilities of CONGEN to implement further constraints to reject undesired structures can be used at any time during the procedure. After cyclization and rearrangement, the quench command yields a group of structures which can be examined (utilizing CONGEN) for the presence of known skeletons. The saturate and quench commands are used if only simple testing the structures for the presence of known skeletons is desired. For example, quenching of **5** yields the pinane skeleton **6**, while one of the possible 1,2-alkyl shifts exercised on **5**, followed by quenching, yields the fenchane skeleton **7**. Duplicate structures can



arise in a variety of ways from this procedure. In general, a given structure can be arrived at using different combinations of processes. The program tests the structures at each step to eliminate automatically any duplicates formed in this way.

RESULTS AND DISCUSSION

Monoterpanes

The results for monoterpanes obtained using Method I.A are summarized in Table 2. The results obtained using Method I.B for an assumed linkage are summarized in Table 3. Note that there is complete agreement with Table 2 in instances where the results overlap.[11]

These results, even considering only skeletons based on a head-to-tail linkage, are surprising, if not staggering. We present below a summary of the mono- and bicyclic head-to-tail skeletal types constructed by the program. Some of the types may appear implausible based on chemical stability or present knowledge of cyclization processes, but most would be stable (if not natural) products.

Table 3. Possible monocyclic and bicyclic monoterpenoid skeletons obtained using Method I.B

Class	Assumed Linkage[a]	Number of Skeletons
Monocyclic	H-T	29
	H-H	20
Bicyclic	H-T	342
	H-H	247

[a] H-T and H-H mean head-to-tail and head-to-head, respectively.

Monocyclic monoterpanes

The 29 possible head-to gail skeletons are presented in Fig. 1,. The structures numbered **8–13** (Fig. 1) are common, naturally occurring skeletons.[13]

Several important monocyclic monoterpanes are not found in Fig. 1, for example, the tropane, "ortho" and "meta" menthane skeletons and other less common skeletons.[13,18] These skeletons are not head-to-tail isop-

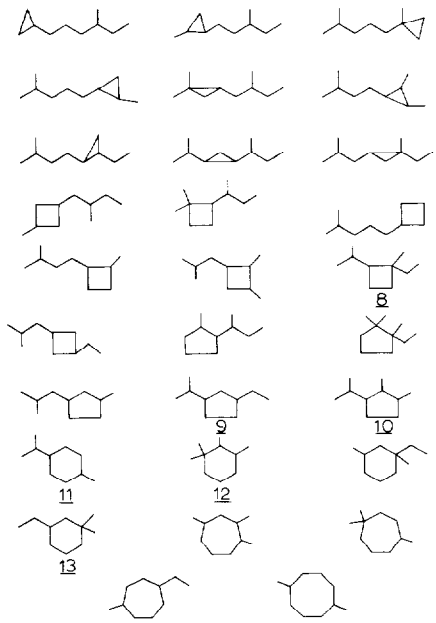


Fig. 1. The 29 possible monocyclic monoterpane skeletons based on a head-to-tail isoprene linkage (from Method I.A and I.B).

Table 2. Possible C_{10} acyclic, monocyclic and bicyclic skeletons, obtained using Method I.A

Empirical Formula	Class	Total Number of Isomers[a]	Isoprene Linkage[b,c]			
			Any	H-H	T-T	H-T
$C_{10}H_{22}$	Acyclic	75	10	1	1	1
$C_{10}H_{20}$	Monocyclic	475	122	20	13	29
$C_{10}H_{18}$	Bicyclic	1792	797	247	139	342

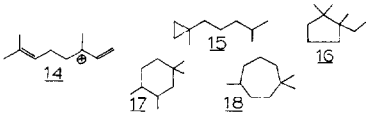
[a] Excluding those with multiple bonds.

[b] The linkage code "Any" means only that the structure can be decomposed into two disjoint isopentyl groups; the linkage may be in any way. H-H, T-T and H-T mean head-to-head, tail-to-tail and head-to-tail, respectively.

[c] There is some overlap among the three categories H-H, H-T and T-T because some structures can be decomposed into isoprene units in more than one way.

renoids. Rearrangements must have occurred, assuming that they were formed originally from a head-to-tail precursor. Method II (below) allows an exploration of transformations which might occur. This procedure can suggest origins for these skeletons, and will also suggest additional skeletons which might arise from similar rearrangements.

The mechanistic model, Method II is relatively straightforward to apply to obtain the monocyclic monoterpenes. There are two structures to consider for cyclization and rearrangement, **3** and **14**, the latter representing linaloyl pyrophosphate.



Simple cyclization of **3** and **14** must necessarily yield a subset of the 29 possible monocyclic monoterpenes based on a head-to-tail isoprene linkage (Fig. 1). Subsequent transformations (Table 1) exercised on these skeletons allow consideration of the results of various possible rearrangements. Initial cyclization yields five structures, including **8**, **11**, **13** (Fig. 1) after quenching and saturation. The other two structures (**15**, **16**) do not occur naturally, to our knowledge. The other three natural products which formally obey the head-to-tail isoprene rule (**9**, **10**, **12**) can be pictured as arising from cyclization of an open chain precursor different from **3** or **14**. Alternatively, **9**, **10** and **12** might arise from **8** in a sequence of rearrangements **8**→**9**, **8**→**10**→**12** but only after complex hydride shifts to place the carbonium ion site (usually regarded as the driving force[3,8]) at the proper point to initiate the ring expansions. In fact, there seem to be no obvious pathways for formation of several of the monocyclic natural products which would not also yield many other structures if allowed as general rearrangements.

For example, beginning with the five initial cyclization products from **3** and **14**, an alkyl shift (under constraints given in the caption to Fig. 3) yields 13 structures (including **8**, **11** and **13**), but no new, naturally occurring skeletons result. A hydride shift followed by an alkyl shift yields thirteen structures, none of which are naturally

Table 4. Numbers of bicyclic and spiro monoterpenoid skeletons, all possessing a head-to-tail linkage, derived from **2** utilizing Method I.B

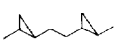
Bicyclic Skeleton	Number[a]	Example
Bicyclo[1.1.0]butanes	19	19
Bicyclo[1.1.1]pentanes	8	20
Bicyclo[2.1.0]pentanes	33	21
Bicyclo[2.2.0]hexanes	11	22
Bicyclo[3.1.0]hexanes	29	23
Bicyclo[2.1.1]hexanes	18	24
Bicyclo[3.2.0]heptanes	24	25
Bicyclo[4.1.0]heptanes	28	26
Bicyclo[2.2.1]heptanes	14	27
Bicyclo[3.1.1]heptanes	12	28
Bicyclo[3.3.0]octanes	4	29
Bicyclo[4.2.0]octanes	5	30
Bicyclo[5.1.0]octanes	13	31
Bicyclo[3.2.1]octanes	15	32
Bicyclo[2.2.2]octanes	2	33
Bicyclo[4.1.1]octanes	11	34
Bicyclo[5.2.0]nonanes	2	35
Bicyclo[6.1.0]nonanes	1	36
Bicyclo[3.3.1]nonanes	1	37
Bicyclo[3.2.2]nonanes	3	38
Bicyclo[4.2.1]nonanes	2	39
Bicyclo[5.1.1]nonanes	1	40
Bicyclo[3.3.2]decanes	1	41
Total bicyclics	257	
Spiro Skeleton[b,c]		
3-3	7	42
3-4	11	43
3-5	8	44
3-6	5	45
3-7	1	46
4-4	4	47
4-5	3	48
4-6	1	49
Total spiro	40	
Total spiro + bicyclics[d]	297	

[a] There are no structures for other bicyclic systems not mentioned, such as bicyclo[4.3.1], [4.4.0]decanes and [4.3.0]nonanes.

[b] The entries in this column are the sizes of the spiro rings.

[c] There are no structures for 3-8, 4-7, 5-5 and 5-6.

[d] The total differs from the complete set of 342 by the structures which consist of two separate ring systems, e.g.



occurring. If one allows a much less probable 1,3 hydride shift preceeding the alkyl shift, then skeletons **10**, **17** and **18** result, but thirteen other skeletons are constructed also. Obviously, any application of a sequence of rearrangement processes, under reasonable constraints, will predict far more skeletons than those which commonly occur. Even though a scheme can always be conceived to transform one skeleton into another, many new skeletons will also be created on general application of the same mechanism. This conclusion is modulated by the fact that we consider no stereochemical constraints, which might serve to reduce the possible cyclizations and rearrangements at each stage (however, see Conclusions).

Bicyclic monoterpanes

We have again focused our attention on the head-to-tail linked structures simply because of the already vast numbers of possibilities considering only ten carbons and two rings (Tables 2 and 3). We present in Table 4 the numbers of the various bicyclic and spiro ring systems which formally obey the isoprene rule. Representative structures for each skeletal type are presented in Fig. 2, including the naturally occurring skeletons thujane (**23**), filifolane (**25**), carane (**26**), camphane (**27**) and pinane (**28**). The other skeletal types have no naturally occurring examples; the other common naturally occurring skeletons are non-head-to-tail or not isoprenoid. Some of the skeletal types presented in Table 4 and Fig. 2 are strained ring systems (for example, the bicyclo[1.1.0] butanes) and less plausible as natural products on that basis. Many of the remaining skeletons, however, are expected to be stable, if not natural products.

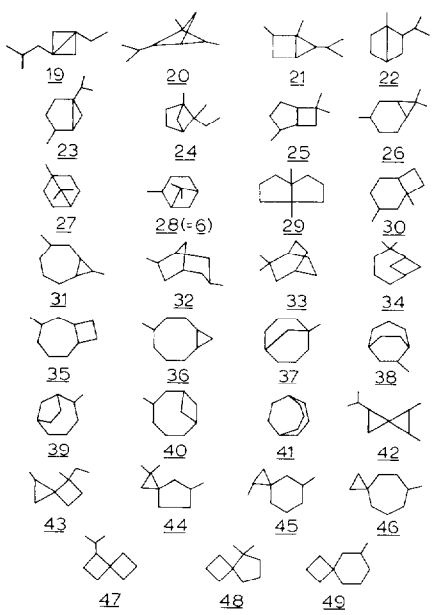


Fig. 2. Representative head-to-tail linked, bicyclic, monoterpenoid skeletons for the skeletal types summarized in Table 4.

In view of the numbers of possibilities presented above, the selectivity of the cyclization and rearrangement processes in nature is remarkable indeed. Schemes presented previously[3,8] have invoked hydride and alkyl shifts to rationalize formation of common skeletons. We have emulated these proposals using Method II to form doubly cyclized skeletons from **3** and **14**. There are eight such skeletons, depicted in Fig. 3, including the naturally

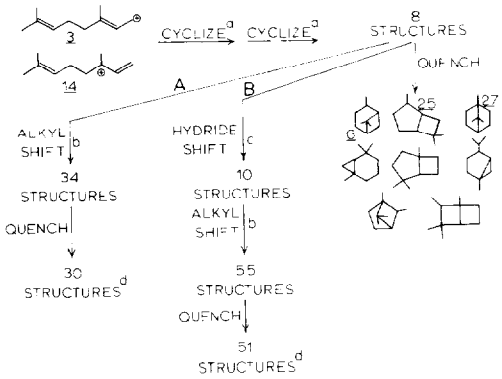


Fig. 3. Examples of rearrangement schemes used to form bicyclic monoterpenoid skeletons. Constraints were used as follows: (a) Cyclization—Markovnikov and anti-Markovnikov, no allylic rearrangement; (b) alkyl shifts—1,2 or 1,3 shifts allowed from carbon atoms of degree three or four (i.e., no formation of primary carbonium ions), no formation of vinylic carbonium ions, ring or chain bonds allowed to shift; (c) hydride shifts—1,2 shifts allowed from secondary or tertiary carbon atoms, no vinylic carbonium ions allowed; (d) fewer structures result on quenching because some unique ionic structures are made equivalent when the charged site is removed.

occurring filifolane (**25**), camphane (**27**) and pinane (**6–28**) skeletons. We have explored various combinations of alkyl and hydride shifts beginning with these eight skeletons. The number of structures resulting from such combinations increases rapidly. For example, a hydride shift followed by two alkyl shifts (under constraints summarized for similar shifts, Fig. 3) yields 91 unique skeletons.

Interestingly, the entire set of common, naturally occurring skeletons[18] can be rationalized by initial cyclizations followed by an alkyl shift (Path A, Fig. 3) or a hydride/alkyl shift combination (Path B, Fig. 3). The source of each common skeleton is summarized in Table 5. Although Paths A and B (Fig. 3) account for all common skeletons, they also predict 61 additional, unique structures (some equivalent structures are constructed by Paths A and B). It appears on manual examination that stereochemical arguments are insufficient to eliminate these other possibilities.

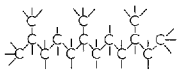
Table 5. Possible origins of various bicyclic monoterpene skeletons using Method II

Skeleton	Initial Cyclization	Path[a] A	B
thujane	-	-	✓
filofolane	✓	✓	-
carane	-	-	✓
camphane	✓	✓	-
pinane	✓	✓	✓
isocamphane	-	✓	-
fenchane	-	✓	✓
α-fenchane	-	✓	✓
β-fenchane	-	-	✓
"B001"[18]	-	✓	-
"B020"[18]	-	-	✓

[a] See Figure 3.

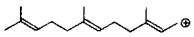
Sesquiterpanes

The scope of potential skeletal isomerism in the C_{15} terpenoids is so vast as to preclude use of Method I.A for the simple reason of efficiency. In fact, Method I.B is also highly impractical for the higher n-cyclics. Using the basic farnesane skeleton (**50**) in conjunction with Method I.B., we obtain 79 monocyclic skeletons. We estimate many hundreds of bicyclics and thousands of tricyclics using this method.

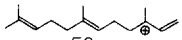


50

The common known sesquiterpane skeletons[17b,18] are heavily represented by skeletons which no longer retain the head-to-tail linkages exemplified by **50**. However, they can be visualized as arising from initial cyclization of farnesyl or nerolidyl pyrophosphate. We have, therefore, explored isomerism of sesquiterpene skeletons using Method II, beginning with **51** and **52**, representing farnesyl and nerolidyl pyrophosphate, respectively.



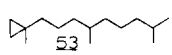
51



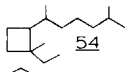
52

Monocyclic sesquiterpanes

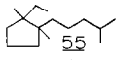
Initial cyclization of **51** and **52** yields nine skeletons, **53–61**. This group of structures includes the known skeletons bisabolane (**56**), germacrane (**60**) and humulane (**61**). Skeletons **56**, **60** and **61**, together with **57**, are



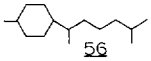
53



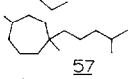
54



55

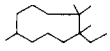


56



57

58



59

60

61

mentioned as initial cyclization products of **51** and implicated as precursors for further cyclization (see structures **6–9** of Ref. 17*b*). However, neither **57** nor any of **53–55**, **58** or **59** have been invoked in subsequent cyclization schemes used to rationalize the interrelationships among important sesquiterpenoid skeletons.[3,8] Other known monocyclic skeletons are thought to arise from rearrangements of the above skeletons or from ring opening of a bicyclic system. We have not explored sources of additional skeletons based on rearrangements because of the greater importance of the bicyclic and tricyclic systems.

Bicyclic sesquiteranes

We have explored several cyclization and rearrangement pathways as routes to formation of bicyclic systems. Some results are presented in Table 6, including known skeletons predicted by each pathway. We also explored the pathway involving a single cyclization, a hydride shift, followed by a second cyclization (41 structures) followed by an alkyl shift (182 structures). The resulting structures have fewer representatives of known skeletons than those summarized in Table 6.

The predictive ability of these pathways is poor. We have not explored more complex pathways which manual examination reveals are necessary to arrive at other known skeletons because we know that a wide variety of additional skeletal types will be produced. Note, for example, the dramatic increase in the number of structures with a single alkyl shift (Table 6).

Tricyclic sesquiterpenes

The predictive power of cyclization and rearrangement processes for tricyclic sesquiterpane skeletons is discouraging. For example, three cyclizations of **51** and **52** yield 56 skeletons, of which "006" and "005"[18] are naturally occurring, indicating that even a severely constrained cyclization model is not capable of focusing on known skeletons. A single alkyl shift from the set of 56 yields 238 structures, while a hydride shift followed by an alkyl shift yields more than 450 structures. None of the common tricyclic skeletons are accessible via these routes. Almost all of the tricyclic skeletons summarized by Devon and Scott[18] arise from more complex rearrangements. Again, if such rearrangements were used to transform all of the possible precursors, far too many structures would be predicted to be meaningful.

CONCLUSIONS

Our results illustrate the enormous selectivity of the natural processes toward formation of only a very small percentage of the possible mono- and sesquiterpenoid skeletons. This, of course, is responsible for the success

Table 6. Cyclization and rearrangement pathways to bicyclic sesquiterpane skeletons and predicted, known skeletons[18]

Pathway[a]	Number of Predicted Skeletons	Predicted Common Skeletons
Double Cyclization (and subsequent)	31	carotane, cuparane, caryophyllane[b]
Alkyl Shift	132	laurane, santalane, carotane, cuparane, caryophyllane
(or)		
Hydride Shift/ Alkyl Shift	216	acorane, chamigrane

[a] For constraints, see caption to Figure 3.

[b] The double cyclization pathway predicts only two of the twenty-five, fifteen carbon "less common"[18] skeletons.

of structural analysis based in part on analogy with known terpenoids. But our results should also serve as a warning that many other plausible structures are also possible and this variety should be kept in mind when new structures are elucidated.

Our hope was that Method II would help focus on the naturally occurring skeletons, suggest other likely candidate skeletons which are as yet unknown, and quantify the utility of such rearrangement schemes applied to elucidation of new structures. But cyclization and rearrangement processes as exemplified by Method II are inadequate and the concept fails as an hypothesis for prediction of plausible structures because it predicts far more possibilities than are observed. Although more detailed mechanistic models may be used by those active in terpene research, there is no precise codification of the structural and stereochemical requirements which favor one pathway among several alternatives. Stereochemical constraints might serve to reduce the number of possibilities, but probably not enough to modify our basic conclusion. For example, we have studied rearrangements of only known cyclized skeletons and still obtain tremendous numbers of alternatives to those observed. The inadequacy of the rearrangement and cyclization hypothesis is understood, as we stated at the outset. We only stress again that extreme caution must be used if it is used in any predictive sense in structure elucidation. However, some of our results might provide a guide to search for new structures or to rationalize why other alternatives are not observed.

There are many other questions which could be explored concerning the scope of structural isomerism of terpenoids under various constraints. We invite persons interested in any sets of structures discussed in this paper to contact us. Alternatively, the CONGEN program plus extensions used in this work are available for outside users (see Experimental). This approach can be used to explore other aspects of the specific problems introduced in our study, or to assist in the determination of new structures. In addition, CONGEN can be used to substantiate arguments based on biogenetic considerations, for example, to survey large groups of structural alternatives for those which obey the isoprene rule, or for those which are based on a known skeleton.

EXPERIMENTAL

The CONGEN program runs on a Digital Equipment Corporation KI-10 computer at the SUMEX computer facility[2] at Stanford University. The program is available to an outside community of users via a nationwide computer network, to the limits of available resources. For additional information on access to SUMEX or to details of results, contact the authors or Prof. J. Lederberg, SUMEX, Dept. of Genetics, Stanford University Medical School, Stanford University, Stanford, CA 94305.

REFERENCES

[1]Part XXIV, *Applications of Artificial Intelligence for Chemical Inference.* For Part XXII, see B. G. Buchanan, D. H. Smith, W. C. White, R. G. Gritter, E. A. Feigenbaum, J. Lederberg and C. Djerassi, *J. Am. Chem. Soc.* in press.

[2]We wish to thank the National Institutes of Health for their generous financial support (RR 0612-05A1) and for their support of the Stanford University Medical Experimental Computer Facility, SUMEX (RR-785) on which the CONGEN program[5] is implemented.

[3]L. Ruzika, *Experientia* **9**, 357 (1953).

[4a]A. Stoessl, J. B. Stothers and E. W. B. Ward, *J. Chem. Soc. Chem. Commun.* 709 (1974); [b]S. Takahashi, H. Iinuma, T. Takita, K. Maeda and H. Umezawa, *Tetrahedron Letters* 4663 (1969); [c] M. Tada, Y. Moriyama, Y. Tanahashi and T. Takahashi, *Ibid.* 4007 (1971); [d]S. K. Paknikar, S. V. Bhatwadekar and K. K. Chakravarti, *Ibid.* 2973 (1975); [e]I. A. Southwell, *Ibid.* 1885 (1975); [f]S. C. Taneja and H. P. Tiwari, *Ibid.* 1995 (1975); [g]B. J. Burreson, F. X. Woolard and R. E. Moore, *Ibid.* 2155 (1975); [h]N. H. Fischer, R. A. Wiley, Jr., H-N. Lin, K. Karimian and S. M. Politz, *Phytochemistry* **14**, 2241 (1975); [i]J. P. Campello and S. F. Fonesca, *Ibid.* 2299 (1975); [j]R. D. Stipanovic, P. J. Wakelyn and A. A. Bell, *Ibid.* 1041 (1975).

[5]R. E. Carhart, D. H. Smith, H. Brown and C. Djerassi, *J. Am. Chem. Soc.* **97**, 5755 (1975).

[6]E. E. van Tamelen, *Accts. Chem. Res.* **8**, 152 (1975).

[7a]R. B. Clayton, *Aspects of Terpenoid Chemistry and Biochemistry* (Edited by T. W. Goodwin), p. 1. Academic Press, New York (1971) and Refs cited; [b]M. J. O. Francis, *Ibid.* 29 (1971); and Refs cited.

[8a]J. H. Richards and J. B. Hendrickson, *The Biosynthesis of Terpenes, Steroids and Acetogenins.* Benjamin, New York (1964); [b]G. Rücker, *Angew. Chem.* **85**, 895 (1973).

[9]E. Zavarin, *Phytochemistry* **9**, 1049 (1970).

[10a]L. M. Masinter, N. S. Sridharan, J. Lederberg and D. H. Smith, *J. Am. Chem. Soc.* **96**, 7702 (1974); [b]L. M. Masinter, N. S. Sridharan, R. E. Carhart and D. H. Smith, *Ibid.* 7714 (1974).

[11]Methods I.A and I.B use completely different parts of the CONGEN program. Agreement on results helps build confidence in the correctness of the program.

[12]Other computer programs can model certain mechanisms of cyclizations and/or molecular rearrangements such as 1,2-shifts. See, for example [a]W. T. Wipke and T. M. Dyott, *J. Am. Chem. Soc.* **96**, 4825 (1974); [b]C. J. Collins, C. K. Johnson and V. F. Raaen, *Ibid.* 2524 (1974); [c]T. M. Gund, P. v. R. Schleyer, P. H. Gund and W. T. Wipke, *Ibid.* **97**, 743 (1975).

[13]Statements throughout this paper about the occurrence of certain skeletons are based on the skeletons presented in books by Simonsen,[14] de Mayo,[15] Overton,[16] Newman,[17] and Devon and Scott.[18]

[14]J. L. Simonsen, *The Terpenes* (2nd Edition, revised), Vol. II. Cambridge University Press (1949).

[15]P. de Mayo, *Mono- and Sesquiterpenoids*, in *The Chemistry of Natural Products*, (Edited by K. W. Bentley) Vol. II, Interscience, New York, N.Y. (1959).

[16a]A. F. Thomas, *Terpenoids and Steroids*, Vol. 1, Specialist Periodical Reports, p. 7. K. H. Overton, Sr., Reporter, The Chemical Society, London (1971); [b]J. S. Roberts, *Ibid.* p. 51.

[17a]D. Whittaker, *Chemistry of Terpenes and Terpenoids* (Edited by A. A. Newman), p. 11. Academic Press, New York (1971); [b]J. S. Roberts, *Ibid.* p. 88.

[18]T. K. Devon and A. I. Scott, *Handbook of Naturally Occurring Compounds*, Vol. II, *Terpenes.* Academic Press, New York (1972).

[19]The groupings can be (and were) done by the program under the control of the chemist so that subsets of the results can be isolated and examined.