

Conformations of Cycloheptadecane. A Comparison of Methods for Conformational Searching

Martin Saunders,^{*,†} K. N. Houk,^{*,†} Yun-Dong Wu,[†] W. Clark Still,^{*,§} Mark Lipton,[§] George Chang,[§] and Wayne C. Guida[§]

Contribution from the Department of Chemistry, Yale University, New Haven, Connecticut 06511, Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90024, and Department of Chemistry, Columbia University, New York, New York 10027. Received December 16, 1988

Abstract: As a test of the effectiveness of various methods for searching the conformational space of highly flexible molecules, we conducted a series of conformational searches on cycloheptadecane. The methods studied included systematic and random conformational searches in both internal (torsion angle) and external (Cartesian) coordinate frames. Distance geometry and molecular dynamics were also investigated. While none of the methods examined found all the low-energy conformations in a single search, all methods except distance geometry located the same global minimum. Some methods performed better than others at finding the other low-energy conformers. In all, 262 conformations of cycloheptadecane having MM2 energies within 3 kcal/mol of the global minimum were discovered. Among highly symmetrical structures, cycloheptadecane lies close to the boundary distinguishing problems which can and cannot be adequately addressed by using contemporary methodology and resources.

What is the preferred conformation of cycloheptadecane? Better, what conformers are significantly populated at room temperature or within, say, 3 kcal/mol of the global minimum? These questions are representative of those currently being asked throughout chemistry, biochemistry, and molecular biology. By addressing these questions using recently developed conformational search techniques, this paper shows how the conformational multiple minimum problem can be solved for a reasonably complex molecule.

A number of different methods have been proposed as solutions to the conformational multiple minimum problem.¹ In this paper we discuss the strengths and weaknesses of several promising conformational search procedures and compare the efficiency of the various methods in solving a challenging problem in conformational analysis—finding all the conformational minima of cycloheptadecane within 3 kcal/mol of the ground state on the MM2² potential surface.³ With its 51 atoms, cycloheptadecane is a large ring hydrocarbon formally having 147 ($3N - 6$) degrees of freedom and hence is expected to be rich in low-energy conformational minima. To convey the magnitude of the problem, we note that restricting a 17-carbon chain to gauche or anti torsion angles alone yields 3¹⁴ or 4 782 969 conformations. Most of these geometries do not place the chain ends near enough to close the ring and many are unreasonably high in energy due to the proximity of nonbonded atoms. However, thousands of other conformations composed of approximately gauche and anti torsion angles do yield rings lacking high-energy nonbonded contacts. There are also many conformations having one or more torsion angles which depart substantially from ideal gauche and anti geometries. The objective of this work is to find all of these low-energy conformers.

Most approaches to conformation searching operate as follows. First, a crude starting geometry is produced, and its structure is optimized by molecular mechanics energy minimization. The resulting minimum energy conformer is then compared with previously found conformers to test for possible duplication. If the conformer thus generated is a previously undiscovered one, it is added to an accumulating list of unique conformers. The cycle is then repeated by obtaining a new crude starting geometry, energy minimizing, etc. At some point, for example, when all given starting geometries have been used or when new minima cease to be found, the search is terminated. Since the energy minimization part of the process simply refines starting geometries

to nearby local minima, it is the starting geometry generation algorithm which most directly controls the overall effectiveness of the search in reaching convergence. While starting geometry generation controls the number of structures which must be minimized to reach convergence, the total time also depends upon the speed at which each structure is energy minimized.

The methods used to generate starting geometries can be divided into two broad categories—deterministic searches which cover all areas of conformational space systematically and stochastic methods which use a random element in exploring space. The systematic or grid searches are typically conducted by altering torsion angles and operate by generating all combinations of selected values for some or all rotatable torsion angles to produce starting geometries distributed throughout conformational space. Although the systematic procedures provide guaranteed coverage of all regions of space, they generate an extraordinarily large number of starting structures when the number of possible values for each torsion angle is large. If the number of values for each torsion is small, then the resolution with which space is covered may be inadequate to provide a starting geometry in the vicinity of each minimum. Consequently low-energy minima may be missed. The stochastic or Monte Carlo searches, on the other hand, generate starting geometries using random or pseudorandom variations of molecular geometry. Such methods do not search conformational space in a completely random manner in that they start with stable conformers and limit their conformational space explorations to variations in selected internal (torsion angle) coordinates or to small (<3.5 Å) Cartesian displacements of atoms. The random methods will cover all regions of conformational space, however, if they are unconstrained and allowed to run for a sufficiently long period of time. Since the probability of finding new conformers at any point in a random search depends upon the number of undiscovered conformers, the yield of new conformers decreases as the search progresses.

Another distinguishing feature of various conformational search methods is the coordinate system in which they operate. Com-

(1) For a recent review of conformational searching, see: Howard, A. E.; Kollman, P. A. *J. Med. Chem.* **1988**, *31*, 1669.

(2) Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127.

(3) This project began with a wager between K.N.H. and M.S. concerning the best method for locating the global minimum and other low-energy conformations of large ring molecules. According to the original terms of the bet, M.S. was the winner; however, we leave it to the reader to judge for himself which is actually the best of the tested methods. The studies at each university were carried out independently. The Monte Carlo internal coordinate searches and the improved Cartesian stochastic searches were carried out after the original searches by the other methods were completed and compared.

[†]Yale University.

[‡]University of California.

[§]Columbia University.

monly used systems are external (Cartesian) coordinates, internal (bond lengths, bond angles, and torsion angles) coordinates, and distance geometry⁴ representations (matrices of internuclear distances). In the paragraphs below, we describe potential advantages of the various coordinate systems.

One other technique, molecular dynamics, has also been used for exploring conformational space. While molecular dynamics is effective at searching local conformational space, it requires significantly more computer time than the other methods to generate significantly different geometries and is slow to cross large energetic barriers separating minima. Thus we do not expect it to be competitive with the other methods for global search problems like this one.⁵

We describe here our evaluations of conformational space searching by using both systematic and random methods operating in both internal and external coordinate systems. We will show that some methods are more efficient than others at solving the cycloheptadecane problem. We will also show that other issues, such as hardware speed and the effectiveness of the energy minimization algorithm in converging rapidly, have a major impact on what kind of problem can be solved in a reasonable length of time. Finally we note that the cycloheptadecane problem is simplified by the symmetry of the cycloalkane ring and is comparable in difficulty to a conformational search of a saturated 12-membered unsymmetrical ring compound. Both of these conformational analysis problems appear to lie just at the boundary separating problems which can and cannot be solved by using contemporary methodology and computing resources.

The Case for External Coordinate Searches

A list of Cartesian coordinates for all the atoms of a molecule is by far the simplest of all representations of a molecular geometry to manipulate. The stochastic method employs a random translation algorithm to generate starting geometries. It operates by applying limited, random translations or "kicks" to each atom of known conformations to give new starting geometries for energy minimization. Examples of its application to cycloalkanes up to cyclopentadecane have already been published by one of us (M.S.).^{6a,b} An important advantage of sets of external as opposed to internal coordinates is that exactly the same algorithm can be used for acyclic, monocyclic, and polycyclic structures. The algorithm for the kick requires only eight lines of FORTRAN to implement. Thus, the Cartesian search procedure is both simple and general.

After translating the atoms to new positions, the steric energy rises substantially. However, optimization using a simple two-body force field (STRFIT) lowers the steric energy rapidly and effectively.⁷ Energy minimization with a standard molecular mechanics program then gives a local minimum which is either the same as or is different from the original conformation. Only a method for storing the conformers and a procedure for searching previously found conformers are necessary to produce a complete search procedure. It has been found that a maximum kick size of around 3 Å returns to the starting conformer infrequently and yet gives a structure which can be energy minimized rapidly.

Cartesian Stochastic Search

Eight separate runs were made with the stochastic search method.^{6a,b} The maximum size of the random kick, which is applied to each atom at each cycle, was varied in the different runs but was in the range of 2.7–3.1 Å. No appreciable difference

in performance of the search was found as a function of the kick size in this range. If the maximum kick size selected was much smaller (<1.5 Å), the probability of falling back to the starting structure became appreciable. The disadvantage of a maximum kick size which is much larger than 3 Å is that a larger fraction of the structures found have steric energies greater than the limit selected and the average time for optimization increases.

As previously described, each stochastic kick was followed by optimization of the resulting cycloheptadecane starting geometry, first using the STRFIT two-body force field and then by MM2, before the structure was compared with all the structures found previously in that run. The kick-refinement cycle took an average of 15 CPU min on a MicroVAX II computer. The first eight runs (using a total of 94 days of CPU time) totalled 9015 kicks and yielded a total of 766 structures (including duplicates) within a 3-kcal/mol range above the lowest energy structure.

The structures provided by the eight runs were combined and sorted by energy. Duplicate structures were then removed. After this process (and the refinement described below), 222 unique conformers were left in the range of 3 kcal/mol above the lowest energy one (see Table I, method 1a).

Since the first eight runs using the stochastic search method were done over a year ago, substantial improvements have been made in the stochastic search program (M.S.). The present program no longer uses the most recently found structure as the starting point for the next kick. Instead, the program randomly chooses one of the structures previously found by using a function which chooses the lower energy ones with greater probability. After each structure is chosen, the program increases the recorded energy by a small amount (~0.05 kcal/mol). The purpose of this is to make it slightly less probable that this structure is chosen again. This approach is similar to the "usage-directed search" described below.¹¹ Higher energy structures are still chosen as starting points but with lower probabilities.

Important changes were also made in the two-body force program which is used as a prefitter directly after the kick and before using MM2. The geometry optimization procedure is now divided into three separate stages. First, only bonding distances are adjusted, and then 1,3-interactions are included to improve bond angles. Finally, the nonbonded forces are turned on as well. The force field parameters and the convergence procedure were also very substantially improved (using a stochastic method to search for better parameter sets). The object was to make the program work more quickly and yield structures with lower MM2 strain energies (so as to save time in the subsequent MM2 optimization). Indeed, the average time for a complete cycle (kick-prefit-MM2) was reduced from around 15 to 10–11 min. However, the overall program efficiency measured by the number of low-energy structures found per unit time improved by a much greater factor. A likely reason for this substantial improvement is that, in the earlier program, some kicked structures which should have refined to low-energy conformations ended up as relatively high-energy structures due to less than optimal prefitter parameters. This is much less likely with the present program which uses a force law much closer to that of MM2. The fraction of conformers found at low energy is consequently much higher in the improved version of the search program (see Table I, method 1b).

Since the MM2 program often does not refine to the precise minimum of the MM2 force field, and sometimes stalls close to transition states, another operation was carried out. A program was written which repeatedly rerefines each structure, followed by a stochastic search for the best structure by using very small random kicks.^{6c} The object was to bring each conformation as close to the MM2 minimum energy structure as possible. After many cycles of application of this program, no additional decreases in energy occurred. We will comment on energy minimization convergence in a later section.

In considering the efficiency of any search procedure, one might ask how long it takes to find all of the conformers. Unfortunately, there is no simple answer. The total number of conformers of most molecules is usually unknown, and it is difficult to prove in

(4) Crippen, G. *Distance Geometry and Conformational Calculations*; Research Studies Press, Wiley: New York, 1981. Weiner, P. K.; Profeta, S.; Wipff, G.; Havel, T.; Kuntz, I. D.; Langridge, R.; Kollman, P. A. *Tetrahedron* **1983**, *39*, 1113.

(5) A referee has taken issue with this statement and believes that molecular dynamics may be more efficient as a global search method than we suggest here, see: Noguta, T.; Go, N. *Biopolymers* **1985**, *24*, 427.

(6) (a) Saunders, M. *J. Am. Chem. Soc.* **1987**, *109*, 3150. (b) Saunders, M. *J. Comput. Chem.* **1989**, *10*, 203. (c) For a related method, see: Ferguson, D. M.; Raber, D. J. *J. Am. Chem. Soc.* **1989**, *111*, 4371.

(7) Saunders, M.; Jarret, R. M. *J. Comput. Chem.* **1986**, *7*, 578.

Table I. Unique Conformers Found versus CPU Time Used during Conformational Searches^a

kcal/mol	MicroVAX II CPU days								
	3	6	9	12	15	20	30	60	90
1a. Saunders Cartesian Coordinate Stochastic Search ^l									
1			9				8	10	11
2			37				50	64	67
3			92				137	207	222
1b. Saunders Improved Cartesian Coordinate Stochastic Search ^l									
1			11		11 ^e				
2			50		59 ^e				
3			140		178 ^e				
2. Houk/Wu Torsional Tree-Search ^l									
1	4		7, ^c 7 ^d						
2	26		36, ^c 42 ^d						
3	85		114, ^c 138 ^d						
3. Still/Lipton Torsional Tree-Search ^l									
1	4			9				9	
2	31			60				63	
3	95			190				211	
4. Still/Chang/Guida Torsional Monte Carlo Search, Random Walk ^l									
1	9	9	10	11	11	11	11	11	
2	41	46	56	61	61	63	67	67	
3 ^b	102	135	179	200	205	218	237	237	
5. Still/Chang/Guida Torsional Monte Carlo Search, Usage-Directed (Run A) ^l									
1	10	10	10	10	10	11	11	11	
2	44	53	57	61	63	66	69	69	
3 ^b	110	152	183	203	215	232	249	249	
5. Still/Chang/Guida Torsional Monte Carlo Search, Usage-Directed (Run B) ^l									
1	11								
2	68								
3	260								
results from a single 72-h run on a Convex C210 computer									
6. Distance Geometry ^{f,j}									
1					2, ^g 6 ^h		7 ^h		
2					30, ^g 35 ^h		44 ^h		
3					100, ^g 122 ^h		176 ^h		
7. Molecular Dynamics ^{i,l}									
1							9, ^j 7 ^k		
2							57, ^j 46 ^k		
3							169, ^j 156 ^k		

^a The numbers of structures reported correspond to unique minima with energies relative to the global minimum up to that given in the table; data entries are cumulative over the CPU time given. ^b These numbers are minimum values since 6–8 Å van der Waals soft cutoffs were used during the search with a 3 kcal/mol cutoff, and the resulting structures were then reminimized by full matrix Newton–Raphson with all nonbonded interactions. ^c Results from 1330 starting geometry run. ^d Results from 1794 starting geometry run. ^e Results from 4.5 CPU days on a VAX 3500 (speed ~3.5 times a MicroVAX II). ^f Using the DGEOM program written by Jeff Blaney and G. Crippen to generate starting geometries and the BATCHMIN program (see text) to energy minimize and eliminate duplicates. ^g Solving for carbons only. ^h Solving for carbons and hydrogens. ⁱ Using BATCHMIN to generate and sample 4000 starting geometries over the course of a constant temperature molecular dynamics simulation and energy minimization with elimination of duplicates. ^j Simulation temperature 500 K, CH bonds constrained by SHAKE, timestep = 1.25 fs. ^k Simulation temperature 1000 K, CH bonds constrained by SHAKE, timestep = 0.33 fs. ^l Method.

a conformational search the point at which no more structures remain to be found. Even when the number of conformers is given, the time necessary for any particular search to find them all will vary from run to run depending upon the initial structure and random number sequence. Thus it would be necessary to carry out multiple searches and average the results, an option unfortunately precluded by the magnitude of the cycloheptadecane problem. An advantage of this stochastic search procedure is that, with assumptions, one can estimate the probability of finding (or missing) any structure.^{6b}

While we present evidence below that 262 conformations of cycloheptadecane lie within the first 3 kcal/mol, the rate at which this limit is approached is more important in the context of this paper than the time necessary to find them all. In our stochastic Cartesian searches, one run of 435 steps found 7 of the 15 lowest energy conformers. A somewhat longer 1000-step run found 11 of the 15. In Table I, the cumulative number of distinct conformers found by a series of searches are given as a function of

total CPU time. The stochastic Cartesian searches are given as methods 1a and 1b. The full cumulative results for method 1a covers 90 CPU days on a MicroVAX II during which time 9015 structures were generated and energy minimized.

The Case for Internal Coordinate Searches

Regardless of the coordinate system used in a search of conformational space, the number of possible three-dimensional structures expands geometrically with the number of degrees of freedom, and the molecular $3N - 6$ degrees of freedom make the problem of finding all minima by a full-space search intractable for all but the simplest of molecules. Nevertheless, some coordinate systems appear better than others. A coordinate system which distinguishes energetically constrained degrees of freedom (e.g., bond lengths and angles) from relatively unconstrained ones (e.g., torsion angles) where the major variations occur is particularly useful. With such a distinction, the total number of degrees of freedom can be reduced to the most relevant, least constrained

ones. For searching the conformational space of molecules, we therefore use one of the well-known properties of conformers (i.e., that they differ primarily in torsion angles) to confine our searches to torsion space alone. There, an otherwise R^{3N-6} problem is reduced to at most a R^N one (R is the number of points sampled along each torsional coordinate, N is the total number of atoms and N' is the number of independent torsions).

The more the low-energy geometrical features of molecules are incorporated into starting structures, the more the conformational search should give low-energy final conformers. Such geometrical features can be localized, e.g., idealized C–C and C–H bond lengths, tetrahedral bond angles, and staggered torsional arrangements. Additionally, they can consist of larger substructures such as extended π -systems or specific rings. Internal coordinates make it simple to install the features which are associated with molecular stability into starting geometries and also make it easy to explore widely separated regions of conformational space by orchestrating the concerted atom movements of large molecular substructures. Thus internal coordinates in general and torsion angles in particular would seem to provide the most appropriate coordinate frame in which to conduct conformational searches.

Internal Coordinate Tree-Search

Several of us (W.C.S. and M.L.) recently described a method for searching the torsional space of molecules by using a tree-search algorithm which efficiently and uniformly generates starting geometries throughout all of conformational space.⁸ The method operates in two stages: first, we generate starting geometries by using MACROMODEL V2.0, and, second, we minimize the structures with deletion of the duplicate or high-energy conformers.⁹ The structure generation phase begins with selection of a series of possible values (e.g., 0°, 60°, 120°, 150°, ...) for each torsion angle and then formation of all possible combinations of the values to generate starting geometries which are compatible with nonbonded contact and ring closure constraints.

Applied to the cycloheptadecane problem, the search parameters found to be adequate for the medium rings (i.e., 60° dihedral angle resolution, 0.5–5.0 Å ring closure distance, 1.0 Å nonbonded rejection distance)¹⁰ could not be used because of the excessive number of starting geometries generated ($>10^5$). The magnitude of the problem at hand is too great for the tree-search applied as described for the medium rings.

To extend the tree-search method to larger systems, we investigated ways to limit the number of starting geometries while retaining as complete as possible a set of the low-energy final conformers. We began by reducing the dihedral angle resolution to 120° and then further restricted the search by adding a series of conformationally reasonable constraints. First, we required that the initial two dihedral angles in the ring have values +60° and +60°. We therefore began all conformations of the ring with a +gauche/+gauche (+g/+g) "corner" on the assumption that at least one such corner should appear somewhere within each conformation. The symmetry of cycloheptadecane rendered the position of the corner in the ring and its absolute configuration irrelevant. The 120° search with the described constraint gave ~20 000 starting geometries. Since this number of structures was more than we could afford to energy minimize, we sought additional constraints to further limit the number of structures generated.

In conformational analysis, +gauche/–gauche (+g/–g, syn) pentane fragments are generally associated with high steric energies due to the proximity of the nonbonded terminal carbons. In *cis*-1,3-dimethylcyclohexane-like systems, for example, the 1,3-diaxial methyl/methyl interaction is repulsive by more than 3 kcal/mol. In the more flexible acyclic systems, however, the energetic penalty for +g/–g over +g/+g arrays is only 1–2 kcal/mol. Thus elimination of structures having two or more +g/–g pentane arrays would be expected to simplify the problem by eliminating conformers expected to fall outside our energetic upper bound (3 kcal/mol). Restricting the number of +g/–g arrays did indeed reduce the number of starting geometries. Allowing no +g/–g arrays in the starting geometries gave 405 structures leading to 58 unique minima after geometry optimization within 3 kcal/mol of the lowest energy conformer. Allowing no more than one +g/–g array, the search gave 7265 structures yielding ~200 unique minima within the 3 kcal/mol energetic bound. Allowing two or more +g/–g arrays in the starting geometries produced too many structures to minimize in a reasonable length of time. Thus, while removing starting structures having multiple +g/–g torsion angles simplifies the search, it was not helpful here because few of the starting geometries had the energetically disallowed two or more +g/–g torsional arrays.

Another way to control the number of structures generated is to vary the ring closure distance constraint. The more lenient this constraint, the more structures one will generate. However, as the length of the ring closure bond increases, starting geometries become more likely to undergo major conformational changes as the minimization algorithm rapidly adjusts the closure distance to that of a normal bond. Enlarging the ring closure distances in starting geometries thus makes the search more random. By enlarging the maximum closure distance from 5 to 7.5 Å, we generated ~2500 new starting geometries which led to many previously found conformers plus 11 new ones. A total of 211 unique conformers were found for all the Still/Lipton tree-searches combined. Over the course of 30 CPU days, a total of 10 170 starting geometries were minimized to yield 9, 63, and 211 conformers within 1, 2, and 3 kcal/mol, respectively, of the global minimum (see Table I, method 3).

The above attempts to favor the generation of low-energy structures at the minimal 120° dihedral angle resolution are compared with the other methods in Tables I and II. At this resolution, the tree-searches did not find the 262 (0–3 kcal/mol) known minima of cycloheptadecane. Additional minima could be found by using nonstaggered torsion angles or by relaxing constraints to generate additional starting geometries. However, it is clear that our 120° resolution tree-search is inadequate for large conformational problems and would appear to have little advantage over the more random methods described herein.

The Houk/Wu tree-search method is similar to the Still/Lipton search. Starting conformers are generated selectively by using geometrical restrictions, such as allowable dihedral angles, closure distance, and transannular distances. A ring structure is constructed sequentially by adding groups to a growing chain. The groups in the present case are methylenes but could also be *cis* or *trans* double bonds, triple bonds, carbonyls, phenyls, biphenyls, etc. After each addition of a group in all allowable conformations, a geometry check was carried out to eliminate redundant structures and any geometries having short transannular distances and long ring closure distances. The transannular distance constraint used was 2.7 Å which avoids generation of +60°/–60° (+g/–g) pentane local structures. As discussed later, the +60°/–90° type pentane substructure is common in medium ring molecules and also in the cycloheptadecane conformers. In order to construct geometries having this type of local structure, 90° dihedral angles were included for some separated bonds. A key difference between this search and the ones described above is the use of $\pm 90^\circ$ torsion angles. In general, +60°/–90° type local structures do not occur contiguously. Therefore, +90° and –90° dihedral angles were allowed only when separated by two bonds. These angles were used for the bonds after the fourth atom of the chain, because more conformational space could be covered. For the addition

(8) Lipton, M.; Still, W. C. *J. Comput. Chem.* **1988**, *9*, 343. For related methods, see: Bruccoleri, R.; Karplus, M. *Biopolymers*, **1987**, *27*, 137. Dammkoehler, R. A.; Karasek, S. F.; Shands, E. F. B.; Marshall, G. R. *J. Computer-Aided Molec. Design* **1989**, *3*, 3.

(9) BATCHMIN is the noninteractive modeling program used with the MACROMODEL interactive molecular modeling program: Still, W. C.; Mohamadi, F.; Richards, N. J. G.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T. MACROMODEL V2.0, Department of Chemistry, Columbia University, New York, NY 10027.

(10) 60° dihedral angle resolution defines possible values of each torsion as 0°, 60°, 120°, 180°, 240°, and 300°. A ring closure distance of 0.5–5.0 Å rejects any starting geometry having the ends of the cleaved ring (the ring closure atoms) closer than 0.5 Å or more widely separated than 5.0 Å. A 1.0 Å nonbonded exclusion distance constraint rejects any starting geometry having any pair of atoms (except the ring closure atoms) <1.0 Å.

of the last atom of the chain, both staggered and eclipsed dihedral angles were allowed. Otherwise, only staggered bonds were permitted. The maximum ring closure distance used was 2.5 Å. In our first run, we generated 1330 starting geometries. After MM2 refinement of the structures, we found 7, 36, and 114 conformers within 1, 2, and 3 kcal/mol, respectively, of the global minimum. When 1794 starting geometries were minimized over ~9 CPU days, 7, 42, and 138 conformers within the previous energetic bounds were found (see Table I, method 2).

There is one final point we wish to make about tree-search methods (2 and 3). As we have implemented them, our methods operate as if all atoms in the structure were distinguishable. While some elements of starting structure symmetry can be incorporated into the search by building ring conformations starting with one or more gauche torsion angles, taking full advantage of symmetry is not easy with the tree-search methods. Our tree-searches therefore provide redundant coverage of many regions of the conformational space of high symmetry structures such as cycloheptadecane. In contrast, the random methods implicitly make allowances for symmetry in that they can be allowed to run until unique (by all appropriate symmetry operations) structures cease to be found. We do not consider this to be a significant disadvantage of the tree-search approach. Most structures of interest are not fully symmetrical.

Internal Coordinate Monte Carlo Search

The Monte Carlo searches of torsion space operate much like the internal coordinate tree-search methods except that torsional arrays and angular variations are chosen randomly and the structure generation and minimization are done as a single operation by the BATCHMIN V2.1 software. The internal coordinate Monte Carlo search has been recently described,¹¹ but we summarize the operations used for cycloheptadecane as follows.

We began by building an arbitrary conformer of cycloheptadecane by using interactive molecular mechanics and defining a ring closure bond as described above. We next assigned a 0.5–3.5 Å ring closure distance constraint to allow selection of starting geometries having the two ends of the cleaved ring close to one another. An iterative conformational search was then initiated. BATCHMIN first selected a random subset of the rotatable torsions and a random rotational increment between 0° and 360° for each. After applying the random torsional rotations, the structure was tested for ring closure and high-energy nonbonded contacts. If the structure had no pair of atoms separated by less than one fourth of the sum of their van der Waals radii, it was energy minimized to a gradient of 10⁻² kcal/Å·mol by using the MM2 force field. A conformer thus produced was kept only if its energy was within 3 kcal/mol of the instant global minimum¹² and did not duplicate any previously stored conformer. By choosing an initial geometry from among those previously stored, the next search cycle began. We monitored the progress of the search by periodically examining the job log file in which was recorded the disposition of each structure minimized. The time for each full Monte Carlo search cycle on a MicroVAX II computer averaged 5.1 min.

The variables which have the most significant effects on the outcome of the search are the number of torsion angles rotated and the choice of the starting geometry used in each step of the search. In previous studies we found the efficiency of the search to be dependent upon the number of torsion angles varied in each step. Although the general case of an unsymmetrical structure is best searched by variation of up to all possible rotatable torsion angles, the simpler conformational space of symmetrical structures like cycloheptadecane is more effectively explored by using only a fraction of the available torsions.¹¹ In the cycloheptadecane searches, we varied 3–5 randomly selected torsion angles at each Monte Carlo step.

In previous studies, we also found that low-energy final conformers are favored by selecting low-energy starting geometries for each Monte Carlo step. The argument for such a heuristic is that, with large molecules at least, low-energy full structures are composed of low-energy substructures, and a limited number of random variations of internal coordinates leaves many of the energetically favorable substructures intact. Thus in the searches below, we allowed only structures within 3 kcal/mol of the instant global minimum to serve as starting geometries for Monte Carlo search steps.

The two runs reported in Table I (methods 4 and 5 (run A)) correspond to different methods of selecting a starting structure for a Monte Carlo search step. In the entry labeled random walk, a step beginning with structure N produced structure N+1. If the energy of structure N+1 was within the specified energetic range of 3.0 kcal/mol of the instant ground state, then structure N+1 was used as the starting geometry for the next Monte Carlo step; otherwise, structure N was reused. In such a search, some structures are found more frequently than others and are thus used more often as starting geometries. Conformers encountered infrequently are rarely used as starting geometries, and any conformers deriving selectively from them are systematically disfavored. A solution to this problem is the usage-directed search.¹¹

In the entries of Table I labeled usage-directed, we employed a usage criterion to choose a starting structure from those previously stored instead of simply accepting the most recent structure as described for the random walk method. In this search, we use each conformer within 3 kcal/mol of the global minimum an equal number of times as a starting geometry. Thus, after a Monte Carlo step beginning with structure N and yielding structure N+1, we select the next starting geometry by finding the subset (M) of the N+1 saved unique structures having energy within 3 kcal/mol of the instant global minimum. Within subset M, we then select the structure which has been used the least as the next starting geometry. It can be seen from the data in Table I that the usage-directed search is more efficient at finding distinct low-energy minima than is the random walk. In the 30-day usage-directed run (Table I, method 5, run A) on a MicroVAX II, we found all 69 of the known conformers within 2 kcal/mol of the global minimum. In a separate 3-day usage-directed run (Table I, method 5, run B) on a Convex C210 computer, we found 260 out of 262 known conformers within 3 kcal/mol.

Distance Geometry Searches

To evaluate the distance geometry method, we used the DGEOM program written by Jeff Blaney and Gordon Crippen to generate starting geometries of cycloheptadecane. The distance bounds were set to the default values by the program. We used the "Eclipsed" option so that the lower distance bound for 1,4-interactions allowed 0° torsion angles. We generated ~3000 random starting geometries in two different ways. First, we used DGEOM to solve for the carbon backbone positions only to yield 3080 random starting geometries. Second, we used DGEOM to solve for the positions of both carbons and hydrogens. By using this method, 3500 random starting geometries were produced. This number of starting geometries corresponds approximately to the number of structures which can be energy minimized in ~12 CPU days on a MicroVAX II computer. For the 30 CPU day run, we generated 8000 starting geometries by using DGEOM to solve for both carbons and hydrogens. As with the other methods, the time required to generate the starting geometries was insignificant (~4% of the total time) in comparison to that necessary for energy minimization.

The starting geometries were then energy minimized by using the BATCHMIN program in the same way that we minimized structures in the internal coordinate Monte Carlo searches described above. In the case of the backbone-only starting geometries, hydrogens were added by using analytical geometry with standard bond lengths and angles prior to energy minimization. The results of the two 12 and one 30 CPU day runs are given in Table I as the entries labeled method 6. None of the distance geometry searches found as many conformers as the other methods

(11) Chang, G.; Guida, W. C.; Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 4379.

(12) We define the lowest energy structure found during any incomplete conformational search as the "instant global minimum" to distinguish it from the final global minimum energy structure.

examined for the same number of starting geometries minimized. One of the 12-day distance geometry runs missed the true global minimum (+19.06 kcal/mol). Instead, the lowest energy structure it found had steric energy +19.83 kcal/mol. The +19.83 kcal/mol structure was also found by the other search methods, but it is higher in energy than the seven lowest energy conformers. The other 12 and the 30 CPU day distance geometry runs found the same global minimum (+19.06 kcal/mol) as the other methods. Although the results with generating carbons and hydrogens are better in our studies, we note that these are results from single runs and any general advantages of generating both hydrogen and heavy atom positions simultaneously have yet to be established.

Molecular Dynamics Searches

Our molecular dynamics searches were carried out starting with an arbitrarily selected conformer of cycloheptadecane which was ~1.5 kcal/mol above the global minimum. To favor the rapid exploration of conformational space, we carried out our simulations at relatively high temperatures. Temperature was controlled by a thermal bath¹³ with an 0.2 ps time constant, and all CH bond lengths were constrained by SHAKE.¹⁴ The MM2 force field was used, and the simulation was carried out by using BATCHMIN. One simulation was conducted at 500 K. At that temperature, we could use a timestep of 1.25 fs and cover 1000 ps in 14 CPU days on a MicroVAX II. A second 14-day simulation was carried out at 1000 K and ran for 250 ps by using a timestep of 0.33 fs. Four thousand structures were sampled uniformly in time throughout each simulation. These structures served as starting geometries for energy minimization and duplicate elimination. The two simulations/minimizations required ~30 CPU days each, and the results are listed as method 7 in Table I.

As the data in Table I show, molecular dynamics as described above did find many conformations of cycloheptadecane, but it was less effective at finding all conformations than any other method studied. On examining the disposition of each energy-minimized structure from the 500 K simulation, it could be seen that many conformers had lifetimes in the 1–3 ps range, and several persisted for as long as 10 ps. Thus the only way to cover more of the conformational space would be to run the simulation for a longer time or at a higher temperature. In the 1000 K simulation, the lifetime of conformers was reduced, but energy minimization of sampled structures gave a larger proportion high-energy conformers. Overall, the run at 500 K was somewhat better than the one at 1000 K which found fewer low-energy conformers and even missed the global minimum.

Efficiency of Structure Generation, Energy Minimization, and Duplicate Elimination

Although the best way to compare the various methods of searching is to contrast results for a given number of starting geometries minimized, the computer time necessary to generate the starting structures, to minimize their energies, and to search for duplicates is an important practical issue. As for starting structure generation, the various methods have been described above, and we should add only that the time by any of the methods other than molecular dynamics is insignificant in comparison with the time necessary for energy minimization. Thus with cycloheptadecane and a MicroVAX II computer, the stochastic Cartesian method generates roughly 10³ structures/CPU min, and the slower internal coordinate torsional methods generates approximately 10² structures/CPU min, while energy minimization of a single structure takes 5–15 min. Energy minimization time aside, the efficiencies of the various search methods at finding low-energy minima starting from a fixed number of starting geometries is summarized in Table II. The information there indicates that all methods included proceed with similar efficiency during the early stages of the search.

Table II. Rate of Conformational Search

method used	percentage of total minima found/100 starting geometries		
	1 kcal/mol ^a	2 kcal/mol ^b	3 kcal/mol ^c
1b ^d	8.3	6.7	4.5
2 ^e	6.1	6.3	5.4
3 ^f	6.1	7.5	5.5
5 ^g	9.1	6.4	4.4

^a Assuming 11 minima total. ^b Assuming 69 minima total. ^c Assuming 262 minima total. ^d Based on a single run of 1194 stochastic steps using the improved method. ^e Based on single run of 598 starting geometries. ^f Based on single run of 405 starting geometries. ^g Average of five 100-step runs using different starting points for random number generation and an initial structure which was 10 kcal/mol above the global minimum.

Table III

method	hardware	software	CPU min/structure
1a	MicroVAX II	STRFIT+MM2	~15
1b	MicroVAX II	new STRFIT+MM2	10–11
2	MicroVAX II	MM2	11
3, 4, 5	MicroVAX II	BATCHMIN	5.1
6	MicroVAX II	DGEOM+BATCHMIN	5.5
7	MicroVAX II	BATCHMIN	11
5	Convex C210	BATCHMIN	0.2

Duplicate detection and elimination are also relatively rapid operations. For cycloheptadecane, the time to eliminate duplicates was similar to that necessary to generate the initial structures and corresponds to <1% of the total search time. In the work described above, Houk and Wu used a comparison of ordered dihedral angles, Saunders used a comparison of selected nonbonded distances, and Still, Lipton, Chang, and Guida used a least-squares superimposition¹⁵ of Cartesian coordinates. All methods speeded the structural comparisons by testing for duplication only when the energies of potential duplicates were similar. By using any of these methods, comparisons are straightforward provided the structures being tested were adequately refined by the energy minimization process. For duplicate structures which have converged to the MM2 criterion (energy stabilization using block diagonal Newton Raphson optimization), dihedral angles should match within 10° (Houk/Wu), nonbonded distances should match to within 0.5 Å (Saunders), and the least-squares superimposition should give no atomic superimposition error greater than 0.25 Å (Still/Lipton/Chang/Guida). As a final test, Saunders retested all conformers for duplication by a comparison of torsion angles. Because of the symmetry of cycloheptadecane, all methods tested duplications with all possible numbering systems by permuting the atom-numbering systems around the ring in both directions. The Houk/Wu method, the final Saunders dihedral test, and Still/Lipton/Chang/Guida methods required an additional test to detect enantiomeric conformations involving changing the sign of the torsion angles or atomic z-coordinates.

Excepting molecular dynamics, the most time-consuming part of the searches was the energy minimization of the various starting geometries to minimum energy conformers. Depending on hardware and software used, the energy minimization time varied substantially as summarized in Table III. The CPU times above are only approximately comparable since different energy minimization schemes were used with the different search methods and since the same degree of convergence was not achieved by all methods. Saunders' Cartesian method used a double minimization scheme in which a rapid approximate force field (STRFIT) was used prior to final minimization with the standard MM2 field. Many of the duplicates could be eliminated prior to minimization with the relatively slow MM2 program. The other methods also included schemes to avoid time-consuming minimization to full convergence of structures destined to become duplicates or

(13) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.

(14) Ryckaert, J.-P.; Ciccoliti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327. van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *51*, 1243. Ryckaert, J. P. *Mol. Phys.* **1985**, *55*, 549.

(15) Kabsh, W. *Acta Crystallogr.* **1976**, *A32*, 944. Kabsh, W. *Acta Crystallogr.* **1978**, *A34*, 837.

high-energy conformers. Both the Houk/Wu and the Still/Lipton tree-search methods accomplished this goal by conducting the energy minimization in several passes. After the first 100 iteration minimization pass, for example, some duplicates and many high-energy structures (>5 kcal/mol of the ground state) could be eliminated from further processing. A similar scheme was used during the Still/Chang/Guida Monte Carlo searches by having the energy minimization program BATCHMIN fully minimize only those structures apparently converging to conformers within the 3 kcal/mol energetic upper bound. This selection was made by terminating the minimization of any structure which was more than 3.3 kcal/mol above the instant ground state after the first 75 iterations of the minimization. These simple tests were quite effective at improving overall search speed and are responsible for approximately doubling the speed of the Houk/Wu and Still/Lipton/Chang/Guida cycloheptadecane searches.

With structures the size of cycloheptadecane or larger, convergence to a stable final geometry can be a problem with MM2. Thus apparently distinct conformers (according to the duplicate elimination criteria above) found during the search occasionally collapsed to a single minimum upon further minimization with a more exacting convergence criterion. By using the standard convergence criterion, the MM2 program terminates minimization when the energy stabilizes to yield final structures having often rms gradients > 1 kcal/Å, whereas gradients of 10^{-3} kcal/Å or less should be attainable at the energetic minimum by using single precision arithmetic with structures of this size.

Convergence with BATCHMIN was not a problem. With the MM2 force field and a 20% successive over-relaxation block-diagonal Newton-Raphson algorithm, convergence to a final rms gradient of 10^{-2} kcal/Å-mol was complete in 200 or fewer iterations in approximately 98% of the cycloheptadecane structures minimized. Complete convergence to a gradient of 10^{-4} kcal/Å-mol was achieved by using a final 1-5 iterations of full-matrix Newton-Raphson minimization. The question of saddle point structures was answered unambiguously by examining the second derivative matrix for negative eigenvalues. Structures characterized by one or more negative eigenvalues correspond to an energetic maximum. Approximately 0.5% of the original structures produced were shown by the eigenvalue test to be energetic saddle points. Without the tests to verify a low gradient ($<10^{-3}$ kcal/Å-mol) and all positive eigenvalues, it was not possible to distinguish true minima from poorly converged or saddle point structures. The 262 structures reported herein were all true minima by these tests.

Finally, we should note that dramatic enhancements in searching speed are possible with faster hardware. Using BATCHMIN v2.1 with a vectorized nonbonded derivative routine running on a CONVEX C210 computer, a 24-fold speed enhancement over a MicroVAX II was measured for the usage-directed Monte Carlo search. Thus, any of the Houk/Wu or Still/Lipton/Chang/Guida runs reported in this paper can be accomplished in just over 1 day with an existing hardware/software system. By using a contemporary supercomputer, a highly optimized energy minimization routine, and the search methods described here, it appears that >98% of the conformers of cycloheptadecane within 3 kcal/mol of the ground state could be found within a single day.

Conformations of Cycloheptadecane

Combining the results of all our searches, we found 11, 69, and 262 conformations of cycloheptadecane within 1, 2, and 3 kcal/mol, respectively, of the ground state by using the MM2 force field without distance limits for nonbonded interactions. All of these conformations were shown to be true minima by calculation of their vibrational frequencies after a final full-matrix Newton-Raphson energy minimization. While we cannot be sure that all conformations within the 3 kcal/mol energetic bounds were discovered, we note that each of the 262 conformations reported here was found by three or more independent searches by using different search methods. We believe that this finding provides convincing evidence that the vast majority if not all of the min-

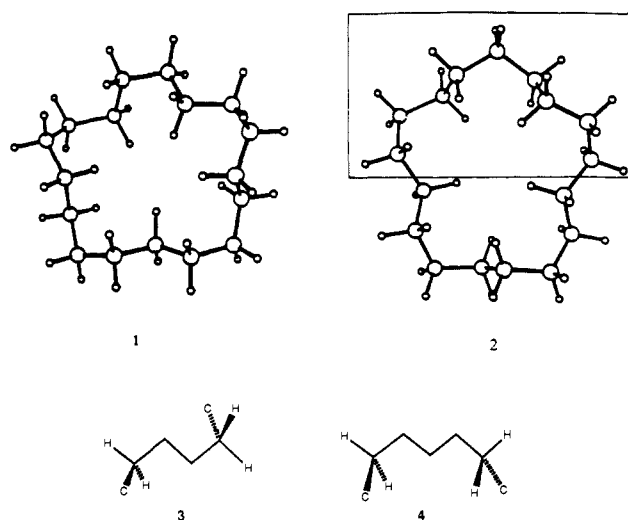


Figure 1.

Table IV. Distributions of Dihedral Angles of Cycloheptadecane Conformers

dihedral range	within 1 kcal/mol	within 2 kcal/mol	within 3 kcal/mol
0-40	0	0	0
41-50	0	16	106
51-60	46	247	812
61-70	35	198	701
71-80	16	104	402
81-90	0	16	91
91-100	4	42	184
101-110	0	5	33
111-120	0	1	2
121-130	0	2	16
131-140	1	2	5
141-150	0	4	23
151-160	4	19	104
161-170	30	169	615
171-180	51	348	1326

imum energy conformers have been located.

Our conformational searches show that cycloheptadecane, like other flexible structures, is a molecule whose properties are not primarily dependent upon its global minimum energy conformer. Indeed, the Boltzmann population of the most stable conformation would be only 8% at 300 K if the entropies of all the conformers were the same. The second conformer with a symmetry number of 2 would comprise 4% of the conformational mixture. Conformers within 0-1 and 0-2 kcal/mol account for 34% and 67%, respectively, of the total population.

Concerning structure, odd-membered hydrocarbon ring structures can have one 2-fold axis of symmetry. Although a plane of symmetry bisecting an atom and a bond is theoretically possible, no structure of this symmetry was found. This is not surprising, since such a structure would require one torsion angle to equal the highest energy *n*-butane torsion angle of 0° . By comparing dihedral angles, it was possible to examine all structures for approximate 2-fold symmetry axes. Seventeen of the structures were found to be nearly symmetrical by this test.

Figure 1 shows the two lowest energy conformers found in this study. The global minimum, 1, which has a steric energy of +19.06 kcal/mol, could be assigned as [133433] following Dale's nomenclature.¹⁶ It does not possess symmetry but could be converted to C_1 symmetry if the dihedral angle about the top bond in the picture is compressed to 0° . The second best conformer, 2, has almost identical steric energy (+19.08 kcal/mol) to that of the global minimum and has nearly C_2 symmetry. It is designated as [34343]. Torsion angles and MM2 energies for all 262 minima and stereo pair plots for the 11 lowest energy conformers

(16) Dale, J. *Acta Chem. Scand.* 1973, 27, 1115.

Table V. Distributions of Three Contiguous Dihedral Angles in Cycloheptadecane Conformers

dihedral pattern	within 1 kcal/mol	within 2 kcal/mol	within 3 kcal/mol
a a a	2	17	96
a a g	60	376	1478
a g a	12	93	418
g a g	24	150	516
g a g-	0	1	15
a g g	77	410	1326
a g g-	7	82	378
g g g	4	26	90
g g g-	1	18	132
g g- g	0	0	5

are provided as Supplementary Material.

Table IV summarizes the distributions of dihedral angles of cycloheptadecane conformers. Several features are worthy of note. (1) Dihedral angles below 40° are not found in any conformer within 3 kcal/mol of the global minimum. For *n*-butane, ab initio and MM2 indicate that the conformer with a 40° dihedral angle is about 1 kcal/mol less stable than the one with a 60° dihedral angle.¹⁷ (2) A dihedral angle in the range of 101°–150° occurs in only one of the conformers within 1 kcal/mol but becomes more common in higher energy conformers. (3) Although the majority of the dihedral angles in the conformers found are approximately gauche or anti, dihedral angles of 90°–100° do occur frequently. They appear in four of the 11 conformers within 1 kcal/mol but are more common in higher energy conformers. All of these are coupled with a normal gauche angle, that is, 65°/–90° or –65°/90°, as found in medium ring and acyclic molecules.^{6,8,18}

To learn more about the conformational features of the conformers, we determined the frequency of occurrence of local structures composed of three and four contiguous dihedral angles.¹⁹ For the purpose of the analysis, all of the dihedral angles were divided into three groups: gauche (g), 1°–119°; minus gauche (g-), –1°–(–119°), and anti (a), 121°–(–120°). Table V gives the distributions of three contiguous dihedral angles. There are ten nonequivalent arrangements. Two of them appear only rarely. One is the arrangement of g a g-. It occurs only in one of the conformers within 2 kcal/mol of the ground state at +1.67 kcal/mol. As shown by 3, the two terminal carbons in this fragment are on different sides of the approximate plane defined by the four central atoms. A long chain is required to connect the two termini without introducing significant distortion. Thus this arrangement does not occur in rings smaller than cyclopentadecane (conclusion from dihedral angles for conformers of cycloheptane–cyclohexadecane within 2 kcal/mol of global minimum), but it becomes possible when the ring becomes larger.

Another high-energy fragment is g g-. This is the perfect arrangement in the chair conformation of cyclohexane; however, it causes significant steric destabilization when the termini are not bonded. For example, the g g- conformer of *n*-hexane is 5.5 kcal/mol less stable than the all anti conformer according to MM2 calculations.

Table VI summarizes the distributions of arrangements of four contiguous dihedral angles. There are 25 nonequivalent arrangements. These structures incorporate several interesting features. (1) Although the arrangement of three contiguous gauche dihedral angles (g g g) is relatively low in energy, the quadruple gauche arrangement g g g g was not found in any of the conformers of cycloheptadecane. This arrangement corresponds to helical local geometry. Elongation of the chain from the two termini in anti fashion, that is a g g g g a, causes the two

Table VI. Distributions of Four Contiguous Dihedral Angles in Cycloheptadecane Conformers

dihedral pattern	within 1 kcal/mol	within 2 kcal/mol	within 3 kcal/mol
a a a a	0	2	18
a a a g	4	30	156
a a g a	20	123	538
a a g g	35	186	646
a a g g-	5	67	294
a g a g	4	62	283
a g a g-	0	1	15
a a g g-	34	170	513
a g g g	8	52	174
a g g g-	1	18	126
a g g- a	3	34	146
a g g- g	0	0	10
a g- g g	1	14	76
g a a g	2	32	152
g a a g-	26	141	509
g a g g	42	223	665
g a g g-	2	15	84
g a g- g	0	0	0
g- a g g	0	1	15
g g g g	0	0	0
g g g g-	0	0	6
g g g- g	0	0	0
g g g- g-	0	2	28
g g- g g-	0	0	0
g g- g- g	0	0	0

termini to separate to such an extent that the ring cannot be closed. If the chain elongation occurs in gauche fashion, a further gauche interaction is introduced. (2) Except for the arrangement of g g g-, which occurs 2 and 28 times, respectively, in the conformers within 2 and 3 kcal/mol of the global minimum, all other arrangements with four contiguous gauche dihedral angles rarely occur. (3) Similar to the arrangement of g a g-, the arrangement of g a a g also has the two terminal carbon atoms on different sides of the central approximate plane, as shown by 4. While the former appears very rarely, the latter occurs more frequently. This is presumably due to the fact that in the former the two termini are pointed in opposite directions, while in the latter they point in the same direction.

Conclusion

As the data in Table I show, the effectiveness with which the cycloheptadecane conformational search problem is solved varies significantly with the search method used. While all of the methods were able to locate the global minimum, no single run found all the minima within 3 kcal/mol though some of the runs came close to doing so. Two of the most popular methods for conformational space searching, distance geometry and molecular dynamics, performed relatively poorly and located considerably fewer minima than did the other methods in the same CPU time. Molecular dynamics methods aside, the time for starting geometry generation and structural comparisons is insignificant (<5% of the total) in comparison with the time necessary to minimize the structures. Thus the principal determinants of the time necessary to complete the search are the intrinsic speed of the energy minimization algorithms and the effectiveness with which starting geometries are produced which yield unique, low-energy minima. Some of the differences in the CPU times given in Table I reflect differing efficiencies of the energy minimization algorithms used with the different search methods. If minimization speed is ignored and the results considered on a per starting geometry basis, it is found (Table II) that all of the methods generate low-energy minima at a similar rate, at least early in the search when few duplicates are being produced.

In comparing the utility of the random and systematic approaches to conformational searching, the random methods provide significant advantages. Primary among these is that the random methods lend themselves naturally to implementation as continuous search processes. These search-minimize-search-minimize... procedures can be used to search conformational space as ex-

(17) Jorgensen, W. L.; Buckner, J. K. *J. Phys. Chem.* **1987**, *91*, 6083 and references therein.

(18) Ab initio quantum calculations of *n*-pentane give a minimum quite close to 65°/–90° at the 6-31G* level of theory. Wiberg, K. B. Private communication.

(19) Uiterwijk, J. W. H. M.; Harkema, S.; Waal, B. W. V. D. *J. Chem. Soc., Perkin Trans. 2* **1983**, 1843. Uiterwijk, J. W. H. M.; Harkema, S.; Feil, D. *J. Chem. Soc., Perkin Trans. 2* **1987**, 721.

tensively as necessary to reach (in principle) any desired degree of search convergence. With the current objective, the high symmetry of cycloheptadecane reduces the number of possible conformers substantially. Consequently, the continuous process searches take advantage of the fact that the cycloheptadecane conformational searches converge (cease to find new, distinct conformers) sooner than they would if the molecule being analyzed was less symmetrical. Systematic search methods like the tree-search methods described above make it difficult to take advantage of all symmetry elements and tend to approach all conformational problems as if they were the more complex (but more general) unsymmetrical ones.

We believe that we have found most if not all of the conformers within 3 kcal/mol of the global minimum. Among these low-energy conformers are a significant number having high energy substructures such as +g/-g pentane segments and nearly eclipsed C-C bonds. In defining the set of low-energy conformers, we have taken care to be sure that all structures found by the searches and subsequent energy minimizations represent true minima. The best way to verify structures as true minima was to reminimize structures by using the full-matrix Newton-Raphson technique to gradients <0.001 kcal/Å-mol and then examine the vibrational frequencies for imaginary modes. Without such a test, it was not

possible to distinguish true minima from saddle point structures or unconverged geometries lying in relatively flat areas of the MM2 potential energy surface.

In conclusion, the cycloheptadecane problem, even though simplified by symmetry, appears to be a challenging but solvable problem in conformational analysis. It has served here as a test for several newly developed techniques for searching conformational space. Since this 17-membered ring lies just at the limit of what can be accomplished in a reasonable length of time, it should provide a good test of the effectiveness of new methods as they are developed.²⁰

Registry No. Cycloheptadecane, 295-97-6.

Supplementary Material Available: A listing of all conformations found within 3 kcal/mol of the ground state giving MM2 energies and torsion angles and stereopair plots of all conformers within 1.0 kcal/mol of the global minimum (11 pages). Ordering information is given on any current masthead page.

(20) We thank the following agencies for their support of this work: NIH (Grant GM40544 (K.N.H.)), NSF (Grants CHE 8512785 (K.N.H.), and CHE 8605891 (W.C.S.)). We thank Dr. Jeff Blaney at DuPont for his advice on the use of DGEOM for the distance geometry searches.

Empirical and ab Initio Estimates of the Stabilities of Fluorine-Containing Cations¹

Dion A. Stams,[†] T. Darrah Thomas,[‡] Duncan C. MacLaren,[‡] De Ji,[‡] and Thomas Hellman Morton*[†]

Contribution from the Department of Chemistry, University of California, Riverside, California 92521, and Department of Chemistry, Oregon State University, Corvallis, Oregon 97331. Received December 27, 1988

Abstract: Heats of formation of two classes of monofluorinated cations, fluoronium ions and α -fluoro carbocations, are estimated ab initio and empirically. $C_3H_6F^+$ structures that are isoelectronic to stable C_3H_6O isomers are examined on the SCF potential energy surface (6-31G** basis set). The cations isoelectronic to alcohols all correspond to ion-molecule complexes between $C_3H_5^+$ and hydrogen fluoride where the equilibrium C-F bond distances are calculated to be >2 Å. The cation isoelectronic to propylene oxide is predicted to be unstable with respect to 1-fluoroisopropyl structures with planar skeletons. The remaining cations resemble their isoelectronic neutrals. Empirical estimates are based on core-ionization energies by using the method of equivalent cores. The two methods provide estimates within 15 kJ mol⁻¹ ($= 3.6$ Kcal mol⁻¹ $= 0.16$ eV) of each other in a number of cases so long as appropriate isodesmic reactions are used. For nonisodesmic reactions equivalent-cores approximations can be substantially in error. The agreement between equivalent cores and ab initio estimates are evaluated for the two isodesmic reactions $R_1OR_2 + (CH_3)_2F^+ \rightarrow R_1FR_2^+ + (CH_3)_2O$ for fluoronium ions and $O=CXY + CH_3CHF^+ \rightarrow FCXY^+ + CH_3CHO$ for α -fluoro carbocations. The SCF electronic energy changes, ΔE_1 , when the ions are constrained to the geometries of their isoelectronic neutrals are found, in most cases, to agree with experimental values. By using correlations based on ΔE_1 , core-ionization energies for oxygen-containing molecules are predicted: 539.4 eV for acetone enol, >539.5 eV for the lowest triplet state of acetone, and 538.9 eV for formamide with the nitrogen lone pair twisted out of conjugation with the carbonyl. ΔH_f values computed for ions in their relaxed geometries agree, for the most part, with equivalent cores estimates. Substituent effects are evaluated ab initio by computing the energy barrier for rotation about the C=N bond of $RCH=NH_2^+$: SCF values of ΔH^* for R = H, F, and vinyl at 6-31G** are 261, 190, and 166 kJ mol⁻¹, respectively. For cyclic fluoronium ions, the four-member ring ion fluoretanium is estimated to have $\Delta H_f = 660$ kJ mol⁻¹, within 10 kJ mol⁻¹ of two other $C_3H_6F^+$ isomers, $FCH_2CHCH_3^+$ (**14**) and $CH_3CH_2CHF^+$ (**2**). The five-member ring fluorolanium and the six-member ring fluoranium are estimated to have heats of formation of 555 and 515 kJ mol⁻¹, respectively.

Although fluorine is the most electronegative element, the consequences of substituting a cation with fluorine are not obvious. In recent years we have described gas-phase chemistry of the 2-fluoroisopropyl cation, **1**, which can be easily formed.^{2,3} When it has sufficient internal energy **1** expels hydrogen fluoride in a

simple 1,2-elimination without undergoing any rearrangement. This raises the question as to whether **1** might represent the only

(1) Portions of this work were presented at the 194th National Meeting of the American Chemical Society, Los Angeles, CA, September 1988.

(2) (a) Redman, E. W.; Johri, K. K.; Lee, R. W. K.; Morton, T. H. *J. Am. Chem. Soc.* **1984**, *106*, 4639-4640. (b) Redman, E. W.; Johri, K. K.; Morton, T. H. *J. Am. Chem. Soc.* **1986**, *108*, 780-784.

[†]University of California.

[‡]Oregon State University.