# Low Mode Search. An Efficient, Automated Computational Method for Conformational Analysis: Application to Cyclic and Acyclic Alkanes and Cyclic Peptides

**István Kolossváry*,† and Wayne C. Guida*,‡**

*Contribution from the Department of Chemical Information Technology, Technical University of Budapest, Szt. Gellért tér 4, 1111 Budapest, Hungary, and Research Department, Pharmaceuticals Division, Ciba-Geigy Corporation, Summit, New Jersey 07901*

*Received July 24, 1995. Revised Manuscript Received March 21, 1996*[⊗]

**Abstract:** The location of energy minima on the conformational energy surface of molecules by computational methods (conformational searching) continues to play a key role in computer-assisted molecular modeling. Although a number of conformational search procedures have been devised over the past several years, new more efficient methods are urgently needed if molecules with increased complexity are to be treated in a quantitative manner. In this paper we describe a method, termed low-mode search (LMOD), which is based on eigenvector following (or mode following), for the exhaustive exploration of the potential energy hypersurface of molecules. It is particularly efficient at searching the conformational space of both cyclic and acyclic molecules, and we describe its effectiveness for a number of conformational search problems including acyclic, monocyclic, and bicyclic hydrocarbons and cyclic pentapeptides. No special treatment of rings in cyclic molecules is necessary, nor is it necessary to define rotatable bonds. LMOD generates structures "automatically" with minimum input from the user. We demonstrate that LMOD is one of the most efficient procedures yet devised for conformational searching of small- to medium-sized molecules.

## Introduction

Location of the stationary points, particularly energy minima[1] on the potential energy hypersurface of stable molecules, continues to play a central role in computational chemistry. Over the past several years, a multitude of conformational search methods have been devised for generating the low-energy (and hence relevant) conformers of both cyclic and acyclic molecules, and the practicing computational chemist can now choose from a wide variety of procedures for conformational searching.[2] These methods include procedures that involve the systematic variation of torsion angles,[3−8] the stochastic (sometimes referred to as Monte Carlo) variation of torsion angles[9−12] (including methods that rely on the genetic algorithm[13−15]), methods based on Metropolis Monte Carlo[16,17] (including simulated annealing[18−22]), the stochastic variation of Cartesian coordinates,[23−27] the stochastic variation of internuclear distances (distance geometry),[28−32] methods that employ molecular dynamics,[33,34] and the flapping,[35,36] flipping,[37] or flexing[38] of rings or mapping of rings onto generic shapes.[39]

* Correspondence may be addressed to either author.
† Technical University of Budapest. E-mail: kolos@dumbo.inc.bme.hu.
‡ Ciba-Geigy Corp. E-mail: wcg@ussu.ciba.com.
⊗ Abstract published in *Advance ACS Abstracts,* May 1, 1996.

(1) The importance of locating saddle points, as well as minima, has been described in the literature. See, for example: (a) Anet, F. A. L. *J. Am. Chem. Soc.* **1990**, *112*, 7172. (b) Kolossváry, I.; Guida, W. C. *J. Am. Chem. Soc.* **1993**, *115*, 2107.
(2) For reviews, see: (a) Howard, A. E.; Kollman, P. A. *J. Med. Chem.* **1988**, *31*, 1669. (b) Leech, A. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH: New York, 1991; Vol. 2, pp 1−47.
(3) White, D. N. J.; Kitson, D. H. *J. Mol. Graphics* **1986**, *4*, 112.
(4) Smith, G. M.; Veber, D. F. *Biochem. Biophys. Res. Commun.* **1986**, *134*, 907.
(5) Fukazawa, Y.; Usui, S.; Uchio, Y.; Shiobara, Y.; Kodama, M. *Tetrahedron Lett.* **1986**, *27*, 1825.
(6) Lipton, M.; Still, W. C. *J. Comput. Chem.* **1988**, *9*, 343.
(7) Dammkoehler, R. A.; Karasek, S. F.; Shands, E. F. B.; Marshall, G. R. *J. Comput.-Aided Mol. Des.* **1989**, *3*, 3.
(8) Goodman, J. M.; Still, W. C. *J. Comput. Chem.* **1991**, *12*, 1110.
(9) Fine, R. M.; Wang, H.; Shenkin, P. S.; Yarmush, D. L.; Levinthal, C. *Proteins: Struct., Funct. Genet.* **1986**, *1*, 342.
(10) Shenkin, P. S.; Yarmush, D. L.; Fine, R. M.; Wang, H.; Levinthal, C. *Biopolymers* **1987**, *26*, 2053.
(11) Chang, G.; Guida, W. C.; Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 4379.
(12) Weinberg, N.; Wolfe, S. *J. Am. Chem. Soc.* **1994**, *116*, 9860.
(13) Blommers, M. J. J.; Lucasius, C. B.; Kateman, G.; Kaptein, R. *Biopolymers* **1992**, *32*, 45.
(14) McGarrah, D. B.; Judson, R. S. *J. Comput. Chem.* **1993**, *14*, 1385.
(15) Judson, R. S.; Jaeger, E. P.; Treasurywala, A. M.; Peterson, M. L. *J. Comput. Chem.* **1993**, *14*, 1407.
(16) Li, Z.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 6611.
(17) von Freyberg, B.; Braun, W. *J. Comput. Chem.* **1991**, *12*, 1065.
(18) Wilson, S. R.; Cui, W. *Tetrahedron Lett.* **1988**, *29*, 4373.
(19) Wilson, S. R.; Cui, W. *Biopolymers* **1990**, *29*, 225.
(20) Wilson, S. R.; Cui, W.; Moskowitz, J. W.; Schmidt, K. E. *J. Comput. Chem.* **1991**, *12*, 342.
(21) Guarnieri, F.; Cui, W.; Wilson, S. R. *J. Chem. Soc., Chem. Commun.* **1991**, 1542.
(22) Guarnieri, F.; Wilson, S. R. *Tetrahedron* **1992**, *48*, 4271.
(23) Saunders, M. *J. Am. Chem. Soc.* **1987**, *109*, 3150.
(24) Saunders, M. *J. Comput. Chem.* **1989**, *10*, 203.
(25) Saunders, M. *J. Comput. Chem.* **1991**, *12*, 645.
(26) Ferguson, D. M.; Raber, D. J. *J. Am. Chem. Soc.* **1989**, *111*, 4371.
(27) Ferguson, D. M.; Glauser, W. A.; Raber, D. J. *J. Comput. Chem.* **1989**, *10*, 903.
(28) Weiner, P. K.; Profeta, S.; Wipff, G.; Havel, T.; Kuntz, I. D.; Langridge, R.; Kollman, P. A. *Tetrahedron* **1983**, *39*, 1113.
(29) Crippen, G. M.; Havel, T. F. *Distance Geometry and Molecular Conformation*; John Wiley: New York, 1988.
(30) Havel, T. F. *Prog. Biophys. Mol. Biol.* **1991**, *56*, 43.
(31) Peishoff, C. E.; Dixon, J. S. *J. Comput. Chem.* **1992**, *13*, 565.
(32) Crippen, G. M. *J. Comput. Chem.* **1992**, *13*, 351.
(33) Sun, Y.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 33.
(34) Byrne, D.; Li, J.; Platt, E.; Robson, B.; Weiner, P. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 67.
(35) Goto, H.; Osawa, E. *J. Am. Chem. Soc.* **1989**, *111*, 8950.
(36) Goto, H.; Osawa, E. *Tetrahedron Lett.* **1992**, *33*, 1343.
(37) Goto, H.; Osawa, E. *J. Chem. Soc., Perkin Trans. 2* **1993**, 187.
(38) Kolossváry, I.; Guida, W. C. *J. Comput. Chem.* **1993**, *14*, 691.
(39) Gerber, P. R.; Gubernator, K.; Müller, K. *Helv. Chim. Acta* **1988**, *71*, 1429.

S0002-7863(95)02478-4 CCC: $12.00 

In spite of the substantial progress that has been made thus far, conformational searching of molecules containing greater than approximately 12 rotatable bonds still presents a seemingly impregnable challenge. Yet, the ability to accurately enumerate all of the relevant (e.g., low-energy) conformations of a molecule is one of the crucial operations that one must perform whenever quantitative molecular modeling is to be accomplished. Moreover, conformational searching is critical to the interpretation and understanding of experimental results (e.g., NMR spectroscopic data[40−45]). Thus, conformational search procedures with increased efficiency are urgently needed if computer-assisted molecular modeling is to be of utility for complex molecular or macromolecular systems.[46] In this paper, we describe a novel method for conformational searching that can be applied with equal facility to cyclic and acyclic molecules. We demonstrate that our method is significantly more efficient at locating low-energy conformers than the best methods currently available.

Our approach relies upon the fact that a number of eigenvector-following techniques have been devised for locating saddle points on molecular potential energy surfaces.[47−51] These methods operate by initiating the saddle point search at or near a local minimum. Diagonalization of the Hessian matrix affords its eigenvalues and the associated eigenvectors (the so-called "normal modes" of vibration). One of these (generally low-frequency) modes is selected and followed uphill (i.e., in a direction that leads to an increase in potential energy). Since the structural vibrational modes carrying a molecule from one conformation to another are typically the low-frequency modes, the fundamental concept behind eigenvector-following, or mode-following as it is frequently called, is to start from a minimum and maximize the energy along a selected low-frequency mode ($\leq$250 cm$^{-1}$) while minimizing the energy along the remaining modes. The resulting, so-called minimum-energy path leads to a saddle point on *well-behaved* potential energy surfaces. The potential energy gradient and the Hessian matrix are evaluated at every step until the gradient vanishes and the Hessian matrix possesses one and only one negative eigenvalue, at which point the saddle point has been located.

Although the efficiency and reliability of the eigenvector-following procedures are matters of debate, once a saddle point emanating from a local minimum has been found, a second minimum, which is connected to the first, can be located as well. This task can be accomplished by continued movement from the saddle point along the potential energy hypersurface by small increments further away from the first minimum, in the direction of negative canonical curvature. The movement along this direction lowers the potential energy, and subsequent energy minimization *generally* affords the other minimum.[1b]

Since the potential energy hypersurface is a network of *interconnected* minima and saddle points, we reasoned that one could utilize a procedure that relies on eigenvector following for conformational searching. Thus, one could initiate the search by starting with any local minimum. By using one of the eigenvector-following techniques, one could locate a saddle point associated with this minimum and then the other minimum associated with this saddle point. By application of the eigenvector-following technique to the second minimum or to a different eigenvector of the first minimum, additional minima could be located which could then be used to find additional saddles, etc. Although such a technique could work in principle, there are at least two reasons why it might be inefficient as a conformational search procedure. First, it is known that the eigenvector-following techniques frequently fail to converge on saddle points. Second, the need to evaluate both the gradient and Hessian, and also to diagonalize the Hessian, at each step would render such a procedure hopelessly slow.

In this paper we describe a method that is based "in spirit" on the mode-following technique described above but overcomes the inefficiencies that would arise from precise implementation of mode following as a conformational search procedure. We have termed this new conformational search method low-mode search (LMOD), and we demonstrate that it is remarkably efficient at locating the low-energy minima of a number of cyclic as well as acyclic molecules.

## Computational Methodology

**Algorithmic Details.** The basic tenet of our LMOD procedure is to utilize a "brute force" approach based on the mode-following concept. It operates as follows: an initial arbitrary minimum-energy conformer is subjected to normal mode analysis (as described above), and the low-frequency modes are stored as an array of eigenvectors of the *non*-mass-weighted Hessian matrix. The number of low-frequency modes considered is determined by a user-defined frequency threshold (typically 250 cm$^{-1}$). LMOD searches the low-frequency modes systematically along the corresponding eigenvectors which are searched in both directions. The initial structure is continuously perturbed along one of its low-mode eigenvectors in discrete steps until the increase in potential energy exceeds a user-defined threshold during a single step. On very rare occasions, we have found that during these steps the energy first increased but then decreased, indicative of a nearly quadratic potential energy hypersurface where LMOD actually crossed the barrier at or near a saddle point. The resulting perturbed initial structure is subsequently subjected to energy minimization. Although there is no guarantee that this subsequent energy minimization will in fact cross a potential energy barrier, in our experience it crosses barriers most of the time. Of course, it can happen that energy minimization will carry the starting structure to a minimum that is not connected to the minimum used to initiate the mode-following procedure, but essentially, LMOD typically focuses the search to the local neighborhood of a minimum on the potential energy hypersurface. Note that in LMOD the eigenvector search direction is never updated, and hence *no reevaluation of the Hessian is performed*.

As the search progresses, an ensemble of conformers is collected, and these are used as starting structures for structural perturbation along their low-frequency modes. Furthermore, each of the low-frequency modes associated with each structure is employed for structural perturbation. The new minima found during an LMOD search become new focal points, and thus LMOD necessarily explores the entire potential energy surface (although inclusion of higher frequency modes may be necessary to locate high-energy conformers).

LMOD, as we have described it above, is a systematic search procedure, but as such, it is bounded by the number of low-frequency

(40) Ghose, A. K.; Sanghvi, Y. S.; Larson, S. B.; Revankar, G. R.; Robins, R. K. *J. Am. Chem. Soc.* **1990**, *112*, 3622.
(41) Nikiforovich, G. V.; Prakash, O.; Gehrig, C. A.; Hruby, V. J. *J. Am. Chem. Soc.* **1993**, *115*, 3399.
(42) Gonnella, N. C.; Zhang, X.; Jin, Y.; Prakash, O.; Paris, C. G.; Kolossváry, I.; Guida, W. C.; Bohacek, R. S.; Vlattas, I.; Sytwu, T. *Int. J. Peptide Protein Res.* **1994**, *43*, 454.
(43) Mierke, D. F.; Kurz, M.; Kessler, H. *J. Am. Chem. Soc.* **1994**, *116*, 1042.
(44) Burgess, K.; Ho, K.-K.; Pettitt, B. M. *J. Am. Chem. Soc.* **1995**, *117*, 54.
(45) Cicero, D. O.; Barbato, G.; Bazzo, R. *J. Am. Chem. Soc.* **1995**, *117*, 1027.
(46) For a discussion of emerging challenges in computational chemistry, see: Stillinger, F. H.; Andersen, H. C.; Auslander, L.; Beveridge, D. L.; Davidson, E. R.; Guida, W. C.; Kollman, P. A.; Lester, W. A.; Martin, Y. C.; Schatz, G. C.; Schlick, T.; Scott, L. R.; Sumners, D. L.; Wolynes, P. G.; Houk, K. N.; Weidman, S. T. *Mathematical Challenges from Theoretical/Computational Chemistry*; National Academy Press: Washington, DC, 1995.
(47) Cerjan, C. J.; Miller, W. H. *J. Chem. Phys.* **1981**, *75*, 2800.
(48) Simons, J.; Jorgensen, P.; Taylor, H.; Ozment, J. *J. Phys. Chem.* **1983**, *87*, 2745.
(49) O'Neal, D.; Taylor, H.; Simons, J. *J. Phys. Chem.* **1984**, *88*, 1510.
(50) Banerjee, A.; Adams, N.; Simons, J.; Shepard, R. *J. Phys. Chem.* **1985**, *89*, 52.
(51) Baker, J. *J. Comput. Chem.* **1986**, *7*, 385.

modes considered. Therefore, when the systematic search directions are exhausted for a particular minimum, LMOD switches to a stochastic or Monte Carlo procedure. In the Monte Carlo mode, random directions comprised of a random mixture of the low-mode eigenvectors are searched in exactly the same manner as in the systematic mode. Note that, in the Monte Carlo mode, the search direction is never updated as is the case for the systematic mode. The Monte Carlo mode could, in principle, suffer from the well-known deficiencies of stochastic search methods when applied to exhaustive conformational searches. LMOD, however, has proven in our hands to be equally effective in the systematic mode and the Monte Carlo mode. In our opinion, LMOD preserves its efficiency in the Monte Carlo mode toward the end of the search simply because the magnitude of the search, whether it is systematic or stochastic, is small compared to that of other search techniques. The LMOD search can be considered to be a "divide and conquer" procedure where numerous *local* searches are conducted in a *low-dimensional* subspace of the (local) nuclear configuration space.

In our hands, the unbounded systematic multiple minimum search technique[8] (SUMM) has proven to be superior to other methods for the exhaustive searching of conformational space.[38] SUMM was shown[8] to be superior to a torsion angle Monte Carlo procedure[11] that was shown to be the most efficient conformational search method tested when compared to other procedures for conformational searching of cycloheptadecane.[52] The SUMM procedure operates in torsion space by selecting values for torsion angle variation from a fixed set of torsion angles appropriate for a systematic search initially conducted at 120° resolution. When this set has been exhausted, torsion angles are selected from those appropriate for increasingly higher resolution. Multiple torsion angles are typically varied simultaneously. SUMM, when applied to cyclic molecules, follows the so-called "ringmaker" approach[53] by opening the ring(s), followed by systematic variation of the remaining torsion angles of the temporarily acyclic structure, and then reclosure of the ring(s). The resulting starting or probe structures are then subjected to energy minimization.

**Computational Details.** All the computations were done on either a Hewlett-Packard 9000/705 or a Hewlett-Packard 9000/730 workstation running our modified version of BatchMin 3.5.[54] SUMM is available in MacroModel/BatchMin, and was used with the default options including torsional memory and ring closure preoptimization.[8] The LMOD procedure has been integrated into BatchMin's conformational search system, and it was used with the same option settings in all of the test analyses. The low-frequency modes ($\leq$250 cm$^{-1}$) were searched by multiple 2.5 Å steps in a search direction comprised of pure eigenvectors or a random mixture of the eigenvectors of the non-mass-weighted Hessian matrix until the energy increase exceeded 10 000 kJ/mol during a single step. Preliminary studies indicated that a frequency threshold of 250 cm$^{-1}$ was sufficient to generate approximately 20 modes for each cycloheptadecane conformer. Moreover, the efficiency of the search did not appear to be highly sensitive to this threshold. Likewise, preliminary studies indicated that a step size of 2.5 Å and an energy threshold of 10 000 kJ/mol in a single step would be appropriate for the conformational searches we describe. Exhaustive "tuning" of LMOD with respect to these parameters was not done.

The starting or probe structures were subjected to energy minimization with the truncated Newton conjugate gradient (TNCG) minimizer.[55,56] Generation of the probe structures and the subsequent energy minimization (EM) is termed an MC/EM step,[38] although "MC" as implemented in LMOD (or SUMM) is not necessarily a Monte Carlo procedure. The so-called usage-directed structure selection scheme[11] was employed to select the input structure for each MC/EM step. For searches in which the number of low-energy conformers was "known" *a priori* (or could be reasonably estimated from a systematic search), the search was terminated when all of the known conformers were located. However, for all other searches, a fixed number of MC/EM steps was employed, the magnitude of which was based solely upon the desire to conduct as complete a conformational search as possible balanced against the finite computational resources at hand. The fully minimized starting structures (rms($G$) < 0.01 kJ/(mol Å)) were compared with those already found during previous MC/EM steps based on superposition of their non-hydrogen atoms, and only unique conformers were saved. The structure comparison included symmetry operations to detect whether two conformers were only subject to a reflection and/or rotation of the numbering of their non-hydrogen atoms. In the case of the bicyclic molecules, equivalent alternative atom-numbering systems were also checked. The final set of conformers was also subjected to normal mode analysis, and only the true minima were kept. Furthermore, when conformational searches were performed using different methods, the unique conformers generated by each technique were also compared to each other. The force fields employed were the MacroModel[54] variants (MM2* and MM3*) of the authentic Allinger MM2 and MM3 force fields,[57−59] and the MacroModel[54] variant (AMBER*) of the authentic Kollman AMBER force field.[60]

## Results and Discussion

**Cycloheptadecane MM2.** Our first test revisited cycloheptadecane, whose conformational space has been used as a standard test terrain for the comparison of new conformational search techniques.[12,36,38,52] In a previous paper we reported the remarkable performance of our torsional flexing algorithm (TFLEX) and the SUMM procedure for this conformational analysis problem.[38] In our hands, the SUMM procedure was able to locate all 262 known[52] low-energy conformers (within 12.6 kJ/mol (3 kcal/mol) above the global minimum) of cycloheptadecane on the MM2 potential energy hypersurface in a single search. Therefore, the cycloheptadecane problem was chosen as a first challenge to SUMM by our new LMOD procedure.

SUMM found all 262 cycloheptadecane conformers in 16 851 MC/EM steps, which required 186.6 ks of cpu time on a Hewlett-Packard 9000/705 workstation, whereas LMOD needed only 11 631 MC/EM steps in 93.4 ks of cpu time. In the cycloheptadecane test, LMOD outperformed SUMM by a factor of 2. Table 1 shows the details that elucidate the 100% increase in speed. The time required for energy minimization decreased by about 25%, but 75% of the speed increase was due to the elimination of the ring closure problem when using LMOD. SUMM generated over 150 000 probe structures subject to ring closure violation, which were therefore rejected, and which consumed almost 60 ks of cpu time.

Recently, a constrained stochastic search (CSS) procedure has been reported to generate only cyclic probe structures devoid of ring closure violations.[12] The authors of that paper also found all of the 262 low-energy cycloheptadecane conformers by minimizing 60 000 probe structures (60 000 MC/EM steps by our terminology). Unfortunately, their data do not allow a direct comparison of their CSS methodology with our LMOD procedure. Nonetheless, CSS minimized more than 4 times as many probe structures as SUMM and more than 5 times as many probe structures as LMOD to find all 262 low-energy cycloheptadecane conformers. It is true, however, that, in contrast to SUMM and LMOD (and many other conformational search techniques), CSS has been implemented in such a way that it explores the *entire* potential energy hypersurface. Thus, it searches for *all* minima rather than for *all of the low-energy* minima.

(52) Saunders, M.; Houk, K. N.; Wu, Y.; Still, W. C.; Lipton, M.; Chang, G.; Guida, W. C. *J. Am. Chem. Soc.* **1990**, *112*, 1419.
(53) Still, W., C.; Galynker, I. *Tetrahedron* **1981**, *37*, 3981.
(54) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440. MacroModel/BatchMin is available from W. Clark Still, Columbia University, New York, NY.
(55) Ponder, J. W.; Richards, F. M. *J. Comput. Chem.* **1987**, *8*, 1016.
(56) Schlick, T.; Overton, M. *J. Comput. Chem.* **1987**, *8*, 1025.
(57) Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127.
(58) Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551.
(59) Lii, J.-H.; Allinger, N. L. *J. Am. Chem. Soc.* **1989**, *111*, 8566.
(60) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.

**Table 1.** Comparison of the Performance of LMOD and SUMM for Conformational Searching of Cycloheptadecane

| method | no. of MC/EM steps | no. of unique conformers[a] | cpu time[b] (ks) | no. of probe structures rejected |
|---|---|---|---|---|
| | | MM2 | | |
| LMOD | 11 631 | 262 | 91.7[c] | 0[e] |
| | | | 1.7[d] | 4481[f] |
| | | | Σ 93.4 | Σ 4481 |
| SUMM | 16 851 | 262 | 127.5[c] | 150 039[e] |
| | | | 59.1[d] | 15 064[f] |
| | | | Σ 186.6 | Σ 165 103 |
| | | MM3 | | |
| LMOD | 25 000 | 134 | 288.9[c] | 0[e] |
| | | | 5.8[d] | 6679[f] |
| | | | Σ 294.7 | Σ 6679 |
| SUMM | 25 000 | 134 | 290.0[c] | 288,616[e] |
| | | | 96.1[d] | 24,586[f] |
| | | | Σ 386.1 | Σ 313 202 |

[a] Energy window 12.6 kJ/mol (∼3 kcal/mol) above the global minimum. [b] Measured on a Hewlett-Packard 9000/705 workstation running our modified BatchMin 3.5 software. [c] cpu time used for energy minimization. [d] cpu time used for miscellaneous geometry operations: generating probe structures, checking close van der Waals contacts, relieving ring closure violations, and eliminating duplicates. [e] Probe structures violating the ring closure distance constraint 0.5−3.5 Å (recommended value in the BatchMin documentation). [f] Probe structures with bad van der Waals contacts.

**Cycloheptadecane MM3.** The MM3 potential energy hypersurface of cycloheptadecane is quite different from that of MM2. Reminimization of the 262 MM2 minima with MM3 yielded only 134 cycloheptadecane conformers within 12.6 kJ/mol (3 kcal/mol) above the global minimum. It is instructive to note that the order of the corresponding conformers is different with the two different force fields. The most striking difference is that the order of the two lowest energy conformers is reversed and the energy difference between the global minimum and the "runner-up" is only 0.112 kJ/mol with MM2 but much larger, 3.61 kJ/mol, with MM3. This result cast some doubt on whether there are, in fact, only 134 MM3 cycloheptadecane conformers within 12.6 kJ/mol above the global minimum. Therefore, we conducted a conformational search using the MM3 force field for energy minimization and allowing an excessive number of MC/EM steps (25 000) to conduct a presumably complete exploration of the low-energy conformers. Indeed, since SUMM and LMOD both found the same 134 cycloheptadecane conformers, we are reasonably confident that the search has been exhaustive.

The MM3 results in Table 1 are similar to those with MM2. In particular, the cpu time used for energy minimizations is virtually identical with the two conformational search methods since an equal number of MC/EM steps was employed. This means the increase in speed with LMOD can be fully attributed to the elimination of the ring closure problem when LMOD is used. SUMM generated close to 300 000 probe structures subject to ring closure violation, which required almost 100 ks of cpu time. Each of the 134 low-energy conformers was located multiple times; however, the average duplication rate was 39 with LMOD, but only 17 with SUMM. In addition, LMOD found all the conformers at least five times, whereas SUMM located five particular conformers only two or three times.

**Cyclododecane MM2.** The conformational space available to cyclododecane has been studied very extensively.[1b,12,25,35,37] Cyclododecane is the largest cycloalkane whose *entire* conformational space can be searched exhaustively using contemporary desktop hardware and software resources. In a previous study[1b] we located 121 cyclododecane conformers on the MM2 potential energy hypersurface using TFLEX[38] and SUMM.[8] The recently devised CSS procedure has also been reported to find 121 conformers.[12,61] Therefore, it seemed that cyclododecane possessed only 121 conformers on the MM2 potential energy hypersurface.[62] LMOD, however, found a 122nd conformer within a 25 000 MC/EM step limit (see Table 2). Thus, there could well be additional cyclododecane conformers yet to be discovered. This is particularly likely since the new 122nd conformer is only 49th in the series (34.1 kJ/mol above the global minimum). This relatively-high-energy structure possesses interesting symmetry with torsion angles of 176°, −111°, 64°, −111°, 176°, −111°, 64, −111°, 176°, −111°, 64°, and −111° for the carbon atoms. The structure is shown in Figure 1. The fact that this conformer possesses six torsions that are nearly −120°, and thus involved in eclipsed interactions, explains its high energy. It is also not surprising that SUMM failed to locate this conformer within 25 000 MC/EM steps since it would have to have achieved a resolution of 60°, which is not likely given the number of possible starting structures, and the fact that multiple torsions were allowed to vary simultaneously.

**Table 2.** Comparison of the Performance of LMOD and SUMM for Conformational Searching of Cyclododecane

| method | no. of MC/EM steps | no. of unique conformers[a] | cpu time[b] (ks) | no. of probe structures rejected |
|---|---|---|---|---|
| | | MM2 | | |
| LMOD | 25 000 | 122 | 109.5[c] | 0[e] |
| | | | 2.8[d] | 14 721[f] |
| | | | Σ 112.3 | Σ 14 721 |
| SUMM | 25 000 | 121 | 107.6[c] | 3 406 881[e] |
| | | | 192.1[d] | 662 721[f] |
| | | | Σ 299.7 | Σ 4 069 602 |
| | | MM3 | | |
| LMOD | 25 000 | 98 | 157.3[c] | 0[e] |
| | | | 3.4[d] | 11 722[f] |
| | | | Σ 160.7 | Σ 11 722 |
| SUMM | 25 000 | 97 | 145.1[c] | 2 738 135[e] |
| | | | 158.4[d] | 547 999[f] |
| | | | Σ 311.5 | Σ 3 286 134 |

[a] Energy window infinity. [b] Measured on a Hewlett-Packard 9000/705 workstation running our modified BatchMin 3.5 software. [c] cpu time used for energy minimization. [d] cpu time used for miscellaneous geometry operations: generating probe structures, checking close van der Waals contacts, relieving ring closure violations, and eliminating duplicates. [e] Probe structures violating the ring closure distance constraint 0.5−3.5 Å (recommended value in the BatchMin documentation). [f] Probe structures with bad van der Waals contacts.

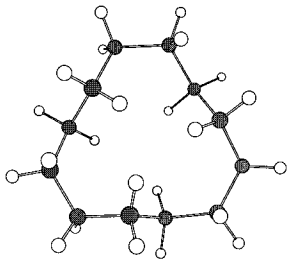

**Figure 1.** Ball and stick representation of the additional conformer of cyclododecane located by LMOD that was not found by SUMM.

The results in Table 2 clearly indicate that the ring closure problem is more severe with the smaller cyclododecane than with the larger cycloheptadecane. The cpu time used for energy

(61) The authors of the CSS paper misquoted our TFLEX/SUMM search results on cycloundecane and cyclododecane in Table 7 of ref 12.

(62) The 121 conformers do not include structures containing pyramidal carbon atoms. In contrast to the study described in ref 12, we discarded such structures. Nonetheless, our present study found a number of such fully converged structures with the BatchMin warning message "rejected by distorted sp3 atom".

minimizations is virtually identical, but SUMM actually consumed a lot more cpu time to generate and eliminate several million probe structures subject to ring closure violation. As a net result, LMOD outperformed SUMM by a factor of over 2.5. The direct comparison with CSS is not straightforward in this case either. Nonetheless, CSS minimized 30 000 probe structures to obtain "*all*" of the cyclododecane conformers,[12] which is comparable to our 25 000 MC/EM steps, notwithstanding that LMOD found an additional low-energy conformer.

**Cyclododecane MM3.** Similar to the cycloheptadecane case, the MM3 potential energy hypersurface of cyclododecane displays significantly fewer minima than the one of MM2. The MM3 results in Table 2 are qualitatively very similar to the MM2 results. There are, however, two notable differences. Energy minimization of the cyclododecane probe structures requires almost 50% more cpu time with MM3 than with MM2, and the ring closure problem is less severe for structures generated with MM3 than with MM2. As a consequence, LMOD outperformed SUMM by *only* a factor of 2 in this case, and LMOD found an additional unique conformer not located by SUMM. This additional conformer is of very high energy (216.8 kJ/mol above the global minimum), but still *much much* lower than the recently discovered "trefoil knot" structure, which is more than 4000 kJ/mol higher in energy than the global minimum.[12] It is instructive to note that the trefoil knot or other potential knot structures are expected to become physically meaningful only in very large rings ($N > 50$).[12,63] Therefore, although LMOD did not locate the trefoil knot in the present study, it can be set to explore higher frequency vibrational modes, which could possibly lead to the trefoil knot. Our goal in this study has been to search for "normal" structures (i.e., physically relevant structures). In this respect, when compared to CSS, LMOD found 98 normal MM3 cyclododecane conformers excluding the trefoil knot, whereas CSS found 97 including the trefoil knot.

***n*-Octane MM2.** The successful tests on cycloalkanes encouraged us to challenge SUMM on a problem for which it is particularly well suited. The conformational search problem for *n*-alkanes is ideal for the SUMM procedure because *n*-alkanes possess almost exclusively *anti* and *gauche* torsion angles, and thus SUMM should be able to locate all conformers containing these torsions at its lowest operative resolution. First, *n*-octane was subjected to systematic conformational searching with an infinite energy window using the MULTIC mode of MacroModel.[6,54] A 60° grid search seemed appropriate for this problem, but as a safeguard, a higher resolution (36°) was employed. The resulting 100 000 *n*-octane structures were pruned to 44 666 by eliminating structures with bad van der Waals contacts. The remaining structures were subjected to energy minimization and compared with each other, affording 98 unique *n*-octane conformers. Thus, it appears that there are 98 conformers of *n*-octane. LMOD and SUMM were then instructed to terminate either when 98 unique conformers were found or when 10 000 MC/EM steps had been completed.

LMOD completed the task in 3966 steps, having located 98 conformers that were identical to those found with the systematic search, but SUMM could only locate 97 *n*-octane conformers during the full length of the 10 000 step search (see Table 3). The missing conformer is the 68th in the sequence (21.4 kJ/mol above the global minimum). The most likely reason that SUMM could not locate this conformer is that it has an unusual torsion angle (∼90°) in the middle of the chain. SUMM would not be expected to achieve a resolution higher than 120° for the 10 000 MC/EM steps in which multiple torsions were allowed to vary simultaneously, and therefore, the ∼90° torsion angle escaped detection.

**Table 3.** Comparison of the Performance of LMOD and SUMM for Conformational Searching of *n*-Octane

| method | no. of MC/EM steps | no. of unique conformers[a] | cpu time[b] (ks) | no. of probe structures rejected |
|---|---|---|---|---|
| | | MM2 | | |
| LMOD | 3966 | 98 | 9.3[c] | 0[e] |
| | | | 0.1[d] | 458[f] |
| | | | ∑ 9.4 | ∑ 458 |
| SUMM | 10 000 | 97 | 19.7[c] | 96 121[e] |
| | | | 3.2[d] | 8562[f] |
| | | | ∑ 22.9 | ∑ 104 683 |
| | | MM3 | | |
| LMOD | 686 | 99 | 2.4[c] | 0[e] |
| | | | 0.0[d] | 54[f] |
| | | | ∑ 2.4 | ∑ 54 |
| SUMM | 1542 | 99 | 3.9[c] | 18 869[e] |
| | | | 0.2[d] | 2304[f] |
| | | | ∑ 4.1 | ∑ 21 173 |

[a] Energy window infinity. [b] Measured on a Hewlett-Packard 9000/705 workstation running our modified BatchMin 3.5 software. [c] cpu time used for energy minimization. [d] cpu time used for miscellaneous geometry operations: generating probe structures, checking close van der Waals contacts, and eliminating duplicates. [e] Probe structures rejected by a torsional memory used by SUMM to prevent the procedure from revisiting the same neighborhoods of conformational space defined by the same sets of torsional angles only subject to symmetry relationships. [f] Probe structures with bad van der Waals contacts.

It is instructive to note that SUMM generated a large number of probe structures that were rejected by a torsional memory[8,64] (see Table 3). The torsional memory option is employed by SUMM to prevent the procedure from revisiting the same neighborhoods of conformational space defined by the same sets of torsional angles only subject to symmetry relationships. *n*-Octane is, of course, subject to a great degree of redundancy due to symmetry relationships between the *anti* and *gauche* torsion angles.

***n*-Octane MM3.** Minimization of the 44 666 *n*-octane structures from the grid search afforded 99 unique MM3 conformers. In this test, LMOD and SUMM both found all of the 99 conformers, but LMOD was considerably more efficient than SUMM (686 and 1542 MC/EM steps, respectively).

**Bicyclo[5.5.1]tridecane MM2.** In order to compare LMOD against SUMM in systems for which multiple ring closures are possible, our final test of hydrocarbons included two bicyclic systems originally investigated computationally by Saunders.[24] Bicyclo[5.5.1]tridecane can be subjected to a complete or a nearly complete conformational search. LMOD was tested against SUMM in a 5000 step search which was indicative of a nearly complete search, because a few conformers were found less than five times. Bicyclic hydrocarbons are comprised of three families of "in−in", "in−out", and "out−out" conformers, respectively.[24] This notation of the three different families of structures is explained by the inward or outward orientation of the hydrogen atoms attached to the bridgehead carbon atoms. Formation of the in−in conformers is unfavorable for small systems with short bridges, but the in−in configuration is favored beyond a certain ring size.[24] Our test searches started with an in−out conformer. The chirality of the bridgehead atoms was not fixed, allowing the generation of all three families of bicyclo[5.5.1]tridecane conformers.

(63) The trefoil knot structures for higher cycloalkanes were described as "to be published" in ref 17 of the study described by Weinberg and Wolfe.[12]

(64) The torsional memory option of SUMM[8] was also employed for cyclic structures, but the number of probe structures rejected by torsional memory was negligible compared to the large number of probe structures subject to ring closure violation.

**Table 4.** Comparison of the Performance of LMOD and SUMM for Conformational Searching of Bicyclo[5.5.1]tridecane

| method | no. of MC/EM steps | no. of unique conformers[a] | cpu time[b] (ks) | no. of probe structures rejected |
|---|---|---|---|---|
| | | MM2 | | |
| LMOD | 5000 | 64 | 19.9[c] | 0[e] |
| | | | 0.4[d] | 2227[f] |
| | | | Σ 20.3 | Σ 2227 |
| SUMM | 5000 | 64 | 20.7[c] | 2 358 515[e] |
| | | | 16.3[d] | 8470[f] |
| | | | Σ 37.0 | Σ 2 366 985 |
| | | MM3 | | |
| LMOD | 5 000 | 56 | 29.7[c] | 0[e] |
| | | | 0.6[d] | 1704[f] |
| | | | Σ 30.3 | Σ 1704 |
| SUMM | 5000 | 56 | 29.5[c] | 3 452 386[e] |
| | | | 23.9[d] | 10 552[f] |
| | | | Σ 53.4 | Σ 3 462 938 |

[a] Energy window infinity. [b] Measured on a Hewlett-Packard 9000/705 workstation running our modified BatchMin 3.5 software. [c] cpu time used for energy minimization. [d] cpu time used for miscellaneous geometry operations: generating probe structures, checking close van der Waals contacts, relieving ring closure violations, and eliminating duplicates. [e] Probe structures violating the ring closure distance constraint 0.5−2.5 Å (recommended value in the BatchMin documentation). [f] Probe structures with bad van der Waals contacts.

LMOD and SUMM both found the same 64 structures during the search including Saunders' best in−in, in−out, and out−out conformers.[24] The structure with lowest energy is an out−out conformer followed by an in−out conformer 3.7 kJ/mol higher in energy. The lowest in−in conformer is 14th in the sequence with an energy that is 20.9 kJ/mol above the lowest energy structure. Table 4 shows that the almost 2-fold speed advantage of LMOD over SUMM can be, once again, fully attributed to the ring closure problem. The cpu time used for energy minimizations is virtually identical, but SUMM generated over two million probe structures subject to ring closure violation. Of course, the ring closure problem is further exacerbated with bicyclic systems where, instead of only one, two rings must be closed simultaneously.

**Bicyclo[5.5.1]tridecane MM3.** Similar to the single-ring systems tested, the MM3 potential energy hypersurface of bicyclo[5.5.1]tridecane possesses fewer minima than the one of MM2. LMOD and SUMM both located only 56 structures. The lowest energy out−out conformer is a lower minimum on the MM3 potential energy hypersurface than on the potential energy hypersurface for MM2. The succeeding in−out conformer is 5.8 kJ/mol higher in energy. The lowest in−in conformer is now 15th in the series with an energy that is 24.9 kJ/mol above the global minimum. Table 4 shows that, similar to cyclododecane, energy minimization requires almost 50% more cpu time with MM3 than with MM2, but interestingly the ring closure problem in this case is more severe with MM3 than with MM2. As a net result, the almost 2-fold speed advantage of LMOD over SUMM prevails.

**Bicyclo[6.5.5]octadecane MM2.** The large bicyclo[6.5.5]-octadecane molecule escaped our attempts to conduct an exhaustive search. The energy gap between the lowest energy conformer of each of the three different families of structures containing different bridgehead configurations is ∼20 kJ/mol apiece. This means that a search on a molecule of this size that includes a representative set of the higher energy out−out conformers would require an energy window far too wide (>50 kJ/mol) for the search to be exhaustive using the computational resources at hand. Therefore, we split the conformational analysis of bicyclo[6.5.5]octadecane into three separate searches (5000 MC/EM steps, energy window 25 kJ/mol). The chirality of the bridgehead carbon atoms was fixed this time to afford only in−in, in−out, and out−out conformers. However, constraining the configuration of the bridgehead atoms is not enough. Saunders has pointed out in his bicyclic study[24] that in−in conformers can be converted into out−out conformers without the inversion of the *local* configuration of the bridgehead atoms by simply pulling one bridge through the ring formed by the other two bridges. To prevent this *global* inversion, an additional distance constraint was applied to distinguish in−in and out−out conformers on the basis of the distance of the bridgehead hydrogen atoms.

**Table 5.** Comparison of the Performance of LMOD and SUMM for Conformational Searching of In−In Bicyclo[6.5.5]octadecane

| method | no. of MC/EM steps | no. of unique conformers[a] | cpu time[b] (ks) | no. of probe structures rejected |
|---|---|---|---|---|
| | | MM2 | | |
| | | Σ 102[g] | | |
| LMOD | 5000 | 101 | 38.9[c] | 0[e] |
| | | | 1.1[d] | 4992[f] |
| | | | Σ 40.0 | Σ 4992 |
| SUMM | 5000 | 101 | 53.7[c] | 5 326 110[e] |
| | | | 36.5[d] | 18 451[f] |
| | | | Σ 90.2 | Σ 5 344 561 |
| | | MM3 | | |
| | | Σ 108[g] | | |
| LMOD | 5000 | 108 | 51.6[c] | 0[e] |
| | | | 1.4[d] | 4175[f] |
| | | | Σ 53.0 | Σ 4175 |
| SUMM | 5000 | 108 | 60.5[c] | 5 539 914[e] |
| | | | 37.8[d] | 21 195[f] |
| | | | Σ 98.3 | Σ 5 561 109 |

[a] Energy window 25 kJ/mol above the lowest in−in conformer (global minimum). [b] Measured on a Hewlett-Packard 9000/705 workstation running our modified BatchMin 3.5 software. [c] cpu time used for energy minimization. [d] cpu time used for miscellaneous geometry operations: generating probe structures, checking close van der Waals contacts, relieving ring closure violations, and eliminating duplicates. [e] Probe structures violating the ring closure distance constraint 0.5−3.5 Å (recommended value in the BatchMin documentation). [f] Probe structures with bad van der Waals contacts. [g] LMOD and SUMM combined.

The combined results of the three separate searches are as follows. None of the searches were even nearly complete because, in all three cases, the *combined* LMOD−SUMM output files afforded more unique conformers than either LMOD or SUMM alone (see Tables 5−7). As a matter of fact, this situation is typical for practical conformational searches. In many instances, the search must involve a fairly large energy window in order to include relatively-high-energy conformers (e.g., when the binding conformation of an inhibitor to an enzyme is sought), and the use of a wide energy window of 25−30 kJ/mol is typical. In this case, one is interested in using a search procedure that affords the greatest number of conformers within a necessarily limited time period.

Our present study confirmed that the structure of lowest energy possessed the same MM2 energy as the global minimum in−in conformer found by Saunders.[24] The lowest energy in−out and out−out conformers are 20.4 and 40.6 kJ/mol higher in energy, respectively. Saunders' best in−out conformer is the second in our sequence ($\Delta E$ = 1.3 kJ/mol), and his best out−out conformer is tenth ($\Delta E$ = 10.9 kJ/mol). Figure 2 depicts our global minimum in−in structure (i.e., the lowest energy in−in conformer). The results in Tables 5−7 show that, with this particular molecule, LMOD required substantially less cpu time for energy minimization than SUMM. The ring closure problem (especially with the in−in configuration) further added to the advantage of LMOD over SUMM, resulting in a net speed increase of 50% to over 100%.

**Table 6.** Comparison of the Performance of LMOD and SUMM for Conformational Searching of In−Out Bicyclo[6.5.5]octadecane

| method | no. of MC/EM steps | no. of unique conformers[a] | cpu time[b] (ks) | no. of probe structures rejected |
|---|---|---|---|---|
| | | MM2 | | |
| | | ∑ 229[g] | | |
| LMOD | 5000 | 217 | 39.1[c] | 0[e] |
| | | | 1.2[d] | 4793[f] |
| | | | ∑ 40.3 | ∑ 4793 |
| SUMM | 5000 | 218 | 46.7[c] | 1 714 776[e] |
| | | | 13.9[d] | 9059[f] |
| | | | ∑ 60.6 | ∑ 1 723 835 |
| | | MM3 | | |
| | | ∑ 254[g] | | |
| LMOD | 5000 | 244 | 51.2[c] | 0[e] |
| | | | 1.4[d] | 3996[f] |
| | | | ∑ 52.6 | ∑ 3996 |
| SUMM | 5000 | 244 | 57.3[c] | 1 500 898[e] |
| | | | 12.4[d] | 9,119[f] |
| | | | ∑ 69.7 | ∑ 1 510 017 |

[a] Energy window 25 kJ/mol above the lowest in−out conformer. [b] Measured on a Hewlett-Packard 9000/705 workstation running our modified BatchMin 3.5 software. [c] cpu time used for energy minimization. [d] cpu time used for miscellaneous geometry operations: generating probe structures, checking close van der Waals contacts, relieving ring closure violations, and eliminating duplicates. [e] Probe structures violating the ring closure distance constraint 0.5−3.5 Å (recommended value in the BatchMin documentation). [f] Probe structures with bad van der Waals contacts. [g] LMOD and SUMM combined.

**Table 7.** Comparison of the Performance of LMOD and SUMM for Conformational Searching of Out−Out Bicyclo[6.5.5]octadecane

| method | no. of MC/EM steps | no. of unique conformers[a] | cpu time[b] (ks) | no. of probe structures rejected |
|---|---|---|---|---|
| | | MM2 | | |
| | | ∑ 103[g] | | |
| LMOD | 5000 | 98 | 39.1[c] | 0[e] |
| | | | 1.1[d] | 4355[f] |
| | | | ∑ 40.2 | ∑ 4355 |
| SUMM | 5000 | 93 | 45.4[c] | 3 054 617[e] |
| | | | 22.0[d] | 12 197[f] |
| | | | ∑ 67.4 | ∑ 3 066 814 |
| | | MM3 | | |
| | | ∑ 130[g] | | |
| LMOD | 5000 | 122 | 53.1[c] | 0[e] |
| | | | 1.3[d] | 3841[f] |
| | | | ∑ 54.4 | ∑ 3841 |
| SUMM | 5000 | 122 | 57.6[c] | 2 129 659[e] |
| | | | 16.1[d] | 9934[f] |
| | | | ∑ 73.7 | ∑ 2 139 593 |

[a] Energy window 25 kJ/mol above the lowest out−out conformer. [b] Measured on a Hewlett-Packard 9000/705 workstation running our modified BatchMin 3.5 software. [c] cpu time used for energy minimization. [d] cpu time used for miscellaneous geometry operations: generating probe structures, checking close van der Waals contacts, relieving ring closure violations, and eliminating duplicates. [e] Probe structures violating the ring closure distance constraint 0.5−3.5 Å (recommended value in the BatchMin documentation). [f] Probe structures with bad van der Waals contacts. [g] LMOD and SUMM combined.

**Bicyclo[6.5.5]octadecane MM3.** The MM3 searches were not exhaustive either with, perhaps, the exception of the in−in configuration where LMOD and SUMM combined afforded the same 108 conformers found by both LMOD and SUMM alone. With MM3, the lowest energy in−out and out−out conformers were 19.7 and 43.6 kJ/mol higher in energy than the in−in global minimum, respectively. The lowest energy out−out structure was virtually identical to the lowest energy out−out MM2 structure, and in fact, when this structure was subjected to energy minimization with the MM2 force field, it minimized to the MM2 structure. Shown in Figure 3 is the global minimum in−in structure (i.e., the lowest energy in−in conformer).

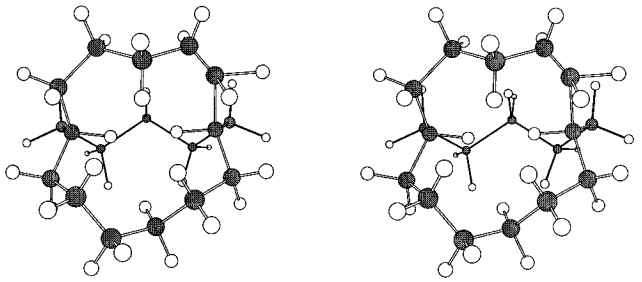

**Figure 2.** Stereoview of the MM2 global minimum in−in structure of bicyclo[6.5.5]octadecane.

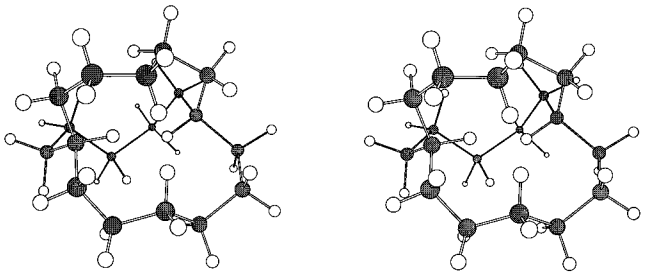

**Figure 3.** Stereoview of the MM3 global minimum in−in structure of bicyclo[6.5.5]octadecane.

Interestingly, this was the only cyclic test molecule whose MM3 potential energy hypersurface possessed more minima than the MM2 potential energy hypersurface. Similar, however, to the other test molecules, energy minimization of bicyclo-[6.5.5]octadecane proceeded significantly more slowly with MM3 than with MM2, but LMOD was still faster than SUMM. The net speed of LMOD slightly decreased compared to that of MM2, but Tables 5−7 show a very respectable 40−80% advantage over SUMM.

**Cyclopenta-L-alanine AMBER.** In order to test the performance of LMOD for a conformational search problem that does not involve a hydrocarbon, we conducted conformational searches with LMOD and SUMM on the cyclic peptide cyclopenta-L-alanine. A related, but more rigid, cyclic pentapeptide, cyclo-D-Pro-L-$Ala_4$, has recently been studied by NMR.[43] The authors of that paper concluded that the NMR data could only be rationalized by averaging over the relevant conformations for this molecule. Thus, the ability to efficiently conduct conformational searches on molecules like cyclic peptides is highly relevant to an understanding of experimental results.

For our searches on cyclopenta-L-alanine, the united atom AMBER force field was employed with a dielectric "constant" of 4*r*, and we conducted searches that involved 10 000 MC/EM steps (see Table 8). Both LMOD and SUMM found 60 conformers within a 35 kJ/mol energy window. We believe that the search was exhaustive since the average duplication rates for LMOD and SUMM were 36 and 57, respectively. As with the previous searches on hydrocarbons, LMOD was faster than SUMM (for the 10 000 step search) primarily due to the additional time required by SUMM for ring closure.

It is worth noting that no constraints were placed on the amide bonds and thus both SUMM and LMOD found the same 60 conformers, which contained both *cis*- and *trans*-amides. The global minimum contained all *cis*-amide bonds, and the lowest energy *all-trans* stereoisomer possessed an energy that was 30.8 kJ/mol higher in energy than the global minimum. Conformations with all other permutations of *cis*- and *trans*-amides were also located. With SUMM the amide torsions were explicitly allowed to vary during the MC step, and the fact that it located both *cis* and *trans* isomers is not surprising. The fact that

**Table 8.** Comparison of the Performance of LMOD and SUMM for Conformational Searching of Cyclopenta-L-alanine Using AMBER

| method | no. of MC/EM steps | no. of unique conformers[a] | cpu time[b] (ks) | no. of probe structures rejected |
|---|---|---|---|---|
| LMOD | 10 000 | 60 | 29.8[c] | 0[e] |
| | | | 0.3[d] | 910[f] |
| | | | Σ 30.1 | Σ 910 |
| SUMM | 10 000 | 60 | 36.1[c] | 78 989[e] |
| | | | 7.9[d] | 7118[f] |
| | | | Σ 44.0 | Σ 86 107 |

[a] Energy window 35 kJ/mol above the global minimum; a distance dependent dielectric constant of 4*r* was employed. [b] Measured on a Hewlett-Packard 9000/705 workstation running our modified BatchMin 3.5 software. [c] cpu time used for energy minimization. [d] cpu time used for miscellaneous geometry operations: generating probe structures, checking close van der Waals contacts, relieving ring closure violations, and eliminating duplicates. [e] Probe structures violating the ring closure distance constraint 0.5−4.0 Å (recommended value in the BatchMin documentation; one of the amide bonds was opened in order to leave the chiral centers unperturbed). [f] Probe structures with bad van der Waals contacts.

LMOD was also able to locate both *cis*- and *trans*-amides is indicative of its utility as a general conformational search method even in the absence of any definition of those torsion angles that are explicitly allowed to vary.

**BQ123 AMBER.** In order to apply LMOD to a conformational search problem on a molecule for which solution NMR data are available, we conducted conformational searches using LMOD on the cyclic pentapeptide BQ123 (cyclo-D-Trp-D-Asp-Pro-D-Val-Leu), which we had studied previously using both NMR spectroscopic and computational methods.[42] BQ123 is a potent antagonist of a specific receptor ($ET_A$) for endothelin, which itself is a potent 21-amino acid peptide vasoconstrictor.

Previously,[42] we had conducted conformational searches on BQ123 with our TFLEX method and the united atom AMBER force field using a modified version of BatchMin. In the present study, we have conducted conformational searches on BQ123 using 5000 MC/EM steps and employing the all-atom AMBER force field. The searches were performed using the GB/SA solvation model[65] and a 25 kJ/mol energy window on an HP 9000/730 workstation. First, we compared the efficiency of LMOD with TFLEX for this conformational search problem. As expected, LMOD was more efficient than TFLEX at locating low-energy conformers of BQ123. Thus, LMOD located 654 low-energy conformers with a cpu time of 153.7 ks, whereas TFLEX only located 521 structures and required 210.7 ks of cpu time.

In order to determine whether LMOD could be effectively employed in conformational searches involving constraints, an additional search was performed in which two hydrogen bond constraints, which were derived from the NMR studies,[42] were used. These hydrogen bond constraints involve hydrogen bonds between the D-Asp NH and the D-Val C═O, and the D-Val NH and D-Asp C═O. The constraints were imposed by using the FXDI command in BatchMin which applies a harmonic flat-well energetic restraint to the potential energy function. No modifications were made to LMOD, and it was simply used with the same options as described above with this modified potential energy function. As before 5000 MC/EM steps were employed, and the constrained search yielded 261 conformers. When these structures were subjected to reminimization with the constraints removed, 209 unique structures were produced.

(65) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127.

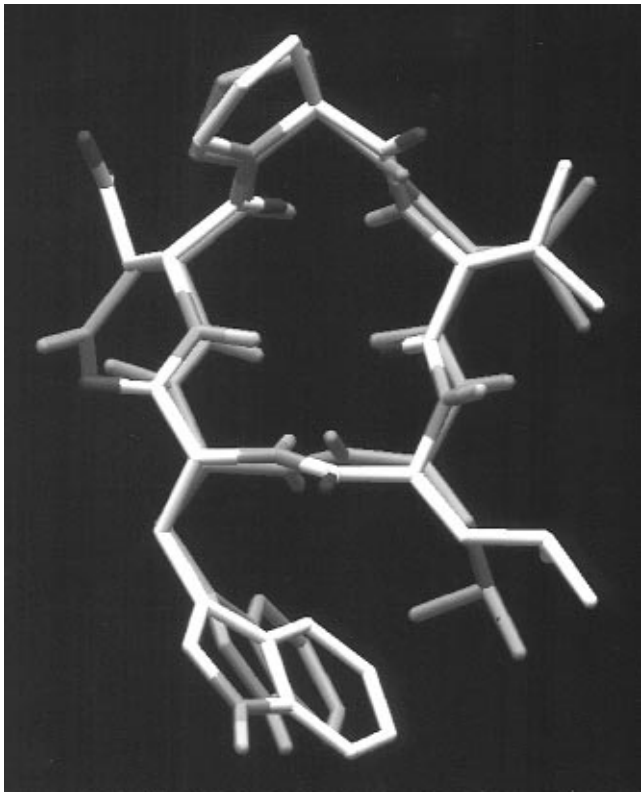

**Figure 4.** Comparison of one of the NMR-derived structures (green) of BQ123 with the lowest energy conformer (carbon atoms, white; nitrogens, blue; oxygens, red), yielding a favorable match. Note that the Asp side chain orientation is poorly defined by the NMR NOE data, and thus neither the NMR-derived structure nor the LMOD-derived structure violates the NMR distance constraints for this side chain.

The reminimized structures were then compared (using techniques as previously described[42]) to an ensemble of 20 refined structures derived from NMR studies performed on BQ123 in aqueous solution.[42] The NMR structures satisfy the hydrogen-bonding distance constraints and an additional 66 NOE-derived distance constraints. Four conformers were located that compared favorably with at least one of the NMR-derived structures. The matching structure of lowest energy was 6.6 kJ/mol higher in energy than the global minimum, and Figure 4 shows a comparison between this structure and the NMR structure with which it possesses the best match.

It is noteworthy that in our prior studies[42] the lowest energy structure in agreement with the NMR data was 22.0 kJ/mol higher in energy than the global minimum. Although in the present study the conformer that afforded the best match to the NMR structures was 15.4 kJ/mol higher in energy than the global minimum, the conformer shown in Figure 4 agrees quite well with the NMR data. The present TFLEX search afforded only one matching structure, which was 20.0 kJ/mol higher in energy than the global minimum. It is true, however, that neither the TFLEX search nor the LMOD search was exhaustive, since for each search, there were a number of conformers that were located only a single time. Thus, it is possible that there are additional conformers yet to be found that agree with the NMR data. It is also possible that a more extensive TFLEX search would locate additional conformers. Nonetheless, we have shown that one can readily apply LMOD to conformational searches performed on modified (e.g., constrained) potential energy surfaces. We have demonstrated that, in addition to unconstrained searches, LMOD can be used to conduct constrained conformational searches, and hence it should find utility

as a conformational search tool that can be used effectively in conjunction with experimental data (e.g., NMR-derived distance constraints).

## Conclusion

Adequate conformational sampling remains one of the key issues that must be addressed if quantitative molecular modeling results are to be attained. For example, the ability to accurately sample conformational space is essential for obtaining converged results in free energy simulations.[66,67] Thus, the development of new robust, highly efficient conformational search procedures continues to be of central importance in computational chemistry. In this paper, we have described a novel conformational search method termed LMOD for the exhaustive location of low-energy minima on the potential energy hypersurface of acyclic, monocyclic, and bicyclic molecules. We have compared LMOD to SUMM, which has been shown to be one of the most efficient conformational search procedures yet devised for the exhaustive conformational analysis of small molecules,[8,38] for a variety of conformational search problems. In each case LMOD performed significantly better than SUMM.

LMOD is equally appropriate for searching the conformational space available to cyclic and acyclic molecules. There is no need for special treatment of rings. Moreover, LMOD operates in neither torsion space nor Cartesian space, nor any other user-defined search space: LMOD generates its own search space! Thus, the user only needs to supply the threshold for the low-frequency modes and the energy threshold for energy minimization to occur, and then LMOD generates structures automatically. This feature of conformational searches performed with LMOD can be highly advantageous. When applied to large-scale 2D−3D structural data base conversions, for example, it affords a distinct advantage over search techniques requiring manual selection of rotatable bonds.

We believe that LMOD is one of the most efficient conformational search procedures yet devised and it should find wide utility as a conformational search method.[68]

JA952478M

(66) Guarnieri, F.; Still, W. C. *J. Comput. Chem.* **1994**,*15*, 1302.
(67) Senderowitz, H.; Guarnieri, F.;Still, W. C. *J. Am. Chem. Soc.* **1995**, *117*, 8211.

(68) Our LMOD procedure is intimately coupled to the BatchMin[54] program. We have provided the code for LMOD to Professor W. Clark Still, Columbia University, for inclusion in future releases of BatchMin.