

# Low-Mode Conformational Search Elucidated: Application to C<sub>39</sub>H<sub>80</sub> and Flexible Docking of 9-Deazaguanine Inhibitors into PNP

ISTVÁN KOLOSSVÁRY, WAYNE C. GUIDA\*

*Novartis Institute for Biomedical Research, Novartis Pharmaceuticals Corporation, 556 Morris Ave., Summit, New Jersey 07901*

*Received 20 May 1999; accepted 21 July 1999*

**ABSTRACT:** We previously described a new conformational search method, termed low-mode search (LMOD), and discussed its utility for conformational searches performed on cycloalkanes and a cyclic penta-peptide.<sup>1</sup> In this report, we discuss a rigorous implementation of mode following (*c*-LMOD) for conformational searching, and we demonstrate that for a conformational search involving cycloheptadecane, this rigorous implementation is capable of finding all of the previously known structures. To the best of our knowledge, this is the first computational proof that mode following can be used for conformational searches conducted on a complex molecular system. We show, however, that, as expected, it is generally inefficient to perform a conformational search in this manner. Nonetheless, *c*-LMOD has been shown to be an excellent method for conducting conformational analyses involving conformational interconversions, where the location of saddle points is important. We also describe refinement to our original LMOD procedure (*l*-LMOD) and discuss its utility for a difficult conformational search problem, namely locating the global minimum energy conformation of C<sub>39</sub>H<sub>80</sub>. For this search, *l*-LMOD combined with limited torsional Monte Carlo movement was able to locate the lowest energy structures yet reported, and significantly outperformed a pure torsional Monte Carlo and a genetic algorithm-based search. Furthermore, we also demonstrate the utility of *l*-LMOD combined with random translation/rotation of a ligand for the extremely difficult problem of docking flexible ligands into flexible protein binding sites on a system that includes 9-deaza-guanine-based inhibitors docked

*Correspondence to:* I. Kolossváry; e-mail: Istvan.Kolossvary@pharma.Novartis.com, or Wayne C. Guida; e-mail: guidawc@eckerd.edu

\*Present address: Schrödinger, Inc., 1 Exchange Place, Jersey City, New Jersey 07302, or Eckerd College, St. Petersburg, Florida 33711

**Keywords:** low-mode search; conformational analysis; mode following; saddle point search; flexible docking

## Introduction

One of the most successful models used in computational chemistry is the potential energy surface (PES).<sup>2</sup> The PES is a remarkable graphical representation of individual molecules, molecular complexes, and conformational interconversions, and allows for an astonishingly simple link between key chemical concepts and basic geometrical features of surfaces. In the context of conformational analysis, minima and saddle points on the PES are the most important features. Minima correspond to minimum energy conformations and saddle points correspond to transition states associated with conformational interconversions. Saddle points, just like passes in a mountain area, provide the easiest way to pass from one valley to another. On the PES, saddle points, therefore, represent transition states providing the lowest energy barriers for conformational interconversions. The goal of conformational analysis is to find the set of interrelated energy minima and saddle points revealing the network of conformational interconversions (NCI).<sup>3</sup> The NCI describes the dynamics of concerted atomic motion characteristic of the conformational behavior of a molecule. Of course, the interesting conformations are those of sufficiently low free energy to be populated at reasonable temperatures.

We recently argued that one could utilize a procedure that relies on the ravine-following concept, often termed mode following (MF) or eigenvector following (EF), for conformational searching.<sup>1</sup> Thus, one could initiate the search by starting with any local minimum. By using one of the MF/EF techniques, one could locate a saddle point associated with this minimum and then the other minimum associated with this saddle point. The MF/EF techniques essentially follow a selected ravine uphill to a saddle point. By passing through the saddle point and applying a local energy minimization (EM) method, a second minimum could be located on the other side of the saddle. By subsequent application of MF/EF to the second minimum or to a different ravine of the first minimum, additional

minima could be located that could then be used to find additional saddles, etc., eventually mapping out the entire NCI.

It should be noted that the ravine-following concept (applied to conformational analysis) was first suggested by Crippen and Scheraga.<sup>4</sup> They named it the method of gentlest ascent and reasoned that “a procedure which suggests itself is one in which a local minimum is found in some relatively low-energy region, and then a path is selected which leads away from this minimum, but along which the energy rises as little as possible, eventually passing through a low-lying saddle point (“mountain pass”); subsequent minimization of  $E$  leads to the next local minimum.” They also pointed out that “the procedure of departing from the local minimum by an easy route can then be repeated until a whole sequence of adjacent minima have been found.” To our knowledge, however, their idea has not been explored any further or used for conformational searching.

A rather special application of the ravine-following concept has been used recently for global minimum searching. Nakamura et al.<sup>5</sup> used a combination of the diffusion equation method<sup>6</sup> (DEM) and MF/EF to locate the global minimum energy conformation of a small DNA fragment and an oligopeptide. DEM proceeds by flattening out the PES by analogy of the physical diffusion process, leaving only a single energy minimum well, which is the global minimum on the flat DEM surface. Reverse diffusion is then applied to map the DEM global minimum back to the original PES. There is no guarantee, however, that this procedure finds the PES global minimum in one step. In practice, diffusion and reverse diffusion is applied iteratively until no further progress can be achieved in lowering the energy on the PES. Reverse diffusion operates on a series of intermediate surfaces between the single-well DEM surface and the original PES. Nakamura et al.<sup>5</sup> showed that a simplified MF/EF technique, similar to our low-mode search procedure (LMOD),<sup>1</sup> can be applied during the reverse diffusion process to explore the nearest neighborhood of the back-propagating DEM global minimum to make sure that the PES global minimum does not remain hidden. For general conformational searching,

however, we believe that the LMOD search procedure represents the first practical application of this method for conformational searching. In this article we describe an in-depth analysis of the method, discuss improvements to the method, and illustrate its utility for a difficult conformational search problem and for the flexible docking problem.

## ALGORITHMS

LMOD has two fundamentally different working modes, one in which the ravine-following concept is fully utilized involving curvilinear atomic movement (*c*-LMOD), and one in which strictly linear atomic displacement is applied (*l*-LMOD) according to the original implementation of LMOD. Curvilinear LMOD is the first conformational search method designed for NCI exploration. It should be stressed, however, that the use of *c*-LMOD is an “overkill,” and it is not recommended if only minimum energy conformations are sought for which *l*-LMOD is the method of choice. Nonetheless, *c*-LMOD is an excellent tool for visualizing conformational interconversions, for which the location of saddle points is essential. The *c*-LMOD algorithm is presented here, with technical details given in parentheses.\* The associated MF/EF procedure is described separately.

### *c*-LMOD

An arbitrary molecular structure is energy minimized, and its low-mode Hessian eigenvectors (associated with the lowest eigenvalues) are computed.<sup>7†</sup> The low-mode eigenvectors are closely related to the low-frequency vibrational modes of a molecule primarily responsible for conformational interconversions.

### *c*-LMOD PROCEDURE

Repeat steps 1–3 until convergence.<sup>‡</sup>

1. Select a single low-mode or multiple low modes<sup>8</sup> of the starting structure and follow the

\**c*-LMOD and *l*-LMOD were implemented in MacroModel/BatchMin 6.5 developed in Professor Clark Still's laboratory at Columbia University, New York, NY 10027.

†The low-mode eigenvectors are computed with the inverse power method, also known as the inverse iteration method, using the EISPACK library.

‡Convergence can be defined in many different ways, for example, by limiting the number of minima to be found, but convergence criteria are left to the user, so it is not an integral part of the *c*-LMOD algorithm.

corresponding Hessian eigenvector(s) uphill, seeking a saddle point. If MF/EF can find a saddle point, proceed to 2, otherwise return to the minimum and try other low mode(s) or follow the same mode(s) in the opposite direction.<sup>§</sup>

2. Compute the Hessian at the saddle point. The Hessian should have at least one negative eigenvalue, two or more if MF/EF converged to a higher order saddle point. Pass through the saddle point along a “negative” eigenvector, to the other side of the energy barrier. The resulting structure has slightly lower energy than the saddle point structure, and subsequent energy minimization will usually carry it down to the bottom of the neighboring minimum-energy well.
3. Save the new minimum-energy structure, select a new starting structure,<sup>9</sup> compute its Hessian, and return to 1.<sup>¶</sup>

The MF/EF procedure employed is a simple implementation of a mode-following saddle point search.<sup>10</sup> Mode following is initiated by a short move along a selected low mode of the Hessian matrix of the starting structure in point 1 above. At the new point (slightly higher than the minimum energy point) the Hessian is reevaluated, and its eigenvectors are calculated. The eigenvector, which is most similar to the starting low-mode eigenvector, i.e., which has the largest overlap with it, is selected as the degree of freedom to be maximized. Maxi-

<sup>§</sup>In principle, some safeguard could be included in the algorithm to prohibit getting stranded in a minimum well in case MF/EF cannot find a saddle point. However, we have not encountered a single such case in numerous tests.

<sup>¶</sup>Although in most cases only low-energy conformations are sought, the minimum-energy structure in step 2 is always kept, because low-energy conformations are often interconnected through high-energy conformations on the PES. Therefore, rejecting high-energy structures in step 2 can easily result in a truncated NCI, potentially isolating and never finding certain low-energy conformations. The most straightforward structure selection is to always use the last energy-minimized structure as the next starting structure. This is essentially a random walk procedure, which mimics the behavior of real molecules to some extent. However, the so-called usage-directed search<sup>9</sup> is significantly more efficient. Random walk explores the NCI following a single path, whereas usage-directed search proceeds along multiple paths, much like putting together a jigsaw puzzle of the NCI. We found that the use of fast inverse iteration routines in EISPACK allowed us to calculate the low-mode eigenvectors “on the fly.” The direct approach for calculating eigenvectors has turned out to be significantly more efficient than the old implementation of LMOD where the eigenvectors were calculated only once for a particular conformation, but had to be stored and then recovered every time they were needed.

mization is accomplished by taking a short FMNR step (see below) in the local coordinate system defined by the eigenvectors of the new Hessian, but reversing the sign of the eigenvalue of the selected eigenvector. This procedure is continued, iteratively, always changing the sign of the eigenvalue of the ravine eigenvector to be negative, while keeping the sign of all the remaining eigenvalues positive, until convergence to a saddle point is achieved.

The mathematical foundation of MF/EF is based on a coordinate transformation applied to the well-known full-matrix Newton–Raphson (FMNR) minimization formula. The FMNR search step ( $\Delta\mathbf{x}_k$ ) in the  $k$ th iteration can be expressed by a simple formula in a local coordinate system spanned by the eigenvectors of the current Hessian matrix ( $\mathbf{H}_k$ ):

$$\Delta\mathbf{x}_k = -\mathbf{g}_k \mathbf{H}_k^{-1} = -\sum_{i=1}^N \frac{\mathbf{u}_i^{(k)\top} \mathbf{g}_k}{\lambda_i^{(k)}} \mathbf{u}_i^{(k)} \quad (1)$$

where  $\lambda_i^{(k)}$  and  $\mathbf{u}_i^{(k)}$  denote, respectively, the eigenvalues and associated eigenvectors of the current Hessian  $\mathbf{H}_k$ ,  $\mathbf{u}_i^{(k)\top} \mathbf{g}_k$  is the projection of the current gradient vector to  $\mathbf{u}_i^{(k)}$ , and  $N$  is the number of degrees of freedom. Equation (1) clearly illustrates that FMNR maximizes the energy along the eigenvectors with negative eigenvalues, and minimizes the energy along the eigenvectors with positive eigenvalues. It should be noted that a projection method based on the Eckart constraints is preferably used with eq. (1) to project out external translation/rotation from the Hessian.<sup>11</sup> The Eckart constraints are expressed in the form of an  $N \times 6$  matrix  $\mathbf{C}$ , whose orthogonal columns are unit vectors of atomic displacements corresponding to the three independent, external translations, and three infinitesimal rotations with respect to the centroid of the molecule. The  $N \times N$  projection matrix  $\mathbf{P} = \mathbf{C}\mathbf{C}^\top$  is employed to project out external translation/rotation from the Hessian by applying  $\mathbf{H}_p = \mathbf{P}^\top \mathbf{H} \mathbf{P}$ . Note, however, that for the sake of simplicity the following simple algorithmic implementation of mode following applied in step 1 above is described without explicit reference to the projected  $\mathbf{H}_p$ .

The MF/EF algorithm is as follows:

1. Select one of the low-mode eigenvectors\*  $\mathbf{u}_s$  (small eigenvalue) of the starting structure and

\*For the sake of simplicity, this algorithmic scheme refers to only a single  $\mathbf{u}_s$ . However, the implementation of c-LMOD in MacroModel/BatchMin 6.5 includes the option of employing Gotō's frontier mode-following method<sup>8</sup> where multiple  $\mathbf{u}_s$  vectors are used in exactly the same way as described for a single  $\mathbf{u}_s$ .

move uphill along this (or the negative) direction to a short distance. Let this point be the starting point  $\mathbf{x}_0$  of the mode-following search.

2. Evaluate the gradient and the Hessian at  $\mathbf{x}_0$ , and compute the eigenvalues and eigenvectors of the Hessian. Find the eigenvector  $\mathbf{u}_s^{(0)}$ , which has the largest overlap with  $\mathbf{u}_s$ .
3. Apply eq. (1) to calculate  $\Delta\mathbf{x}_0$ , but in the denominator substitute  $|\lambda_i^{(0)}|$  for all  $i \neq s$  and  $-|\lambda_s^{(0)}|$  for  $i = s$  to make sure that a maximization step is applied in the direction of  $\mathbf{u}_s^{(0)}$  and minimization along all the remaining eigenvectors.
4. Limit the step size to some maximum coordinate movement and move along  $\Delta\mathbf{x}_0$  to  $\mathbf{x}_1$ . Note that sophisticated MF/EF algorithms in quantum chemical programs utilize a complicated procedure termed trust radius algorithm to determine an ideal step size. However, for molecular mechanics calculations, the application of a simple absolute limit is just as useful.<sup>3a</sup>
5. Reevaluate the gradient and the Hessian at  $\mathbf{x}_1$ , and compute the eigenvalues and eigenvectors of the Hessian. Find the eigenvector  $\mathbf{u}_s^{(1)}$ , which has the largest overlap with  $\mathbf{u}_s^{(0)}$ .
6. Apply eq. (1) to calculate  $\Delta\mathbf{x}_1$ , and substitute  $|\lambda_i^{(1)}|$  for all  $i \neq s$  and  $-|\lambda_s^{(1)}|$  for  $i = s$  in the denominator to ensure that a maximization step is applied in the direction of  $\mathbf{u}_s^{(1)}$  and minimization along all the remaining eigenvectors.

$$\Delta\mathbf{x}_1 = -\sum_{i \neq s}^N \frac{\mathbf{u}_i^{(1)\top} \mathbf{g}_1}{|\lambda_i^{(1)}|} \mathbf{u}_i^{(1)} - \frac{\mathbf{u}_s^{(1)\top} \mathbf{g}_1}{-|\lambda_s^{(1)}|} \mathbf{u}_s^{(1)}$$

7. Step along  $\Delta\mathbf{x}_1$  uphill to a distance determined by the step-size limit, to  $\mathbf{x}_2$ . And so on . . . until  $|\mathbf{g}_k|^2 \leq \epsilon$ . Upon convergence, check the number of negative eigenvalues of the Hessian to determine the character of the saddle point.

Figures 1 and 2, respectively, depict two different ravine-following paths generated by the mode-following algorithm on a simple test surface  $f(x, y) = (x^2 + y - 11)^2 + (x + y^2 - 7)^2$ .<sup>10</sup> The contour surface representation clearly shows the presence of four minima, four saddle points, and a single maximum in the middle. The two different paths start from the same minimum following the two different eigenvectors of the local Hessian matrix. Note

that the two different paths start out orthogonally, but eventually are driven by the shape of the surface to the same saddle point. Also note that a three-dimensional version of Figure 2 is used below to illustrate the basic idea of the *l*-LMOD algorithm.

### ***l*-LMOD**

The basic tenet of the linear LMOD procedure is to utilize a “brute force” approach based on the mode-following concept.<sup>1</sup> Instead of precisely locating a saddle point to cross the barrier between two minima, *l*-LMOD in essence forces the molecule through the barrier by a single leap along one of the low-mode eigenvectors of the starting structure in step 1 of the *c*-LMOD algorithm above. Although there is no guarantee that subsequent energy minimization will, in fact, cross a potential barrier and reach another minimum, it was demonstrated that it crossed barriers a majority of the time. Of course, it can happen that energy minimization will carry the trial structure to a minimum that is not connected to the minimum used to initiate the linear mode-following procedure; but essentially, *l*-LMOD typically focuses the search to the local neighborhood of a minimum on the PES. This situation is demonstrated for the test surface on Figure 3, which is a three-dimensional version of Figure 2 for better visualization of the topography of the surface. The three-dimensional perspective on Figure 3 clearly shows that the linear extension of one of the eigenvectors of the local Hessian crosses or rather “tunnels” the barrier, and emerges in the diagonally opposite minimum well. Note, however, that the *l*-LMOD step finds a minimum, which is not necessarily directly connected to the one used to initiate the search.

The simple *l*-LMOD algorithm is as follows. An arbitrary molecular structure is energy-minimized and its low-mode Hessian eigenvectors are computed in the same way as in the *c*-LMOD algorithm.

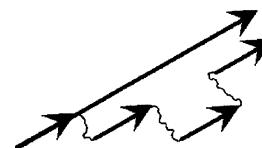
### ***l*-LMOD PROCEDURE**

Repeat steps 1–3 until convergence.

- 1a. Select a single low mode or multiple low modes of the starting structure. Move along the eigenvector of the single mode or along a vector comprised of a randomly generated linear combination of the selected multiple modes by linear displacement, in a single leap. The resulting  $3N$ -dimensional leap vector, which corresponds to the two-dimensional leap vector in Figure 3 is scaled

to move the fastest moving atom of the selected mode(s) to a randomly chosen distance in three-dimensional space in either of the two opposite directions between a user-specified minimum and maximum value.

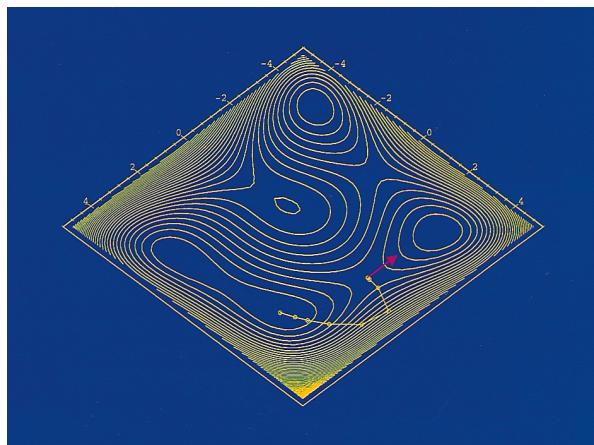
- 1b. An alternative move in *l*-LMOD is based on a similar concept to the SHAKE algorithm often used in molecular dynamics.<sup>12</sup> Note that *l*-LMOD often causes bonds to stretch to unreasonable lengths because of the strictly linear displacement along a long leap vector. As it turns out, this is not a real problem, because subsequent EM quickly restores the bond lengths to their original values. Nonetheless, we implemented a procedure in *l*-LMOD that does not allow bonds to stretch significantly during the *l*-LMOD move. The simple procedure involves splitting the single leap into multiple steps and applying a few steps of steepest descent minimization (SD) between them, as schematically shown below (Scheme 1):



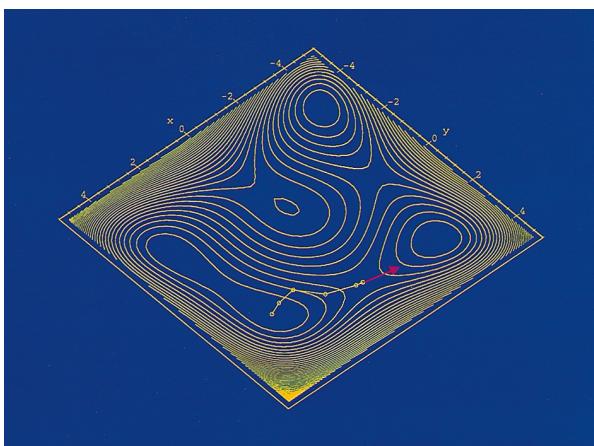
**SCHEME 1.**

where a single *l*-LMOD leap is split into four smaller steps, each being parallel to the long-leap vector. Each of the first three moves are followed by a few steps of SD minimization symbolized by the thin, wavy connecting lines. Note that the tip of the long-leap vector and the tip of the last short-leap vector in this alternative procedure do not coincide on the PES. The corresponding structures, however, are fairly similar in terms of their main structural features. The main difference is that the latter structure is devoid of excessive bond lengths due to the beneficial effect of SD, which usually restores overly stretched bonds and/or bond angles in only a few iterations.

2. Energy minimize the structure corresponding to the tip of either the long-leap vector or the last short-leap vector.
3. Save the new minimum-energy structure, select a new starting structure, compute its Hessian, and return to 1. Note that *l*-LMOD employs the usage-directed search protocol and only saves the low-energy structures.

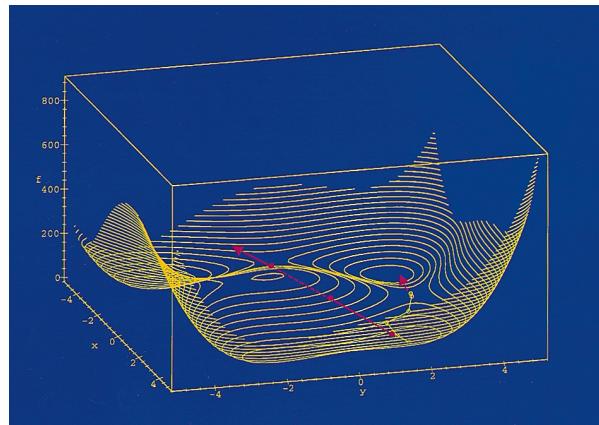


**FIGURE 1.** Contour plot of the Culot test surface<sup>10</sup> with one of the two ravine-following paths shown, starting at the minimum (3.0, 2.0) and ending at the saddle point (0.0867, 2.88). The red arrow shows the downhill direction along the “negative” eigenvector of the Hessian at the saddle point. Reproduced with permission from Gy. Keserü, I. Kolossváry; Molecular Mechanics and Conformational Analysis in Drug Design, Copyright Blackwell Science, Oxford, 1999.

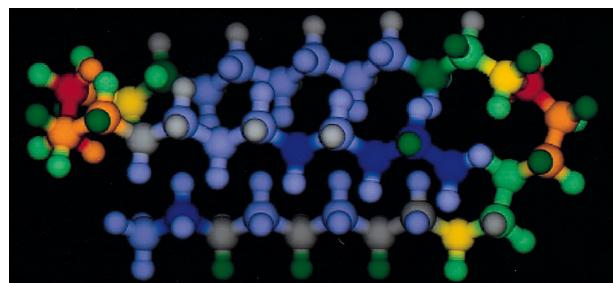


**FIGURE 2.** Contour plot of the Culot test surface<sup>10</sup> with the alternative ravine-following path starting at the minimum (3.0, 2.0) and ending at the saddle point (0.0867, 2.88). The red arrow shows the downhill direction along the “negative” eigenvector of the Hessian at the saddle point. Reproduced with permission from Gy. Keserü, I. Kolossváry; Molecular Mechanics and Conformational Analysis in Drug Design, Copyright Blackwell Science, Oxford, 1999.

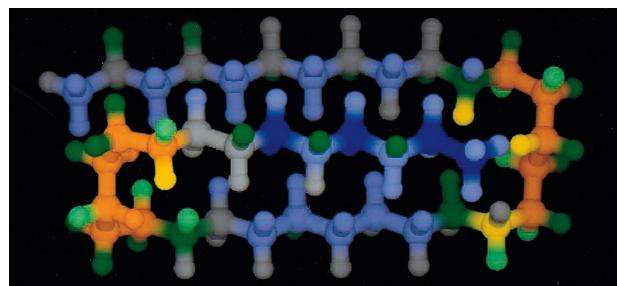
The *l*-LMOD algorithm presented here is essentially identical to an earlier implementation of the method in MacroModel/BatchMin 6.0,<sup>13</sup> including perturbation along a mixture of low-mode eigenvectors, but the application of the alternative structure



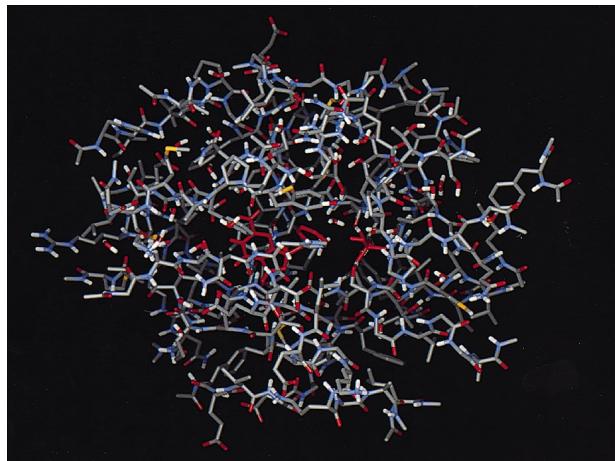
**FIGURE 3.** Three-dimensional contour plot of the Culot test surface<sup>10</sup> showing the ravine-following path in Figure 2. The long, red arrow shows the *l*-LMOD leap from the minimum at (3.0, 2.0), through the “tunnel,” and emerging on the other side of the barrier. Reproduced with permission from Gy. Keserü, I. Kolossváry; Molecular Mechanics and Conformational Analysis in Drug Design, Copyright Blackwell Science, Oxford, 1999.



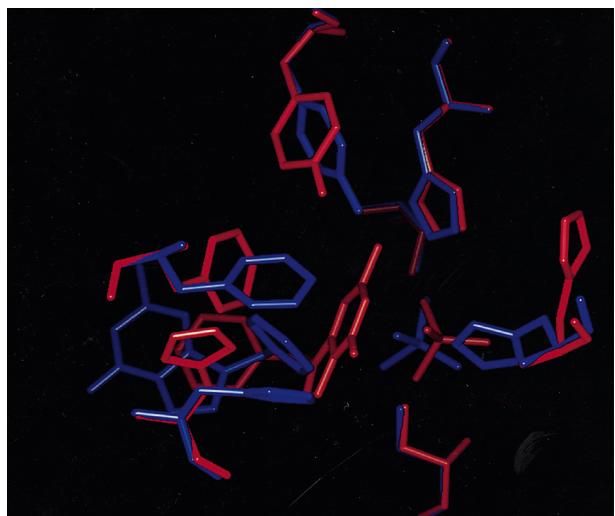
**FIGURE 4.** Global minimum of C<sub>39</sub>H<sub>80</sub> found by *l*-LMOD,  $E = 53.313$  kJ/mol. Colors represent the energy distribution (energy scale: red/stressed  $\rightarrow$  blue/relaxed).<sup>22</sup> Reproduced with permission, Copyright J. M. Goodman, 1998.<sup>19</sup>



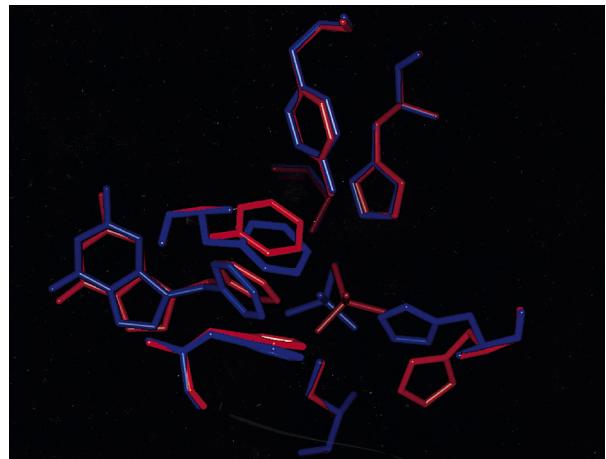
**FIGURE 5.** Highly symmetrical, low-energy conformation of C<sub>39</sub>H<sub>80</sub> found by *l*-LMOD,  $E = 54.716$  kJ/mol. Colors represent the energy distribution (energy scale: red/stressed  $\rightarrow$  blue/relaxed). Reproduced with permission, Copyright J. M. Goodman, 1998.



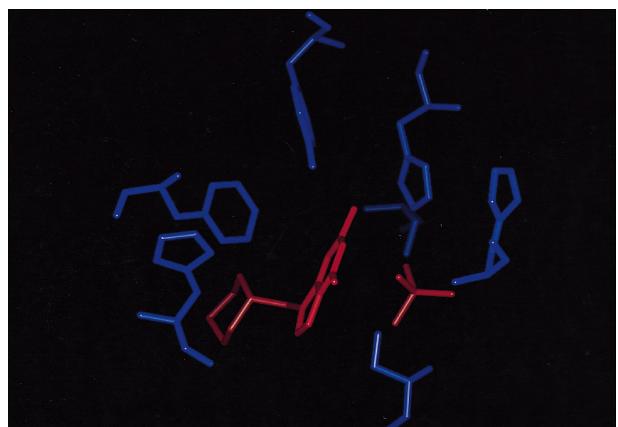
**FIGURE 6.** PNP/9-benzyl-9-deazaguanine model used for flexible docking studies. Inhibitor and phosphate are shown in red. The surrounding 12-Å shell of residues plus water molecules that constitute the model are colored by atom type.



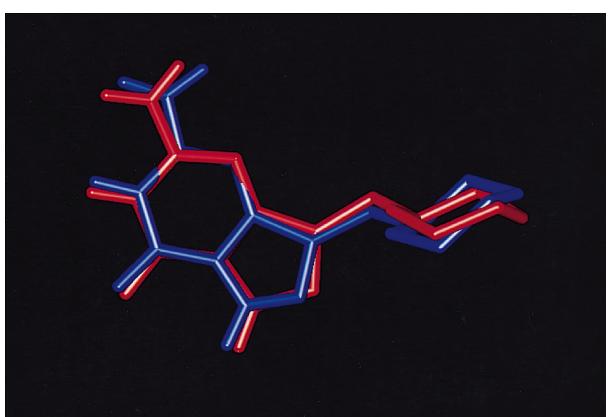
**FIGURE 7.** X-ray structure of the PNP/9-benzyl-9-deazaguanine complex superimposed with the starting structure for the LMOD docking calculation. The X-ray structure is shown in blue, and starting geometry for the docking study is shown in red. For clarity, only the inhibitor, phosphate, and moving residues are shown.



**FIGURE 8.** X-ray structure of the PNP/9-benzyl-9-deazaguanine complex superimposed with the global minimum energy structure for the LMOD docking calculation. The X-ray structure is shown in blue, and global minimum is shown in red. For clarity, only the inhibitor, phosphate, and moving residues are shown.



**FIGURE 9.** Starting structure for the LMOD docking calculation performed on the PNP/9-(cyclohexyl-methyl)-9-deazaguanine complex. The position of the amino acid side chains is identical to the starting structure show in Figure 7. The inhibitor and phosphate are shown in red, and the binding site is shown in blue. For clarity, only the inhibitor, phosphate, and moving residues are shown.



**FIGURE 10.** Superimposition of the crystallographically observed conformation with the seventh lowest energy conformation found by the LMOD search. The X-ray structure is shown in blue, and the docked conformation is shown in red.

perturbation involving SD is a new feature in Version 6.5,\* and also allows for intermediate storage of structures generated along the LMOD path via a BatchMin debug switch. This feature, combined with MacroModel's movie option, is an excellent visual tool for studying conformational interconversions with either *l*-LMOD or, especially, *c*-LMOD.

In all of the following test calculations a truncated Newton conjugate gradient minimizer<sup>14</sup> was applied (TNCG option in MacroModel) that uses an analytical Hessian recalculated at each minimization step. The minimization termination criterion was set to 0.01 kJ/mol Å gradient RMS.

## TEST ON CYCLOHEPTADECANE

Cycloheptadecane has been used as a popular test case for conformational searching.<sup>15</sup> Locating its 262 low-energy conformations within a 3-kcal/mol energy window above the global minimum is still a challenging problem for any conformational search method. The number 262 has been recently called into question by Ngo and Karplus,<sup>16</sup> who reported the existence of an additional conformation. Our first study was aimed at verifying the existence of a 263rd conformation. In our hands, only two methods, LMOD and the systematic unbounded multiple minimum search method (SUMM) by Goodman and Still<sup>17</sup> were able to find 262 conformations of cycloheptadecane within the lowest 3 kcal/mol.<sup>1</sup> Our original LMOD procedure, which is virtually identical to *l*-LMOD was reported to find 262 cycloheptadecane conformations on the MM2 PES within a 12.6 kJ/mol (~3 kcal/mol) energy window in 11,631 search steps, whereas SUMM required 16,851 steps to find the same conformations.<sup>1</sup> However, it should be noted that those results were somewhat misleading; 12.6 kJ/mol is, in fact, slightly higher than 3.00 kcal/mol and, as it turns out, there are at least two additional conformations within the 12.6 kJ/mol energy window. Two independent *l*-LMOD and SUMM searches were run, each seeded with the 262 previously found conformations, to test whether additional conformations can be found in 100,000 step searches, respectively. This time *l*-LMOD and SUMM both found two additional conformations, i.e., 264 cycloheptadecane conformations altogether within the lowest 12.6 kJ/mol (*l*-LMOD and SUMM both found the same 264 structures). This means that our original results<sup>1</sup> inadvertently represented searches that had been aborted prematurely. The

*l*-LMOD and SUMM searches were now repeated with a 12.56 kJ/mol (3.002 kcal/mol) energy window, which corresponds to the 3.002 kcal/mol window used by Ngo and Karplus. *l*-LMOD found the original 262 structures<sup>15</sup> in 27,848 search steps, and SUMM found the same 262 structures in 53,652 steps. The *l*-LMOD search was applied using random mixtures of the first 15 nontrivial modes for structure perturbation, and the distance range for the fastest moving atom was set to 3–8 Å. The *l*-LMOD search took 7.75 CPU hours, and the SUMM search took 13.5 CPU hours on a SGI R10K processor.

Of course, the question remained whether there was an additional conformation missing. Therefore, an additional 500,000 step *l*-LMOD and SUMM searches were run with the 12.56 kJ/mol energy window to search for the 263rd cycloheptadecane conformation. However, neither searches located additional minima on the MM2 PES (infinite nonbonded cutoffs), and each and every known cycloheptadecane conformation was found more than 40 times (several hundred times on average). Unfortunately, we could not obtain the structure of the 263rd conformation reported by Ngo and Karplus<sup>16</sup> from the authors, and therefore, can neither verify nor disprove their results. Structures 261–263 from the 12.6 kJ/mol search were reminimized below 10<sup>−6</sup> kJ mol<sup>−1</sup> Å<sup>−1</sup> RMS gradient, and their double precision energy has been found to be 2.999766, 3.001557, and 3.004018 kcal/mol above the global minimum (19.093039 kcal/mol), respectively.

Let us turn now the focus on *c*-LMOD and the very ambitious test problem of locating all of the 262 original cycloheptadecane conformations by exploring the NCI. It should be stressed that this test was not conducted to determine the performance of *c*-LMOD relative to other conformational search methods. Rather, this scientific test was conducted to ascertain whether it was possible to conduct a conformational search on a complex molecular system using an algorithm that relies upon the precise implementation of the mode-following concept. The *c*-LMOD search was initiated with the global minimum energy structure, and the usage-directed search protocol was applied for structure selection.<sup>†</sup>

<sup>†</sup>We also tried the random walk search, but it failed miserably. Several random walk searches were started from different conformations, but the searches failed without exception to find more than about a hundred of the 262 conformations. Analysis of the search paths revealed that random walk was extremely prone to get trapped in a local region of the PES by revisiting the same set of minima many times over via different saddle points, without any progress in terms of finding new low-energy conformations.

\*Available from Schrödinger, Inc., 1500 SW First Avenue, Suite 1180, Portland, OR 97201; <http://www.schrodinger.com> (1999).

No energy window was applied to intermediate saddle point structures and local minima, but the progress of the search was monitored by counting the number of unique minimum energy conformations within the 12.56 kJ/mol energy window above the global minimum. The search was terminated when all of the 262 original cycloheptadecane conformations were found. The saddle point searches were completed utilizing Gotō's frontier mode-following method<sup>8</sup> applied to the lowest five nontrivial modes. Algorithmic details were given above. The search statistics are as follows. The full search was completed in a total of 116,315 steps, which took 641 CPU hours (almost 4 weeks) on a single SGI R10K processor. 89,508 saddle points and 42,130 unique minima were found. The 262 target conformations were found 33 times on average. It should also be noted that *c*-LMOD failed in only 26,808 attempts to converge to a saddle point, which is a remarkable 77% success rate in the field of saddle point searching. *c*-LMOD has been proven to be particularly successful in this extremely difficult test, and we believe that it should find wide utility for complex conformational searches aimed at studying conformational interconversions. It is instructive to note that utilization of the "DEBG 920" BatchMin V6.5 debug switch will store all of the intermediate structures (including minima, saddle points, and structures generated along the mode-following path moving toward a saddle point) during the *c*-LMOD search in the output file. The snapshots in the output file can be animated using MacroModel's "movie" or "autoread" options. The resulting cartoon provides a spectacular visual perception of the conformational interconversions.

### C<sub>39</sub>H<sub>80</sub>

It has recently been shown that the largest *n*-alkane with a linear, all-*trans* global minimum (MM2) energy conformation is C<sub>17</sub>H<sub>36</sub>.<sup>18</sup> Longer hydrocarbon chains prefer twists and turns due to the increasing flexibility that allows distant methylene groups to be pulled together by stabilizing van der Waals interactions. In fact, long *n*-alkane chains coil up into spectacular, folded structures, and studying this process might give important clues to the famous protein-folding problem.<sup>19</sup> Nair and Goodman<sup>20</sup> have found several folded conformations of C<sub>39</sub>H<sub>80</sub> with lower energy than the all-*trans* conformation (102.55 kJ/mol MM2 energy) utilizing a genetic algorithm. Interestingly, however, chemists were able to construct folded conformations of C<sub>39</sub>H<sub>80</sub> using Dreiding models that were

significantly lower in energy than those generated by the computer.

The conformational space available to C<sub>39</sub>H<sub>80</sub> is vastly complex. It is amazing that the chemist's eye could easily outperform the computer. Therefore, we took up the challenge and tested the low-mode procedure to see whether it could find even lower energy conformations of C<sub>39</sub>H<sub>80</sub>. A known shortcoming of the *l*-LMOD algorithm prompted the use of a mixed-mode search strategy. For low modes, which represent almost pure torsional rotations, application of the strictly linear movement along the leap vector has special consequences. In particular, the "linearized" torsional rotation represented by such low modes is invariably associated with extensive bond stretching. The larger the rotational angle, the longer the leap vector and the more extensive the bond stretching. In fact, the upper limit of rotation by linear movement is 90°, nevertheless, with the high cost of infinite bond stretching. This problem was analyzed in detail from a different perspective by Kolossváry and McMartin.<sup>21</sup> Here we want to stress that such torsional low modes are typical for long-chain molecules and, therefore, *l*-LMOD was used in conjunction with pure torsional Monte Carlo search.

The mixed-mode search strategy involved the application of either *l*-LMOD or torsional Monte Carlo (as implemented in the MCMM—Monte Carlo Multiple Minimum—command in Macro-Model/BatchMin) for structure perturbation. This means that structures selected for perturbation by the usage-directed protocol during any stage of the search were subjected to either a *l*-LMOD step or a MCMM step. The *l*-LMOD step was applied using leap vectors comprised of different random mixtures of the first 10 nontrivial low modes, and the fastest moving atom with respect to a particular leap vector was displaced by a randomly selected distance along the leap vector between 5 and 10 Å. Bond lengths were kept within reasonable limits by the application of partial SD minimization along the leap vector as described in step 1b of the *l*-LMOD algorithm above. The MCMM step involved the simultaneous rotation of up to three randomly selected torsion bonds that were allowed to vary by 0 to 180 degrees. It should be noted that MCMM was utilized only for "bridging" distant regions of the PES that can be reached in a single long jump by simultaneous, >90° torsional rotation about a few backbone bonds. There was no attempt made to rotate more than three bonds and, thus, trying to explore the entire PES via MCMM because preliminary tests clearly showed that MCMM

alone could not find a single C<sub>39</sub>H<sub>80</sub> conformation lower in energy than the all-*trans* conformation (102.55 kJ/mol) starting from an arbitrary, high-energy structure. Preliminary tests also verified that although *l*-LMOD alone was, in fact, capable of traversing large distances on the PES, it required numerous intermediate steps, and was quite expensive in terms of CPU time. The mixed-mode strategy was, therefore, designed to combine the advantages of *l*-LMOD and MCMM, using them for what they do best, i.e., using *l*-LMOD for local exploration and using MCMM for locating distant, low-energy regions on the PES.

Several 10,000 step searches were run from the same high-energy starting structure (148.66 kJ/mol) with different random seeds to gain an appreciation for the complexity of the problem. Unique conformations were stored only if they were within the lowest 15 kJ/mol. The searches took approximately 18–20 CPU hours on a single SGI R10K processor. The global minimum for most of the searches was in the energy range of 55–70 kJ/mol, which is significantly better than any previous computer generated model, but still higher than the “man-made” global minimum with an energy of 54.744 kJ/mol.<sup>19</sup> However, two searches resulted in global minima with an even lower energy of 54.716 and 53.313 kJ/mol, respectively (energies were calculated with infinite nonbonded cutoffs). Note that the term global minimum is used here as a synonym for the lowest energy structure found during a particular search. We still do not know whether we have found the true gas phase global minimum.

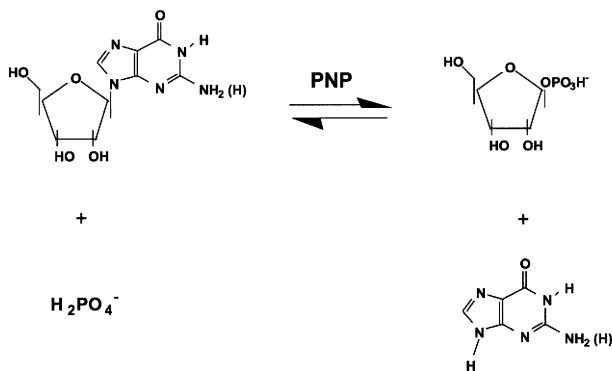
A final 500,000 step search was run starting with our lowest energy structure so far to provide additional computational support to its “title” as the global minimum. The search was run on five SGI R10K processors simultaneously using distributed BatchMin, and took a total of 158 CPU hours. The search revisited the starting structure 191 times and located a total of 631 conformations within the lowest 15 kJ/mol, but could not find a new global minimum. It did, however, find one more structure with lower energy than the previous man-made global minimum. Thus, the computer found three structures that were “better” than the one found by the human experts. The triumvirate of those structures consists of the current global minimum of C<sub>39</sub>H<sub>80</sub> with an energy of 53.313 kJ/mol, followed by two computer models with energies of 54.582 and 54.716 kJ/mol, respectively. The global minimum is shown in Figure 4, and the highest energy structure of the triumvirate is shown in Figure 5. The figures were reproduced with permission

from Goodman's<sup>19</sup> web site. The colors were generated with the MMCOL program, and represent the energy distribution within the molecule (energy scale: red/stressed → blue/relaxed).<sup>22</sup> Note the interesting difference between the structures. The higher energy structure in Figure 5 has a high degree of symmetry, and the energy is evenly distributed throughout the molecule with some stress in the hairpin turns. The global minimum structure in Figure 4, however, is less symmetrical, and it has a fair amount of stress in the hairpin turns, yet it is lower in energy than the symmetrical structure by 1.403 kJ/mol. Two questions arise immediately with answers yet to be given. One, why did the expert chemists not find the symmetrical structure? Two, is it conceivable that the symmetrical structure is, in fact, the global minimum, but some artifact in the MM2 force field biases the energetic order? Note, for example, that MM3<sup>23</sup> and the Merck Molecular Force Field (MMFF)<sup>24</sup> both reverse the energetic order of the structures in Figures 4 and 5, rendering the symmetrical structure lower in energy than the asymmetrical structure by 1.27 and 2.02 kJ/mol, respectively.

## FLEXIBLE DOCKING

We next turned our attention to the problem of docking a flexible ligand into a flexible binding site. We reasoned that in nature such a docking maneuver occurs with highly coupled motion involving low-frequency vibrations of, primarily, active site protein residues, and that LMOD should be capable of capturing the essential features of this physical process. Flexible docking is crucial for structure-based drug design,<sup>25</sup> and it has only recently been addressed by contemporary methodology. Most of the current flexible docking methods only take ligand flexibility into consideration.<sup>26</sup> Other methods have been developed in which limited binding site flexibility is taken into account.<sup>27</sup> Recently, methods have begun to emerge in which both ligand and protein flexibility is considered. For example, Leach has described a method in which the amino acid residues in and surrounding the binding site are allowed to assume different discrete rotameric states during sampling.<sup>28</sup> Moreover, Caflisch et al. have recently described a method in which random starting configurations are subjected to a minimization procedure in which nonbonded interactions are gradually switched on followed by Monte Carlo minimization of the lowest energy structures.<sup>29</sup>

We decided to test the LMOD procedure (*l*-LMOD, to be precise) for the flexible dock-

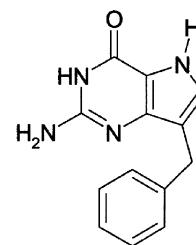
**SCHEME 2.**

ing (flexible ligand and flexible binding site) of inhibitors of purine nucleoside phosphorylase (PNP).<sup>30</sup> This enzyme has been a target for structure-based drug design, its crystal structure both in the apo form and as a complex with numerous inhibitors having previously been determined. PNP catalyzes the reversible phosphorylation of guanosine and inosine (and their 2'-deoxy analogs), and is involved in the purine salvage and catabolism biochemical pathways. The human enzyme is a trimer with identical subunits (and, thus, three identical binding sites), and a molecular weight of ~97 kDa. Each subunit contains 289 amino acid residues (Scheme 2).

Due to its size, molecular modeling of the full trimer of PNP is not feasible, and previous modeling studies used a truncated model in which a shell of residues surrounding one of the binding sites, and lying within 7 Å of the inhibitor and phosphate, was employed.<sup>31</sup> Thus, chains of intact residues were included in the modeling studies if any atom of the residue was within a 7 Å radius of any atom of either the inhibitor or the phosphate. In the present study, we also chose to employ a truncated, but larger, model of PNP.

#### COMPUTATIONAL DETAILS FOR FLEXIBLE DOCKING

We constructed a model for our computational studies from the structure of the bovine PNP/9-deaza-9-benzylguanine complex determined crystallographically to 1.9 Å resolution. Starting with the trimer and constructing the model from one of the binding sites, our initial model consisted of the inhibitor, phosphate, and a shell of residues within 12 Å from either the inhibitor or phosphate. This model includes residues from the adjacent subunit, some of which are known to affect binding. Preliminary results with larger shells resulted in increased



9-Benzyl-9-deazaguanine

**SCHEME 3.**

CPU time for the docking studies but no greater accuracy. All crystallographically determined water molecules, within the 12-Å shell, were included in the calculation. Figure 6 illustrates the truncated model used in the calculations (Scheme 3).

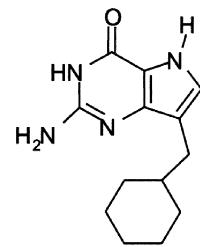
No solvation model was employed for the calculations. However, a distance-dependent dielectric "constant" of  $4r$  was used with the AMBER\* united atom force field. Explicit hydrogens were added to the benzene ring, however. During the LMOD structural perturbation, and during the subsequent energy minimization, seven residues (including backbone) were allowed to move freely. The remaining residues were treated as "frozen atoms" (i.e., they were not allowed to move; only their electrostatic and van der Waals interactions with the moving atoms were included in the calculation). The seven moving residues were: Ser 33, His 64, His 86, Tyr 88, Phe 159 (from the adjacent subunit), Ser 220, and His 257. These amino acids are known to be key residues in the binding site of PNP.<sup>30</sup> The four water molecules that are observed crystallographically to occupy the binding site in PNP were treated as "frozen" atoms. The inhibitor and phosphate were allowed to move freely during the LMOD maneuver, and the subsequent energy minimization. An additional structural perturbation, however, was applied to the inhibitor (but not the phosphate); it was subjected to explicit translation/rotation with respect to the binding site via the MOLS command available in BatchMin. The MOLS move included the simultaneous translation/rotation of the ligand by random variations between 0.1–0.5 Å and 0–90°, respectively. The LMOD perturbation was alternated with the MOLS move to ensure efficient sampling of all of the allowed binding modes. Large energy barriers make ligand translation and rotation extremely difficult for sampling methods that are affected by these high barriers like LMOD or MD. For example, in our hands, MD simulations were highly ineffective for ligand

docking. However, explicit translation/rotation of the ligand is a purely geometric operation unaffected by energy barriers, but perhaps giving rise to extremely crowded structures. In such a case, however, subsequent energy minimization generally relieves these strained structures in only a few steps. Preliminary investigations using LMOD alone showed that some orientational sampling of the ligand does occur, but in most cases we found that mixing LMOD with MOLS resulted in significantly enhanced overall sampling. This situation is comparable to coupling of the TORS command with LMOD, as previously described above for C<sub>39</sub>H<sub>80</sub>.

## Results

To initiate the LMOD docking search, the inhibitor was placed in a random conformation and oriented in the binding site in a manner that was structurally quite different from the crystallographically observed orientation. The phosphate was placed in a random orientation, and five of the seven moving residues were oriented randomly with respect to their side chain conformation. Figure 7 shows the X-ray structure superimposed with the starting structure for the docking search. We decided to attempt a complete orientational/conformational search that would locate all of the low energy binding modes. A 50,000 step search required 2.4 million CPU seconds on a single SGI R10k processor. A 15-kJ window was employed, and 149 structures were found within this energy window. The global minimum energy structure quite closely resembled the X-ray structure, particularly with respect to the placement and orientation of the inhibitor and its binding conformation. With the exception of His 64, all of the remaining residues were found to be in conformations that were very closely related to those observed crystallographically. It is known that His 64 is quite flexible, and reorients quite readily, depending upon the nature of the inhibitor and whether phosphate is present in the binding site.<sup>32</sup> Figure 8 shows the X-ray structure superimposed on the global minimum.

It is likely that even this extremely lengthy docking search was not fully converged. Although most of the structures were found more than 10 times each, a few of the structures were located only one or two times. However, it is rare for converged searches to be carried out when one is performing a docking calculation, because, generally, one is simply interested in performing a "representative" search that can locate a structure that is closely related to the one observed crystallographically. In



9-(cyclohexyl-methyl)-9-deazaguanine

**SCHEME 4.**

fact, this is the docking paradigm that is normally applied to determine whether a docking calculation is successful or not.<sup>26–29</sup> Accordingly, we performed a series of calculations to find the optimal conditions to perform a "representative" search with LMOD. Specifically, we examined the effect of utilizing a smaller shell of residues for the calculation and the effect of using reduced cutoffs. We found that for 9-benzyl-9-deazaguanine, a 6.0 Å shell, with a 3.5 Å van der Waals cutoff and a 6.0 Å electrostatic cutoff produced optimal results in terms of speed, while still being able to locate a structure closely resembling the X-ray structure in a minimum number of steps. The speed up relative to using the 12 Å shell and "standard" BatchMin cutoffs (7.0 Å VDW; 12.0 Å ELE) was >2.5. Thus, a 1000 step job, which required less than 6 CPU hours on a single SGI R10k processor, resulted in a net speed up of more than two orders of magnitude. LMOD is implemented in BatchMin in such a way that distributed BatchMin can be used. Thus, distributing a single LMOD job among multiple processors results in further speed up that scales almost linearly with the number of processors and should find significant utility for structure-based drug design, especially in the lead optimization phase.

To further test LMOD for docking of a more stringent ligand with regard to its conformational preferences, we decided to perform a docking calculation using 9-(cyclohexyl-methyl)-9-deazaguanine. In addition to possessing two rotatable bonds (like 9-benzyl-deazaguanine) this molecule can adopt either an axial or equatorial conformation with respect to the cyclohexane ring. For this PNP inhibitor we also knew the bound conformation from X-ray crystallographic studies.<sup>33, 34 \*</sup> To initiate the LMOD

\*The crystallographically observed bound conformation is based on Fourier difference maps calculated at 3.2 Å resolution for the human enzyme.<sup>33</sup> It is known that the bovine and human enzyme are highly homologous with none of the changes in sequence occurring in the active site.<sup>34</sup>

docking search, the inhibitor was placed in a random conformation with respect to rotation about the single bonds, and placed in an axial orientation with respect to the cyclohexane conformation. Note that the crystallographically observed cyclohexane conformation for this inhibitor is equatorial. The inhibitor was oriented in the binding site in a manner that was structurally quite different from the crystallographically observed orientation. The phosphate was placed in a random orientation. Figure 9 shows the starting structure for the docking search. For this search, the binding site shell and the initial position of the moving residues was identical to the one used for the docking search with 9-benzyl-9-deazaguanine. A 20,000 step LMOD/MOLS search required close to 1 million CPU seconds on a single SGI R10k processor. A 15-kJ window was employed, and 78 structures were found within this energy window, all of which possessed the equatorial orientation for the cyclohexane ring. The global minimum energy structure quite closely resembled the X-ray structure, but the best match shown in Figure was found with conformation #7 (3.7 kJ/mol above the global minimum).

## Summary

In this work we have demonstrated the utility of the LMOD procedure in difficult conformational search/docking problems. We report the first application of the rigorous implementation of mode following (*c*-LMOD) to a conformational search involving cycloheptadecane, and demonstrate that, although feasible, it is inefficient to perform a conformational search in this manner. On the other hand, if one is interested in conformational analyses involving conformational interconversions, *c*-LMOD provides an excellent tool for this purpose when used with "DEBG 920" in BatchMin for visualization. We have also refined our original LMOD procedure (*l*-LMOD). For the very difficult conformational search performed on C<sub>39</sub>H<sub>80</sub>, *l*-LMOD combined with a limited torsional Monte Carlo movement was able to locate the lowest energy structures yet reported, and significantly outperformed a pure torsional Monte Carlo (MCMM) and a genetic algorithm-based search. In addition, we have demonstrated *l*-LMOD's utility combined with random translation/rotation of a ligand for the extremely difficult problem of docking flexible ligands into the flexible binding site of PNP.

It is noteworthy that in most docking problems one does not know *a priori* which residues have an

effect on binding. It is well known, however, that for many proteins a number of residues undergo conformational changes upon binding. We have now performed a number of preliminary calculations with *l*-LMOD in which we have defined the set of moving residues to be all of the ones within a 5–10 Å sphere surrounding the active site. Such a sphere typically includes 50–100 moving residues, and represents a particularly difficult search problem with hundreds of degrees of freedom. Such calculations take several days rather than several hours on a single processor. However, it is feasible with *l*-LMOD to dock 50–100 different ligands in a few weeks using multiple processors, thus rendering *l*-LMOD an excellent tool for lead optimization where detailed knowledge of binding interactions is required. We believe that *l*-LMOD will have wide utility for complex docking problems involving flexible protein binding sites.

---

## Acknowledgments

We are indebted to Professor Steven Ealick of Cornell University for providing us with the coordinates for the 9-benzyl-9-deazaguanine/PNP complex, and to George Lamprinidis and Bryan Marten for invaluable help with the PNP docking calculations.

---

## References

1. Kolossváry, I.; Guida, W. C. *J Am Chem Soc* 1996, 118, 5011.
2. (a) Mezey, P. G. *Potential Energy Hypersurfaces*; Elsevier: New York, 1987; (b) Boswell, D. R.; Coxon, E. E.; Coxon, J. M. *Adv Mol Mod* 1995, 3, 195.
3. (a) Kolossváry, I.; Guida, W. C. *J Am Chem Soc* 1993, 115, 2107; (b) Kolossváry, I.; Guida, W. C. *J Mol Struct (Theochem)* 1994, 308, 91.
4. Crippen, G. M.; Scheraga, H. A. *Arch Biochem Biophys* 1971, 144, 462.
5. Nakamura, S.; Hirose, H.; Ikeguchi, M.; Doi, J. *J Phys Chem* 1995, 99, 8374.
6. (a) Kostrowicki, J.; Piela, L.; Cherayil, B. J.; Scheraga, H. A. *J Phys Chem* 1991, 95, 4113; (b) Kostrowicki, J.; Scheraga, H. A. *J Phys Chem* 1992, 96, 7442.
7. Stoer, J.; Bulirsch, R. *Introduction to Numerical Analysis*; Springer: New York, 1993.
8. Gotō, H. *Chem Phys Lett* 1998, 292, 254.  
Gotō's so-called frontier mode-following method, which involves following multiple low-mode eigenvectors, has proven in our hands to be more efficient in finding saddle points than traditional mode-following in a number of tests.
9. Chang, G.; Guida, W. C.; Still, W. C. *J Am Chem Soc* 1989, 111, 4379.

10. Culot, P.; Dive, G.; Nguyen, V. H.; Ghysen, J. M. *Theoret Chim Acta* 1992, 82, 189.
11. Eckart, C. *Phys Rev* 1935, 47, 552.
12. (a) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J Comput Phys* 1977, 23, 327; (b) Ryckaert, J. P. *Mol Phys* 1985, 55, 549; (c) Tobias, D. J.; Brooks III, C. L. *J Chem Phys* 1988, 89, 5115.
13. Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J Comput Chem* 1990, 11, 440.
14. (a) Ponder, J. W.; Richards, F. M. *J Comput Chem* 1987, 8, 1016; (b) Schlick, T.; Overton, M. *J Comput Chem* 1987, 8, 1025.
15. Saunders, M.; Houk, K. N.; Wu, Y.-D.; Still, W. C.; Lipton, M.; Chang, G.; Guida, W. C. *J Am Chem Soc* 1990, 112, 1419.
16. Ngo, J. T.; Karplus, M. *J Am Chem Soc* 1997, 119, 5657.
17. Goodman, J. M.; Still, W. C. *J Comput Chem* 1991, 12, 1, 110.
18. Goodman, J. M. *J Chem Inf Comput Sci* 1997, 37, 876.
19. Goodman, J. M. <http://www.ch.cam.ac.uk/MMRG/alkanes/> (1998).
20. Nair, N.; Goodman, J. M. *J Chem Inf Comput Sci* 1998, 38, 317.
21. Kolossváry, I.; McMartin, C. *J Math Chem* 1992, 9, 359.
22. Goodman, J. M. <http://www.ch.cam.ac.uk/SGTL/software.htm1#MolMod> (1998).
23. Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *J Am Chem Soc* 1989, 111, 8551.
24. Halgren, T. A. *J Comput Chem* 1996, 17, 490.
25. Bohacek, R. S.; McMartin, C.; Guida, W. C. *Med Res Rev*, 1996, 16, 3.
26. See, for example, (a) Miller, M. D.; Kearsley, S. K.; Underwood, D. J.; Sheridan, R. P. *J Comp Aided Mol Design* 1994, 8, 153; (b) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J Mol Biol* 1996, XX, 470; (c) Caflisch, A.; Niederer, P.; Anliker, M. *Proteins* 1992, 13, 223; (d) Leach, A. R.; Kuntz, I. D. *J Comput Chem* 1992, 13, 730; (e) Judson, R. S.; Tan, Y. T.; Mori, E.; Melius, C.; Jaeger, E. P.; Treasurywala, A. M.; Mathiowetz, A. *J Comput Chem* 1995, 16, 1405; (f) Morris, G. M.; Goodsell, D. S.; Huey, R.; Olson, A. J. *J Comput Aided Mol Design* 1996, 10, 293; (g) Welch, W.; Ruppert, J.; Jain, A. N. *Chem Biol* 1996, 3, 449; (h) Makino, S.; Kuntz, I. D. *J Comput Chem* 1997, 18, 1812; (i) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J Mol Biol* 1996, 261, 470; (j) Nola, A. D.; Roccatano, D.; Berendsen, H. J. C. *Proteins* 1994, 19, 174; (k) Westhead, D. R.; Clark, D. E.; Murray, C. W. *J Comput Aided Mol Design* 1997, 11, 209.
27. See, for example, (a) McMartin, C.; Bohacek, R. S. *J Comput Aided Mol Design*, 1997, 11, 333; (b) Guida, W. C.; Bohacek, R. S.; Erion, M. D. *J Comput Chem*, 1992, 13, 214; (c) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J Mol Biol* 1997, 267, 727.
28. Leach, A. R. *J Mol Biol* 1994, 235, 345.
29. Apostolakis, J.; Plückthun, A.; Caflisch, A. *J Comput Chem* 1998, 19, 21.
30. Babu, Y. S.; Montgomery, J. A.; Bugg, C. E.; Carson, W. M.; Narayana, S. V. L.; Cook, W. J.; Ealick, S. E.; Guida, W. C.; Erion, M. D.; Sechrist III, J. A. In: Veerapandian, P., ed. *Structure-Based Drug Design*; Marcel Dekker: New York, 1997.
31. Montgomery, J. A.; Niwas, S.; Rose, J. D.; Sechrist III, J. A., Babu, Y. S.; Bugg, C. E.; Erion, M. D.; Guida, W. C.; Ealick, S. E. *J Med Chem* 1993, 36, 55.
32. Ealick, S. E., personal communication.
33. Sechrist, J. A.; Niwas, S.; Rose, J. D.; Babu, Y. S.; Bugg, C. E.; Erion, M. D.; Guida, W. C.; Ealick, S. E.; Montgomery, J. A. *J Med Chem* 1993, 36, 1847.
34. Koellner, G.; Luić, M.; Shugar, D.; Saenger, W.; Bzowska, A. *J Mol Biol* 1997, 265, 202.