Conformational Sampling by Self-Organization

Huafeng Xu, Sergei Izrailev, and Dimitris K. Agrafiotis*

3-Dimensional Pharmaceuticals, Inc., 665 Stockton Drive, Exton, Pennsylvania 19341

Received March 25, 2003

A new stochastic algorithm for conformational sampling is described. The algorithm generates molecular conformations that are consistent with a set of geometric constraints, which include interatomic distance bounds and chiral volumes derived from the molecular connectivity table. The algorithm repeatedly selects individual geometric constraints at random and updates the respective atomic coordinates toward satisfying the chosen constraint. When compared to a conventional distance geometry algorithm based on the same set of geometric constraints, our method is faster and generates conformations that are more diverse and more energetically favorable.

INTRODUCTION

Since the physical properties and biological behavior of a molecule usually depend on its accessible, low energy conformations, fast and reliable computational methods for producing such conformations are extremely valuable.[1] For molecules with only a few rotatable bonds, systematic enumeration of discretized torsions can be used to search exhaustively for these low energy conformations.[2−5] However, the exponential growth of the search space with the number of rotatable bonds as well as problems associated with ring closures limit the utility of systematic search as a general conformational sampling technique. Stochastic methods designed to sample low energy conformations represent a viable alternative. In its simplest form, a stochastic method randomly perturbs the current conformation of the molecule, minimizes it in energy, and repeats the process to generate a sequence of minimized conformations.[6−8] Standard simulation techniques, such as molecular dynamics and Monte Carlo methods, have been used to generate an ensemble of conformations that lie in the low energy regions on the potential energy surface.[9−12] All of these methods generate conformations in a continuous trajectory, in that each trial conformation is derived from the preceding one by a relatively small change. Because of this continuity, a large number of conformations are generated between the important low energy ones, and a considerable amount of computer time is spent on the calculation and minimization of potential energies for these transitional conformations.

A rather different approach is to first generate conformations that are chemically sensible, without any direct energy calculation. A sensible molecular conformation must satisfy a set of apparent constraints. The connectivity and common covalent bond lengths and angles require that the distance d_{ij} between any pair of atoms i and j fall between certain bounds, $l_{ij} \leq d_{ij} \leq u_{ij}$. Experimental data such as NOE measurements and contextual chemical intuition, such as contact pairs in a protein−ligand complex, can supply further distance constraints. These are usually supplemented by a set of volume constraints that prevent the signed volume V_{ijkl} formed by four atoms i, j, k, l from exceeding certain limits. Volume constraints are used to enforce planarity of conjugate systems and correct chirality of stereocenters. The distance and volume constraints greatly reduce the number of accessible conformations to a molecule and the search space to be considered in conformational sampling. A class of methods, referred to as distance geometry, attempt to generate candidate conformations that are consistent with such geometric constraints. These candidate conformations can then be subject to energy minimization to identify the low energy ones.

Distance geometry has been successfully applied to a wide range of problems including conformational analysis,[13,14] NMR structure determination,[15−17] protein structure prediction,[18] and ligand docking.[19] Conventional distance geometry methods attempt to minimize an error function that measures the violation of geometric constraints, the definition of which will be given in the Methods section. These methods involve four basic steps: (1) generating the interatomic distance bounds, (2) assigning a random value to each distance within the respective bounds, (3) converting the resulting distance matrix into a starting set of Cartesian coordinates, and (4) refining the coordinates by minimizing distance constraint violations. To ensure that reasonable conformations are generated, the original upper and lower bounds are usually refined using an iterative triangular smoothing procedure. Although this process improves the initial guess, the randomly chosen distances may still be inconsistent with a valid 3-dimensional geometry, necessitating expensive metrization schemes[16,20,21] or higher dimensional embeddings[13] prior to error refinement, or lengthy refinement procedures if random starting coordinates are used.

Previously we introduced a self-organizing algorithm for producing coordinates in a low-dimensional space that best preserve a set of distance constraints.[22] The algorithm has been shown to vastly outperform methods that use conventional techniques to minimize the distance error function. In this article, we extend the method to the problem of conformational sampling.

* Corresponding author phone: (610)458-6045; fax: (610)458-8249; e-mail: dimitris.agrafiotis@3dp.com.

10.1021/ci0340557 CCC: $25.00
Published on Web 06/10/2003

METHODS

Following convention, we assess the violations of the distance and volume constraints by the following error function

$$S = S_d + S_v = \sum_{i<j} f(d_{ij},l_{ij},u_{ij}) + \alpha \sum_k h(V_k,V_k^l,V_k^u)$$

The first sum gives the violation of the distance constraints, where

$$f(d_{ij},l_{ij},u_{ij}) = \left(\frac{d_{ij}^2 - l_{ij}^2}{d_{ij}^2}\right)^2 \text{ if } d_{ij} < l_{ij},$$

$$f(d_{ij},l_{ij},u_{ij}) = \left(\frac{d_{ij}^2 - u_{ij}^2}{u_{ij}^2}\right)^2 \text{ if } d_{ij} > u_{ij}$$

and $f(d_{ij},l_{ij},u_{ij}) = 0$ otherwise, the second sum gives the violation of the volume constraints, where $h(V_k,V_k^l,V_k^u) = (V_k - V_k^l)^2$ if $V_k < V_k^l$, $h(V_k,V_k^l,V_k^u) = (V_k - V_k^u)^2$ if $V_k > V_k^u$, and $h(V_k,V_k^l,V_k^u) = 0$ otherwise, and α is a scaling factor. Convention sets $\alpha = 0.1$. Minimizing the error function S with respect to the atomic coordinates generates conformations that satisfy the distance and volume constraints. Because there may be inconsistencies in the distance constraints, it is often impossible to minimize S to 0.

Recently, we described a self-organizing algorithm for minimizing the distance violation, S_d, in the context of learning nonlinear manifolds. The algorithm, which we called stochastic proximity embedding (SPE), repeatedly selects a random pair of points (atoms) and moves their positions in the direction that minimizes the individual error function $f(d_{ij},l_{ij},u_{ij})$. SPE was shown to rapidly and reliably minimize the total distance error function S_d. We conjecture that the algorithm succeeds for the following reason. If all the distance constraints can be satisfied simultaneously, each individual $f(d_{ij},l_{ij},u_{ij}) = 0$, and the global minimum of S_d is $\min(S_d) = 0$. Thus, repeatedly bringing random individual $f(d_{ij},l_{ij},u_{ij})$ toward their minimum results in the global minimum of S_d. In fact, we have shown that the average movement of each atom is antiparallel to the gradient of a simplified error function, which minimizes to 0 when all the distance constraints are satisfied.[23] By virtue of continuity, we expect the algorithm to work even when the distance constraints have small inconsistencies and cannot be satisfied simultaneously.

Here we extend the algorithm to minimize the total error function S. The volume error function S_v is also comprised of a sum of individual contributions and reaches the minimum $\min(S_d) = 0$ when every individual $h(V_k,V_k^l,V_k^u) = 0$, provided that the constraints are consistent and can be satisfied simultaneously. Each individual $h(V_k,V_k^l,V_k^u)$ involves 4 atoms. Similar to our procedure for minimizing S_d, we randomly select a volume constraint k, and move the positions of the 4 atoms involved in the direction that minimizes the individual error $h(V_k,V_k^l,V_k^u)$. The detailed algorithm proceeds as follows:

1. Randomly place the atoms in a box of appropriate size.
2. Select a distance learning rate λ_d, a volume learning rate λ_v, and a relative frequency for enforcing a distance or a volume constraint, ν.
3. With probability ν, do (4); otherwise, do (5).
4. Randomly select a pair of atoms, i and j, and compute their distance $d_{ij} = ||\boldsymbol{x}_i - \boldsymbol{x}_j||$. If $l_{ij} \le d_{ij} \le u_{ij}$, leave the atomic positions unchanged. Otherwise, update the coordinates $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ by

$$\mathbf{x}_i \leftarrow \mathbf{x}_i + \lambda_d \frac{1}{2} \frac{t_{ij} - d_{ij}}{d_{ij} + \epsilon} (\mathbf{x}_i - \mathbf{x}_j)$$

and

$$\mathbf{x}_j \leftarrow \mathbf{x}_j + \lambda_d \frac{1}{2} \frac{t_{ij} - d_{ij}}{d_{ij} + \epsilon} (\mathbf{x}_j - \mathbf{x}_i)$$

where t_{ij} is the nearest bound to d_{ij} (i.e., $t_{ij} = l_{ij}$ if $d_{ij} < l_{ij}$, or $t_{ij} = u_{ij}$ if $d_{ij} > u_{ij}$), and ϵ is a small number used to avoid division by zero.

5. Randomly select a volume constraint k and the four atoms involved, p, q, s, t. Compute the signed volume V_{pqst} formed by the 4 atoms. If $V_k^l < V_{pqst} < V_k^u$, leave the atom positions unchanged. Otherwise, compute the gradient of the signed volume with respect to the atomic positions, $\mathbf{g}_\mu = \nabla_\mu V_{pqst}$, where $\mu = p,q,s,t$, and update the atomic coordinates by

$$\mathbf{x}_\mu \leftarrow \mathbf{x}_\mu + \lambda_v (V_k^0 - V_{pqst}) \frac{\mathbf{g}_\mu}{\sum_{\beta=p,q,s,t} |\mathbf{g}_\mu|^2}$$

where V_k^0 is the nearest bound to V_{pqst} (i.e., $V_k^0 = V_k^l$ if $V_{pqst} < V_k^l$, or $V_k^0 = V_k^u$ if $V_{pqst} > V_k^u$).

6. Repeat (3)−(5) for a prescribed number of steps, S.
7. Decrease the learning rates λ_d and λ_v by prescribed decrements $\delta\lambda_d$ and $\delta\lambda_v$.
8. Repeat (3)−(7) for a prescribed number of cycles, C.

A common set of parameters for our method are $\lambda_d = \lambda_v = 1$, $C = 50$, $\delta\lambda_d = \delta\lambda_v = 0.9/C$, $S = 50\times N$, and $\nu = \max(0.5, 1-8.0 \times (||V||)/(N(N+1)/2 + ||V||))$, where N is the number of atoms in the molecule, and $||V||$ is the total number of volume constraints. These parameters ensure that the embedding time scales linearly with the number of atoms and have been previously shown sufficient for convergence.[22,23]

The distance and volume constraints are specified by a set of rules, under which atom groups that match specific substructural patterns are assigned the corresponding geometric constraints. The rules are supplied to the program in a separate file and can be easily extended and customized by users. In this work, hydrogen atoms are not included in the distance constraints. The effect of including hydrogens will be explored in future publications.

Because distance geometry methods generate conformations that satisfy only the geometric constraints without regard to any potential energy function, the resulting conformations are not necessarily of low energy. To study the energy distribution of these conformations and identify the low energy ones, each conformation generated by our self-organizing scheme was locally minimized using the Merck Molecular Force Field (MMFF94).[24−28] The role of distance geometry is to filter out nonsensible geometries and limit energy calculation and minimization to only a small

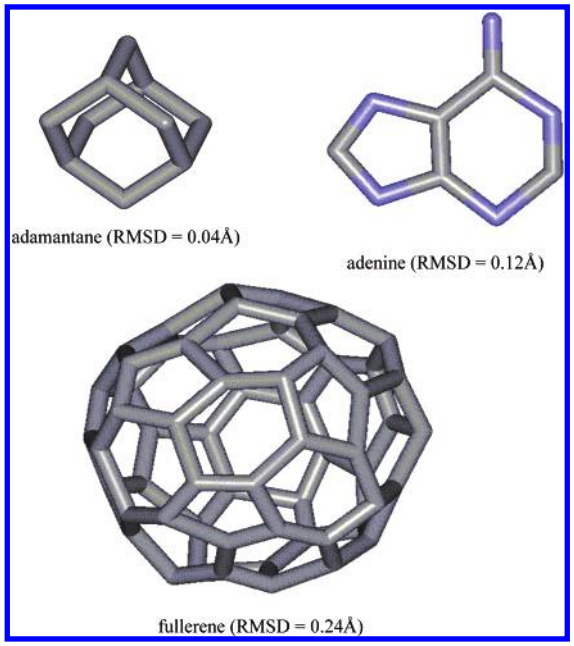

Figure 1. Typical conformations of rigid molecules generated by SPE. The RMSDs from the respective minimized structures are shown in parentheses.

number of candidate conformations. Moreover, when sufficient and proper geometric constraints are constructed, our method can produce raw conformations that are very similar to the energy-minimized structures, as demonstrated by the examples of rigid molecules in the Results section. This greatly reduces the number of iterations necessary for the convergence of energy minimization.

In this work, all calculations were performed on an SGI Origin 2000 workstation equipped with a 250 MHz MIPS IP27 R10000 processor and 1024 Mb of RAM.

RESULTS AND DISCUSSION

An appropriate context to assess molecular conformations is a prerequisite to a fair evaluation of conformational sampling methods. In this work, we intend to demonstrate the "chemical sensibility" of the generated conformations by showing that the geometries generated by SPE are close to the corresponding local energy minima, as measured by RMSD, and that a majority of the conformations minimize to distinct low energy structures. We choose to use an empirical force field to compute the potential energies. Comparison of the conformations to X-ray or NMR data is beyond the scope of this paper; a thorough evaluation of the method in this light is already in progress and will be reported in due course. We are also developing a SPE based molecular docking program, where we will compare the SPE generated conformations to the actual bound conformations as determined by X-ray crystallography.

When applied to rigid molecules, our method always found the correct conformations (Figure 1). This is due to the redundancy of the distance matrix and the cooperative nature of the atomic refinements—moving one pair of atoms toward satisfying their distance constraints simultaneously improves many other distances involving these atoms. The raw conformations generated by our method are very close to the final structures after energy minimization, as evidenced by the small RMSDs between the raw conformations and the minimized structures. A small RMSD value normally signifies that a small number of iterations is required for convergence in the local energy minimization. It also suggests that the raw conformations may be used without energy minimization when crude conformations are adequate. Worth noting is the fullerene molecule, which entails numerous volume constraints, requiring the five- and six-member aromatic rings to be planar. It therefore constitutes a challenge for our extension of self-organization to include volume constraints. As shown in Figure 1, our method correctly generated the normal soccer ball-like conformation of fullerene, with only a small deviation of 0.24 Å in RMSD from the energy-minimized conformation.

For flexible molecules, the global minimum is usually unknown, and the merits of the algorithm can only be assessed by comparison to another method (here, we used the widely used RUBICON[29] distance geometry program). Four celebrated molecules were examined—cycloheptadecane, raloxifene, the free base of Gleevec (imatinib mesylate), and [Met^5]-enkephalin (sequence YGGFM). To ensure statistical significance, 10 000 different conformations were generated by each program using an identical set of rules. Both methods use this set of rules to assign the same geometric constraints (RUBICON may assign additional constraints internally, but we are unable to verify this without access to its source code). Because RUBICON rejects conformations with large constraint violations, it generated only 8086 conformations for raloxifene, 9669 conformations for Gleevec, and 8034 conformations for [Met^5]-enkephalin. The chirality of the L-amino acids in each conformation of [Met^5]-enkephalin was checked, and no violation was found for either method. SPE occasionally generates conformations with large geometric violations, but for conformations that locally minimize to low energy structures, the geometric violations in the conformations generated by SPE and RUBICON are comparable.

The comparison was based on several criteria—the speed of generating the initial conformations, the coverage of energetically favorable conformations, the rate of discovering distinct conformations, and the lowest energy obtained during the entire search. A good method must be fast, must generate more conformations that minimize to unique low energy structures, and must quickly identify the global minimum. As illustrated in Table 1 and Figure 2, our method outperforms RUBICON on all counts. Indeed, SPE was up to an order of magnitude faster in generating the raw conformations, and these consistently minimized in energy to more distinct conformations in all four cases (two conformations were considered distinct if the corresponding minimized structures differed by more than 0.05 Å in RMSD). Molecular symmetry was taken into account. The superimposition program exhaustively checked all possible matches between the atoms in the two molecules and found the best alignment between them. Conformations related by a permutation of topologically equivalent atoms did *not* count as distinct conformations. The ability of a conformational sampling method to discover many distinct low energy conformations is important in a number of situations. For example, in molecular docking,[30] the inhibitor may adopt a different conformation upon binding, and it is critical to include a large number of conformations in the sampling to correctly predict the binding mode.[31] Similarly, in comparative mo-

Table 1. Raw CPU Time t Required To Generate One Conformation, the Number of Distinct Conformations n Discovered within 10 000 Trials, and the Lowest Energy Minimum E_{min} for Each Molecule Found by SPE and RUBICON[a]

molecule	t^{SPE} (sec)	$t^{RUBICON}$ (sec)	n^{SPE}	$n^{RUBICON}$	n_{min}^{SPE} (kcal·mol^{-1})	$E_{min}^{PUBICON}$ (kcal·mol^{-1})
cycloheptadecane	0.049	0.615	5908	4453	4.5105	4.5105
raloxifene	0.242	1.215	1915	689	82.2486	82.2486
Gleevec	0.261	1.677	3482	1228	64.8114	64.8482
[Met^5]-enkephalin	0.306	1.819	9995	8019	71.3763	73.8235

[a] t is computed by dividing the total CPU time by the number of trial conformations and does not include energy minimization. Two conformations are considered distinct if, after local energy minimization, they differ by an RMSD larger than 0.05 Å.

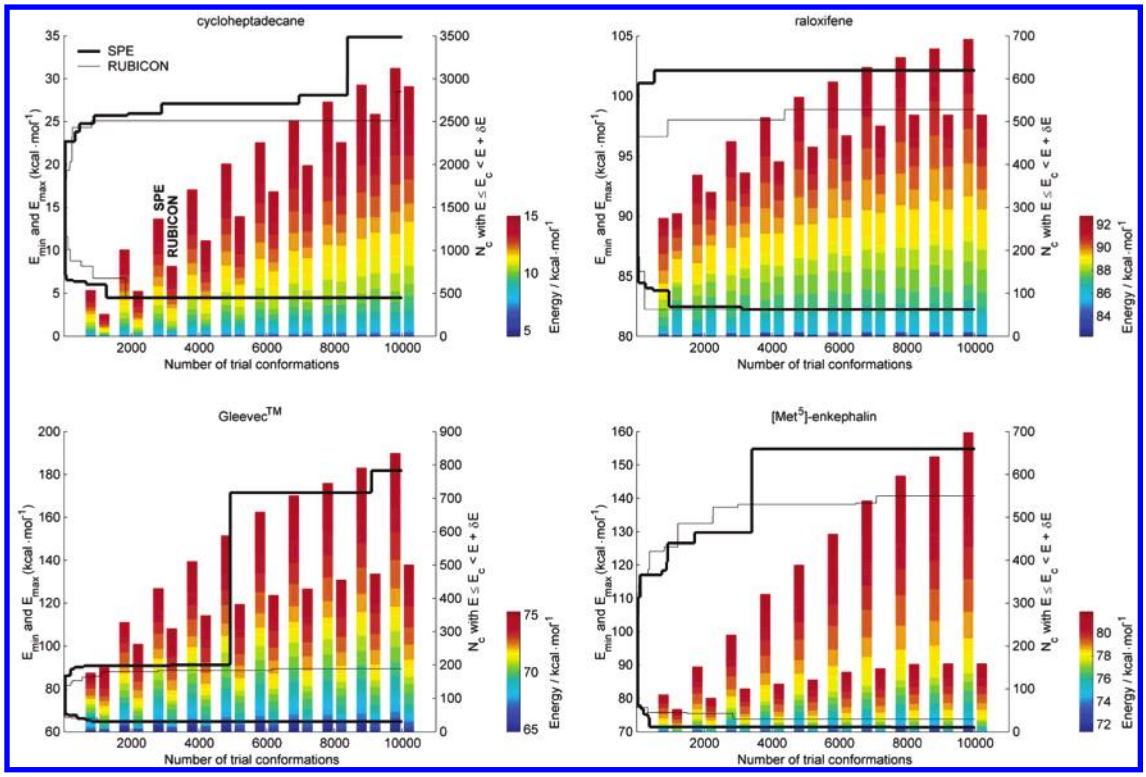

Figure 2. Comparison of sampling efficiency between SPE and RUBICON. The solid lines show the minimum and maximum energy (E_{min} and E_{max}) discovered by the two methods after a number of trials, with the energy values indicated by the left ordinate of each plot (thick lines for SPE, thin lines for RUBICON). The bar graphs show the number of distinct conformations N_C found by each method after a number of trials (SPE on the left, RUBICON on the right), with the numbers listed on the right ordinate of each plot. Since usually only the energetically favorable conformations are of chemical interest, only the conformations whose minimized energies are within 10.0 kcal·mol^{-1} from the global minimum are included. Each bar is further divided into 20 segments that represent nonoverlapping energy intervals of 0.5 kcal·mol^{-1} from the global minimum to 10.0 kcal·mol^{-1} above, and whose corresponding medium energy values are indicated by the color map to the right of each plot. The length of each segment shows the number of distinct conformations whose minimized energies fall within the corresponding energy interval. Two conformations are considered distinct if the corresponding minimized structures differ by more than 0.05 Å in RMSD (see text for detail).

lecular field analysis,[32] a consensus conformation has to be deduced from a number of candidate conformations, demanding the sampling program to produce a diverse set of low energy structures. Our method demonstrates a consistent and significant advantage over RUBICON in this respect. For raloxifene, Gleevec, and [Met^5]-enkephalin, the difference was even more pronounced in the low energy region, as manifested by the significantly longer segments of blue color for SPE compared to RUBICON in the bar graphs in Figure 2. For example, for [Met^5]-enkephalin, SPE discovered 69 distinct conformations with minimized energy within 5.0 kcal·mol^{-1} above the lowest energy minimum, whereas RUBICON discovered only 9. SPE was also superior in locating the lowest energy structure—both methods found the same global energy minima for cycloheptadecane and raloxifene, but RUBICON failed to identify the lowest energy minima of Gleevec and [Met^5]-enkephalin discovered by SPE. These conformations do not seem to violate any apparent geometric constraints imposed by RUBICON, and therefore they should not have been excluded by RUBICON by intention. The more flexible the molecule, the more critical it is to exhaustively sample its accessible conformations in order to identify the lowest energy structure(s). SPE demonstrates superior sampling ability, which enables it to find the low energy conformations that elude RUBICON. In addition, SPE finds the global minimum in a smaller or comparable number of trials (the lowest energy structures discovered after minimization superimposed with the respective raw conformations produced by SPE are shown in Figure 3). The lowest energy structures for the molecules found by SPE are provided in MDL SDF format in the Supporting Information.

A similar method for enforcing distance constraints in molecular conformations has been proposed by de Groot et al.,[33] where random pairs of atoms are selected and their coordinates are adjusted so that the distance constraints

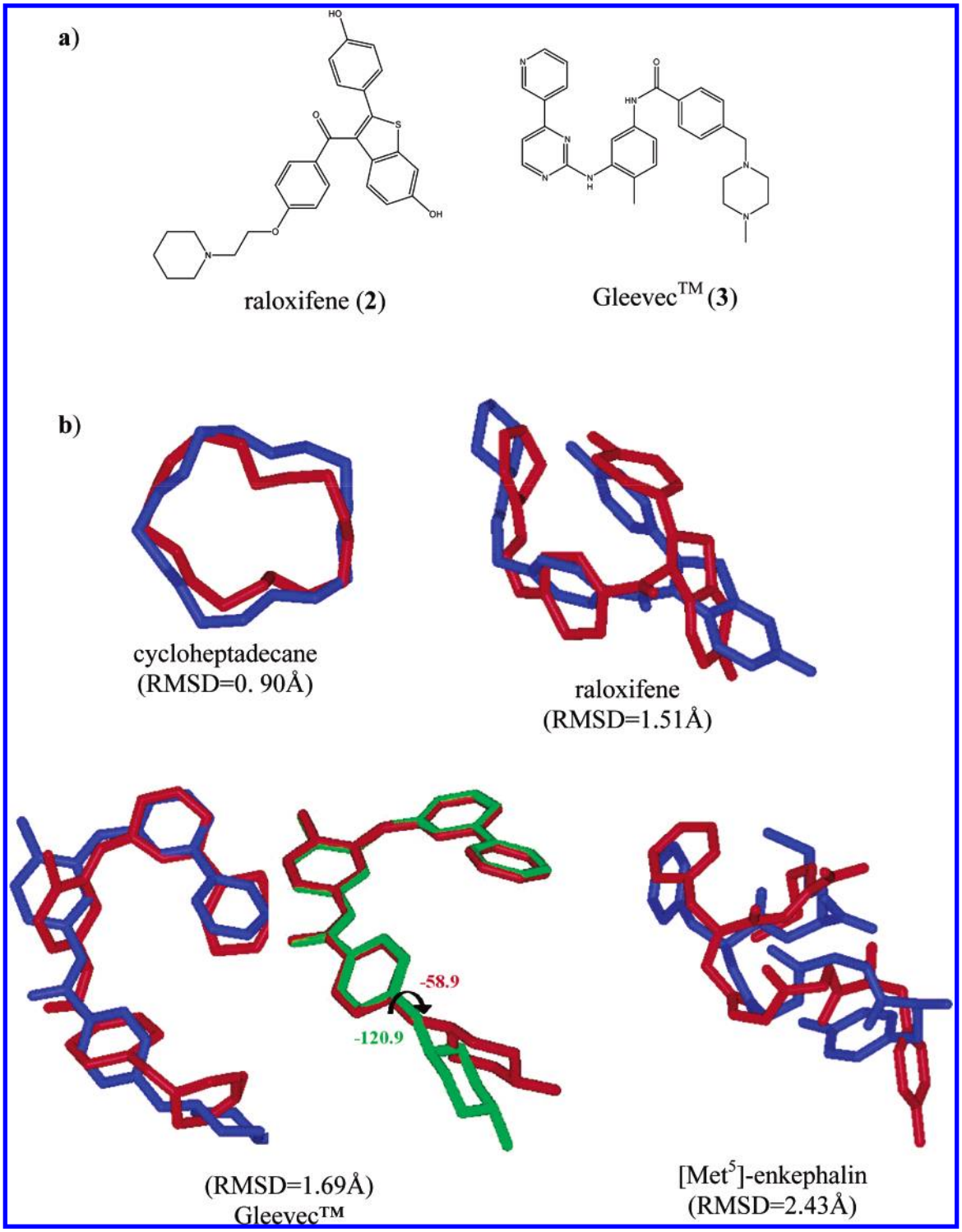

Figure 3. (a) The chemical structures of raloxifene and the free base of Gleevec. (b) Superimposition of the lowest energy structures discovered after minimization (blue) with the respective raw conformations produced by SPE (red). The corresponding RMSDs are shown in the parentheses. For Gleevec, we also show the lowest energy structures found by SPE (red) and RUBICON (green), after local energy minimization. The torsion angles that differ between the two structures are labeled in respective colors.

between them are exactly satisfied. We have shown that by using a decreasing learning rate, our method avoids oscillatory behavior in the final stages of the embedding.[23] More importantly, our method treats the volume constraints during the course of the embedding and therefore ensures the correct chirality of the stereocenters and the planarity of conjugated atoms in the resulting conformations. This is critical for biological molecules, which contain a large number of chiral centers. For example, our method consistently produced the correct chirality of the L-amino acids in [Met^5]-enkephalin, whereas without enforcing the volume constraints, only 1 out of $2^3 = 8$ conformations would have all the correct C_β chiral configuration. For a typical mid-sized protein containing 100 residues, the probability of generating a conformation with all chiral centers properly assigned is abysmally low. A more stringent test is performed on the molecule morphine (Figure 4), which contains five closely positioned chiral carbons. Our method can produce the correct chirality of all

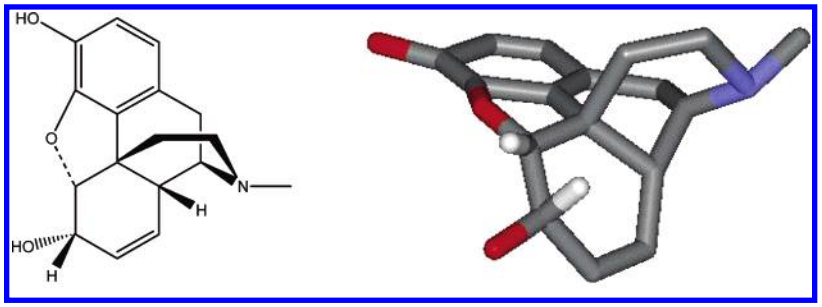

Figure 4. Chemical structure and a generated 3-dimensional conformation of morphine.

five chiral centers in half of the trials, whereas ignoring the volume constraints during the embedding leads to only one out of approximately 27 conformations with all atoms having properly assigned chirality. The alternative of correcting the misassigned chiralities *after* the conformation has been generated is an extremely complicated and computationally intensive task.

CONCLUSIONS

Although only the Merck Molecular Force Field is used to compute the potential energies in this work, we expect the results to be qualitatively similar when other empirical force fields are employed, because SPE does not depend on any specific energy parameters and therefore should not be favored by any particular force field. Although the specific details of the comparison may differ depending on the potential energy function employed, we believe that the main advantages of our method—its raw speed and the diversity of the conformations that it generates—will remain. Since energy minimization is still the rate-limiting step, the speed differential will be most noticeable in applications where crude conformations could suffice (e.g. pharmacophore modeling). We are currently focusing on several enhancements for generating better initial geometries and for detecting and eliminating conformations that are likely to lead to the same local minima as well as on the use of this technique for protein structure prediction, ligand docking, and pharmacophore modeling.

ACKNOWLEDGMENT

We thank Dr. Raymond F. Salemme of 3-Dimensional Pharmaceuticals, Inc. for his insightful comments and support of this work.

Supporting Information Available: The lowest energy structures for the molecules found by SPE in MDL SDF format. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Leach, A. R. A survey of methods for searching the conformational space of small and medium-sized molecules. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH: New York, 1991; Vol. 2.
(2) Lipton, M.; Still, W. C. *J. Comput. Chem.* **1988**, *9*, 343.
(3) Bruccoleri, R. E.; Karplus, M. *Biopolymers* **1987**, *26*, 137.
(4) Bruccoleri, R. E.; Karplus, M. *Macromolecules* **1985**, *18*, 2767.
(5) Go, N.; Scheraga, H. A. *Macromolecules* **1970**, *3*, 178.
(6) Saunders, M. *J. Am. Chem. Soc.* **1987**, *109*, 3150.
(7) Ferguson, D. M.; Raber, D. J. *J. Am. Chem. Soc.* **1989**, *111*, 4371.
(8) Chang, G.; Guida, W. C.; Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 4379.
(9) Auffinger, P.; Wipff, G. *J. Comput. Chem.* **1990**, *11*, 190.
(10) Bruccoleri, R. E.; Karplus, M. *Biopolymers* **1990**, *29*, 1847.
(11) Li, Z.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 6611.
(12) Jorgensen, W. L.; Tirado-Rives, J. *J. Phys. Chem.* **1996**, *100*, 14508.
(13) Spellmeyer, D. C.; Wong, A. K.; Bower, M. J.; Blaney, J. M. *J. Mol. Graphics Modell.* **1997**, *15*, 18.
(14) Feuston, B. P.; Miller, M. D.; Culberson, J. C.; Nachbar, R. B.; Kearsley, S. K. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 754.
(15) Havel, T. F.; Wüthrich, K. *J. Mol. Biol.* **1985**, *182*, 281.
(16) Kuszewski, J.; Nilges, M.; Brünger, A. T. *J. Biomolecular NMR* 1992, 2, 33.
(17) Mumenthaler, C.; Braun, W. *J. Mol. Biol.* **1995**, *254*, 465.
(18) Huang, E. S.; Samudrala, R.; Ponder, J. W. *Protein Sci.* **1998**, *7*, 1998.
(19) Meng, E. C.; Gschwend, D. A.; Blaney, J. M.; Kuntz, I. D. *Proteins: Struc., Funct., Genet.* **1993**, *17*, 266.
(20) Havel, T. F.; Snow, M. E. *J. Mol. Biol.* **1991**, *217*, 1.
(21) Havel, T. F.; Wüthrich, K. *Bull. Math. Biol.* **1984**, *46*, 673.
(22) Agrafiotis, D. K.; Xu, H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15869.
(23) Agrafiotis, D. K. *J. Comput. Chem.* In press.
(24) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 616.
(25) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490.
(26) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 520.
(27) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 553.
(28) Halgren, T. A.; Nachbar, R. B. *J. Comput. Chem.* **1996**, *17*, 587.
(29) DAYLIGHT Chemical Information Systems, Inc. (www.daylight.com).
(30) Muegge, I.; Rarey, M. *Rev. Comput. Chem.* **2001**, *17*, 1.
(31) Kearsley, S. K.; Underwood, D. J.; Sheridan, R. P.; Miller, M. D. *J. Comput.-Aided Mol. Design* **1994**, *8*, 565.
(32) Cramer, R. D. I.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
(33) de Groot, B. L.; van Aalten, D. M. F.; Scheek, R. M.; Amadei, A.; Vriend, G.; Berendsen, H. J. C. *Proteins: Struct., Funct., Genetics* **1997**, *29*, 240.

CI0340557