Contribution of Conformer Focusing to the Uncertainty in Predicting Free Energies for Protein−Ligand Binding

Julian Tirado-Rives and William L. Jorgensen*

Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107

Received June 27, 2006

When a ligand binds to a protein, it is typically not in the lowest-energy conformation for the unbound ligand and there is also a loss of conformational degrees of freedom. The free-energy change for this "conformer focusing" is addressed here formally, and the associated errors with its estimation or neglect are considered in the context of scoring functions for protein−ligand docking and computation of absolute free energies of binding. Specific applications for inhibition of HIV-1 reverse transcriptase are reported. It is concluded that the uncertainties from this source alone are sufficient to preclude the viability of current docking methodology for rank-ordering of diverse compounds in high-throughput virtual screening.

Introduction

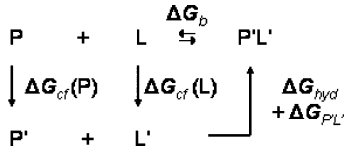
Accurate computation of free energies of binding is central to computer-aided drug design. The process of interest,

$$\mathrm{H(aq)} + \mathrm{L(aq)} \leftrightarrows \mathrm{HL(aq)} \qquad (1)$$

has numerous components for a flexible ligand L binding to a flexible host H in solution to yield the complex HL.[1,2] The problem has been highlighted by the great current interest in virtual high-throughput screening (VHTS) of compound libraries for binding to biomolecular receptors, most typically enzymes, with docking programs.[3] The minimal goal is reasonable rank-ordering of compounds such that if a selection of high-ranking compounds is obtained and assayed, it is likely that some will show activity. From a practical standpoint, the window of activity is narrow. It is limited on the low end by the improbability of finding a library compound that is so well matched with the target that activity is observed below a concentration of ∼50 nM. And it is limited on the high end around 100 μM owing to anticipated difficulty in optimizing weaker leads, normal assay conditions, and the insolubility of many, typical library compounds at higher concentrations in aqueous solution. This 2000-fold window corresponds to a free-energy range of 4.5 kcal/mol at 25 °C. If errors in estimation of free energies of binding are greater than this, the common occurrence of false positives and negatives can be expected.

There have been numerous comparisons of the results of retrospective docking exercises with alternative software and scoring functions.[3,4] The general impression is one of inconsistent performance, a trend toward improvement, and limited use in live applications where compounds are actually acquired and assayed. More quantitative assessment of potential sources of error is warranted to better gauge the theoretical limits on accuracy and to identify areas for improvement. The computational errors can arise in many ways depending on the methodology; some possibilities include errors in the structure of the host, choice of ionization states, structure of the complex, inadequate sampling of internal degrees of freedom, evaluation of changes in solvation, and the host−ligand energetics and/or scoring functions.

Scheme 1



Conformer Focusing: Theory

Presently, just one additional source of error, which has received little quantitative attention, is analyzed. It can be termed "conformer focusing". Most organic molecules have multiple torsional degrees of freedom for rotation about the central bond in W−X−Y−Z quartets of non-hydrogen atoms that lead to a spectrum of conformational states. Analysis of 1711 oral drugs shows a range of 0−34 such rotatable bonds with a mean of 5.5;[5] the expected number of conformers for a druglike molecule then averages $2^{5.5}-3^{5.5} \approx 45-420$. Many of these are high in energy and are not populated; however, a subset of low-energy ones is populated. Upon forming the complex, many conformers of the unbound ligand become sterically inaccessible. The same is true for the host. The preferred bound conformer of the ligand is likely not the lowest-energy one in the unbound state, and it may not even have a significant unbound population, which yields a reorganization penalty.[6] In fact, a recent analysis of 150 protein−ligand complexes found that the bound conformer is 4−5 kcal/mol higher in energy on average than the lowest-energy conformer in solution.[6d]

A thermodynamic cycle representing conformer focusing can be construed as in Scheme 1. The different conformational manifolds for the unbound and bound states are distinguished by the symbols L and L′ for the ligand and P and P′ for the protein. The free-energy change associated with conformer focusing for the ligand can be expressed by considering its conformational partition functions for the unbound and bound conditions, Z_u and Z_b:

* To whom correspondence should be addressed. Phone: 203-432-6278. Fax: 203-432-6299. E-mail: william.jorgensen@yale.edu.

10.1021/jm060763i CCC: $33.50 © 2006 American Chemical Society
Published on Web 09/12/2006

$$\Delta G_{cf} = \beta^{-1} \ln \frac{Z_u}{Z_b} = \beta^{-1} \ln \left[\frac{\sum_i n_i \exp(-\beta \epsilon_i)}{\sum_j n_j \exp(-\beta \epsilon_j)} \right] \quad (2)$$

Here, a simple, discrete model is assumed with conformational states i (unbound) and j (bound), degeneracies n_i and n_j, and $\beta = 1/(k_B T)$. Furthermore, entropy variations arising from differences in vibrational frequencies for the conformers are ignored, and the set of conformers for the bound ligand is taken to be a subset of those for the unbound ligand. Thus, there is a free-energy penalty for the loss of conformational states.

In principle, one needs complete knowledge of the population of conformational states for both the unbound and bound ligand to evaluate ΔG_{cf}. For assessment of potential errors, a further simplification can be made by considering the case where only one conformer k is left for the bound ligand. Equation 2 can then be reduced to eq 3 where ϵ_1 is the energy for the lowest-energy conformer and Z_1 is the modified partition function ratio using the relative energies $\epsilon_i - \epsilon_1$. The first term in eq 3 is the reorganization penalty, and the second is for the loss of conformational states. For example, in a trivial case where the unbound ligand has only three equivalent conformers and the bound ligand populates just one of them, $\Delta G_{cf} = \beta^{-1} \ln 3$. For binding in aqueous solution, the unbound states need to reflect the intrinsic gas-phase energies of the conformers, E_i, and the free energy of hydration, G_i^{aq}, so $\epsilon_i = E_i + G_i^{aq}$. This leads to eq 4, whose evaluation requires a conformational search that computes E_i and G_i^{aq} for all i.

$$\Delta G_{cf} = (\epsilon_k - \epsilon_1) + \beta^{-1} \ln[\sum_i n_i \exp(-\beta(\epsilon_i - \epsilon_1))] = \Delta \epsilon_k + \beta^{-1} \ln Z_1 \quad (3)$$

$$\Delta G_{cf} = (E_k - E_1) + (G_k^{aq} - G_1^{aq}) + \beta^{-1} \ln Z_1 = \Delta E_k + \Delta G_k^{aq} + \beta^{-1} \ln Z_1 \quad (4)$$

Results

Different cases can be considered for evaluation of errors associated with treatment of ΔG_{cf}. If one does not explicitly consider ΔG_{cf}, the potential errors are large; the bound conformer is expected to range from 0 to 15 kcal/mol above the lowest-energy one.[6d] In addition, if there are multiple low-energy conformers that are lost in going to the bound state, the second term in eq 3 can provide several kcal/mol of additional uncertainty. Neglect of ΔG_{cf} is common practice in scoring functions used for docking. For example, the popular ChemScore does not consider the items in eq 2 or eq 3; there is just a penalty of ~0.6 kcal/mol for each rotatable bond in the ligand that becomes "frozen" upon binding.[7] Successful ranking of conformationally diverse compounds with such scoring cannot be expected.

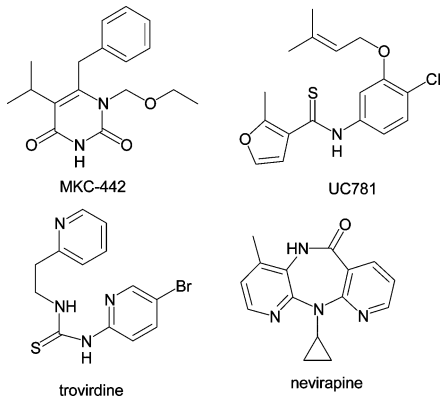
If one does explicitly consider ΔG_{cf}, potential errors are associated with the implied conformational searches and evaluation of E_i and G_i^{aq}. For VHTS, it is likely that E_i would be calculated with a force field (FF); ab initio or density functional theory (DFT) calculations would normally be impractical. For small molecules covered in the training sets, FF errors for conformational energetics are small, 0.0−0.5 kcal/mol.[8] However, for drug-size molecules, the situation is less clear, particularly when there are substructures outside the training sets and when ΔE_k is not small. For G_i^{aq}, a reasonable approach would be to perform the conformational search including hydration modeled by a fast continuum method such as GB/SA.[9] Optimistically, the additional error from that source is 0.5−1.0 kcal/mol.[9b] Examples are now provided for non-nucleoside inhibitors of HIV reverse transcriptase (NNRTIs).

MKC-442 (emivirine) and UC781 both have six rotatable bonds. X-ray crystal structures have been determined, so their conformations in the bound state are known (Figure 1, left).[10,11] They are potent NNRTIs and the binding site is engulfing, so the bound inhibitors have little conformational freedom. The terms in eq 4 were evaluated by first performing a conformational search with the BOSS program[12] using the OPLS/CM1A force field[13] and GB/SA hydration.[9b] The resultant conformer k closest in structure to the bound conformer was identified and compared to the lowest-energy conformer (Figure 1); $\beta^{-1} \ln Z_1$ was evaluated using the energetic results from the conformational search. To obtain additional estimates of ΔE_k, alternative force fields were applied mostly using MacroModel,[14] and ab initio and DFT calculations were performed using Gaussian 03 on conformers 1 and k.[15] HF/6-31G(d) optimizations were executed starting from the OPLS/CM1A structures followed by B3LYP/6-31G(d) and MP2/6-31G(d) energy evaluations (Table 1). Single-point calculations using the FF structures from the conformational search were also carried out; they give similar results with somewhat greater variation in ΔE_k.

Table 1. Conformational Results for NNRTIs[a]

	MKC-442	UC781	trovirdine	nevirapine
no. of conformers[b]	57	102	69	2
conformer k	5	85	5	1
$\beta^{-1} \ln Z_1$	0.77	1.47	1.30	0.41
ΔE_k, OPLS/CM1A	1.49	4.57	−4.07	0.00
ΔG_k, GB/SA	−0.14	1.91	4.85	0.00
ΔG_{cf}	2.12	7.95	2.08	0.41
ΔG_{cf} range[c]	0.3, 3.1	1.6, 8.0	−6.3, 2.6	0.41
ΔE_k,HF/6-31G(d)[d]	2.52	0.98	−8.45[e]	0.00
ΔE_k, B3LYP/6-31G(d)[d]	2.00	0.63	−8.87[e]	0.00
ΔE_k,MP2/6-31G(d)[d]	2.11	−1.80	−9.33[e]	0.00
ΔE_k, OPLS2001	−0.03	2.95	−3.59	0.00
ΔE_k, OPLS2005	−0.27	−0.43	−6.50	0.00
ΔE_k, MMFF	2.44	−1.33	−12.49	0.00

[a] Energies in kcal/mol. ΔE_k is the energy difference between the lowest-energy conformer and the one, k, most similar to that in the crystal structure of the RT complex. [b] Excluding mirror images. [c] Using the alternative estimates of ΔE_k. [d] Using HF/6-31G(d) optimized geometries. [e] Result for desbromotrovirdine.



For MKC-442, 57 unique conformers were found; since none are planar, each has a degeneracy of 2. The bound conformation is similar to unbound conformer 1 (Figure 1); the most similar conformer is number 5, which is higher in energy by 1.49 kcal/mol but slightly better hydrated (0.14 kcal/mol). There are only

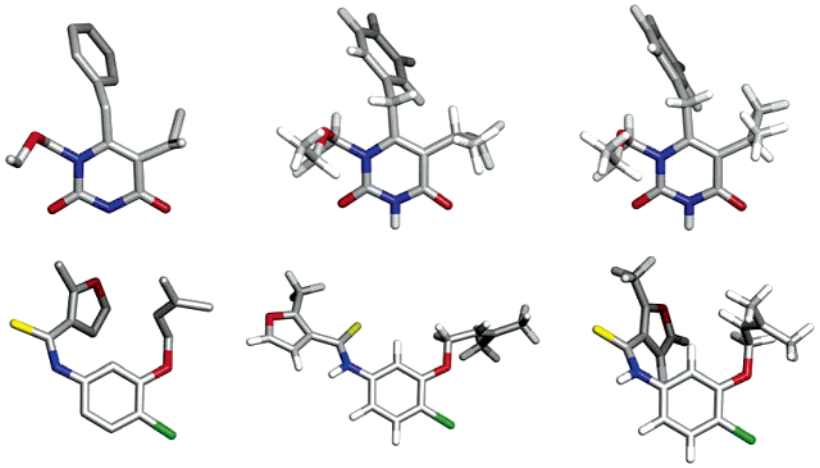


Figure 1. Conformers of MKC-442 (top) and UC781 (bottom): (left) from the 1rt1 and 1rt4 crystal structures in complexes with HIV−RT; (middle) the lowest-energy conformer as found by the conformational search with GB/SA hydration; (right) the conformer *k* from the conformational search most similar to the one in the crystal structure.

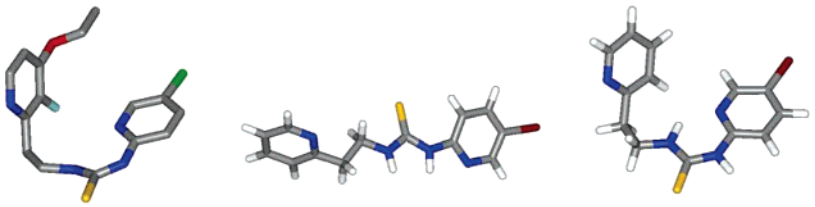


Figure 2. Left: structure of a trovirdine analogue from the 1dtt crystal structure. Middle: lowest-energy structure of trovirdine from the conformational search with GB/SA hydration. Right: conformer most similar to the bound structure for trovirdine from the conformational search.

8 conformers within 2 kcal/mol of conformer 1, so $\beta^{-1} \ln Z_1$ is relatively small, 0.77 kcal/mol. The net ΔG_{cf} is 2.1 kcal/mol. The overall range from the different results in Table 1 for ΔE_k is −0.3 to 2.5 kcal/mol, which gives a range of 0.3 to 3.1 kcal/mol for ΔG_{cf}. This can be considered to be a favorable case owing to the similarity of conformers 1 and *k*.

For UC781, 102 conformers were identified with 24 within 2 kcal/mol of conformer 1, so $\beta^{-1} \ln Z_1$ is now larger, 1.47 kcal/mol. However, the bound conformer and conformer 1 are very different because the thioamide is *E* in the bound state, which permits the characteristic hydrogen bond for most NNRTIs between the NH and the backbone C═O of Lys101. Thus, the conformer most similar to the bound one, number 85, is now higher in energy. From OPLS/CM1A, ΔE_{85} is 4.57 kcal/mol and conformer 85 is also computed to be less well hydrated by 1.91 kcal/mol (the dipole moments for conformers 1 and 85 are 4.32 and 2.45 D from the FF and are 3.97 and 2.12 from the HF/6-31G(d) calculations). So there are both significant reorganization and desolvation penalties, and the net ΔG_{cf} is 7.95 kcal/mol. It is notable that even the ab initio and DFT results for ΔE_{85} span 3 kcal/mol in this case. Other estimates of ΔE_{85} were also obtained using alternative FFs; conformers 1 and 85 were optimized again in each case. Some results in kcal/mol are 5.18 from MM2 in Chem3D, 0.05 from MM+ in HyperChem, and 2.95, −0.43, and −1.33 from OPLS2001, OPLS2005, and MMFF in MacroModel. The OPLS/CM1A results are toward the high end of the range, though it is noted that careful parametrization had been performed for the torsional energetics of *N*-alkyl- and *N*-arylthioamides using MP2 results. For just the *Z* to *E* energy difference for *N*-phenylthioacetamide, a broad range of results is obtained: −5.8 kcal/mol from MMFF, −1.7 from HF/6-31G(d)//HF/6-31G(d), −0.5 from MP2/6-31G(d)//MP2/6-31G(d), −0.4 from B3LYP/6-31G(d)//B3LYP/6-31G(d), 0.2 from MP2/6-311+G(d,p)//HF/6-31G(d), and 0.9 from OPLS/CM1A.

A further example is provided by trovirdine, which again has six rotatable bonds. The 1dtq and 1dtt crystal structures are available for analogues in complex with HIV-1 RT.[16] A conformational search for trovirdine with OPLS/CM1A and GB/SA hydration yielded 69 conformers; conformer 5 shows the closest correspondence for the six dihedral angles about the bonds between the two rings and those in the crystal structures (Figure 2). The lowest-energy structure in GB/SA water is extended with the hydrogen-bonding sites exposed to the solvent, while the bound conformer has an internal hydrogen bond and an *E*,*Z*-conformation for the thiourea fragment. The *E* component overlays with the *E*-thioamide for bound UC781.

As summarized in Table 1, conformer 5 is computed to be 4.07 kcal/mol lower in energy than conformer 1 in the gas phase using OPLS/CM1A, but it is less well hydrated by 4.85 kcal/mol owing to the internal hydrogen bond. The cancellation leads to a relatively small ΔG_{cf}, 2.08 kcal/mol from this FF. The relative energetics of the two conformers is challenging for a force field owing to the difference in both the thiourea conformation and the internal hydrogen bonding. The strength of the latter is sensitive to the assignment of partial atomic charges. Consequently, there is an 8.9 kcal/mol spread in the alternative FF results in Table 1. Again, the FF results in Table 1 are from optimizations for both conformers with each force field. The results are not improved if single-point calculations are performed using the OPLS/CM1A optimized structures; the spread is then 11.0 kcal/mol. The ab initio and DFT results are more consistent with ΔE_5 ranging from −8.45 to −9.33 kcal/mol. The latter calculations were performed for desbromotrovirdine. This is not expected to affect the results significantly; with OPLS/CM1A the ΔE_5 values are −4.07 and −4.18 kcal/mol for trovirdine and its desbromo analogue. Thus, on the basis of the ab initio and DFT results, it is likely that a conformer like 5 with the internal hydrogen bond is in fact the lowest-

energy conformer for troviridine in water rather than the extended conformer 1.

Finally, for nevirapine (Viramune), there is only one rotatable bond and one low-energy conformer, plus its mirror-image, and one that is ~3 kcal/mol higher in energy. The low-energy conformer is the same as observed in crystal structures.[17] This provides a limiting reference point in Table 1.

Discussion

Absolute Free Energies of Binding. Accurate prediction of absolute free energies of binding requires accurate computation for each component of the process. For the conformer focusing component, it is apparent that there are significant uncertainties. In Table 1, variation in the $\beta^{-1} \ln Z_1$ term is relatively modest, 0.4−1.5 kcal/mol, for these cases with one or six rotatable bonds. The uncertainty from computation of relative free energies of hydration of the lowest-energy and bound conformers (1 and *k*) is probably less than 1 kcal/mol in cases where the conformers do not have significantly different hydrogen-bonding capacities.[9b] It can be expected to be greater in cases where differences in internal hydrogen-bonding are present, as for trovirdine. Still, the greatest uncertainty arises from evaluation of the conformational energetics. In a simple case such as for rotation of the isopropyl group in MKC-442, a 2 kcal/mol range of estimates is obtained for ΔE_k. However, in cases such as for UC781 and trovirdine in which the conformers under comparison are quite different owing to changes in multiple dihedral angles, errors can accumulate to yield substantial uncertainties. For UC781, the range of values for ΔE_k is greater than 6 kcal/mol including the FF results and still 2−3 kcal/mol just using the ab initio and DFT data. The situation is worse for trovirdine. One can argue that the uncertainty would be less with higher-level ab initio or DFT methods; however, the current calculations are already impractical for application to other than small data sets, since several hours were required for each HF/6-31G(d) optimization on a 3 Gz Pentium IV. These molecules are also only of average size for drugs, and the costs for such calculations increase exponentially with molecular size.

Considering the above, a typical total uncertainty in computation of ΔG_{cf} is expected to be ~5 kcal/mol, when it is evaluated fully including a conformational search with GB/SA hydration and calculation of the energetics for the key conformers with the present ab initio or DFT methods. If current force fields are used, the uncertainty can be anticipated conservatively to be in the 5−10 kcal/mol range. And, of course, $\Delta G_{cf}(L)$ only represents one component of the binding process (Scheme 1). A key problem here is that typical drug-size molecules are much larger than the small, prototypical molecules used in most FF parametrizations, and errors for torsional and nonbonded energetics can accumulate.

Docking and Scoring. Concerning the accuracy of VHTS using docking methods with typical scoring functions, the present results imply that consistent, successful ranking of diverse library members is inconceivable. In view of the ranges for ΔG_{cf} in Table 1, if ΔG_{cf} is not explicitly evaluated, random noise of 0−10 kcal/mol can be expected to modulate the scores for molecules such as the NNRTIs. The first three examples in Table 1, all with three rotatable bonds, would be penalized by the same amount, ~1.8 kcal/mol, for "frozen" torsions by ChemScore,[7] while the range for ΔG_{cf} from the OPLS/CM1A results is 2−8 kcal/mol. Even if ΔG_{cf} were fully evaluated using current force fields, the uncertainty would be the same as discussed above, about 5−10 kcal/mol. It is worth repeating that the enzyme inhibitors considered here have just an average number of torsional degrees of freedom, and $\Delta G_{cf}(L)$ represents only one component of protein−ligand binding. Enrichment, the better than random separation of active compounds from inactive ones in a molecular library, can still be achieved by docking programs.[3,4] Significant enrichment can also be achieved by just filtering on the basis of size and polarity or by other similarity measures.[18,19] Correct ranking of compounds for binding affinity or activity is a much greater challenge.

The possibility for reasonable rank ordering of compounds by docking/scoring methods is clearly improved if conformer focusing can largely be eliminated. For the ligands, this means comparing relatively rigid molecules with few torsional degrees of freedom. To keep the conformer focusing penalty for the protein relatively constant, it would also help to compare ligands that cover essentially the same region in the binding site such that the same residues of the protein are conformer-focused. Otherwise, the uncertainty from the conformer-focusing penalty for the protein is also likely in the 5−10 kcal/mol range. Furthermore, for rigid as well as more flexible ligands, the situation can be improved by emphasizing relative scores or relative free energies of binding for very similar molecules, as in lead optimization for small substituents on a ring. In this case, the conformer focusing penalties for the lead series should be essentially the same except for the differences associated with the optimized substituents.

Conclusion

The present results point out the current problems with quantitative assessment of just the conformer-focusing component of biomolecule−ligand binding. A key issue is that there is considerable uncertainty in computed relative energies for conformers of drug-size molecules. This fundamentally undermines accurate evaluation of absolute free energies of binding. The only obvious practical solution to the conformer focusing component for high-throughput studies requires the development of very accurate force fields. It is important to realize that both the torsional energetics, which can be parametrized using small molecules, and the intramolecular nonbonded energetics need to be accurate. The latter item is, in turn, very sensitive to the representation of intramolecular electrostatic interactions and raises questions about the adequacy of point-charge models, the choice of partial charges, and the needs for explicit polarization.[20] In any event, for accurate evaluation of conformational energetics and ΔG_{cf}, there is compelling need for improved force fields that have been trained using definitive ab initio or DFT results for prototypical systems and that have been extensively tested on larger drug-size molecules. In addition, the coverage of torsional motifs needs to be massive in order to encompass the possibilities for molecules of pharmacological interest. Meanwhile, expectations for consistent success in rank-ordering diverse compound libraries with high-throughput docking need to be low; the physics of protein−ligand binding is challenging.

Acknowledgment. Gratitude is expressed to the National Institutes of Health (Grant GM32136) and National Institute of Allergy and Infectious Diseases (Grant AI44616) for support.

References

(1) Ajay; Murcko, M. A. *J. Med. Chem.* **1995**, *38*, 4953−4967.
(2) Golhke, H.; Klebe, G. *Angew. Chem., Int. Ed.* **2002**, *41*, 2644−2676.
(3) (a) Shoichet, B. K.; McGovern, S. L.; Wei, B.; Irwin, J. J. *Curr. Opin. Chem. Biol.* **2002**, *6*, 439−446. (b) Muegge, I. *Rev. Comput. Chem.* **2001**, *17*, 1−60. (c) Taylor, R. D.; Jewsbury; Essex, J. W. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 151−166.
(4) (a) Bissantz, C.; Folkers, G.; Rognan, D. *J. Med. Chem.* **2000**, *43*, 4759−4767. (b) Stahl, M.; Rarey, M. *J. Med. Chem.* **2001**, *44*, 1035−1042. (c) Bursulaya, B. D.; Totrov, M.; Abagyan, R.; Brooks, C. L., III. *J. Comput.-Aided Mol. Des.* **2003**, *17*, 755−763. (d) Wang, R.;

Lu, Y.; Wang, S. *J. Med. Chem*. **2003**, *46*, 2287−2303. (e) Perola, E.; Walters, W. P.; Charifson, P. S. *Proteins: Struct., Funct., Bioinf.* **2004**, *56*, 235−249. (f) Kellenberger, E.; Rodrigo, J.; Muller, P.; Rognan, D. *Proteins: Struct., Funct., Bioinf.* **2004**, *57*, 225−242. (g) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem*. **2004**, *47*, 1739−1749. (h) Cummings, M. D.; DesJarlais, R. L.; Gibbs, A. C.; Mohan, V.; Jaeger, E. P. *J. Med. Chem*. **2005**, *48*, 962−976. (i) Chen, H.; Lyne, P. D.; Giordanetto, F.; Lovell, T.; Li, J. *J. Chem. Inf. Model.* **2006**, *46*, 401−415.

(5) Dataset from the following: Proudfoot, J. R. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1087−1090. Analysis was done with QikProp, version 3.0 (Schrödinger LLC). A conservative estimate of the number of rotatable bonds was made by excluding trivial rotors, e.g., methyl groups, double bonds, and bonds in rings under size 16.

(6) (a) Nicklaus, M. C.; Wang, S.; Driscoll, J. S.; Milne, G. A. W. *Bioorg. Med. Chem.* **1995**, *3*, 411−428. (b) Boström, J.; Norrby, P.-O.; Liljefors, T. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 383−396. (c) Vieth, M.; Hirst, J. D.; Brooks, C. L., III. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 563−572. (d) Perola, E.; Charifson, P. S. *J. Med. Chem.* **2004**, *47*, 2499−2510.

(7) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425−445.

(8) (a) Allinger, N. L.; Chen, K.; Lii, J.-H. *J. Comput. Chem.* **1996**, *17*, 642−668. (b) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490−519. (c) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc*. **1996**, *118*, 11225−11236.

(9) (a) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129. (b) Jorgensen, W. L.; Ulmschneider, J. P.; Tirado-Rives, J. *J. Phys. Chem. B* **2004**, *108*, 16264−16270.

(10) Hopkins, A. L.; Ren, J. S.; Esnouf, R. M.; Willcox, B. E.; Jones, E. Y.; Ross, C.; Miyasaka, T.; Walker, R. T.; Tanaka, H.; Stammers, D. K.; Stuart, D. I. *J. Med. Chem.* **1996**, *39*, 1589−1600.

(11) Ren, J.; Esnouf, R. M.; Hopkins, A. L.; Warren, J.; Balzarini, J.; Stuart, D. I.; Stammers, D. K. *Biochemistry* **1998**, *37*, 14394−14403.

(12) Jorgensen, W. L.; Tirado-Rives, J. *J. Comput. Chem*. **2005**, *26*, 1689−1700.

(13) Jorgensen, W. L.; Tirado-Rives, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6665−6670.

(14) *MacroModel*, version 9.1, Schrödinger LLC, New York, 2005.

(15) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision B.03; Gaussian, Inc.: Wallingford, CT, 2004.

(16) Ren, J.; Diprose, J.; Warren, J.; Esnouf, R. M.; Bird, L. E.; Ikemizu, S.; Slater, M.; Milton, J.; Balzarini, J.; Stuart, D. I.; Stammers, D. K. *J. Biol. Chem*. **2000**, *275*, 5633−5639.

(17) Kohlstaedt, L. A.; Wang, J.; Friedman, J. M.; Rice, P. A.; Steitz, T. A. *Science* **1992**, *256*, 1783−1790.

(18) Jorgensen, W. L. *Science* **2004**, *303*, 1813−1818.

(19) (a) Willett, P.; Barnard, J. M.; Downs, G. M. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 983−996. (b) Hert, J.; Willett, P.; Wilton, D. J.; Acklin, P.; Azzaoui, K.; Jacoby, E.; Schuffenhauer, A. *J. Chem. Inf. Model.* **2006**, *46*, 462−470.

(20) (a) Ponder, J. W.; Case, D. A. *Adv. Protein Chem.* **2003**, *66*, 27−85. (b) Maple, J. R.; Cao, Y.; Damm, W.; Halgren, T. A.; Kaminski, G. A.; Zhang, L. Y.; Friesner, R. A. *J. Chem. Theory Comput.* **2005**, *1*, 694−715.

JM060763I