Communications to the Editor

Free Energy Calculations in Molecular Design: Predictions by Theory and Reality by Experiment with Enantioselective Podand Ionophores

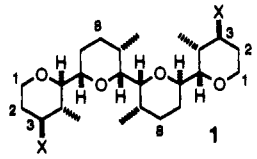
Matthew T. Burger, Alan Armstrong, Frank Guarnieri, D. Quentin McDonald, and W. Clark Still*

Department of Chemistry, Columbia University New York, New York 10027

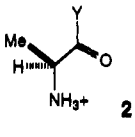
Received December 15, 1993

In recent years, free energy simulations have demonstrated themselves to be a powerful tool for understanding the structural and energetic details of a variety of chemical and biochemical phenomena.[1] Up to now, most such simulations have been directed toward reproducing known data. These are valuable studies which test developing computational methodology against reality and also provide insight into the problems studied. The real value of such simulations, however, is in predicting how new molecular entities should interact with other molecules. If energetic predictions were reliable to within a kcal/mol or so, then simulations would for the first time enable truly rational molecular design. In this communication, we describe retrospective and predictive free energy simulations of the enantioselectivity of simple podand ionophores. As we will show, these simulations not only reproduce known binding data with reasonable accuracy but can also be used to design a highly enantioselective receptor.

Because free energy simulations faithfully mimic reality only when the potential energy force field is sufficiently accurate and when all substantially populated conformational states are adequately sampled, we have chosen a binding experiment which minimizes the errors associated with these issues. This experiment involves measuring the binding selectivity of podand ionophore **1** for enantiomeric α-amino acid-derived substrates **2**. Depending



1



2

on the nature of X and Y, **1** has been reported to bind L-**2** preferentially with enantioselectivities as high as 1.1 kcal/mol.[2] In comparison with the thermodynamically governed binding of structurally distinct substrates, enantioselective binding is relatively simple because it depends only on free energy differences between diastereomeric complexes. The choice of podand ionophore **1** also simplifies binding calculations because its structure is locked by an array of rings and chiral centers into a single binding conformation. This conformational homogeneity facilitates sampling of conformational space and so obtaining converged simulation results. Thus by studying the relative binding of enantiomers and by using a conformationally homogeneous receptor, we maximize our chances for obtaining reliable free energy simulation results.

To compute the difference in the binding energies of **1** for the two enantiomers of **2** (enantioselectivity, $\Delta\Delta G_{L-D}$), we used the standard free energy perturbation (FEP) formula:[3]

$$\Delta\Delta G_{L-D} = \Delta G_L - \Delta G_D$$
$$= -K_B T \ln \langle \exp[-(H_L - H_D)/k_B T] \rangle$$

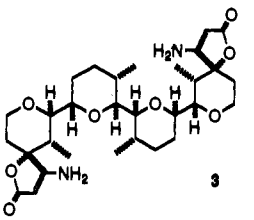
in which the two states L and D designate L- and D-**2** complexed to **1**. The perturbation we used changes the L-complex into the D-complex by interchanging the amino acid α-substituents Me and H. This mutation was carried out over 21 500-ps stages[4] and used double-wide sampling. To generate the required ensemble averages, we used the mixed Monte Carlo/stochastic dynamics method[5] at 300 K. We also used our GB/SA model for chloroform[6] along with Kollman's united atom AMBER force field[7] as implemented in our BatchMin V4.5 computer program.

Table 1. Calculated and Experimental Free Energy Differences ($\Delta\Delta G_{L-D}$, kcal/mol) for Enantioselective Binding of **1** and **2**

X (1)	Y (2)	calcd $\Delta\Delta G_{L-D}$	exptl $\Delta\Delta G_{L-D}$
H	OMe	–0.30 (σ = 0.05)	–0.4 (±0.1)[2c]
NHAc	OMe	–0.64 (σ = 0.04)	–0.8 (±0.1)[2a]
NHAc	NHMe	–0.96 (σ = 0.20)	–1.1 (±0.2)[2a]

The results of our simulations are given in the Table 1 for three **1**/**2** systems for which experimental binding data have been previously reported.[2] These results not only gave the observed preference of **1** for binding L-amino acid derivatives but also reproduced the actual extent of enantioselection within 0.3 kcal/mol. We should add that simpler calculations[8] on the two diastereomeric complexes were much less successful at reproducing experimental enantioselectivities.

Having shown that FEP simulations are in reasonable agreement with observed enantioselectivies of **1**, we used a series of analogous calculations to predict the enantioselectivities for binding **2** (Y = NHMe) with a range of derivatives of **1** substituted at C1, C2, C3, and C8 by various types of hydrogen-bond donating functionalities. According to simple molecular mechanics energy minimization, each of these derivatives could hydrogen bond to the carbonyl of a bound molecule of **2**. However, our FEP calculations predicted that only one of these derivatives, β-aminobutenolide **3**, would bind **2** with particularly high enantioselectivity. The prediction which is in the literature[9] was that L-**2**



3

should be preferentially bound by 2.7 kcal/mol (σ = 0.4) and was made in 1992, prior to initiation of our synthetic work on **3**.

To test this prediction, we synthesized **3** from previously described[2a] ketone **4** (Scheme 1). The aminobutenolide formed

(1) Reviews: Jorgensen, W. L. *Chemtracts: Org. Chem.* **1991**, *4*, 91. Kollman, P. *Chem. Rev.* **1993**, *93*, 2395.

(2) (a) Armstrong, A.; Still, W. C. *J. Org. Chem.* **1992**, *57*, 4580. (b) Wang, X.; Erickson, S. D.; Iimori, T.; Still, W. C. *J. Am. Chem. Soc.* **1992**, *114*, 4128. (c) Iimori, T.; Erickson, S. D.; Rheingold, A. L.; Still, W. C. *Tetrahedron Lett.* **1989**, *30*, 6947.

(3) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420.

(4) Significantly shorter simulation times or fewer stages gave larger free energy uncertainties or a dependence on starting geometry. *Cf.*: Mitchell, M. J.; McCammon, J. A. *J. Comput. Chem.* **1991**, *12*, 271. van Gunsteren, W. F.; Mark, A. E. *Eur. J. Biochem.* **1992**, *204*, 947.

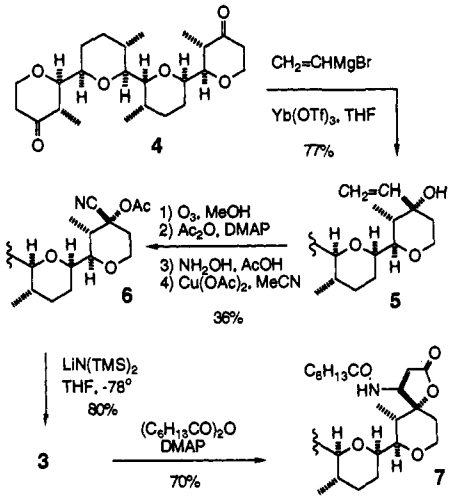
(5) Guarnieri, F.; Still, W. C. *J. Comput. Chem.*, in press.

(6) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127. GB/SA parameter set for $CHCl_3$ from F. Hollinger and W. C. Still utilizes a dielectric offset of 0.9 Å.

(7) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765. Billeter, M.; Howard, A. E.; Kollman, P. A.; Kuntz, I. D. *J. Am. Chem. Soc.* **1988**, *110*, 8385.

0002-7863/94/1516-3593$04.50/0

Scheme 1. Synthesis of Podand β-Aminobutenolide



readily upon base treatment[10] of the cyanohydrin acetate **6**, a material which had to be prepared from **4** via an organometallic addition[11] to provide stereocontrol at C3. Direct cyanohydrin formation from **4** was not possible because it led primarily to the 3-*epi* configuration. Unfortunately, **3** itself turned out to be too insoluble for study; however, we were able to acylate it using heptanoic anhydride to provide the $CDCl_3$-soluble derivative **7**.

Enantioselective binding was measured by a partitioning experiment in which 4 M aqueous racemic alaninium methylamide (**2**, Y = NHMe) hexafluorophosphate was extracted with 2.5 mM **7** in $CDCl_3$. The organic phase bearing extracted **2** was then treated with $(Boc)_2O$, and the L:D ratio of the resulting Boc-alanine derivative was found by chiral HPLC to be 94.5:5.5. Essentially the same enantioselectivity was found with the corresponding benzylamide (**2**, Y = NHBn). These measurements established the actual binding enantioselectivity to be 1.7 ± 0.1 kcal/mol favoring L. Though the measured enantioselectivity with **7** (89 ± 2% ee) turned out to be smaller than we had predicted for **3** (99% ee), our simulations did in fact lead us directly to the most enantioselective podand we have ever prepared.

It should be noted that while our original calculations were carried out on **3**, solubility problems forced us to measure actual binding enantioselectivity using the acylated derivative **7**. We therefore employed Gaussian 6-31G* with CHELPG electrostatic potential fitting to generate a charge set for the β-acetamidobutenolide fragment of **7** and used these charges in a 10.5-ns FEP simulation of **7** (CH_3CO mimicking $C_6H_{13}CO$) and **2** (Y = NHMe). This new simulation could be carried out in <36 h on a network of 10 IBM RS/6000 workstations and gave a calculated enantioselectivity of 1.9 kcal/mol (σ = 0.2) (92% ee) favoring L-**2**. We repeated the simulation but increased the length of each stage to 1000 ps (21 ns total) to verify convergence and obtained a similar final result (2.0 kcal/mol (σ = 0.2), 94% ee). Thus, simulations on **7** itself reproduced our experimental observations to within 0.3 kcal/mol.

We believe that this is the first study where converged free energy simulations of binding selectivity have been used to prioritize alternative receptor designs so that the design predicted to have the highest selectivity could be selected and prepared. While we have yet to show that rejected designs are less selective, we have demonstrated the use of a simulation to find a receptor which turns out to be significantly more selective than any of the dozen or so podands we have studied previously.[2,13]

To apply such a rational approach to molecular design, the accuracy of the calculated energies must be at least as good as the energy differences upon which the selections are based. In the current work, our free energy differences are accurate to better than 0.5 kcal/mol relative to experiment. This level of accuracy is adequate for molecular design efforts. We should note, however, that even with our conformationally simple systems, we had to go to extraordinary lengths[5,12] to converge free energy differences to precisions below 0.5 kcal/mol. Thus, while we have developed a tenable design procedure for predicting the binding properties of conformationally restricted molecules, more flexible molecular systems will require considerably more computational effort or much better methods for sampling conformational space.

Acknowledgment. This work was supported by NIH Grant HL25634 to W.C.S. and an SERC/NATO postdoctoral fellowship to A.A.

Supplementary Material Available: Atomic charges used in simulations of **1** (X = H), **1** (X = NHAc), **2** (Y = NHMe), **3**, and **7** (1 page). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, and can be ordered from the ACS; see any current masthead page for ordering information.

(8) These calculations included simple energy minimization and evaluation of average energy at 300 K (*cf.*: Lipkowitz, K. B.; Raghothama, S.; Yang, J.-a. *J. Am. Chem. Soc.* **1992**, *114*, 1554). They used the same force field, solvation treatment, and ensemble- generating algorithm that was described in the text for the FEP simulations.

(9) Still, W. C. In *Computer Aided Innovation of New Materials II*; Doyama, M., Kihara, J., Tanaka, M., Yamamoto, R., Eds.; North-Holland: Amsterdam, 1993; pp 813–816.

(10) Hiyama, T.; Oishi, H.; Suetsugu, Y.; Nishide, K.; Saimoto, H. *Bull. Chem. Soc. Jpn.* **1987**, *60*, 2139.

(11) Molander, G. A.; Burkhardt, E. R.; Weinig, P. *J. Org. Chem.* **1990**, *55*, 4990.

(12) The precision of our free energy differences depends upon the length of the simulation but is measured using 125-ps block averages to be ≤0.2 kcal/mol (1σ) for our 10-ns simulations. Convergence was verified by comparison with both shorter and longer simulations and with simulations beginning with significantly different starting geometries.

(13) Li, G.; Still, W. C. *J. Org. Chem.* **1991**, *56*, 6964. Li, G.; Still, W. C. *Tetrahedron Lett.* **1992**, *33*, 5929. Li, G.; Still, W. C. *Tetrahedron Lett.* **1993**, *34*, 919. Tang, S.; Still, W. C. *Tetrahedron Lett.* **1993**, *34*, 6701.