Evaluation of the Molecular Configuration Integral in All Degrees of Freedom for the Direct Calculation of Binding Free Energies: Application to the Enantioselective Binding of Amino Acid Derivatives to Synthetic Host Molecules

István Kolossváry

Department of Chemistry, Columbia University
New York, New York 10027
Department of Chemical Information Technology
Technical University of Budapest, Szt. Gellért tér 4
1111 Budapest, Hungary

Received May 15, 1997
Revised Manuscript Received August 25, 1997

A novel theoretical approach has been introduced recently for the direct calculation of conformational free energies without the need for expensive free energy simulations.[1] The new algorithm termed mode integration (MINTA) is based on a particularly efficient implementation of importance sampling Monte Carlo integration. MINTA allows, for the first time, the molecular configuration integral of molecular complexes of real chemical interest to be solved in all degrees of freedom, utilizing a continuum solvation model. The MINTA method was applied here to predict the enantioselective binding of α-amino acid derivatives to podand ionophore hosts, and peptide ligands to C_3-symmetric synthetic receptors. In one particular case, the correct MINTA prediction of a significant, 1.5 kcal/mol entropic stabilization of the L-Ala peptide ligand with respect to its D-Ala enantiomer in binding to the receptor was elucidated by electronic structure calculations.

The statistical-thermodynamic foundation of the calculation of binding affinities of molecular complexes is quite complex,[2] but for most practical problems, the stability of host−guest complexes can be formulated in terms of binding free energy (BFE) differences.[3] For example, one wishes to calculate the BFE difference ($\Delta\Delta G_{L-D} = \Delta G_L - \Delta G_D$) between the L and D enantiomers of a ligand bound to an enantioselective host. The direct calculation of $\Delta\Delta G_{L-D}$, in the classical sense, involves the evaluation of the molecular configuration integral Q:

$$Q_L = \sum_{i=1}^{n_L} \int_{V_i^L} e^{-(E(\mathbf{r})-E_0)/RT}\, d\mathbf{r},$$

$$Q_D = \sum_{i=1}^{n_D} \int_{V_i^D} e^{-(E(\mathbf{r})-E_0)/RT}\, d\mathbf{r} \quad (1)$$

$$\Delta\Delta G_{L-D} = -RT \ln \frac{Q_L}{Q_D} \quad (2)$$

It is assumed throughout that the dominant part of the configuration integral comes from contributions at or near low-energy binding conformations.[1,3,4] Therefore, Q is summed over respectively n_L and n_D conformations, each encompassing different V volumes of the conformational space. $E(\mathbf{r})$ is the molecular mechanics energy with respect to the nuclear coordinates $\mathbf{r}$. $E(\mathbf{r})$ includes the solvation energy as well, preferably in terms of a continuum model which does not introduce new degrees of freedom by explicit solvent molecules. E_0 is the global minimum energy, which is the common reference for both L and D binding conformations. R is the gas constant and T is the absolute temperature.

Table 1. Calculated and Experimental Binding Free Energy Differences ($\Delta\Delta G_{L-D}$, kcal/mol) for Enantioselective Binding of **1** and **2** at 300 K

			$\Delta\Delta G_{L-D}$[a]		
			AMBER*, GB/SA_{CHCl_3}		
X (**1**)	Y (**2**)	no. of confs[b]	FEP[6]	MINTA[c]	exp[6]
H	OMe	292_L, 260_D	−0.30	−0.65	−0.4
NHAc	OMe	358_L, 412_D	−0.64	−0.85	−0.8
NHAc	NHMe	235_L, 270_D	−0.96	−1.31	−1.1
β-acetamido-butenolide	NHMe	28_L, 33_D	−1.90	−1.95	−1.7

[a] Negative values favor the L-**2** enantiomer. [b] The numbers include symmetric doublets (where found) reflecting the C_2 symmetry of **1**. [c] The MINTA integrals were calculated as block averages based on 10 × 1000 independent energy evaluations *per* conformation. The resulting configuration integrals were all subject to less than 15% relative error, which is equivalent to ±0.08 kcal/mol in free energy at room temperature.[1]

Table 2. Calculated and Experimental Binding Free Energy Differences ($\Delta\Delta G_{L-D}$, kcal/mol) for Enantioselective Binding of **3** and **4** with Alanine-Derived Peptides at 300 K

			$\Delta\Delta G_{L-D}$[a]		
			AMBER*, GB/SA_{CHCl_3}		
receptor	ligand	no. of confs[b]	FEP[11]	MINTA[c]	exp[11]
3	**5**	179_L, 226_D	−2.4	−2.3	−2.2
3	**6**	102_L, 53_D	−2.6	−2.0	−2.5
3	**7**	343_L, 265_D	−1.1	−0.30	−0.3
4	**5**	146_L, 306_D	−0.43	−0.17	0.0

[a] Negative values favor the L-Ala enantiomer. [b] The numbers include symmetric triplets (where found) reflecting the C_3 symmetry of hosts **3** and **4**. [c] The error of the MINTA calculations is ±0.08 kcal/mol (see Table 1, footnote *c*).

The first binding calculation reported here involves predicting the binding selectivity of podand ionophore **1** for enantiomeric α-amino acid derivatives **2** in chloroform. **1** is a good test case for BFE calculations, because earlier calculations based on simple energy minimization and evaluation of average energy at 300 K failed to reproduce experimental enantioselectivities.[5] The MINTA results are given in Table 1 for four **1**−**2** systems for which experimental as well as computational (free energy perturbation) binding data have been previously reported.[6] The conformational search for locating the low-energy binding conformations was carried out by our particularly efficient conformational search procedure termed low-mode conformational search (LMOD).[7] The LMOD conformational search included 5000 Monte Carlo energy minimization (MC/EM) steps utilizing the united atom AMBER* force field[8] and the GB/SA continuum solvation model[9] for chloroform in BatchMin V5.5,[10] using a 50 kJ/mol energy window above the global minimum. The individual, conformational contributions to the molecular

(1) Kolossváry, I. *J. Phys. Chem. A* In press.
(2) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047−1069.
(3) Gilson, M. K.; Given, J. A.; Head, M. S. *Chem. Biol.* **1997**, *4*, 87−92.
(4) It is understood that the dominance of the low-energy conformations at room temperature is not generally true, especially for peptides. The conformational search results, however, suggest that this approximation is feasible for the binding calculations presented here.
(5) Lipkowitz, K. B.; Raghothama, S.; Yang, J.-a. *J. Am. Chem. Soc.* **1992**, *114*, 1554.
(6) Burger, M. T.; Armstrong, A.; Guarnieri, F.; McDonald, D. Q.; Still, W. C. *J. Am. Chem. Soc.* **1994**, *116*, 3593−3594.
(7) Kolossváry, I.; Guida, W. C. *J. Am. Chem. Soc.* **1996**, *118*, 5011−5019.
(8) (a) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765−784. (b) McDonald, D. Q.; Still, W. C. *Tetrahedron Lett.* **1992**, *33*, 7747.
(9) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.
(10) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440.

S0002-7863(97)01573-4 CCC: $14.00

configuration integral in eq 1 were calculated by the combined numerical/analytical MINTA algorithm.[1] The combined MINTA procedure involved the numerical integration of the 50 lowest-frequency (soft) vibrational modes and analytical integration of the remaining (hard) modes with use of the harmonic approximation. Note that the soft modes *per se* included contributions from the relative translation and rotation of **2** with respect to **1**. It should also be stressed that the LMOD-MINTA procedure was applied to unconstrained host−guest systems to make sure that both the host and the ligand were fully sampled.

The MINTA results gave the observed preference and affinity of **1** for binding L-amino acid derivatives, and are also in excellent agreement with converged free energy perturbation (FEP) simulations.[6] It should be noted that even though the structure of **1** seems to be frozen by chiral centers and interlocking rings, in fact, the first three moderately enantioselective **1**−**2** complexes afforded several hundred binding conformations within the lowest 50 kJ/mol. Timing data of the FEP simulation of the most strongly binding complex given in ref 6 suggest that MINTA is at least four times faster than FEP for this calculation.

We also applied the MINTA method to calculate the BFE differences of L- and D-alanine-derived peptides **5**−**7** binding to C_3-symmetric synthetic receptors **3** and **4** (hosts **3** and **4** are atropisomers).[11] The LMOD-MINTA calculations were carried out with the same conditions used for the podand calculations above, but a smaller, 25 kJ/mol energy window was applied for the LMOD search to keep the number of conformations within reasonable limits.

The results are given in Table 2 for four systems. **3** is strongly enantioselective with respect to **5**, but its atropisomer **4** binds L-**5** and D-**5** with virtually the same affinity. MINTA reproduced the experimental enantioselectivity of both systems within 0.2 kcal/mol, but FEP was less successful with the weakly binding **4**−**5** complex despite the fact that the FEP simulation was running twice as long (1000 ps) for **4**−**5** than for **3**−**5**.[11] **3**−**6** is the most strongly binding complex, and FEP was slightly better than MINTA for this system. The most interesting system, however, is **3**−**7**, because the moderate enantioselectivity of **3** favoring L-**7** can be rationalized in terms of particularly strong entropic effects. The global minimum energy (enthalpy) binding conformation of **3**−**7** involves the D-**7** ligand, but the global minimum free energy binding conformation involves the L-**7** ligand. The enthalpy difference between them is $\Delta H = 0.7$ kcal/mol favoring D-**7**, but the free energy difference is $\Delta G = 0.8$ kcal/mol favoring L-**7**. The resulting 1.5 kcal/mol entropic stabilization of L-**7** can be rationalized by the entirely different host−guest hydrogen bonding patterns shown in Figure 1. The D-**7** ligand is locked into a rigid binding conformation by three H bonds, while the L-**7** ligand is effectively tethered at only one point,[12] allowing almost free

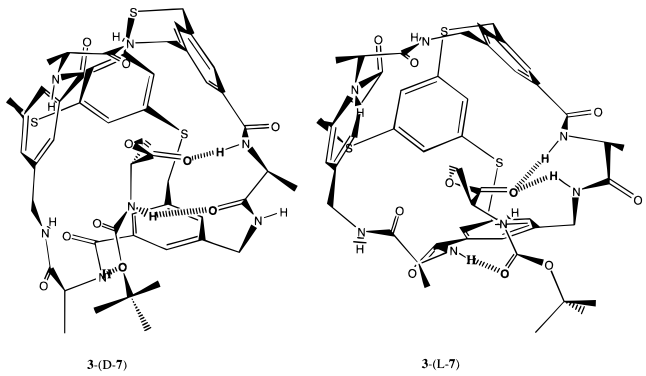
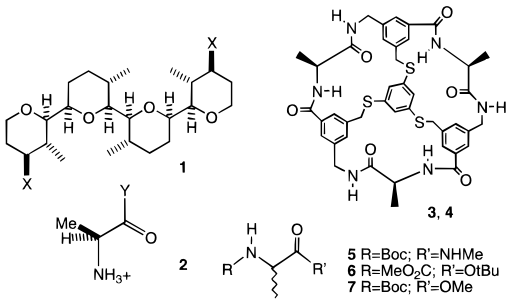


Figure 1. H-bond tethers in **3**-(D-**7**) and **3**-(L-**7**).

internal and external rotational variations for L-**7**.[13] We believe that similar entropic stabilization effects can play an important role in biological ligand−receptor interactions. It should also be noted that the experimental −0.3 kcal/mol enantioselectivity could only be reproduced by MINTA. FEP and the harmonic oscillator model significantly overestimated it (−1.1 and −0.78 kcal/mol, respectively), while a simple calculation that included only the minimized steric energy and GB/SA solvation free energy of the structures incorrectly favored the enantioselective binding of D-**7** by 0.28 kcal/mol.



We have demonstrated the utility of the MINTA method in binding free energy calculations for diverse molecular systems, for which MINTA has performed as well or better than free energy perturbation. Nonetheless, we also have to point out at least two limitations to the utility of the MINTA method. In its current implementation, MINTA cannot include solvation effects via explicit solvent models, and it is limited to approximately 250 freely moving atoms. However, utilizing continuum solvation models such as GB/SA,[9] and the application of grid based methods,[17] can increase the effective number of atoms to several thousand, including realistic solvation treatment. Therefore, we believe that MINTA should find wide utility for calculating relative binding affinities of drug molecules binding to a macromolecular host in biological systems.

Acknowledgment. I am indebted to G. M. Keseru for invaluable help with the H-bond calculations, to H. Senderowitz and D. Q. McDonald for graciously making the original raw data in refs 6 and 11 available, and to W. C. Still and W. C. Guida for carefully reading the manuscript.

JA971573N

(11) McDonald, D. Q.; Still, W. C. *J. Am. Chem. Soc.* **1996**, *118*, 2073−2077.

(12) Single point AM1 calculations were carried out with MOPAC,[14] and the ENPART command was used to partition the energy into its integral subcomponents. The two-center integral contributions spanning between a particular H-bond acceptor and its hydrogen and donor atom, respectively, were summed to model the corresponding H-bond interaction, following a partitioning scheme recently introduced for analyzing hydrogen bonding networks.[15] It has been found that one of the bonds of the bifurcating H bond in the L-**7** complex is 1.5−4 times weaker than the host−guest H bonds in the D-**7** complex and, in fact, the other bond is even repulsive.

(13) At this time there is no direct experimental evidence for the enthalpy-entropy compensation for the complexation of the **3**−**7** system such as complementary binding enthalpy and entropy measurements. However, hydrogen bonding patterns similar to those of **3**−**7** were found via ^{1}H NMR for Boc, amino acid *N*-methylamides bound to **3**,[16] providing at least indirect experimental support for our H-bond analysis.[12]

(14) Stewart, J. J. P. *QCPE J.* **1990**, *11*, 455.

(15) Hajnal, Z.; Keseru, G. M.; Simon, K. *J. Chem. Phys.* Submitted for publication.

(16) (a) Liu, R.; Still, W. C. *Tetrahedron Lett.* **1993**, *34*, 2573−2576. (b) Still, W. C.; Liu, R. *Phil. Trans. R. Soc. London A* **1993**, *345*, 97−104.

(17) Hodge, C. N.; Zacharias, M.; McCammon, J. A. *J. Comput. Chem.* **1995**, *16*, 454-464.