Computer-Assisted Scanning of Ligand Interactions: Analysis of the Fructose 1,6-Bisphosphatase−AMP Complex Using Free Energy Calculations

Mark D. Erion,* Paul D. van Poelje, and M. Rami Reddy*

Metabasis Therapeutics, Inc., 9390 Towne Centre Drive San Diego, California 92121

Received February 23, 2000
Revised Manuscript Received May 10, 2000

High-resolution X-ray structures of protein−ligand complexes reveal the active-site architecture and ligand binding mode as well as the network of electrostatic, hydrogen bond, and van der Waals interactions associated with ligand binding. Unclear from the structure, however, is the strength of each interaction and its contribution to catalysis, binding, and enzyme specificity.[1] Since this information is useful for drug design and protein engineering, site-directed mutagenesis and pseudosubstrate kinetic studies are often performed to assess the role of specific residues in catalytic efficiency and binding affinity.[2,3] While valuable in many cases, both strategies are labor intensive and frequently fail to produce data for the complete set of interactions. Moreover, these studies provide no information on solvation effects and other factors that contribute to the net binding free energy.[4] Herein, we demonstrate the utility of free energy calculations as an alternative strategy for scanning ligand binding sites and identifying interactions important for drug design.[5]

Computational analysis of binding interactions is achieved by calculating the relative binding free energy ($\Delta\Delta G_{bind}$) for a ligand L with an enzyme E relative to either a modified ligand L′ with E or a mutated enzyme E′ with L wherein both L′ and E′ differ from L and E, respectively, by a simple structural modification that primarily affects the interactions of the mutated group. Accurate results are obtained using the free energy perturbation (FEP) methodology[6] which computationally transforms L into L′ or E into E′ in solvent and in the complex to calculate ΔG_{comp} and ΔG_{aq}, the difference of which is $\Delta\Delta G_{bind}$.[7] Satisfactory convergence is achieved, however, only when the transformation entails a small structural change or in some cases a slightly larger change coupled with long simulation times.[8] Since hydrogen bonds formed between E and L are usually eliminated by a simple structural modification in either the protein or ligand, free energy calculations are useful for determining the relative intrinsic strength of each hydrogen bond and the effect of solvation on its overall contribution to ligand binding affinity.

To illustrate the value of free energy calculations for scanning binding site interactions, we targeted the adenosine monophosphate (AMP) binding site of fructose 1,6-bisphosphatase (FBPase). FBPase represents a highly attractive drug target for type II diabetes whose high-resolution X-ray structure showed 11 hydrogen bonds between FBPase and AMP (Figure 1).[9a] Since hydrogen bond strength is not readily predictable and can vary between 1 and 5 kcal/mol depending on the heavy atom pair, distance, angle, local electrostatic environment, and solvent accessibility,[1a] free energy calculations were used to identify the individual hydrogen bonds that contribute most to ligand binding affinity. These efforts were part of our ongoing quest to discover AMP mimetics[10] that exhibit high binding affinity, high enzyme specificity, and good cell penetration. AMP mimetics targeted at intracellular enzymes represent a challenging problem since a large proportion of the total binding affinity of AMP is attributed to hydrogen bonds and electrostatic interactions formed between the negatively charged phosphate and residues in the phosphate binding site. Accordingly, the difficulty in the design of AMP mimetics stems from the need for negatively charged ligands to achieve high binding affinity and the inability of charged molecules to diffuse across cell membranes and penetrate cells.

Relative free energy calculations were conducted using the structure of the human FBPase−AMP complex[9a] and free energy perturbation methodology.[11] The calculated results were compared with inhibition constants determined for AMP analogues and previously reported IC_{50}s for AMP with FBPase mutants[12] (Table 1). Consistent with the experimental data, mutation of the purine base nitrogens, i.e., ^{1}N, ^{3}N, ^{7}N, and ^{9}N, showed that replacement of ^{1}N, ^{3}N, and ^{9}N with CH has little effect on binding affinity whereas replacement of ^{7}N leads to a loss of 2.8 kcal/mol. The 0.6 kcal/mol gain in affinity for the 1-deaza and 3-deaza AMP analogues was consistent with the hydrophobic nature of this portion of the binding site cavity and the absence of hydrogen bond donors in the vicinity of either heteroatom. In contrast, the structural basis for the large loss in binding affinity for 7-deaza AMP was less apparent from X-ray structures of FBPase−AMP complexes, since the interaction between ^{7}N and the hydroxyl of Thr31 (O_g) either spanned a distance slightly outside the normal range for an optimal hydrogen bond[9a] (3.4 Å) or involved an intervening water molecule.[9b] Similar to ^{7}N, the importance of ^{9}N to AMP binding affinity was not readily predicted from the X-ray structure.[9a] Replacement of ^{9}N with a carbon atom converts ^{7}N from a hydrogen bond acceptor to a hydrogen bond donor. Unlike the $^{7}N \rightarrow CH$ mutation, this structural change is tolerated which is unusual since reversals in hydrogen bonding usually lead to a dramatic loss in binding affinity or to significant restructuring of the binding site.[13] Similar to experimental data, 9-deaza AMP (formycin A monophosphate) showed no significant difference relative to AMP in binding affinity or protein interactions, which suggests that Thr31 can act as both a hydrogen bond acceptor and a donor with ^{7}N and that the strengths of the interactions are approximately the same. In addition to ^{7}N, the Thr31 hydroxyl forms a hydrogen bond with $^{6}NH_2$. Consistent with experimental data,[12] mutation of Thr31 to Ala resulted in a 3.1 kcal/mol loss

(1) (a) Fersht, A. R. *Trends Biochem. Sci.* **1987**, *12*, 301−304. (b) Erion, M. D.; Takabayashi, K.; Smith, H. B.; Kessi, J.; Wagner, S.; Honger, S.; Shames, S. L.; Ealick, S. E. *Biochemistry* **1997**, *36*, 11725−11734.

(2) (a) Kati, W. M.; Wolfenden, R. *Science* **1989**, *243*, 1591−1593. (b) Straus, D.; Raines, R.; Kawashima, E.; Knowles, J. R.; Gilbert, W. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 2272−2276.

(3) (a) Erion, M. D.; Stoeckler, J. D.; Guida, W. C.; Walter, R. L.; Ealick, S. E. *Biochemistry* **1997**, *36*, 11735−11748. (b) Stoeckler, J. D.; Poirot, A. F.; Smith, R. M.; Parks, R. E., Jr.; Ealick, S. E.; Takabayashi, K.; Erion, M. D. *Biochemistry* **1997**, *36*, 11749−11756.

(4) Bash, P. A.; Singh, U. C.; Brown, F. K.; Langridge, R.; Kollman, P. A. *Science* **1987**, *235*, 574−576.

(5) For a recent review of free energy calculations in ligand design see: Reddy, M. R.; Erion, M. D.; Agarwal, A. In *Reviews in Computational Chemistry*; Wiley: New York, 2000, Vol. 16, pp 217−304.

(6) (a) Daggett, V.; Brown, F.; Kollman, P. *J. Am. Chem. Soc.* **1989**, *111*, 8247−8256. (b) Pearlman, D. A.; Rao, B. G. In *Encyclopedia of Computational Chemistry*; Wiley: New York, 1999; Vol. 2, pp 1036−1061.

(7) Thermodynamic cycles used to calculate relative binding and solvation free energies are described in the Supporting Information.

(8) (a) Mitchell, M. J.; McCammon, J. A. *J. Comput. Chem.* **1991**, *12*, 271−275. (b) Pearlman, D. A.; Kollman, P. A. *J. Chem. Phys.* **1991**, *94*, 4532−4545. (c) Reddy, M. R.; Erion, M. D. *J. Comput. Chem.* **1999**, *20*, 1018−1027.

(9) (a) Ke, H. M.; Zhang, Y. P.; Lipscomb, W. N. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 5243−5247. (b) Iversen, L. F.; Brzozowski, M.; Hastrup, S.; Hubbard, R.; Kastrup, J. S.; Larsen, I. K.; Naerum, L.; Norskov-Lauridsen, L.; Rasmussen, P. B.; Thim, L.; Wiberg, F. C.; Lundgren, K. *Protein Sci.* **1997**, *6*, 971−982.

(10) Erion, M. D.; Kasibhatla, S. R.; Bookser, B. C.; van Poelje, P. D.; Reddy, M. R.; Gruber, H. E.; Appleman, J. R. *J. Am. Chem. Soc.* **1999**, *121*, 308−319.

(11) Computational details are similar to those used previously, see: Erion, M. D.; Reddy, M. R. *J. Am. Chem. Soc.* **1998**, *120*, 3295−3304.

(12) Gidh-Jain, M.; Zhang, Y.; van Poelje, P. D.; Liang, J. Y.; Huang, S.; Kim, J.; Elliott, J. T.; Erion, M. D.; Pilkis, S. J.; Raafat El-Maghrabi, M.; Lipscomb, W. N. *J. Biol. Chem.* **1994**, *269*, 27732−27738.

(13) Ealick, S. E.; Babu, Y. S.; Bugg, C. E.; Erion, M. D.; Guida, W. C.; Montgomery, J. A.; Secrist, J. A. *3rd. Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 11540−11544.

10.1021/ja000651v CCC: $19.00
Published on Web 06/08/2000

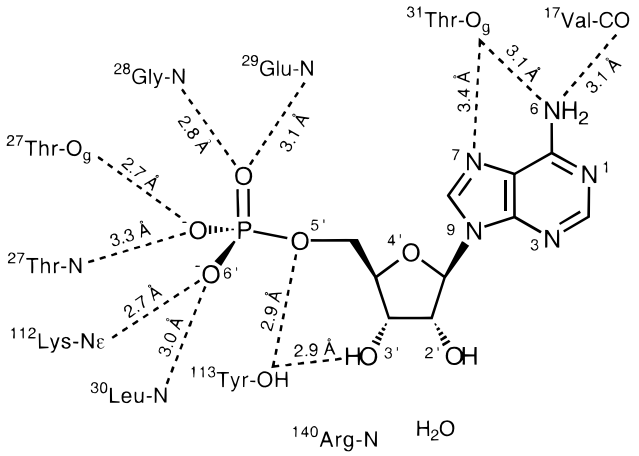

Figure 1. AMP interactions with human FBPase (Site 2).

Table 1. Relative Free Energies for AMP and FBPase Mutations

	free energy differences (kcal/mol)[b]		
mutation[a]	$\Delta\Delta G_{sol}$	$\Delta\Delta G_{bind}$(calc)	$\Delta\Delta G_{bind}$(exp[c])
$^{1}N \rightarrow CH$	0.7 ± 0.4	-0.6 ± 0.5	ND[d]
$^{3}N \rightarrow CH$	1.1 ± 0.4	-0.5 ± 0.5	ND
$^{7}N \rightarrow CH$	0.8 ± 0.5	2.8 ± 0.6	3.3[e]
$^{9}N \rightarrow C$	0.5 ± 0.4	0.6 ± 0.5	0.3[f]
$^{6}NH_2 \rightarrow H$	4.0 ± 0.6	2.3 ± 0.8	2.7[g]
$^{2'}OH \rightarrow H$	2.9 ± 0.7	-0.6 ± 0.7[l]	0.0 [h]
$^{3'}OH \rightarrow H$	2.1 ± 0.6	0.8 ± 0.7[l]	ND
$^{4'}O \rightarrow CH_2$	1.8 ± 0.5	1.1 ± 0.7	0.6[i]
$^{5'}O \rightarrow CH_2$	1.1 ± 0.6	4.6 ± 0.8	>5.4[j]
$^{113}Tyr \rightarrow Phe$	2.9 ± 0.7[m]	3.9 ± 0.9	>4.2[k]
$^{31}Thr \rightarrow Ala$	3.3 ± 0.7[m]	3.1 ± 0.8	2.9[k]

[a] Mutation denoted as AMP or FBPase residue to the corresponding analogue or residue with the indicated change. [b] Free energies are relative to AMP or wild type FBPase. [c] Experimental values are derived from the ratio of IC_{50}(AMP)/IC_{50}(AMP analogue); IC_{50} for AMP: 1.2 ± 0 μM. Other IC_{50}s: [d] ND = not determined; [e] Tubercidin monophosphate: 290 ± 6 μM; [f] Formycin A monophosphate: 2.0 ± 0.2 μM; [g] Purine riboside monophosphate: 96.7 ± 3.3 μM; [h] Aristomycin monophosphate: 3.0 ± 0.1 μM; [i] 2′-Deoxy AMP : 1.1 ± 0.1, μM; [j] 5′-Phosphonic acid: >10000 μM. [k] Reference 12. [l] Site 2. [m] ΔG_{sol} (inter).

in AMP binding free energy. This decrease most likely reflects the loss in interactions to both ^{7}N and $^{6}NH_2$. Last, mutation of the $^{6}NH_2$ to H showed that although the $^{6}NH_2$ forms two hydrogen bonds with FBPase, the overall binding free energy contribution is 0.5 kcal/mol less than ^{7}N due to its significantly larger desolvation costs ($\Delta\Delta G_{sol} = -4.0$ kcal/mol).

Free energy calculations also provided insight into the contributions of the ribosyl heteroatoms to AMP binding affinity. Mutation of the 2′-hydroxyl as well as the 3′-hydroxyl to hydrogen led to a large loss in binding affinity (>5 kcal/mol) which was inconsistent with the experimental data (≈0 kcal/mol). An analysis of the interactions present in the simulation showed that the large loss in binding affinity was most likely due to decreased interactions with the positively charged Arg140 guanidino group, which was in hydrogen bond contact with the 3′ hydroxyl directly and with the 2′ hydroxyl via an intervening water molecule (site 1).[9a] Since the Arg140Ala mutant showed only a modest decrease in AMP binding affinity,[12] the Arg140 contacts appeared to be misrepresented in the modeled structure. Analysis of the AMP interactions in the three remaining subunits showed that the side chain orientations were nearly superimposable across the four subunits except for the Arg140 side chain, which was pointed away from the ribosyl hydroxyls in subunits 2−4. Repeating the FEP calculations using site 2 gave nearly identical relative binding free energies for the ^{7}N and $^{6}NH_2$ mutations while producing much lower values for the ribosyl hydroxyl mutations (Table 1).

Scanning the remaining heteroatoms of the ribosyl phosphate ($^{4'}O$, $^{5'}O$, and $^{6'}O$) produced relative binding free energies consistent with the experimental results. The large decrease in binding free energy for the $^{6'}O \rightarrow H$ mutation was not readily quantified, however, since the reduction in atomic charge led to large changes in the solvation and complex free energies and poor overall convergence. The H-phosphonate exhibits a substantially lower binding affinity relative to AMP due to the loss of interactions with the phosphate oxygen as well as the reduced interaction strength of the remaining three oxygens, which as part of an H-phosphonate exhibit less negative charge. The $^{5'}O \rightarrow CH_2$ mutation resulted in a loss (4.6 kcal/mol) less than the experimental value (>5.4 kcal/mol) possibly due to an under estimation of the electrostatic differences between a phosphate and the less acidic phosphonic acid. Mutation of Tyr113 to Phe led to a >700-fold decrease in the calculated IC_{50} for AMP, which is similar to the experimental data.[12] Since Tyr113 hydrogen bonds to both the 5′ oxygen and the 3′ hydroxyl, it is unclear how much each hydrogen bond contributes AMP binding affinity.

The results suggest that free energy calculations can be used to scan the hydrogen bonds identified in the protein−ligand X-ray structure and accurately assess their ligand binding contribution. New analogues could therefore be designed with a strategy of preserving the strong hydrogen bonds while modifying regions that provide little binding energy with groups that enhance affinity and/or specificity.[14,10] This strategy is particularly appealing for carbohydrates, nucleosides, and nucleotides, which are heteroatom-rich molecules recognized by proteins and polynucleotides through complex hydrogen bond networks. Possibly as a consequence of nature's extensive use of electrostatic interactions to bind these compound classes, few compounds have been transformed into high affinity ligands. Computational methods that map binding site surface properties are of limited use for these compounds, since these methods primarily describe cavity boundaries and hydrophobic interactions.[15] Moreover, methods that use molecular mechanics or other approximate methods to predict the contribution to the binding free energy of individual interactions present in the ligand−protein complex are often inaccurate, since they fail to account for both entropic and solvation contributions. In contrast, relative binding and solvation free energy calculations are highly accurate, especially when the transformation entails small structural changes. Accordingly, scanning individual interactions using free energy calculations provides a set of data that quantitatively maps the binding site cavity in a manner that is readily exploited in the design of new and more effective nucleotide mimetics.

Acknowledgment. We thank Dr. William Lipscomb (Harvard University) for human FBPase crystallographic data and Juergen Schanzer for determining the FBPase inhibition constants.

Supporting Information Available: Protocols for FBPase inhibition assay, solvent and protein-complex simulations, and lists of final atomic coordinates and CHELPG charges for 11 inhibitors (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA000651V

(14) Morales, J. C.; Kool, E. T. *Nat. Struct. Biol.* **1998**, *5*, 950−954.
(15) (a) Pattabiraman, N. *J. Med. Chem.* **1999**, *42*, 3821−3834. (b) Bohacek, R. S.; McMartin, C. *J. Med. Chem.* **1992**, *35*, 1671. (c) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Venkataraghavan, R. *J. Med. Chem.* **1986**, *29*, 2149−2153. (d) Goodford, P. J. *J. Med. Chem.* **1985**, *28*, 849−857.