31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Trypsin-Ligand Binding Free Energy Calculation with AMOEBA

Yue Shi, Dian Jiao, Michael J. Schnieders, and Pengyu Ren

Abstract—**The binding free energies of several benzamidine -like inhibitors to trypsin were examined using a polarizable molecular mechanics potential. All the computed binding free energies are in good agreement with the experimental data. From free energy decomposition, electrostatic interaction was indicated to be the driving force for the binding. MD simulations show that the ligands form hydrogen bonds with trypsin and water molecules nearby in a competitive fashion. While the binding free energy is independent of the ligand dipole moment, it shows a strong correlation with the ligand molecular polarizability.**

I. Introduction

THE discovery of a lead molecule that binds to a targeted protein with high affinity is a major preoccupation of early-stage drug design [1, 2]. Accurate calculation of binding free energies is a must in this process. Treatments of protein-ligand binding, ranging from simple empirical scoring functions to thermodynamic free energy simulations with explicit solvent and full atomic details are widely used [3]. In principle, free energy perturbation (FEP) provides formally rigorous means to compute free-energy changes [4]. Although there have been numerous successful applications [3, 5, 6], calculating biomolecule-ligand affinities remains challenging for the highly polarized or charged system. Both the potential energy functions and sampling efficiency need improvement. In a previous work, we reported the calculations of the absolute binding free energy of benzamidine to trypsin and the relative binding free energy of diazamidine compared to benzamidine, using a polarizable potetinal[7, 8]. All the calculated binding free energies are well within the accuracy of experimental measurement. In this work, we systematically study the relative binding free energies of five ligands of trypsin compared to benzamidine, using a polarizable potential via explicit solvent molecular dynamics simulations. The free energies were decomposed into electrostatic and vdW components to examine the importance of different energy contributions. Also we investigated the relationship between the dipole moment, polarizability and binding free energy of the ligand.

Manuscript received April 7, 2009. This work was supported by Grant Number R01GM079686 from the National Institute of General Medical Sciences and The Robert A. Welch Foundation (grant number F-1691).

Yue Shi, Dian Jiao and Pengyu Ren* are with the Department of Biomedical Engineering, University of Texas at Austin, Austin, TX, 78712 USA (corresponding author: 512-232-1832; fax: 512-417-0616; e-mail: pren@mail.utexas.edu). Michael J. Schnieders is with Department of Chemistry, Stanford University School of Medicine, Stanford, CA 94305

II. Methods

A. Atomistic Model

The benzamidine-trypsin crystal structure (1BTY)[9] was used to generate new structures for the other ligands. Relative free energy changes of five ligands to benzamidine were investigated. Ligand B and C replace the phenyl ring of benzamidine with a 1,3-diazine and 1,4-diazine respectively. Ligand D includes an amino group at 4-position of the phenyl ring. Ligand E is the only ligand in this study with an amine group instead of amidinium group. Ligand F is a derivative of ligand D with two carbon atoms in the ring substituted by nitrogen atoms. Ligand B and C were mutated from benzamidine in the trypsin binding pocket with the crystal complex structure whereas ligand D to F were superimposed on the benzamidine in the pocket and new structures were saved. For each ligand, we soaked the protein in an octahedron box with 4515 water molecules and 58 Å on each side.

B. Force Field and Parameterization

The potential function for the entire system, including trypsin, ligand and water, is given by

$$E = E_{ele} + E_{vdW} + E_{bond} + E_{angle} + E_{torsion} + E_{oop} \quad (1)$$

In AMOEBA force field, the electrostatic interaction composes of permanent atomic charges, dipoles, quadrupoles and the polarization effect by atomic induced dipole [10-12]. The van der Waals interaction is described by a buffered-14-7 function [13]. The electrostatic parameters are derived from quantum mechanical calculation. Each ligand was first optimized with Gaussian03 package at the level of HF/6-31G* [14]. Then the single point calculation was run at MP2/6-311++G(2d,2p) and multipoles of the ligands were calculated with GDMA v2 [15]. The van der Waals (vdW), bond, angle, and atomic polarizability parameters of the ligands were transferred from AMOEBA potential (amoebapro.prm) in TINKER package[16].

C. Free Energy Perturbation

Free Energy Perturbation was used to compute the relative binding free energies between the different ligands. One ligand was perturbed from another ligand in both bulk water and the protein complex. The relative binding free energy between these two ligands can be computed as:

$$\Delta\Delta A_{bind}(L1 \to L2) = \Delta A_{pro}(L1 \to L2) - \Delta A_{wat}(L1 \to L2) \quad (2)$$

The free energy simulations were performed by changing electrostatic and van der Waals parameters between the ligands in steps. When it comes to the annihilation of atoms, the a soft-core buffered-14-7 vdW function used between the

dummy atoms and all other atoms in the system [17]:

$$U_{ij} = \lambda^5 \varepsilon_{ij} \frac{1.07^7}{\left[0.7(1-\lambda)^2 + (\rho + 0.07)^7\right]} \left(\frac{1.12}{0.7(1-\lambda)^2 + \rho^7 + 0.12} - 2 \right) \quad (3)$$

The free energies between two neighboring steps were calculated using the Bennett Acceptance Ratio estimator [18].

MD simulations were performed in parallel for all steps using PMEMD in AMBER v9. NVT dynamics simulations for 1 ns were run at each step, with a1 fs time step, and 9 Å vdW cutoff. Particle Mesh Ewald (PME) was used to treat the electrostatic interactions, with a real-space cutoff of 7.0 Å. We used the Bennett acceptance ratio of 10-5 D per atom as the convergence criterion.

III. Results and Discussion

A. Relative Binding Free Energies

Unlike the absolute binding free energy, relative binding free energy is more likely to be predicted accurately due to the small structural change and the systematic error cancellation. Ligands B through E were perturbed from ligand A whose absolute binding affinity was obtained in our previous work and ligand D was then transformed into ligand F. The calculated absolute and relative binding free energies are in excellent agreement with experimental measurements (Figure 1). The experimental binding free energies are based on inhibition constants determined by spectrophotometry or isothermal titration calorimetry under various assay conditions [19-23]. The existence of multiple experimental values for single ligand indicates that the experimental uncertainty is almost 1 kcal/mol in binding affinity.

B. Electrostatic Interaction as Driving Force for Binding

Although the separation of electrostatics and vdW contribution to the binding free energy is somewhat artificial because their values may vary in different perturbation path, the decomposition of the free energy change may provide valuable illustrations of the driving force of the binding of the ligands. Figure 2 shows the decomposition of the binding affinity for the 6 ligands we calculated. Deng *et al.* [24] reported that the repulsive and dispersive interaction contribute significantly to the binding free energy from WCA decomposition, while the electrostatic interaction is slightly unfavorable. However, these computations were limited to nonpolar ligands such as benzene, toluene and phenol. In contrast, the benzamdine ligands carry net charges and form a salt bridge with the trypsin. For these systems, the electrostatic contributions range from -4.95 to -7.97 kcal/mol, while the contributions from other interactions are only from -0.50 to 2.60 kcal/mol. Thus the electrostatic interaction is indicated as the driving force of the binding of these highly charged ligands to trypsin.

C. Molecular Dipole Moments of the Ligands

Electrostatic interactions are important factors to the trypsin-ligand recognition as the presence of the charged group is crucial. In our previous work, we computed the "polarization free-energy" in both bulk water and trypsin by turning off polarization between the ligand and trypsin. The results showed that polarization works to diminish the effect of permanent electrostatics in driving the binding of ligand to trypsin. It is not surprising as the benzamidine (+) cancels the polarization effect of Asp 171(-) in the binding pocket while forming strong electrostatic attraction. In Figure 3, we showed the molecular dipole moments and polarizability of each ligand and their correlation with binding free energy.

Essex *et al*. [25] and Talhout *et al*. [22] suggested a correlation between the molecular polarity and the binding affinity. They argued that the more polar ligand is better solvated in water and therefore has lower affinity binding to trypsin. However, the scattering plot of binding affinities and ligand dipole moments in Figure 3 does not imply any of such correlation, with a poor R square value of 0.026. Ligand B has the smallest dipole moment among the six ligands, yet its solvation free energy is the largest. The significant free energy change in bulk water (-25.51 kcal/mol) is compensated by that in complex (-23.76 kcal/mol) so that the binding affinity is no stronger than some other ligands. At the same time, ligand E, which bears the largest dipole moment (-10.80 Debye) only has a binding free energy of -5.0 kcal/mol. The electrostatic details beyond the molecular dipole moment play the important role.

Interestingly, Figure 3 shows a reasonable correlation between binding free energy and polarizability, indicating that the stronger the polarizability, the weaker the binding affinity will be. Ligand B, which has the smallest dipole moment, takes a polarizability (10.33 $Å^3$) close to ligand C (10.37 $Å^3$) and has a similar value of binding affinity with ligand C (-4.97 kcal/mol versus -4.87 kcal/mol). According to the study by Brian W. Matthews *et al*. [26], the strength of attraction is directly proportional to both the polarizability and ionization potential of the interacting molecules. Considering the limited range of values for the ionization potential, the attractive force mainly depends on the polarizability. The outcome of our calculations supports this view point. The R square value of linear fit is 0.85, RMSE 0.52, indicating a good linearity between binding affinity and polarizability.

D. Structural Analysis

There are a number of hydrogen bonds between ligands and trypsin, including the amidinium group with Asp 171, Gly 196, Ser 172 and water molecules and the amino group with Ser 177. In the crystal structure (1TBY), Asp 171 forms double hydrogen bonding with the two nitrogen atoms of benzamidine. However, it is not always the case as shown in the simulations. Take ligand A, C, D and E for example, only one hydrogen bond between Asp 171 and amidinium group was consistently observed in the simulations. This is due to the competitive interaction from a water molecule in the binding pocket. Whenever double hydrogen bonding between the ligand and the protein is missing, there is always a water molecule nearby forming a hydrogen bond with the amidinium. Ligand B forms more stable salt bridge

throughout the simulation, with the both N-O distances within 3.5 Å. As for the internal water molecule seen in the crystal structure, our simulation demonstrated the existence of this crystal water was no accident. It interacts with one of the nitrogen atoms of the ligands constantly except for ligand E which has an amine group.

It is worth noting that making the ring less hydrophobic does not improve the binding affinity. On the contrary, the ligands with nitrogen atoms in place of oxygen atoms in the phenyl ring have relatively weaker binding to the trypsin. To be more specific, ligand B and ligand C have higher binding free energies than ligand A. Moreover, the amidinium group (ligand A) has been proved to provide more interactions in the binding pocket than amine group (ligand E) and hence stronger binding. For ligand D, the amino group at 4-position of the phenyl ring formed an additional hydrogen bond with Ser 177 at the catalytic site which enhanced binding by 0.36 kcal/mol.

IV. Conclusion

In this study, the binding affinities of five positively charged benzamidine analogs to trypsin were calculated with polarizable AMOEBA force field. The relative binding free energies were computed by mutating each ligand to benzamidine in both water and protein from MD simulations. The calculated binding free energies are well within the experimental uncertainty. Our results also indicate that electrostatic interaction is the dominant force of the binding of all the ligands. Although the correlation between dipole moments and binding free energies as other group has argued were completely invisible, there is a negative correlation between the polarizability and binding free energy. The structures of the binding complexes and hydrogen bonding dynamics were also examined carefully from molecular dynamics simulations. The presence of water seems play an important role.

Acknowledgment

NIH Grant R01GM079686 and the Robert A. Welch Foundation (grant number F-1691).

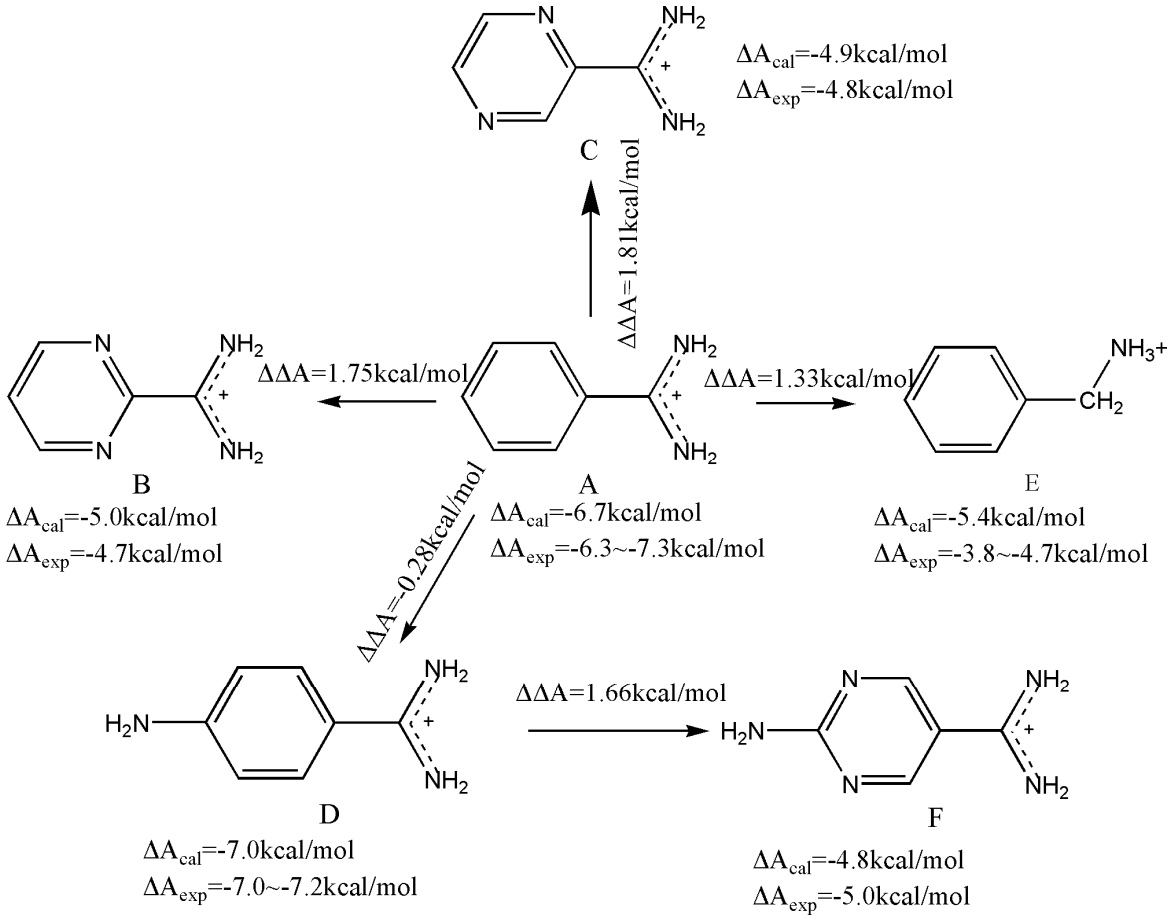

Figure 1. Relative binding free energies between ligands.

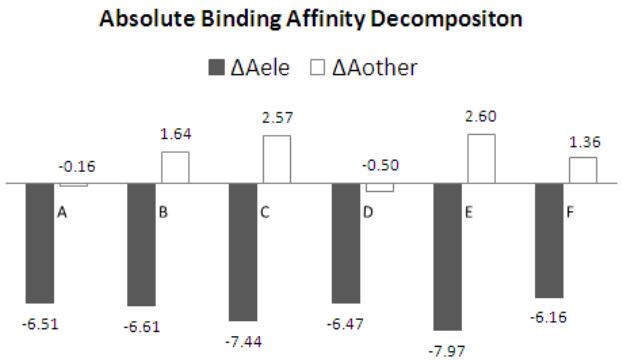

Figure 2. Decomposition of binding free energies (kcal/mol). Grey column is the electrostatic free energy and white column is the contribution of other free energy components including vdW and geometry.

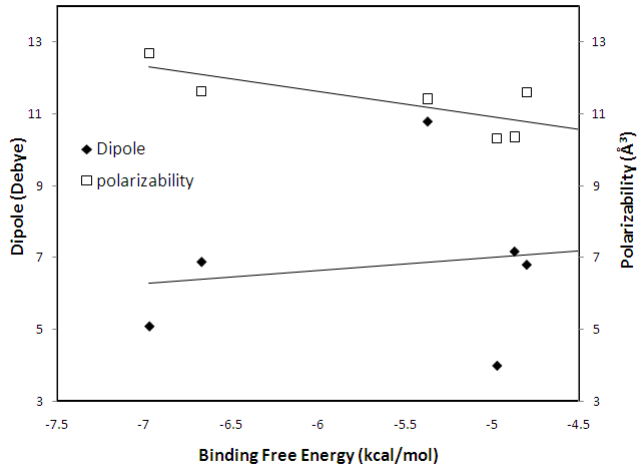

Figure 3. Correlation between dipole/polarizability of the ligands and binding free energy. Molecular dipole moments are in black diamond while polarizabilities are in open squares.

REFERENCES

[1] M. K. Gilson and H. X. Zhou, "Calculation of protein-ligand binding affinities," *Annual Review of Biophysics and Biomolecular Structure,* vol. 36, pp. 21-42, 2007.

[2] W. L. Jorgensen, "The many roles of computation in drug discovery," *Science,* vol. 303, pp. 1813-1818, MAR 19 2004.

[3] H. Gohlke and G. Klebe, "Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors," *Angewandte Chemie-International Edition,* vol. 41, pp. 2645-2676, 2002.

[4] W. L. Jorgensen, "Interactions between amides in solution and the thermodynamics of weak binding," *Journal of American Chemical Society,* vol. 111, p. 2, 1989.

[5] P. Kollman, "Free-Energy Calculations - Applications to Chemical and Biochemical Phenomena," *Chemical Reviews,* vol. 93, pp. 2395-2417, NOV 1993.

[6] M. L. Lamb and W. L. Jorgensen, "Computational approaches to molecular recognition," *Current Opinion in Chemical Biology,* vol. 1, pp. 449-457, Dec 1997.

[7] D. Jiao, P. A. Golubkov, T. A. Darden, and P. Ren, "Calculation of protein-ligand binding free energy by using a polarizable potential," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 105, pp. 6290-6295, Apr 2008.

[8] D. Jiao, J. Zhang, R. E. Duke, G. Li, M. J. Schnieders, and P. Ren, "Trypsin-ligand binding free energies from explicit and implicit solvent simulations with polarizable potential," *J Comput Chem,* Apr 27 2009.

[9] B. A. Katz, J. Finermoore, R. Mortezaei, D. H. Rich, and R. M. Stroud, "Episelection - Novel K-I-Similar-to Nanomolar Inhibitors of Serine Proteases Selected by Binding or Chemistry on an Enzyme Surface," *Biochemistry,* vol. 34, pp. 8264-8280, Jul 4 1995.

[10] J. W. Ponder and D. A. Case, "Force fields for protein simulations," *Protein Simulations,* vol. 66, pp. 27-+, 2003.

[11] P. Y. Ren and J. W. Ponder, "Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations," *Journal of Computational Chemistry,* vol. 23, pp. 1497-1506, DEC 2002.

[12] P. Y. Ren and J. W. Ponder, "Polarizable atomic multipole water model for molecular mechanics simulation," *Journal of Physical Chemistry B,* vol. 107, pp. 5933-5947, JUN 19 2003.

[13] T. A. Halgren, "Representation of van der Waals (vdW) Interactions in Molecular Mechanics Force Fields: Potential Form, Combination Rules, and vdW Parameters," *Journal of the American Chemical Society,* vol. 114, pp. 7827-7843, 1992.

[14] M. J. e. a. Frisch, "Gaussian 03," *(Pittsburgh, PA: Gaussian Inc.),* 2003.

[15] A. J. Stone, "Distributed multipole analysis: Stability for large basis sets," *Journal of Chemical Theory and Computation,* vol. 1, pp. 1128-1132, Nov-Dec 2005.

[16] P. Ren and J. W. Ponder, "Tinker polarizable atomic multipole force field for proteins.," *Abstracts of Papers of the American Chemical Society,* vol. 224, pp. U473-U473, Aug 18 2002.

[17] T. C. Beutler, A. E. Mark, R. C. Vanschaik, P. R. Gerber, and W. F. Vangunsteren, "Avoiding Singularities and Numerical Instabilities in Free-Energy Calculations Based on Molecular Simulations," *Chemical Physics Letters,* vol. 222, pp. 529-539, Jun 3 1994.

[18] C. Bennett, "Efficient Estimation of Free Energy Differences From Monte Carlo Data," *Journal of Computational Physics,* vol. 22, pp. 245-268, 1976.

[19] S. M. Schwarzl, T. B. Tschopp, J. C. Smith, and S. Fischer, "Can the calculation of ligand binding free energies be improved with continuum solvent electrostatics and an ideal-gas entropy correction?," *Journal of Computational Chemistry,* vol. 23, pp. 1143-1149, Sep 2002.

[20] F. Grater, S. M. Schwarzl, A. Dejaegere, S. Fischer, and J. C. Smith, "Protein/ligand binding free energies calculated with quantum mechanics/molecular mechanics," *Journal of Physical Chemistry B,* vol. 109, pp. 10474-10483, May 26 2005.

[21] N. Ota, C. Stroupe, J. M. S. Ferreira-da-Silva, S. A. Shah, M. Mares-Guia, and A. T. Brunger, "Non-Boltzmann thermodynamic integration (NBTI) for macromolecular systems: Relative free energy of binding of trypsin to benzamidine and benzylamine," *Proteins-Structure Function and Genetics,* vol. 37, pp. 641-653, Dec 1 1999.

[22] R. Talhout and J. B. F. N. Engberts, "Thermodynamic analysis of binding of p-substituted benzamidines to trypsin," *European Journal of Biochemistry,* vol. 268, pp. 1554-1560, Mar 2001.

[23] B. A. Katz, K. Elrod, C. Luong, M. J. Rice, R. L. Mackman, P. A. Sprengeler, J. Spencer, J. Hataye, J. Janc, J. Link, J. Litvak, R. Rai, K. Rice, S. Sideris, E. Verner, and W. Young, "A novel serine protease inhibition motif involving a multi-centered short hydrogen bonding network at the active site," *Journal of Molecular Biology,* vol. 307, pp. 1451-1486, Apr 2001.

[24] Y. Q. Deng and B. Roux, "Calculation of standard binding free energies: Aromatic molecules in the T4 lysozyme L99A mutant," *Journal of Chemical Theory and Computation,* vol. 2, pp. 1255-1273, Sep 12 2006.

[25] J. W. Essex, D. L. Severance, J. TiradoRives, and W. L. Jorgensen, "Monte Carlo simulations for proteins: Binding affinities for trypsin-benzamidine complexes via free-energy perturbations," *Journal of Physical Chemistry B,* vol. 101, pp. 9663-9669, Nov 13 1997.

[26] M. L. Quillin, W. A. Breyer, I. J. Griswold, and B. W. Matthews, "Size versus polarizability in protein-ligand interactions: Binding of noble gases within engineered cavities in phage T4 lysozyme," *Journal of Molecular Biology,* vol. 302, pp. 955-977, Sep 29 2000.