Structural Chemistry, Vol. 7, No. 2, 1996

Calculations of the Electrostatic Free Energy Contributions to the Binding Free Energy of Sulfonamides to Carbonic Anhydrase

Jeffry D. Madura,[1,4] Yasushi Nakajima,[1,3] Rodney M. Hamilton,[1] Andrzej Wierzbicki,[1] and Arieh Warshel[2]

Received July 6, 1995; revised September 6, 1995; accepted September 6, 1995

The interactions between biologically important enzymes and drugs are of great interest. In order to address some aspects of these interactions we have initiated a program to investigate enzyme-drug interactions. Specifically, the interactions between one of the isozymes of carbonic anhydrase and a family of drugs known as sulfonamides have been studied using computational methods. In particular the electrostatic free energy of binding of carbonic anhydrase II with acetazolamide, methazolamide, *p*-chlorobenzenesulfonamide, *p*-aminobenzenesulfonamide and three new compounds (MK1, MK2, and MK3) has been computed using finite-difference Poisson-Boltzmann (FDPB) [1] method and the semimacroscopic version [2, 3] of the protein dipole Langevin dipole (PDLD) method [4]. Both methods, FDPB and PDLD, give similar results for the electrostatic free energy of binding even though different charges and different treatments were used for the protein. The calculated electrostatic binding free energies are in reasonable agreement with the experimental data. The potential and the limitation of electrostatic models for studies of binding energies are discussed.

INTRODUCTION

Electrostatic calculations have become a very useful tool in calculating solvation energies and electrostatic free energies of binding for a variety of systems [1]. There are several different methods in which one can calculate the electrostatic free energy. In one approach the macroscopic Poisson-Boltzmann equation is solved using finite differences (FDPB) [1], while another uses a semimacroscopic version of the microscopic protein-dipole Langevin-dipole (PDLD) [4] and its semimacroscopic variant [3]. The present paper uses these two methods in calculating the free energy of binding between carbonic anhydrase and sulfonamide inhibitors.

Carbonic anhydrase is a metaloenzyme which reversibly catalyzes the interconversion between carbon dioxide (CO_2) and bicarbonate anion $(HCO_3)^-$ [5]. There are several well-known human erythrocyte carbonic anhydrase isozymes. One of the more active forms of these isozymes is human carbonic anhydrase II (HCA II). Binding of a variety of ligands to HCA II has been extensively studied kinetically and structurally [6, 7]. As far as the structure is concerned there exists several X-ray structures of the wild-type and various mutants with and without bound ligands. Figure 1 shows active site of HCA II with the bound inhibitor *p*-aminobenzensulfonamide. Molecular mechanics and free-energy per-

[1] University of South Alabama, Department of Chemistry, Mobile, Alabama.

[2] University of Southern California, Department of Chemistry, Los Angeles, California.

[3] Permanent address: Department of Biochemistry, Saitama Medical School, 38 Morohongo, Moroyama-cho, Iruna-gun, Saitama 350-04, Japan.

[4] Correspondence should be directed to Jeffry Madura, University of South Alabama, Department of Chemistry, Mobile, Alabama 36688-0001.

1040-0400/96/0400-0131$09.50/0

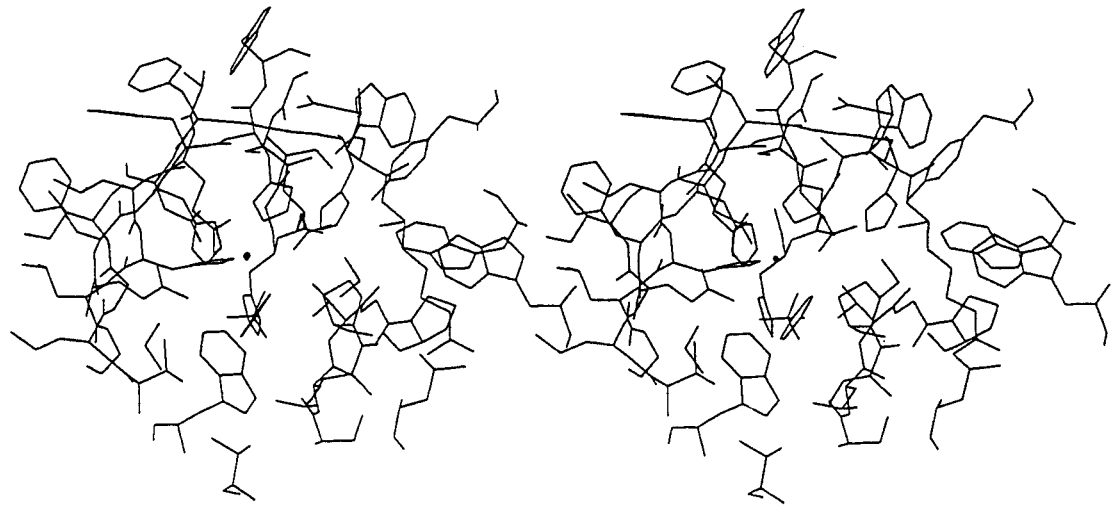


Fig. 1. Stereoview of the active site of carbonic anhydrase II with the bound inhibitor, *p*-aminobenzenesulfonamide. The zinc atom is a small dot in the center of the figure.

turbation (FEP) methods, have been performed to determine the binding strengths of various ligands to HCA II [8, 9]. Although the FEP procedure provides in principle a rigorous tool for study of these types of interactions [10, 11], it is not a method that can be used as a routine tool to evaluate several inhibitors in a short period of time. A powerful alternative for full FEP calculations is provided by the linear response approximation (LRA) which has been used in successful studies of binding free energies [2, 3, 12]. This method, however, is still expensive to be used in fast screening of different drugs (see discussion in Ref. 13). Thus it is tempting to examine the performance of electrostatic models that may provide an effective way of obtaining approximated binding free energies.

This paper focuses on the carbonic anhydrase–sulfonamide system as a model to compare FDPB and PDLD method in studying enzyme–inhibitor interactions and to demonstrate the capabilities of these methods as rapid evaluation tools to study enzyme–substrate interactions.

Seven different inhibitors bound to HCA II have been selected for this study (Fig. 2). We analyzed acetazolamide (ACZ), methazolamide (MTH), *p*-chlorobenzenesulfonamide (CHL), *p*-aminobenzenesulfonamide (AMS) and three Merck inhibitors labeled MK1, MK2, and MK3. The first four inhibitors will be used to establish charges for the sulfonamide functional group (SO_2NH^-) and protocol for the calculation while the last three will be used to test the protocol and the accuracy in which the free energy of binding can be calculated.

In the following section of this paper a brief overview is given for the FDPB and PDLD methods. The next section describes the computational approach used in this paper while the final section discusses the results.

OVERVIEW

The electrostatic free energy of binding can be calculated using equation

$$\Delta G^{elec}_{binding} = G^{elec}_{complex} - (G^{elec}_{enzyme} + G^{elec}_{drug}) \quad (1)$$

The first term is the electrostatic free energy in which the drug is bound to the enzyme, while the second and third terms are the electrostatic free energy of the isolated enzyme and electrostatic free energy of the isolated ligand respectively. In this paper we calculate the electrostatic free energy using two methods. The first method, FDPB, uses continuum electrostatics and a

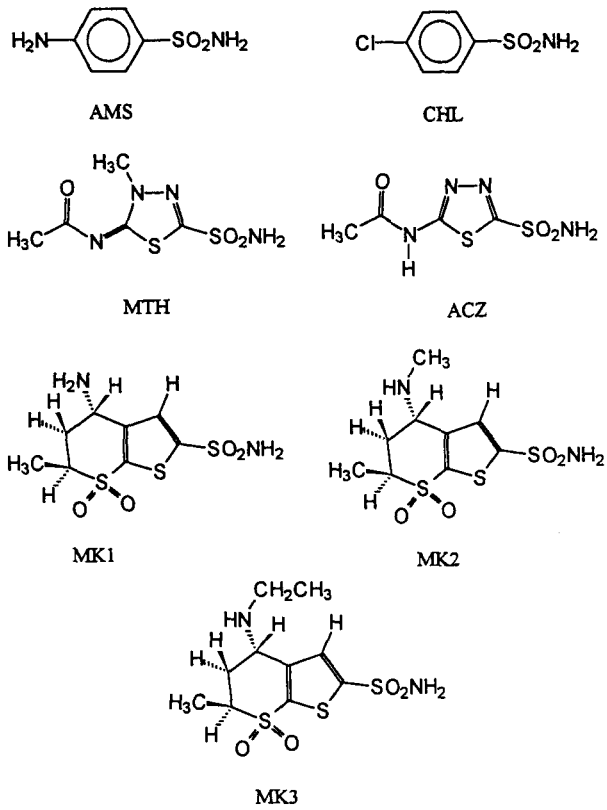


Fig. 2. Sulfonamide inhibitors of carbonic anhydrase. They are *p*-aminobenzene sulfonamide (AMS), *p*-chlorobenzenesulfonamide (CHL), methazolamide (MTH), acetazolamide (ACZ) and three Merck compounds MK1, MK2, and MK3.

simple statistical mechanical description of mobile ions in solution to calculate the electrostatic potential surrounding a solute composed of charges and an internal dielectric immersed in a continuum. The electrostatic potential is calculated by solving the Poisson-Boltzmann equation (Eq. (2)) using finite-difference techniques.

$$\nabla \cdot \epsilon \nabla \phi = \rho^f \lambda \sum_i c_i q_i \exp\left(\frac{-q_i \phi}{k_B T}\right) \qquad (2)$$

Here ϵ is a position dependent dielectric constant, ϕ is the electrostatic potential, ρ^f is the fixed charge distribution of the solute, λ is a function which is 1 in ion-accessible regions and 0 everywhere else, c is the mobile ion concentration, q is the mobile ion charge, k_B is Boltzmann's constant, and T is the absolute temperature. The details of FDPB calculations are discussed in a recent review article [14]. Once the electrostatic potential has been obtained, the electrostatic free energy is calculated using the relationship

$$G^{\text{elec}} = -\tfrac{1}{2} \sum_i q_i \phi_i \qquad (3)$$

where q_i is the ith atomic charge and ϕ_i is the calculated electrostatic potential at i. The electrostatic binding free energy, $\Delta G^{\text{elec}}_{\text{bind}}$, from Eq. (1) is determined by first calculating the electrostatic free energy of the complex, $G^{\text{elec}}_{\text{complex}}$, by solving Eq. (2) for the potential, ϕ, and using that potential in Eq. (3) to calculate $G^{\text{elec}}_{\text{complex}}$. Next the electrostatic free energy of the enzyme, $G^{\text{elec}}_{\text{enzyme}}$, is determined without the drug. Finally the electrostatic free energy of the drug, $G^{\text{elec}}_{\text{drug}}$, is determined without the enzyme. The second method used in this paper, PDLD, is a more microscopic approach. The standard PDLD approach avoids a potential problem in the assignment of a "dielectric constant" for a protein which can be different in various regions of a protein. The protein is surrounded with a solvent described by Langevin dipoles while the protein is treated as a set of charges and point dipoles. The system is divided into three regions. Region I contains the reference region which contains the solute while Region II are the rest of the protein atoms around region I and region III contains the solvent molecules. The electrostatic free energy G^{elec} is expressed as a sum of macroscopic contributions:

$$G^{\text{elec}} = U_{\text{QQ}} + U_{\text{Q}\mu} + U_{\text{Q}\alpha} + G_{\text{Qw}} \qquad (4)$$

where U_{QQ} is the Coulombic interaction energy within region I, $U_{\text{Q}\mu}$ is the interaction between the charges of region I and the permanent dipoles of region II, $U_{\text{Q}\alpha}$ is the energy of polarization of the protein-induced dipoles by the charges of region I, and G_{Qw} is the free energy of polarization of the surrounding solvent molecules (see Refs. 2-4 for more details).

We used a semimicroscopic version of the PDLD approach which is referred to as the PDLD/S approach (see Refs. 2, 3 for details). The approach represents the effect of the protein induced dipoles and a part of the effect of the reorganization of the protein permanent dipoles and penetration by an effective dielectric constant (see Ref. 15 for the meaning of this "dielectric"). The PDLD/S converges faster than the PDLD method and therefore should be more effective for our purpose. The PDLD/S method as implemented in the program POLARIS [3] involves automatic averaging over protein configurations generated by MD calculations [3] and the corresponding binding energy includes in addition to the pure electrostatic term a field dependence hydrophobic term and consistently determined reorganization energy term that reflects the effect of change of protein structure

upon ligand binding. The relevant thermodynamic cycle is described in detail in Ref. 2.

METHODS

In this section the development of parameters for the ligands, along with the process in which coordinates for the enzyme–ligand complexes were derived, and the protocol used in the FDPB and PDLD/S calculation are presented.

Inhibitor Charges

In order to carry out the electrostatic calculations atomic charges and atomic radii need to be known. The approach taken here was to use inhibitors 1–4 to establish a set of charges for the sulfonamide functional group which could be then used for inhibitors 5–7. The initial sulfonamide charges were obtained from 6-31G*/CHELPG [16] calculations for $H_2SO_2NH^-$. These charges were then adjusted to reproduce the experimental free energy of binding for the inhibitors 1–4 using the PDLD method as implemented in the program POLARIS [3]. The final charges used in these calculations for all inhibitors are given in Figs. 3 and 4. UHBD could have been used instead of POLARIS to determine the final set of charges.

Enzyme-Inhibitor Complexes

Since X-ray structures for all of the enzyme–ligand complexes were not available, we used QUANTA 4.0 [17] to generate the initial coordinates for our calculations. The X-ray coordinates for the *p*-aminobenzenesulfonamide/HCA-II (3ca2.pdb) from the Brookhaven Protein Data Bank [18] were used as a basis for the construction of all inhibitor/enzyme complexes. The enzyme–substrate complexes for which there were no coordinates were constructed by overlaying the sulfonamide group of the new inhibitor with the sulfonamide group of *p*-aminobenzensulfonamide. The new complex was then minimized using 200 steps of the steepest descent algorithm of CHARMm22 [17]. These structures were used as a starting point for the electrostatic calculations. The enzyme atomic charges and atomic radii needed for the FDPB calculations were those used by CHARMm22 and QUANTA 4.0 while the ENZYMIX force field parameters [3] were used in the PDLD calculations. The charge for zinc in both the UHBD and POLARIS calculations was +1.

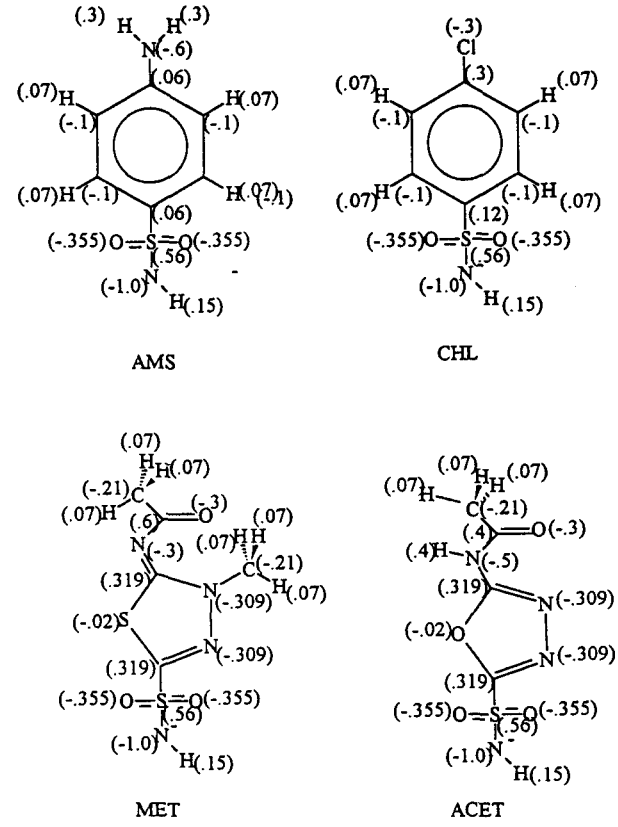


Fig. 3. Atomic charges for *p*-aminobenzenesulfonamide (AMS), *p*-chlorobenzenesulfonamide (CHL), methazolamide (MET) and acetazolamide (ACET).

The UHBD Method

The FDPB electrostatic free energies were computed using the QUANTA/UHBD [17] interface. This interface was used to construct the enzyme–inhibitor coordinates, charges, and radii files. A separate UHBD script was used to calculate the electrostatic free energy of binding. The electrostatic free energy was obtained by calculating the free-energy solvation for the enzyme using a coarse grid with a grid spacing of 1.6 Å followed by a fine grid with a grid spacing of 0.25 Å. The use of a grid with larger spacing followed by the use of a grid with smaller spacing, known as "focusing," is done in order to obtain a better representation of the electrostatic potential in the region of interest, e.g., the active site [19]. The same procedure was repeated for the inhibitor and finally the enzyme–inhibitor complex. The electrostatic free energies of binding were calculated for five enzyme/inhibitor configurations. Each configuration was obtained by running short (0.5 ps)

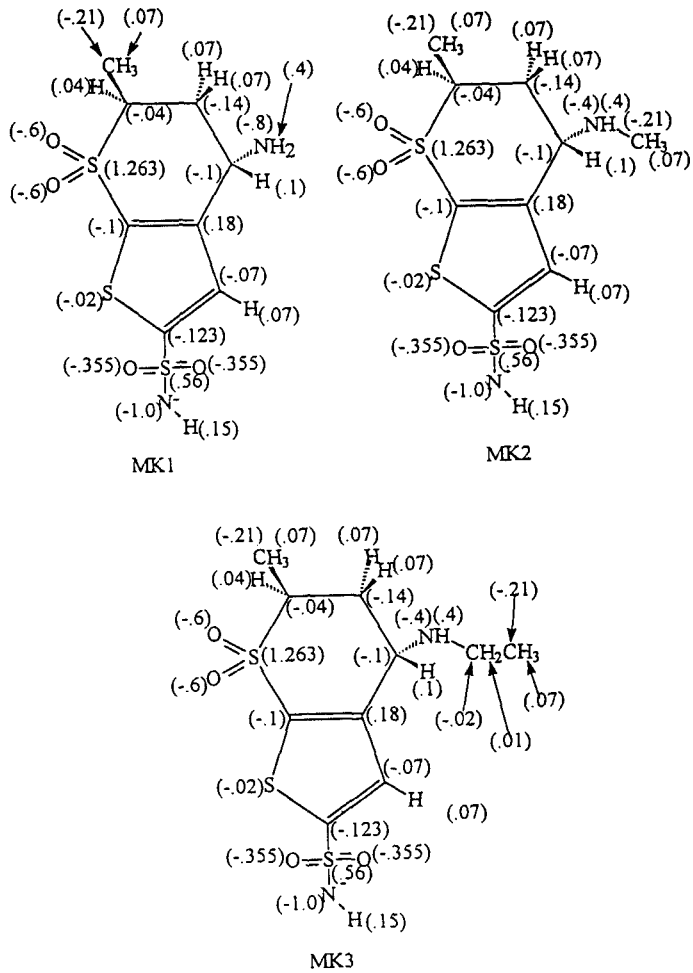


Fig. 4. Atomic charges for inhibitors MK1, MK2, and MK3.

molecular dynamics calculation and then using that configuration to create new coordinate files for the electrostatic calculation by UHBD. The molecular dynamics calculations were performed using the QUANTA/CHARMm interface. The number of grid points for all of these calculations was 60^3. The enzyme and inhibitor dielectric constant was 2.0, while the water dielectric constant used was 80.

The PDLD Method

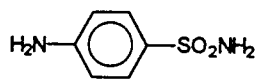
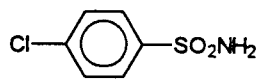
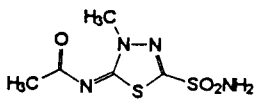
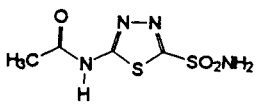
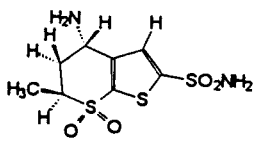
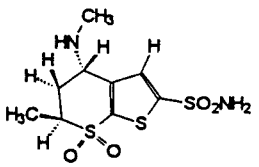
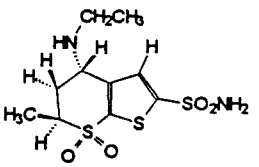
The PDLD/S electrostatic calculations were performed using the program POLARIS [3]. Calculating the electrostatic free energy of binding using POLARIS involves two steps. The first was a preparation step in which the enzyme–inhibitor system is broken into the three regions while the second step involved the calculation of the electrostatic free energy of binding. In the preparation phase all residues beyond 15 Å of the zinc center were trimmed and converted to glycine residues. This was done to reduce the size of the explicit system which is, however, surrounded by a bulk of high dielectric constant (see Ref. 3). Region I was taken as the inhibitor, region II included all the protein residues within 15 Å cutoff radius. Region III was given a radius of 16 Å. Once the system was prepared the electrostatic free energy was obtained automatically by a sequence of eight MD simulations (of 1.0 ps each) where every MD step is followed by a PDLD/S calculation. The MD simulation was performed automatically by a built in ENZYMIX [3] force field of the solvated enzyme substrate complex with the local reaction field long-range treatment [3]. Four of the MD simulations were done with the actual ligand residue charges and four where

the residue charges set to zero. This procedure allows for a consistent evaluation of reorganization energies [3].

RESULTS AND DISCUSSION

Table I summarizes the results of the FDPB and PDLD/S calculations as well as the corresponding experimental binding free energies [20, 21]. The FDPB calculations used a "protein dielectric" $\epsilon = 4$ while the best PDLD/S results were obtained with $\epsilon = 6$. As discussed in Ref. 15 the value of the dielectric used in semimacroscopic calculations reflects all the effects that are not included explicitly (including penetration of solvent upon change of charges) and the value of this dielectric "constant" is expected to increase for highly charged sites such as the zinc ion site. As seen from the table, the calculated PDLD/S and FDPB energies approach the corresponding observed binding free enegies in a qualitative way. It should be noted that the FDPB does not include hydrophobic contributions while the PDLD/S does include hydrophobic and reorganization

Table I. Summary of Absolute Free Energy of Binding Calculations for Carbonic Anhydrase with Sulfonamide Inhibitors Using UHBD (FDPB) and POLARIS (PDLD/S).[a]

Structure	Name	Expt.[a]	PDLD/S Electro $\Delta\Delta G_{bind}$[a]	FDPB Electro $\Delta\Delta G_{elec}$[a]
	*p*-aminobenzene sulfonamide	−8.3	−10.4	−8.1
	*p*-chlorobenzene sulfonamide	−9.4	−12.7	−9.7
	methazolamide	−10.7	−11.1	−7.2
	acetazolamide	−11.1	−9.9	−7.4
	MK-1	−12.0	−15.3	−14.0
	MK-2	−11.9	−14.5	−13.5
	MK-3	−12.8	−14.6	−14.0

[a] Units are kcal/mol. The PDLD/S calculations include hydrophobic and reorganization contributions in addition to the pure electrostatic term. The assumed protein dielectrics were 6 and 4 for the PDLD/S and FDPB respectively.

energy contributions, and both the PDLD/S and FDPB models do not include configurational entropies. Interestingly both approaches give similar results although the charges of the sulfonamide inhibitors 1–4 were developed using the PDLD approach and although the PDLD/S calculations involved a "trimmed" protein which only includes amino acids within a 15 Å sphere from the zinc atom were the original amino acids while in the UHBD calculation all of the amino acids of the enzyme were used. It is important to note however that previous POLARIS and ENZYMIX studies have demonstrated that using PDLD/S truncated protein models in conjunction with proper spherical boundary conditions give reliable results (e.g., Refs. 22, 23), although systems of more than 15 Å are usually recommended.

In both methods, FDPB and PDLD/S, several enzyme-substrate configurations were necessary to obtain an electrostatic free energy of binding. The use of a short molecular dynamics simulations (0.5–1.0 ps) between each configuration appears to be sufficient to generate possible alternate configurations. In the FDPB approach, focusing is necessary to obtain accurate electrostatic free energy of binding. Initially we observe good free energy of binding using a coarse grid (grid spacing of 1.6 Å) but found that when we decreased the grid spacing the calculated electrostatic free energies of binding had not converged. The electrostatic free energy of binding converged as the grid spacing approached 0.25 Å. In addition this spacing has been shown to give accurate electrostatic free energies of solvation [24]. As noted earlier, a charge of +1 for the zinc atom was used. When the zinc atom was given a charge of +2 the binding free energies were too large. Another issue which could effect the calculated results in this work but not considered was the pK_a of the ligand. In all of the ligands except methazolamide, the pK_a for the sulfonamide group is approximately 9–10 while for methazolamide it is 7.4. As mentioned earlier we did not include configurational entropy in the binding free energy. This contribution represents the entropic effects associated with the restriction of the configurational space of the ligand upon binding.

The time required to obtain one free energy of binding by the FDPD approach is approximately 90 min on an IBM RISC 6000 Model 350. This value includes the time to perform the 0.5 ps molecular dynamics calculation with QUANTA 4.0 and the time necessary to perform the UHBD calculation with focusing. The PDLD/S approach takes approximately 60 min to perform the same calculation.

In conclusion, the results from this work illustrates that both the FDPB and PDLD methods can be used as a relatively rapid tool to estimate electrostatic free energy of binding for enzyme/substrate interactions.

ACKNOWLEDGMENTS

The computers used in this paper were supported by a NSF/EPSCoR Equipment grant. JDM would like to thank Cray Research and the Alabama Supercomputer Authority for providing computer time and financial assistance. Dr. Mike Gilson is acknowledged for his comments and the UHBD script to perform the focusing calculations.

REFERENCES

1. Madura, J. D.; Davis, M. E.; Gilson, M. K.; Wade, R. C.; Luty, B. A.; McCammon, J. A. In *Rev. Comp. Chem.* Boyd, D., Lipkowitz, K., Ed.; VCH: New York, 1994; Vol. 5, p. 229.
2. Lee, F. S.; Chu, Z.-T.; Bolger, M. B.; Warshel, A. *Protein Eng.* **1992,** *5*, 215.
3. Lee, F. S.; Chu, Z.-T.; Warshel, A. *J. Comput. Chem.* **1993,** *14*, 161.
4. Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; Wiley: New York, 1991.
5. Kannan, K. K.; Petef, M.; Fridborg, K.; Cid-Dresdner, H.; Lovgren, S. *FEBS Lett.* **1977,** *73*, 115.
6. Taylor, P. W.; King, R. W.; Burgen, A. S. V. *Biochem.* **1970.**
7. Eriksson, A. E.; Kylsten, P. M.; Jones, T. A.; Liljas, A. *Proteins* **1988,** *4*, 283.
8. Merz, K. M.; Murcko, M. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1991,** *113*, 4484.
9. Vedani, A.; Meyer, J. E. F. In *Molecular Dynamics and Protein Structure*; Hermans, J., Ed.; Polycrystal Book Service: Western Springs, 1984.
10. Straatsma, T. P.; McCammon, J. A. In *Annu. Rev. Phys. Chem.*; Strauss, H. L.; Babcock, G. T.; Leone, S. R., Ed.; Annual Reviews, Inc.: Palo Alto, 1992; Vol. 43, p 407.
11. Warshel, A.; Aqvist, J. In *Ann. Rev. Biophys. Biophys. Chem.* Annual Reviews, Inc.: 1991; Vol. 20; p 267.
12. Aqvist, J.; Medina, C.; Samuelsson, J.-E. *Protein Eng.* **1994,** *7*, 385.
13. Warshel, A.; Tao, H.; Fothergill, M.; Chu, Z.-T. *Isr. J. Chem.* **1994,** *34*, 253.
14. (a) Davis, M. E.; McCammon, J. A. *Chem. Rev.* **1990,** *90*, 509. (b) Nicholls, A.; Honig, B. *J. Comput. Chem.* **1991,** *12*, 435.
15. King, G.; Lee, F. S.; Warshel, A. *J. Chem. Phys.* **1991,** *95*, 4366.
16. Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987,** *8*, 894.
17. Quanta 4.0 molecular modeling software from Molecular Simulations, Inc. Burlington, MA, 1994.
18. Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, J. E. F.; Brice, M. D.; Rogers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, *J. Mol. Biol.* **1977,** *112*, 535.
19. (a) Gilson, M. K., Sharp, K. A., Honig, B. H. *J. Comput. Chem.* **1987,** *9*, 327. (b) Madura, J. D.; Briggs, J. M.; Wade, R. C.; Davis, M. E.; Luty, B. A.; Ilin, A.; Antosciewicz, J.; Gilson, M. K.; Bagheri, B.; Scott, L. R.; McCammon, J. A. *Comput. Phys. Commun.* **1995,** *91*, 57.
20. Maren, T. H. *Mol. Pharm.* **1991,** *41*, 419.

21. American Chemical Society Presidential Satellite Television Seminar: Molecular Modeling in the Discovery of New Drugs: American Chemical Society: Washington, DC, 1993.
22. Parson, W. W.; Chu, Z.-T.; Warshel, A. *Biochim. Biophys. Acta* **1990,** *1017*, 251.
23. Aqvist, J.; Leucke, H.; Quiocho, F. A.; Warshel, A. *Proc. Natl. Acad. Sci., USA* **1991,** *88*, 2026.
24. Mohan, V.; Davis, M. E.; McCammon, J. A.; Pettitt, B. M. *J. Phys. Chem.* **1992,** *96*, 6428.