THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Quantum Mechanics/Molecular Mechanics Restrained Electrostatic Potential Fitting

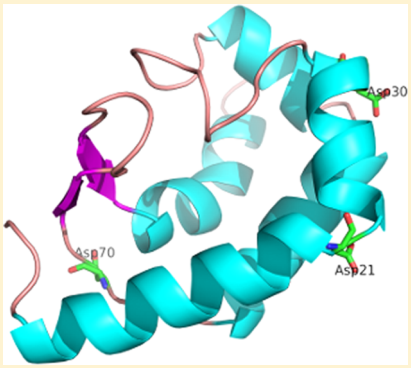
Steven K. Burger,*[,†] Jeremy Schofield,[†] and Paul W. Ayers[‡]

[†]Department of Chemistry, University of Toronto, 80 St. George Street, Toronto, Ontario M5S 3H6, Canada

[‡]Department of Chemistry and Chemical Biology, McMaster University, 1280 Main Street, West, Hamilton, Ontario L8S4L8, Canada

Supporting Information

ABSTRACT: We present a quantum mechanics/molecular mechanics (QM/MM) method to evaluate the partial charges of amino acid residues for use in MM potentials based on their protein environment. For each residue of interest, the nearby residues are included in the QM system while the rest of the protein is treated at the MM level of theory. After a short structural optimization, the partial charges of the central residue are fit to the electrostatic potential using the restrained electrostatic potential (RESP) method. The resulting charges and electrostatic potential account for the individual environment of the residue, although they lack the transferable nature of library partial charges. To evaluate the quality of the QM/MM RESP charges, thermodynamic integration is used to measure the pK_a shift of the aspartic acid residues in three different proteins, turkey egg lysozyme, beta-cryptogein, and Thioredoxin. Compared to the AMBER ff99SB library values, the QM/MM RESP charges show better agreement between the calculated and experimental pK_a values for almost all of the residues considered.



INTRODUCTION

Molecular dynamics has consistently proved to be a reliable tool for computing properties of biochemical systems.[1,2] Recent advances in computing power and software have extended the time scales of such simulations to the experimentally interesting millisecond range where rare events of protein binding, folding, and conformation changes can be regularly observed.[3,4] Often now, the limiting factor for computing properties is not the ability to do long sampling but the quality of the model and the underlying force field.

For determining molecular conformations and interaction energies, the most important elements of a force field are the dihedral, van der Waals (vdW), and electrostatic potential terms. Dihedrals tend to be transferable between systems, as do vdW interactions, which are usually derived from bulk properties. The electrostatics are more variable and cannot be reliably determined based on the type of atom.[5] This lack of transferability hampers the reliability of force fields based on fixed charges such as UFF,[6] AMBER,[7] and CHARMM.[8] While semiempirical methods such as density functional based tight binding (DFTB)[9] and the Gaussian electrostatic model (GEM)[10] avoid the problem by calculating the Coulombic integrals involved directly, they are too computationally expensive to be used with protein systems. Polarizable models such as NEMO,[11] AMBER FF02,[12] OPLS/PFF,[13] SIBFA,[14,15] and AMOEBA[16,17] seek to improve upon fixed charge models by introducing higher-order multipoles and a polarizable term. These models tend to give a much better representation of the electrostatics but are less efficient by at least an order of magnitude.

For fixed charge models, electrostatic potential (ESP) fitting[18,19] to an ab initio calculation is the most commonly used method for determining partial charges at atomic centers. ESP fitting suffers some problems though because the least-squares approach to solving for the partial charges tends to be ill-conditioned[20] and is known to be heavily dependent on the conformation and the basis set.[21−24] To deal with these problems, the restrained electrostatic potential (RESP) fitting was developed in the Kollman group.[24,25] RESP fitting introduces a hyperbolic penalty restraint on the charges, which easily allows many different conformations to be used in the fitting. It also allows for equivalent charges and for charged groups. The potential downside of the restraint is that it tends to reduce the magnitudes of the partial charges, which has been shown to lead to errors in binding energies.[26,27]

RESP fitting has been used to improve the partial charges for amino acid residues used in MM simulations by using a range of statistically probable conformations.[26,28] The charges are usually developed with small capped peptides and then tested on larger proteins to see if properties such as the relative population of 3_{10} to π-helices[29,30] are close to experimentally measured values, which is important for transferability. While partial charges derived in this way may be transferable, Thomas et. al[26,31,32] have shown with semiempirical calculations that the partial charges can vary dramatically in different local environments within a protein[33] and particularly so for buried

Received: September 25, 2013
Revised: October 31, 2013
Published: October 31, 2013

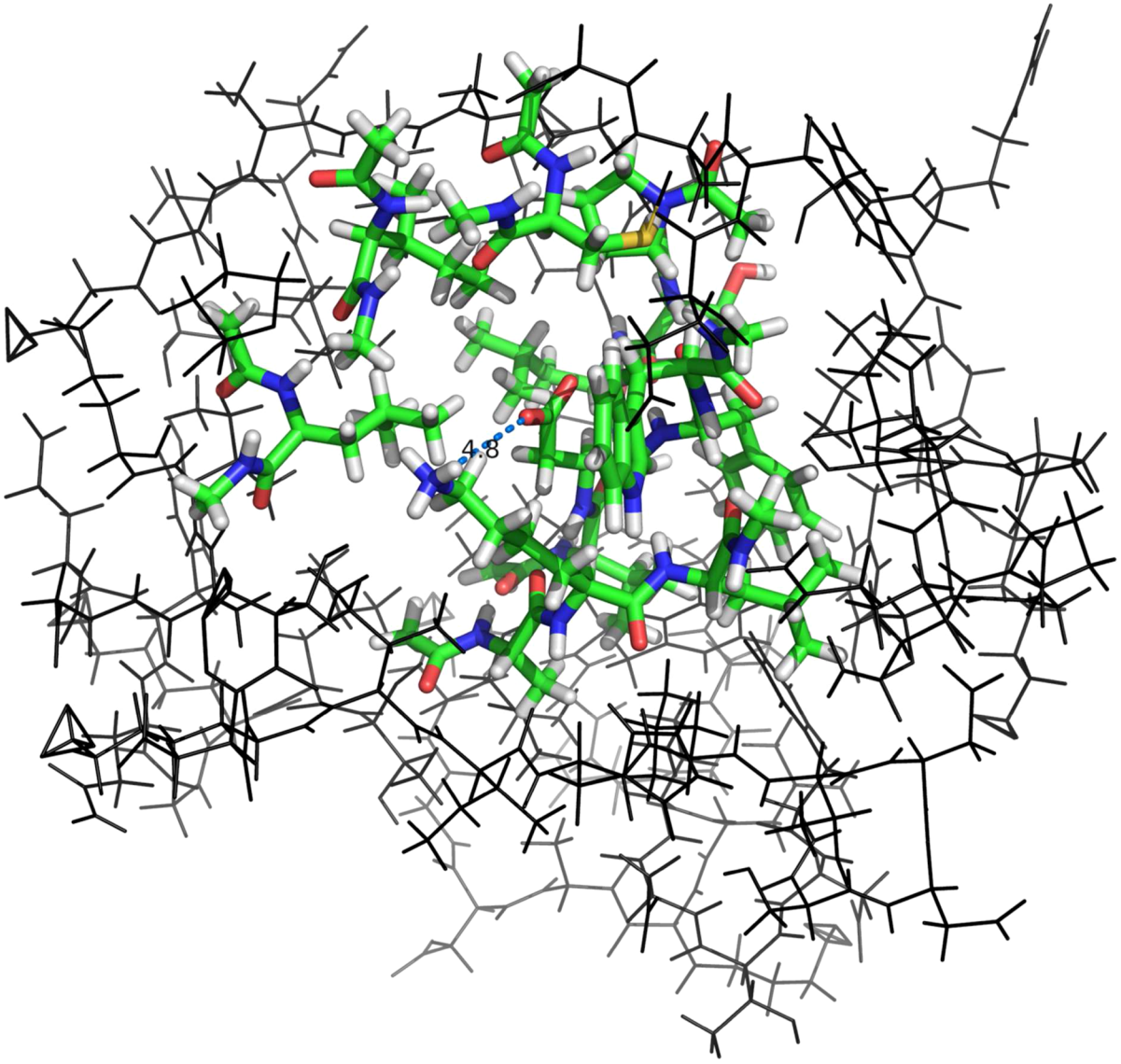


Figure 1. QM cluster of 291 atoms of all residues within 4 Å of Asp26 embedded in the MM charges (shown in black) that constitute the rest of the protein for Thioredoxin (PDB ID: 2TRX).

residues where the pK_a can experience a large shift.[34] In this work, we attempt to derive charges that are customized to the particular protein environment by fitting the ESP using a quantum mechanics/molecular mechanics (QM/MM) calculation. This is an approach that has previously been used in the context of molecular docking.[35] Here, we use the methodology to examine the pK_a shifts of aspartic acid residues.

The pK_a calculations were performed using thermodynamic integration (TI), combined with MD sampling, using an implicit generalized Born (GB) model to approximate the solvent effects. This approach has been used to successfully predict pK_a values of Thioredoxin and RNase A[36,37] and has also been used in the QM/MM context with free-energy perturbation theory.[38] We show that a significantly better correlation to experimental pK_a's can be achieved when using QM/MM-derived charges for the aspartic acid residues under consideration. In particular, we study Asp26 on the same Thioredoxin protein structure as in ref 34. We study Asp residues on the traditional proteins, turkey lysozyme protein[39] (PDB: 1LZ3), and beta-cryptogein[40] (PDB: 1BEO), both of which have been used in previous benchmarking pK_a calculations.[41−44]

METHODS

QM/MM models were constructed for each residue in order to obtain an ab initio ESP. From this, the partial charges were assigned by RESP. To determine if these charges produced a more accurate representation of the ESP, we evaluated the pK_a of residues using TI and compared these against the values obtained using the standard library charges.

RESP QM/MM Fitting. To refit the charges of each ionizable residue, the protein was divided into a QM and MM region. The QM region was defined as all residues within 4 Å of the ionizable residue (see Figure 1). Here, we have excluded explicit water molecules from the calculation because their inclusion would necessitate using a large number of different conformations in the fitting due to the highly mobile nature of water.[45] The residues on the boundary of the QM region were changed into either an acetyl group (ACE) or an *N*-methyl amide (NME). The MM atoms were included as a background

dx.doi.org/10.1021/jp409568h | J. Phys. Chem. B 2013, 117, 14960−14966

set of point charges. In order to neutralize the system, Cl^- or Na^+ ions were added using the AMBER software LEaP. In the rare case where an ion was placed in the QM system, it was subsequently removed.

To maintain consistency with the ff99SB force field, the QM calculations were done at the HF/6-31G(d) level using Gaussian09.[46] Once the QM and MM regions were defined, a QM geometry minimization was done in the presence of the MM field for a maximum of 40 cycles, holding the heavy atoms of the nonionizable residues fixed. The MM atoms were included as fixed charges using the values from the ff99SB force field. After the minimization, the ESP of the QM region was evaluated using the default of four layers per atom and one point per unit area using the Merz–Kollman[19] scheme to sample points. This ESP data were then used for RESP[24] fitting.

RESP fitting was done with the charges on each of the ACE and NME cap atoms restrained to their standard charges.[7] The charges on the surrounding residues were left unrestrained, and their charge values were not used in fitting. For the ionizable residues, three different types of restraints were attempted for the unprotonated state: (1) allowing all the partial charges to change on the residue while restraining the overall charge to be a unit value, (2) allowing the partial charges of the $-CH_2COO$ part of the aspartic acid to change while restraining the backbone atoms to their ff99SB library values, and (3) only allowing the charges of the carboxyl group to change. In each case, the oxygen atoms on the carboxylic group were restrained to have equivalent charges, and the overall the residue was restrained to −1. Comparisons of partial charge values were made against the standard library charges used by Simonson et al.,[36] as shown in Figure 2. The standard library charges were used for protonated ASP residue.

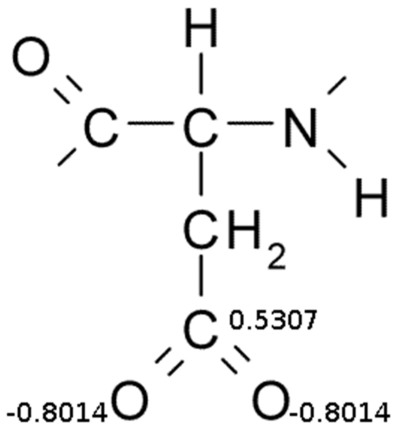


Figure 2. Partial charges varied by RESP fitting. The values used by Simonson et al.[36] are shown for the carboxylate group.

Thermodynamic Integration (TI). TI is a standard method[47] for determining free-energy changes and normally involves a small change in parameters between two different states. In the case of pK_a values, usually the partial charges on the residue are changed as the proton's force field parameters are removed. For Asp/Glu residues, there are no vdW terms associated with the proton, and the charges are the only terms that need to be changed. Unlike the vdW terms, which require special functions, such as the soft-core potential, for the coupling parameter λ in order to obtain a smooth integration, electrostatic charges tend to give a smooth transition in energy.

If the energy is linearly changed between the two states A (protonated) and B (unprotonated), $V(\lambda) = (1 - \lambda)V_A + \lambda V_B$ with a coupling parameter λ, the thermodynamic integral can be written as

$$\Delta G = \int_0^1 \left\langle \frac{\partial V(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda \quad (1)$$

The integration can be done with a Gaussian quadrature scheme. We choose to use the three-point integration, which requires evaluating $\langle \partial V(\lambda)/\partial\lambda \rangle$ at λ = {0.11270, 0.50000, 0.88729} with the weights {0.27777, 044444, 0.27777}. The next two higher-order Gaussian quadrature schemes were attempted, but the results did not change beyond the error bars introduced from the MD simulations. To justify the use of high-order Gaussian integration schemes, significantly more sampling would have to be employed.

From eq 1, the shift in the pK_a can be determined as

$$pK_a^{shift} = \frac{1}{2.303kT}\Delta\Delta G \quad (2)$$

where $\Delta\Delta G$ is the difference in free energy between the ΔG for the model system and ΔG for the protein. To this term, the standard value for the model compound, pK_a^{model}, must be added to get the total value, $pK_a = pK_a^{model} + pK_a^{shift}$. There is some debate as to which is the best value to use for the model compound.[48−50] For this study, we simply use the commonly accepted $pK_a^{model} = 4.0$ for Asp.[51]

To determine the error in the pK_a values, standard error propagation was performed with the error in the $\partial V/\partial\lambda$ values determined by block averaging

$$\sigma^2(A) = \frac{1}{M_b - 1}\sum_{b=1}^{M_b}(A_b^2 - A_b^2) \quad (3)$$

using a correlation time of 10 ps (M_b = 600), where M_b is the number of blocks, A_b is the average of the individual block, and $\langle A \rangle_b$ is the average over the full ensemble. For the TI used in this work, $A = \partial V/\partial\lambda$.

MD Protocol. Simulations were run using the AMBER 11[52] software suite with the GB implicit solvation model.[53] The SHAKE algorithm was used to restrain bonds containing hydrogen atoms. Nonbonded interaction distance cutoffs were set to values beyond the size of the system. The Langevin thermostat was used to maintain the temperature at 300 K after an initial warming phase.

Apart from the modified atomic charges, the same partial charges from the AMBER ff99SB[54] force field were used for the protonated and unprotonated states. An initial minimization and 2 ns of equilibration were performed on each protein with the backbone was held fixed using a 2.0 Å/mol restraint. This was followed by 2–4 ns of unrestrained dynamics until the RMSD of the backbone remained stable. Finally, a 6 ns production run for each lambda value was performed with a time step of 1 fs using the MPI version of Sander to obtain the $\partial V/\partial\lambda$ values necessary to evaluate the integral in eq 1.

RESULTS AND DISCUSSION

Thioredoxin. Thioredoxin has been studied at length with computational simulations using both explicit and implicit water models.[36] Generally, the explicit and implicit models show the same trends, and it has been shown that there is little difference between using the CHARMM or AMBER force field

dx.doi.org/10.1021/jp409568h | J. Phys. Chem. B 2013, 117, 14960−14966

parameters. The most interesting residue is Asp26, which has an unusually high pK_a of 7.5.[55]

In studies that focused purely on the static structure of the protein,[55,56] the principle determining factors for the pK_a were shown to be the positioning of Lys57 in relation to Asp26 and whether this particular group forms a salt bridge. For structures in which a salt bridge is present, the interaction between the groups is strong, and the resulting pK_a is generally calculated to be too low. At larger separations, there is little interaction, and the dielectric value must be changed to match the calculated and experimental pK_a values.

By using TI with MD to calculate the pK_a, one would hope that the problems associated with using static structures would not be present. However, at least with the implicit model, the barrier to convert between Lys57-Asp26 interacting structures and conformations in which Lys57 is freely mobile in the solution is too large to be a regular event on the nanosecond time scale simulations considered in this study. As a result, the calculated pK_a's are very much dependent on whether the TI is done "forwards" or "backwards", that is, if the equilibrated starting structure is started with the Asp26 protonated or not.

The QM/MM cluster used to determine the RESP charges is shown in Figure 1. The pK_a values were calculated starting from an equilibrated structure in which the Asp26-Lys57 formed a salt bridge and from the structure without it. Figure 3 shows that over the course of the simulation, the distance between the Lys57 ε-amino group and the Asp26 carboxyl group remains relatively constant for the smallest lambda value.

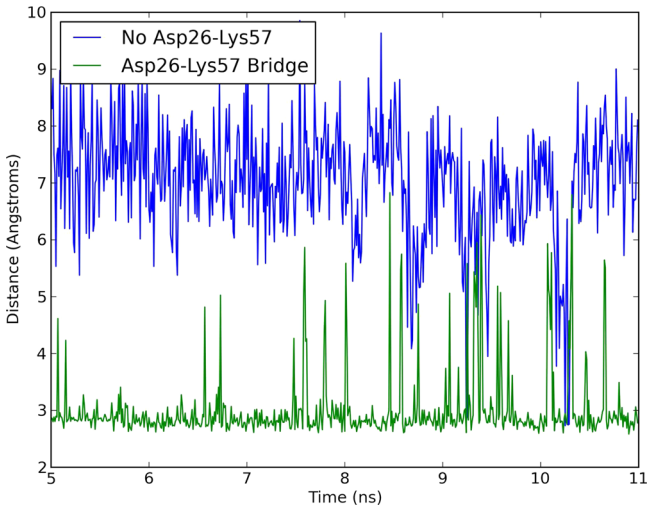


Figure 3. Distance between the closest oxygen on Asp26 and the nitrogen on Lys57 of Thioredoxin for λ = 0.1127. The Asp26-Lys57 salt bridge only forms for the lambda values chosen if the MD simulation is started with an equilibrated structure in which the salt bridge is already formed.

The calculated pK_a results are shown in Table 1. The results show how sensitive the pK_a is to the initial configuration. Although the charges are only derived once from a single crystal structure, the average value is in significantly better agreement with experimental values in both cases. The difference can be attributed to the change in electrostatics. The standard method uses the library charges, 0.5307 and −0.8014 (Figure 2) on the C and O atoms, respectively, while the values from the QM/MM RESP fitting are 0.3052 and −0.6884. These smaller partial charge values correct for the overestimate of the binding

Table 1. Comparison of Calculated pK_a's Starting with an Equilibrated Structure in Which a Salt Bridge Has Formed between Asp26 and Lys57 and with the Residues Being Widely Separated[a]

	standard method	fitting COO	expt. pK_a
No Asp26-Lys57 bridge	9.84 (0.38)	8.4 (1.0)	7.5
Asp26-Lys57 bridge	3.9 (1.1)	6.27 (0.94)	7.5

[a]Errors are given in parentheses.

affinity that gives too low of a pK_a in the case where the Asp26-Lys57 salt bridge is formed. When Lys57 is further away, the lower partial charges and resulting dipole give the residue a more nonpolar nature, which results in a calculated pK_a closer to a more neutral pH.

It should be noted that for Table 1 and the other results, the error bars for the TI calculations are often large because these are nanosecond length trajectories. Running significantly longer trajectories[57] would give a better bound on the values but at a large computational cost. As a result, we can only compare the mean values with the knowledge that if more sampling were done then the conclusions could be affected.

Beta-Cryptogein. Three Asp residues were studied from the enzyme beta-cryptogein (PDB ID: 1BEO) and are shown in Figure 4. All three residues are located on the surface of the

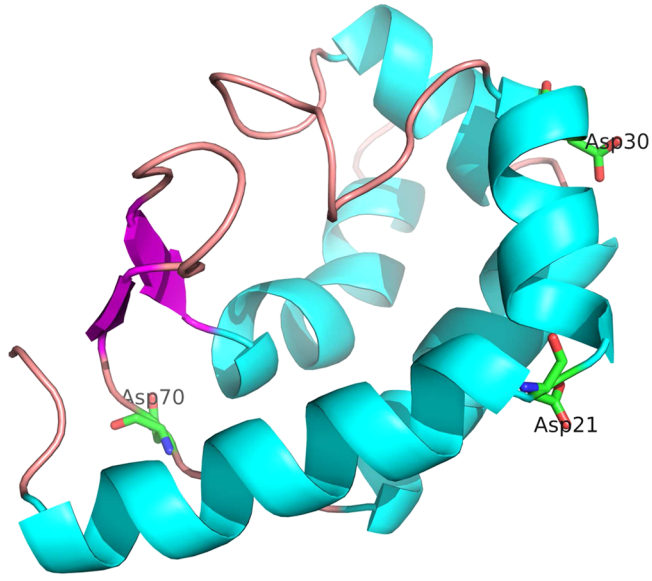


Figure 4. Beta-Cryptogein[64] (PDB ID: 1BEO) and the aspartic acid residues under consideration.

enzyme and are not expected to significantly change the conformation of the enzyme upon protonation. However, even though the Asp resides are located on the surface and have a large exposure area to the solvent, their pK_a's are shifted significantly lower than the free amino acids in solution. This is likely due to local interactions with polar and charged side chains in close proximity to these residues. During the MD runs, salt bridges were observed for the pairs Asp21-Lys62, Asp30-Gln26, and Asp72 with the N-terminus on Thr1. As shown in Table 2, using the AMBER ff99SB library charges consistently leads to overbinding between the charged residues, shifting the pK_a values more than is experimentally observed.

For the QM/MM RESP method, in addition to letting the charges on the carboxyl group change, the methylene group

dx.doi.org/10.1021/jp409568h | J. Phys. Chem. B 2013, 117, 14960−14966

Table 2. pK_a Values for Aspartic Acid Residues of 1BEO Using Different Methods to Assign Partial Charges to the Carboxylate Ion

	standard method	fitting CH_2COO	fitting COO	expt. pK_a
Asp21	1.31 (0.48)	3.65 (0.35)	3.34 (0.44)	2.5
Asp30	2.10 (0.61)	2.14 (0.80)	2.49 (0.42)	2.5
Asp72	0.75 (0.57)	−0.30 (0.48)	3.03 (0.61)	2.6

charges were also varied, with the group restrained having a charge that maintained the overall unit charge on the residue. While the agreement with experimental pK_a values was better than that using standard library charges, confining the changes in partial charges to the carboxyl group resulted in the best agreement with the experimental values. This may be explained by Table 3, which shows the partial charges on the carboxyl

Table 3. Partial Charge Values for the O and C Atoms Part of the Carboxylate Ion on the Protein 1BEO

	atom	standard method	fitting CH_2COO	fitting COO
Asp21	C	0.5307	0.3561	0.1619
	O	−0.8014	−0.7139	−0.6168
Asp30	C	0.5307	0.4359	0.2410
	O	−0.8014	−0.7538	−0.6563
Asp72	C	0.5307	0.3794	0.8983
	O	−0.8014	−0.7255	−0.9850

group atoms. When the methylene group is included, the magnitude of the charges is less than the standard library charges, but the group as a whole seems to be overpolarized by the environment. This same overpolarization problem was even more problematic when we allowed the charges on all of the atoms of the Asp residue to vary (results not shown). As a result, we did not pursue this approach any further and instead confined our studies to varying the atoms closest to the functional group.

Best agreement with the experimental pK_a values is obtained when only the charges on the carboxyl atoms are allowed to change. Only for Asp72 does this not work. For Asp72, QM/MM RESP fitting seems to result in a gross overpolarization of the electron density on the carboxyl group, which is likely due to the proximity of the N-terminus. However, the overbinding that occurs with the larger charges is compensated for in the model system, and the fortuitous cancelation of error results in a pK_a that is closer to the experimental value.

Turkey Egg Lysozyme. Similar to the beta-cryptogein case, the lysozyme enzyme has a number of Asp residues located on the surface on the protein that participate in local interactions that reduce the pK_a. While many authors[58−61] have had success in predicting the pK_a values for this protein, the lysozyme enzyme is a problematic case for the implicit solvent model. While the direction of each pK_a shift is correct, the magnitude of the shift is significantly overestimated. Using QM/MM RESP fit charges improves the estimates of three out of the four cases (Table 4) but is not able to account for the full overestimate of the shift. Like the other proteins, the overall dipole of the carboxyl group is reduced with the QM/MM RESP fit (partial charges are in the Supporting Information, Table S1). However, unlike many of the aspartic acid residues in other proteins studied, there are no strong salt bridges formed between the Asp18/48/66/87 residues and oppositely charged Lys and Arg residues. Instead, most of the polar

Table 4. Calculated pK_a Values for Aspartic Acid Residues of 1LZ3 Using the Standard ff99SB Library Charges Compared with the QM/MM RESP Fit Charges

	standard method	fitting COO	expt. pK_a
Asp18	3.58 (0.32)	2.36 (0.46)	2.7
Asp48	−2.0 (1.1)	−1.21 (0.96)	2.5
Asp66	−3.8 (1.1)	−2.68 (0.70)	2
Asp87	−0.48 (0.66)	−0.50 (0.26)	2.1

interactions are formed with backbone atoms and neutral but polar side-chain groups that may not be described well with the implicit solvent model.

CONCLUSIONS

Rather than using library charges for the residues in a protein, the ESP of each residue can be estimated by a QM/MM calculation followed by RESP fitting. The resulting partial charges give a more accurate representation of the ESP of the local environment of the residue in the protein. However, because the protein changes conformations during the course of the MD simulation, it is not clear how transferable such charges are over an ensemble of states. To see if QM/MM RESP charges derived from the crystal structure are better than the library charges derived from single amino residues, we focused on the pK_a values obtained using MD TI using the implicit solvent model. For the three proteins studied, we have shown that QM/MM RESP charges give significantly better agreement with experimental pK_a values compared to using the fixed library charges. The improvement seems to be due to reducing overly strong binding between Asp residues and nearby polar residues. While the fit charges helped improve the estimates of the pK_a with TI, determining pK_a values in general by use of the implicit water model may be limited by not having explicit waters present, as observed for the lysozyme protein. In future studies, we plan to look at how QM/MM RESP fit charges affect the estimates of properties such as ligand binding, protein folding, and the pK_a's determined at constant pH, which would clarify the utility of QM/MM RESP fitting of all of the residues in a protein before running a simulation rather than using a standard set of library charges. Also, we would like to evaluate the effectiveness of the Hirshfeld[62,63] approaches for QM/MM fitting rather using the MK/RESP approach.

ASSOCIATED CONTENT

Supporting Information

Charges used for each Asp of PDB ID: 1LZ3 are given. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: steven.burger@utoronto.ca.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC) and computing time on SHARCNET are great appreciated.

■ ABBREVIATIONS

ESP, electrostatic potential; MM, molecular mechanical; MD, molecular dynamics; FF, force field; QM, quantum mechanical

■ REFERENCES

(1) Kollman, P. Free-Energy Calculations — Applications to Chemical and Biochemical Phenomena. *Chem. Rev.* **1993**, *93*, 2395−2417.

(2) Karplus, M.; McCammon, J. A. Molecular Dynamics Simulations of Biomolecules. *Nat. Struct. Biol.* **2002**, *9*, 646−652.

(3) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; et al. Atomic-Level Characterization of the Structural Dynamics of Proteins. *Science* **2010**, *330*, 341−346.

(4) Dror, R. O.; Arlow, D. H.; Maragakis, P.; Mildorf, T. J.; Pan, A. C.; Xu, H.; Borhani, D. W.; Shaw, D. E. Activation Mechanism of the β2-Adrenergic Receptor. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 18684−18689.

(5) Verstraelen, T.; Pauwels, E.; De Proft, F.; Van Speybroeck, V.; Geerlings, P.; Waroquier, M. Assessment of Atomic Charge Models for Gas-Phase Computations on Polypeptides. *J. Chem. Theory Comput.* **2011**, *8*, 661−676.

(6) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024−10035.

(7) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.

(8) MacKerell, A. D.; Bashford, D.; Bellott; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.

(9) Seifert, G.; Joswig, J.-O. Density-Functional Tight Binding—An Approximate Density-Functional Theory Method. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 456−465.

(10) Cisneros, G. A.; Piquemal, J. P.; Darden, T. A. Generalization of the Gaussian Electrostatic Model: Extension to Arbitrary Angular Momentum, Distributed Multipoles, And Speedup with Reciprocal Space Methods. *J. Chem. Phys.* **2006**, *125*, 184101.

(11) Engkvist, O.; Astrand, P. O.; Karlstrom, G. Accurate Intermolecular Potentials Obtained from Molecular Wave Functions: Bridging the Gap between Quantum Chemistry and Molecular Simulations. *Chem. Rev.* **2000**, *100*, 4087−4108.

(12) Cieplak, P.; Caldwell, J.; Kollman, P. Molecular Mechanical Models for Organic and Biological Systems Going beyond the Atom Centered Two Body Additive Approximation: Aqueous Solution Free Energies of Methanol and N-Methyl Acetamide, Nucleic Acid Base, and Amide Hydrogen Bonding and Chloroform/Water Partition Coefficients of the Nucleic Acid Bases. *J. Comput. Chem.* **2001**, *22*, 1048−1057.

(13) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A.; Cao, Y. X.; Murphy, R. B.; Zhou, R.; Halgren, T. A. Development of a Polarizable Force Field for Proteins via Ab Initio Quantum Chemistry: First Generation Model and Gas Phase Tests. *J. Comput. Chem.* **2002**, *23*, 1515−1531.

(14) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J.-P. Anisotropic, Polarizable Molecular Mechanics Studies of Inter- and Intramolecular Interactions and Ligand−Macromolecule Complexes. A Bottom-Up Strategy. *J. Chem. Theory Comput.* **2007**, *3*, 1960−1986.

(15) Piquemal, J.-P.; Chevreau, H.; Gresh, N. Toward a Separate Reproduction of the Contributions to the Hartree−Fock and DFT Intermolecular Interaction Energies by Polarizable Molecular Mechanics with the SIBFA Potential. *J. Chem. Theory Comput.* **2007**, *3*, 824−837.

(16) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A., Jr.; et al. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114*, 2549−2564.

(17) Ren, P.; Wu, C.; Ponder, J. W. Polarizable Atomic Multipole-Based Molecular Mechanics for Organic Molecules. *J. Chem. Theory Comput.* **2011**, *7*, 3143−3161.

(18) Chirlian, L. E.; Francl, M. M. Atomic Charges Derived from Electrostatic Potentials — A Detailed Study. *J. Comput. Chem.* **1987**, *8*, 894−905.

(19) Besler, B. H.; Merz, K. M.; Kollman, P. A. Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.* **1990**, *11*, 431−439.

(20) Hu, H.; Lu, Z.; Yang, W. Fitting Molecular Electrostatic Potentials from Quantum Mechanical Calculations. *J. Chem. Theory Comput.* **2007**, *3*, 1004−1013.

(21) Reynolds, C. A.; Essex, J. W.; Richards, W. G. Errors in Free-Energy Perturbation Calculations Due to Neglecting the Conformational Variation of Atomic Charges. *Chem. Phys. Lett.* **1992**, *199*, 257−260.

(22) Chipot, C.; Angyan, J. G.; Ferenczy, G. G.; Scheraga, H. A. Transferable Net Atomic Charges from a Distributed Multipole Analysis for the Description of Electrostatic Properties — A Case-Study of Saturated-Hydrocarbons. *J. Phys. Chem.* **1993**, *97*, 6628−6636.

(23) Koch, U.; Stone, A. J. Conformational Dependence of the Molecular Charge Distribution and Its Influence on Intermolecular Interactions. *J. Chem. Soc., Faraday Trans.* **1996**, *92*, 1701−1708.

(24) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. Application of RESP Charges to Calculate Conformational Energies, Hydrogen-Bond Energies, and Free-Energies of Solvation. *J. Am. Chem. Soc.* **1993**, *115*, 9620−9631.

(25) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.

(26) Genheden, S.; Soderhjelm, P.; Ryde, U. Transferability of Conformational Dependent Charges from Protein Simulations. *Int. J. Quantum Chem.* **2012**, *112*, 1768−1785.

(27) Soderhjelm, P.; Ryde, U. Conformational Dependence of Charges in Protein Simulations. *J. Comput. Chem.* **2009**, *30*, 750−760.

(28) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; et al. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **2003**, *24*, 1999−2012.

(29) Feig, M.; MacKerell, A. D.; Brooks, C. L. Force Field Influence on the Observation of π-Helical Protein Structures in Molecular Dynamics Simulations. *J. Phys. Chem. B* **2003**, *107*, 2831−2836.

(30) Sung, S. S.; Wu, X. W. Molecular Dynamics Simulations of Synthetic Peptide Folding. *Proteins: Struct., Funct., Genet.* **1996**, *25*, 202−214.

(31) Thomas, A.; Milon, A.; Brasseur, R. Partial Atomic Charges of Amino Acids in Proteins. *Proteins* **2004**, *56*, 102−109.

(32) Thomas, A.; Deshayes, S.; Decaffmeyer, M.; Van Eyck, M. H.; Charloteaux, B.; Brasseur, R. Prediction of Peptide Structure: How Far Are We? *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 889−897.

(33) Avbelj, F.; Kocjan, D.; Baldwin, R. L. Protein Chemical Shifts Arising from α-Helices and β-Sheets Depend on Solvent Exposure. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17394−17397.

(34) Isom, D. G.; Castañeda, C. A.; Cannon, B. R.; García-Moreno, E., B. Large Shifts in pKa Values of Lysine Residues Buried Inside a Protein. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 5260−5265.

(35) Cho, A. E.; Guallar, V.; Berne, B. J.; Friesner, R. Importance of Accurate Charges in Molecular Docking: Quantum Mechanical/Molecular Mechanical (QM/MM) Approach. *J. Comput. Chem.* **2005**, *26*, 915−931.

(36) Simonson, T.; Carlsson, J.; Case, D. A. Proton Binding to Proteins: pK_a Calculations with Explicit and Implicit Solvent Models. *J. Am. Chem. Soc.* **2004**, *126*, 4167−4180.
(37) Archontis, G.; Simonson, T. Proton Binding to Proteins: A Free-Energy Component Analysis Using a Dielectric Continuum Model. *Biophys. J.* **2005**, *88*, 3888−3904.
(38) Riccardi, D.; Schaefer, P.; Cui, Q. pK_a Calculations in Solution and Proteins with QM/MM Free Energy Perturbation Simulations: A Quantitative Test of QM/MM Protocols. *J. Phys. Chem. B* **2005**, *109*, 17715−17733.
(39) Harata, K. X-ray Structure of Monoclinic Turkey Egg Lysozyme at 1.3 A Resolution. *Acta Crystallogr., Sect. D* **1993**, *49*, 497−504.
(40) Boissy, G.; de La Fortelle, E.; Kahn, R.; Huet, J. C.; Bricogne, G.; Pernollet, J. C.; Brunie, S. Crystal Structure of a Fungal Elicitor Secreted by Phytophthora cryptogea, a Member of a Novel Class of Plant Necrotic Proteins. *Structure* **1996**, *4*, 1429−1439.
(41) Burger, S. K.; Ayers, P. W. A Parameterized, Continuum Electrostatic Model for Predicting Protein pKa Values. *Proteins* **2011**, *79*, 2044−2052.
(42) Huang, R. B.; Du, Q. S.; Wang, C. H.; Liao, S. M.; Chou, K. C. A Fast and Accurate Method for Predicting pKa of Residues in Proteins. *Protein Eng., Des. Sel.* **2010**, *23*, 35−42.
(43) Davies, M. N.; Toseland, C. P.; Moss, D. S.; Flower, D. R. Benchmarking pK_a Prediction. *BMC Biochem.* **2006**, *7*, 18.
(44) Burger, S. K.; Ayers, P. W. Empirical Prediction of Protein pK_a Values with Residue Mutation. *J. Comput. Chem.* **2011**, *32*, 2140−2148.
(45) Papoian, G. A.; Ulander, J.; Wolynes, P. G. Role of Water Mediated Interactions in Protein−Protein Recognition Landscapes. *J. Am. Chem. Soc.* **2003**, *125*, 9170−9178.
(46) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A., et al. *Gaussian 09*, rvision B.01; Gaussian, Inc.: Wallingford, CT, 2009.
(47) Warshel, A.; Sussman, F.; King, G. Free Energy of Charges in Solvated Proteins: Microscopic Calculations Using a Reversible Charging Process. *Biochemistry* **1986**, *25*, 8368−8372.
(48) Gladysheva, T.; Liu, J. Y.; Rosen, B. P. His-8 Lowers the pK_a of the Essential Cys-12 Residue of the ArsC Arsenate Reductase of Plasmid R773. *J. Biol. Chem.* **1996**, *271*, 33256−33260.
(49) Keim, P.; Vigna, R. A.; Morrow, J. S.; Marshall, R. C.; Gurd, F. R. N. C-13 Nuclear Magnetic-Resonance of Pentapeptides of Glycine Containing Central Residues of Serine, Threonine, Aspartic and Glutamic Acids, Asparagine, and Glutamine. 53. *J. Biol. Chem.* **1973**, *248*, 7811−7818.
(50) Thurlkill, R. L.; Grimsley, G. R.; Scholtz, J. M.; Pace, C. N. pK Values of the Ionizable Groups of Proteins. *Protein Sci.* **2006**, *15*, 1214−1218.
(51) Edgcomb, S. P.; Murphy, K. P. Variability in the pKa of Histidine Side-Chains Correlates with Burial within Proteins. *Proteins: Struct., Funct., Genet.* **2002**, *49*, 1−6.
(52) Case, D.; Darden, T. A.; Cheatham, T. E.; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Crowley, M.; Walker, R.; Zhang, W.; et al. *Amber 11*; University of California: San Francisco, CA, 2010.
(53) Onufriev, A.; Bashford, D.; Case, D. A. Modification of the Generalized Born Model Suitable for Macromolecules. *J. Phys. Chem. B* **2000**, *104*, 3712−3720.
(54) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **2006**, *65*, 712−725.
(55) Langsetmo, K.; Fuchs, J. A.; Woodward, C.; Sharp, K. A. Linkage of Thioredoxin Stability to Titration of Ionizable Groups with Perturbed pKa. *Biochemistry* **1991**, *30*, 7609−7614.
(56) Dillet, V.; Dyson, H. J.; Bashford, D. Calculations of Electrostatic Interactions and pKas in the Active Site of Escherichia coli Thioredoxin. *Biochemistry* **1998**, *37*, 10298−10306.
(57) Shaw, D. E.; Dror, R. O.; Salmon, J. K.; Grossman, J. P.; Mackenzie, K. M.; Bank, J. A.; Young, C.; Deneroff, M. M.; Batson, B.; Bowers, K. J.; et al. Millisecond-Scale Molecular Dynamics Simulations on Anton. In *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis*; ACM: Portland, OR, 2009; pp 1−11.
(58) Tanford, C.; Roxby, R. Interpretation of Protein Titration Curves. Application to Lysozyme. *Biochemistry* **1972**, *11*, 2192−2198.
(59) Kuehner, D. E.; Engmann, J.; Fergg, F.; Wernick, M.; Blanch, H. W.; Prausnitz, J. M. Lysozyme Net Charge and Ion Binding in Concentrated Aqueous Electrolyte Solutions. *J. Phys. Chem. B* **1999**, *103*, 1368−1374.
(60) Bashford, D.; Gerwert, K. Electrostatic Calculations of The pKa Values of Ionizable Groups in Bacteriorhodopsin. *J. Mol. Biol.* **1992**, *224*, 473−486.
(61) Swails, J. M.; Roitberg, A. E. Enhancing Conformation and Protonation State Sampling of Hen Egg White Lysozyme Using pH Replica Exchange Molecular Dynamics. *J. Chem. Theory Comput.* **2012**, *8*, 4393−4404.
(62) Bultinck, P.; Van Alsenoy, C.; Ayers, P. W.; Carbó-Dorca, R. Critical Analysis and Extension of the Hirshfeld Atoms in Molecules. *J. Chem. Phys.* **2007**, *126*, 144111.
(63) Verstraelen, T.; Ayers, P. W.; Van Speybroeck, V.; Waroquier, M. Hirshfeld-E Partitioning: AIM Charges with an Improved Trade-off between Robustness and Accurate Electrostatics. *J. Chem. Theory Comput.* **2013**, *9*, 2221−2225.
(64) Krooshof, G. H.; Ridder, I. S.; Tepper, A. W. J. W.; Vos, G. J.; Rozeboom, H. J.; Kalk, K. H.; Dijkstra, B. W.; Janssen, D. B. Kinetic Analysis and X-ray Structure of Haloalkane Dehalogenase with a Modified Halide-Binding Site. *Biochemistry* **1998**, *37*, 15013−15023.