Accessing the Applicability of Polarized Protein-Specific Charge in Linear Interaction Energy Analysis

Xiangyu Jia,[a] Juan Zeng,[a] John Z. H. Zhang,[a,b] and Ye Mei*[a]

The reliability of the linear interaction energy (LIE) depends on the atomic charge model used to delineate the Coulomb interaction between the ligand and its environment. In this work, the polarized protein-specific charge (PPC) implementing a recently proposed fitting scheme has been examined in the LIE calculations of the binding affinities for avidin and β-secretase binding complexes. This charge fitting scheme, termed delta restrained electrostatic potential, bypasses the prevalent numerical difficulty of rank deficiency in electrostatic-potential-based charge fitting methods via a dual-step fitting strategy. A remarkable consistency between the predicted binding affinities and the experimental measurement has been observed. This work serves as a direct evidence of PPC's applicability in rational drug design. © 2014 Wiley Periodicals, Inc.

DOI: 10.1002/jcc.23547

Introduction

Computer simulation is now widely used in pharmaceutical industry, especially for the prediction of substrate-protein binding mode and strength.[1] Due to the large scale of biological molecules, studies of the structure-function relationship directly at quantum mechanical (QM) level are prohibitively demanding nowadays. Instead, molecular mechanics (MM) has been widely used in these studies and quite a lot of success has been seen in the past few decades. Nevertheless, MM methods often suffer from the force field limitations and model approximations.[2] In contemporary force fields for biological systems, such as AMBER, CHARMM, OPLS, the electrostatic polarization effect is not taken into account explicitly but is implemented in a mean-field fashion. However, protein is a typical heterogeneous system, which can be accommodated in various physiological environments. This limitation has impaired the reliability of classical molecular modeling. Linear scaling QM methods have opened a new avenue for subtle simulations of biological molecules.[3–5] In the last decade, we have developed a linear scaling QM method termed molecular fractionation with conjugate caps, with which the electronic structure of the whole protein can be represented as an assembly of the electronic structures of (residue-based) fragments. Heterogeneous perturbation to the electronic density can be well captured and the electrostatic polarization effect from the environment can also be accounted for in this method.[6] A new set of atomic charges can be acquired from the electronic-density distribution using the restrained electrostatic potential (RESP) fitting method.[7,8] As this new charge scheme can reflect the specific charge distribution of protein and the atomic charges are no longer residue specific but depend on the conformation and orientation of the residue in the protein, we name it the polarized protein-specific charge (PPC).[9] Its efficacy has been proved in some of our previous studies.[10–13] Because RESP fitting method is used in this charge scheme, the numerical difficulty of rank deficiency in solving the linear equations of least-square fit, which is prevalent in electrostatic potential-based charge schemes,[14] is also inherited. This difficulty can be alleviated by several techniques such as singular value decomposition[15] or adding restraints based on the electronic density[16] to prevent the atomic charges from getting abnormally large. Recently, we proposed an improved charge fitting method termed delta RESP (dRESP) that can bypass this numerical difficulty.[17] In dRESP fitting method, we do not attempt to fit the atomic charges directly. Instead, we only fit the perturbations to the atomic charges from an initial guess, for example, AMBER charge. The variation of the fitted atomic charges along with conformational change is significantly suppressed as compared to the RESP charge.

Another difficulty that impedes the applicability of molecular modeling is that the time scale of computer simulations that can be reached by modern computers is generally several orders shorter than the time scales of realistic bioprocesses. Prerequisite of phase space ergodicity is always assumed to be satisfied, but is not always met in practice. A variety of computational methods have been developed to-date to accelerate the simulation of long time scale biological processes. Prediction of ligand binding affinities to protein has strong significance in rational drug design. Various methods have been proposed so far. Some of them are rigorously formulated, whereas others trade accuracy for efficiency. The rigorous methods in this spectrum include thermodynamic integration (TI)[18] and free energy perturbation (FEP).[19] Although based

[a] X. Jia, J. Zeng, J. Z. H. Zhang, Y. Mei
State Key Laboratory of Precision Spectroscopy, Department of Physics, Institute of Theoretical and Computational Science, East China Normal University, Shanghai 200062, China
E-mail: ymei@phy.ecnu.edu.cn

[b] J. Z. H. Zhang
NYU-ECNU Center for Computational Chemistry at NYU Shanghai
Shanghai, 200062, China

Contract grant sponsor: National Natural Science Foundation of China; Contract grant number: 10974054, 20933002, 21173082; Contract grant sponsor: The Shanghai PuJiang Program; Contract grant number: 09PJ1404000

on rigorous formulation, the limitation of TI and FEP is obvious.[20,21] For each intermediate state, a long time molecular dynamics simulation must be carried out to avoid convergence failure, which may not be easy even for some simple systems.[22] At the other end of this spectrum are some empirical methods,[23,24] which can usually give decent result without demanding conformational sampling. In these methods, the binding affinity, or equivalently the scoring function, is expressed as a sum of scaled interacting potentials, such as hydrogen bonding, hydrophobic energy and so forth. Parameters used in scoring functions are usually obtained by fitting to experimental data for a training set. Transferability of these methods to systems outside the training set is not guaranteed. Therefore, they are usually used to predict the relative binding affinities of a series of analogs.

Among them, linear interaction energy (LIE) method[25,26] was first proposed by Åqvist and is still improving.[27] In addition to its initial aim to predict the substrate-protein binding energy, LIE has also been used in the study of point mutation induced binding loss in protein–protein complex.[28,29] In contrary to some stringent methods like FEP and TI, LIE needs only two windows corresponding to the bounded state and the free state. In the free state, the environment of the ligand is the solvent molecules soaking it. Whereas in the bounded state, the environment composes both the protein and the solvent molecules. In the original LIE method, the binding affinity is calculated as the differences of the scaled electrostatic and van der Waals (VDW) interaction between the ligand and its environment in the bounded and the free states, which can be written as

$$\Delta G_{\text{bind}} = \alpha(\langle V^{\text{ele}}_{\text{ligand}-\text{protein}}\rangle - \langle V^{\text{ele}}_{\text{ligand}-\text{solvent}}\rangle) + \beta(\langle V^{\text{vdw}}_{\text{ligand}-\text{protein}}\rangle - \langle V^{\text{vdw}}_{\text{ligand}-\text{solvent}}\rangle).$$

Initially, α is fixed to 0.5, which is derived from the linear response assumption. β is the only parameter and is usually system dependent.[30] Even though the linear response assumption[31] has been justified by many small molecules in water,[32] larger systems like ligand–protein compound could hardly obey the linear response rule because of the evident gap when a macroscopic theory was applied into a microscopic world.[27] Therefore, finding a set of transferability parameters independent of receptor binding site for LIE method is very difficult, if not impossible, even for enzymes of the same class.[33] It is more practical to calibrate both α and β for each system. Besides, Jones-Hertzog and Jorgenson[34] found it reasonable to append a third term, which was related to the solvent accessible surface. Some other variations of LIE method have also been proposed.[33,35–38]

As a rule of thumb, the interaction potential in LIE method requires careful calibration. Many efforts have been dedicated to finding an appropriate force field to improve the quality of LIE predictor. For the fully automated LIE-based virtual screening method, Wallin et al.[39] studied variant charge schemes and found that the semiempirically derived CM1A charge was a fast and reliable alternative for the ligands in working with the OPLS-AA force field for the protein. However, Almlöf et al.[40] compared various force fields and showed that the coefficients of LIE method were independent of the force fields used. Åqvist[41] suggested charged ligands should be treated carefully otherwise the truncated long range solute–solvent electrostatic interaction and the overpolarization effect from solvent–solvent interaction cutoff may introduce notable errors. Some studies[33,35,36,42,43] used implicit water model to address the problem of charged ligands. Zhou et al.[36] utilized generalized Born model to study HIV-1 reverse transcriptase binding set and obtained a correlation coefficient of 0.774 and root mean square (RMS) errors of 1.07 kcal/mol. Huang and Caflisch[33] used the numerical solution of the Poisson equation by the finite-difference technique and found it suitable for ranking large libraries of structurally diverse compounds. As the Coulomb interaction is the dominant term in LIE, atomic charges should be elaborately determined. In this work, we used dRESP charge into the LIE calculation of two protein–ligand binding systems. The quality of this LIE predictor was measured by the correlation coefficient, the root mean square deviation (RMSD) of the predicted binding affinities from the experimental measurements and the leave-one-out cross-validation (LOOCV). By comparing with nonpolarized AMBER charge, we inferred that implementation of polarization effect into LIE method can improve its reliability for the electrostatic interaction-dominated systems.

Avidin–biotin[44] complex (Fig. 1a) is one of the most remarkable protein–ligand binding systems in nature, with a strong binding affinity of 20 kcal/mol. Elucidating the binding mechanism of this classical system has motivated a variety of experimental and theoretical investigations.[45,46] The binding site of this tetrameric enzyme avidin is located at the center of each peptide chain with a cleft exposed to solvent. Biotin and its analogs interact with avidin through several hydrogen bonds and VDW contacts. Structures of biotin and its analogs studied in this work, which are denoted as a1 through a14 hereafter, are listed in Table 1. Their binding affinities to avidin differ by 15 kcal/mol. Wang et al.[47] studied the same system using two-term LIE utilizing nonpolarized charge model. In their study, biotin and the potentially charged analogies a2, a3, a6, and a7 were kept neutral and two different means were adopted to deal with the titration state of avidin. The equilibrium time was 50 ps followed by a 150-ps production run. For a4, a5, a12, and a13, large deviations from experimental measurements were observed. In this study, a different strategy was adopted, which will be formulated in the next section. Our previous experience with (strept)avidin–biotin systems suggests that PPC can provide more reliable structure ensemble and thermodynamics than AMBER charge does.[11,12,48]

β-secretase (BACE)[49] (Fig. 1b) mainly cleaves amyloid precursor protein[50] and it has attracted much attention due to its close relationship with Alzheimer's disease.[51–53] Ligands (listed in Table 2) bound to BACE are large in molecular size with a number of rotatable bonds. Neglecting the differences in entropy penalties may lead to significant deviations in the predicted binding affinities. Using a simple way in counting the entropy change,[54,55] the reliability of LIE is seen to be

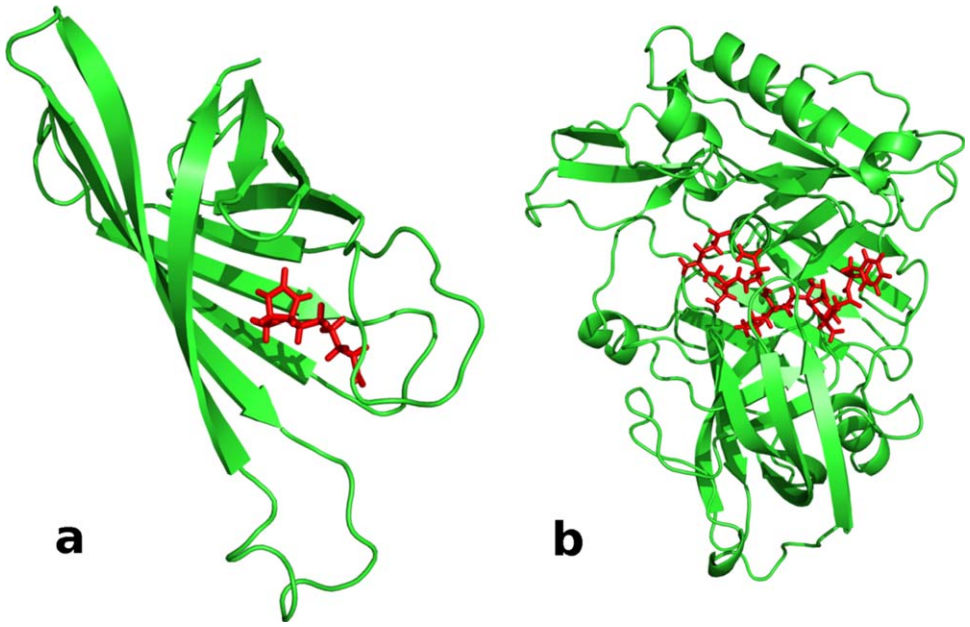


Figure 1. Structures of ligated a) avidin and b) β-secretase. Proteins and ligands are colored in green and red, respectively.

improved. Besides, some of them are heavily charged, which also poses a challenge to LIE calculation.

The remainder of this manuscript is organized as follows. The general description of PPC scheme, the setups of molecular dynamics simulations and the methods for the calculation of the energy components are explained in Method Section. Results are presented and discussed in Results and Discussion Section. We summarize this study in Conclusions Section.

Method

Structure preparation

The crystal structure of biotin–avidin complex was downloaded from the protein data bank (entry: 1AVD). The tetrameric structure was built by imposing the symmetry operation on the dimer. All other complexes with biotin analogs were generated based on this structure. Hydrogen atoms were added using the LEaP module in AmberTools package. The coordinates of missing residues were predicted by Modloop server.[56] H++ software[57] was utilized to determine the protonation states of titratable residues. Each chain of the protein had a net charge of $+5e$. The structures of biotin and its analogs were optimized at B3LYP/6-31G* level. Ligands were initially assigned RESP charge.[7,58] Other force field parameters were taken from the general AMBER force field (GAFF).[59] Processing of BACE was quite similar to that of avidin, except that the side chain of Arg32 in BACE was manually protonated.[42,52,60] Ligands were built by manually modifying OM99-2. The AM1-bcc (Austin Model 1 with bond charge correction)[61,62] charge was initially assigned to the ligands. Each complex was soaked in a periodic truncated octahedral box of TIP3P water with 13 Å buffer. Counterions were added to neutralize the whole system.

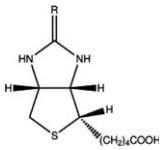
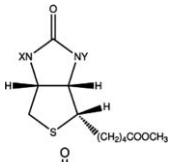
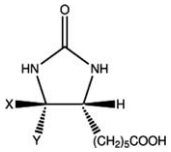
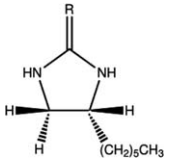
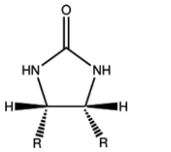
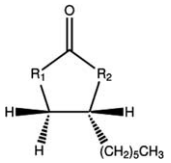
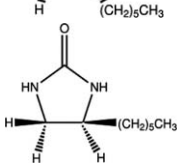
Polarized PPC

The polarized PPC using dRESP fitting scheme was calculated for each protein–ligand complex. With an initial guess of the atomic charges (AMBER 94, RESP, and bcc charge) for the proteins, biotin analogs, and BACE inhibitors, the induced charges on the solute–solvent interface were calculated by solving the Poisson–Boltzmann (PB) equations utilizing delphi program.[63] The internal dielectric constant was set to unity and that of the solvent to 80. Grid density was set to 3.5 grids/Å. Each complex was then divided into the ligand and residue-based fragments. A pair of conjugate caps was added to each fragment to saturate covalent bonds and also mimic the immediate chemical environment. QM calculation was performed for each fragment in the electric field generated by other residues and the induced surface charges. RESP program in AmberTools suite package with in-house modifications was applied to fit the dRESP charge for the protein–ligand complex. The new atomic charges were used for the next cycle of PB calculation. Atomic and induced charges were calculated iteratively until convergence was reached. The convergence was determined by the variation of the dipole moments of the protein and the RMSD of surface charges. Solvation and many-body effects like hydrogen bond cooperativity had been naturally included via iterations. All the QM calculations were carried out at B3LYP/6-31G* level utilizing Gaussian 09 package.[64]

Molecule dynamics simulation

In addition to PPC, mean-field AMBER03 force field, and GAFF were also used in this study for comparison. Altogether 52 simulations were conducted, 28 of them for avidin–ligand complexes, and the other 24 for BACE-ligand systems. All molecular dynamics simulations were carried out by Amber 11

Table 1. Avidin inhibitors studied in this work.

	A1	R = O
	A2	R = S
	A3	R = NH
	A4	X = CH_3O, Y = H
	A5	X = H, Y = CH_3O
	A6	X = CH_3, Y = H
	A7	X = H, Y = CH_3
	A8	R = O
	A9	R = NH
	A10	R = H
	A11	R = CH_3
	A12	R_1 = O, R_2 = NH
	A13	R_1 = NH, R_2 = O
	A14	

package.[65] In the first minimization stage, the protein was restrained and all other molecules were relaxed using steep descent method followed by conjugate gradient minimization. In the second stage, the whole system was thoroughly optimized. The fully relaxed structures were used for PPC fitting. A six-stage heating-up in 1.2 ns altogether was applied to avoid blowing up of the systems. The time step was 1 fs during heating. A 2-ns simulation in canonical ensemble was carried out to equilibrate the whole system to the target temperature, which is followed by an 8-ns production simulation in NPT ensemble with a 2-fs time step. Temperature was regulated by Langevin dynamics[66] with a collision frequency of 1.0 ps^{-1}. The pressure was regulated using Berendsen's barostat.[67] SHAKE algorithm was applied to constrain the covalent bonds involving hydrogen atoms. Particle mesh Ewald[68] with a cut-off of 9.0 Å in real space was utilized to calculate long range Coulomb interaction. VDW interaction was truncated at 9.0 Å. Snapshots from the last 6-ns trajectory were evenly extracted for further analysis.

LIE combined with MM-PBSA

Similar to that of Zhou et al.,[36] a three-term LIE function was used in this study, which can be expressed as

$$\Delta G = \alpha(\langle V_{l-c}^{ele}\rangle - \langle V_{l-f}^{ele}\rangle) + \beta(\langle V_{l-c}^{vdw}\rangle - \langle V_{l-f}^{vdw}\rangle) + \gamma,$$

where l-c stands for ligand–environment interaction in the bound state and l-f for that in the free state. γ is a ligand-independent constant for absolute binding free energy.[28,40] Energy terms in the LIE formula were calculated by MM with PB and surface area (MM/PBSA) analysis.[69] The polar solvation energy can be expressed as

$$G_{sol} = \frac{1}{2}\sum_i q_i \varphi_i,$$

where q_i is the atomic charge and φ_i is the electrostatic potential located the ith atom. For the free state, the solute–solvent solvation energy can be easily obtained by solving the PB equation, and the electrostatic term for ligand–solvent interaction is twice the polar solvation energy. But for the bound state, decomposition analysis was applied to obtain the ligand's interaction with its surroundings.[70] The ligand–receptor electrostatic energy can be calculated by simple Coulomb equation. Therefore, V_{l-c}^{ele} and V_{l-f}^{ele} can be expressed as

$$V_{l-c}^{ele} = 2G_{ele}^{demp} + V_{l-r}^{coul}$$

and

$$V_{l-f}^{ele} = 2G_{ele},$$

respectively. When solving the PB equation, the internal dielectric constant was set to 1 for PPC and 2 for AMBER03 charge. This is due that the polarization effect is included in PPC but not in AMBER03 charge.[71–73] In the free state, the ligand–solvent VDW interaction is approximated to the nonpolar solvation free energy of the ligand (G_{npol}^{lig}). While in the bound state, the VDW interaction is calculated as the sum of the ligand–receptor VDW interaction (V_{l-r}^{vdw}) and the difference in the nonpolar solvation free energies for the complex (G_{npol}^{com}) and the receptor (G_{npol}^{rec}). For ligand–avidin complex, four chains were treated individually in MM/PBSA calculation and were averaged to give the final results.

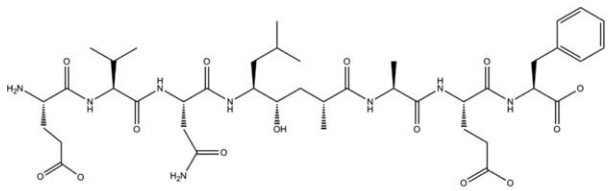
Results and Discussion

Ligand–avidin complex

The RMSD of all the binding set during MD simulations utilizing PPC are shown in Supporting Information Figure S1. After heating and a subsequent 2-ns simulation in the canonical (NVT) ensemble, the structures of all the complexes were stable, regardless of whether the ligand was charged or not. The variations of the VDW and the electrostatic interaction energies over time were shown in Supporting Information Figure S2. Both of the interaction energies underwent only small fluctuations and the convergence of their ensemble averages was

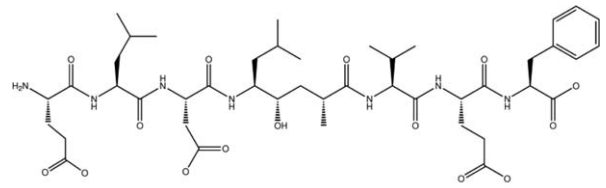
Table 2. *β*-secretase inhibitors studied in this work.

b1

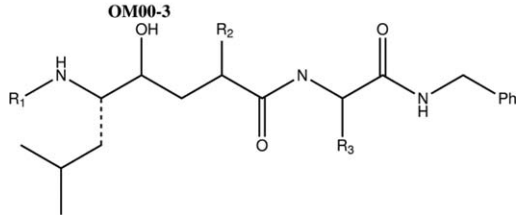


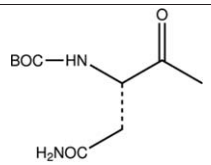
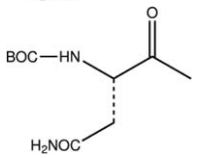
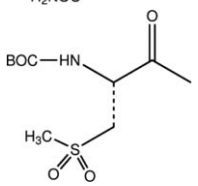
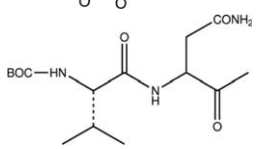
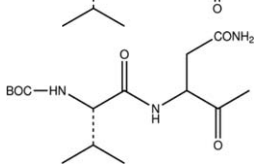
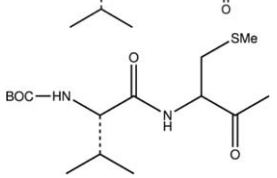
b2

OM99-2

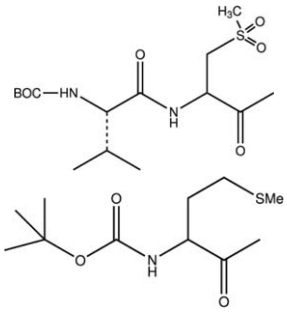
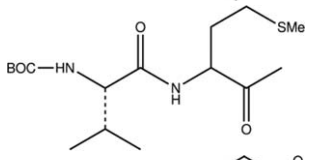
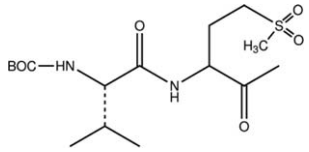


OM00-3



Ligand	R_1	R_2	R_3
b3		Me	Me
b4		Me	$CHMe_2$
b5		Me	$CHMe_2$
b6		Me	Me
b7		Me	$CHMe_2$
b8		Me	$CHMe_2$
b9		Me	$CHMe_2$

(Continued)

b10		Me	$CHMe_2$
b11		Me	$CHMe_2$
b12		Me	$CHMe_2$

guaranteed. For neutral ligands, the fluctuations of VDW and electrostatic energy were lower than 5 and 10 kcal/mol, respectively. But for charged ligands, although the variations of VDW item were comparable with those of neutral ligands, the fluctuations of electrostatic energy were about 5 kcal/mol larger. If a short MD trajectory was chosen, like several picoseconds, large uncertainty may appear due to the slow convergence of the averaged electrostatic energy. In this study, the calculations of the energy components were based on the last 6 ns of each trajectory.

The LIE components for the ligand–avidin binding set are listed in Table 3. From these data, parameters α, β, and γ in LIE equations are fitted in a least-square manner, which are 0.421, 0.274, and 4.362 kcal/mol, respectively. The coefficient before the electrostatic term is a little bit smaller than 0.5, indicating that the linear response is not strictly held. A direct visual comparison of the computed and experimental binding free energies is plotted in Figure 2 (left panel). Most of the ligands fall in the vicinity of the diagonal. However, there are still three outliers, that is, a2, a7, and a12. The predicted binding affinities for these three ligands differ from the experimental measurements by about 2.0, −2.4, and −2.9 kcal/mol. The correlation coefficient is 0.957 and the RMSD between the calculated and experimental binding free energies of the whole set is 1.37 kcal/mol. In view of the wide range covered by the binding affinities of this training set, this result is quite encouraging. To investigate if the ranks of the experimental binding energies were accurately predicted, Spearman's rank correlation coefficient

$$\rho = 1 - \frac{6\sum_i d_i^2}{n(n^2-1)}$$

was also calculated, where n is the number of the ligands used in the training set and d_i is the rank difference between the calculated set and the experimental set for the ith ligand–receptor complex. ρ equals to 0.94 for this binding set, indicating a high correspondence between the calculated and experimental ranks. Large rank differences are seen for the outliers. For a2 and a12, the rank differences are two. For a12, which has the largest binding energy deviation, the rank difference is three. LOOCV analysis[74] was also applied. In this method, ligands were taken out from the binding set one at a time and the theoretical binding energy of the excluded ligand was calculated by parameters fitted from the ligands remained in the training set. All the calculated results are compared with the experimental measurements to validate the robustness of the model used. The results are listed in Table 4. The RMSD was 1.65 kcal/mol and the correlation coefficient was 0.940. The reliability of this charge model in LIE for avidin–biotin analogs is validated.

Due that the crystallographic structures of these avidin–biotin analog complexes are not all available, computation of the interaction energy is based on the MD simulated structure ensemble starting from the experimental structure of avidin–biotin complex. The protein may undergo large scale hydrogen bond rearrangement, if a certain hydrogen bond candidate is eliminated. This conformational change cannot be easily captured by a short simulation. The oxygen atom in the ureido group is changed to a sulfur atom in a2, which is not a good hydrogen bond acceptor. One of the N atom in the ureido group is changed to O atom in a12, which eliminates the hydrogen bond with D128 and causes strong repulsion if the receptor remains in the structure as in the avidin–biotin complex. We speculate that the receptor has large structural difference, when binding with a2, a12, and a1. Although there is only small difference between a6 and a7 in the direction of the methyl group, their binding affinities to avidin differ by 2.5 kcal/mol. This large difference can hardly be interpreted

Table 3. Computed interaction energy components and the computed and experimental binding free energies for the avidin–biotin set.

Ligand	$\langle E_{elec} \rangle$	$\langle E_{vdW} \rangle$	ΔG_{bind}^{calc}	ΔG_{bind}^{exp}	$\lvert \Delta G_{bind}^{calc} - \Delta G_{bind}^{exp} \rvert$
a1	−38.49	−27.97	−19.51	−20.40	0.89
a2	−27.24	−28.33	−14.87	−16.90	2.03
a3	−26.62	−27.81	−14.46	−14.30	0.16
a4	−4.85	−45.37	−10.11	−8.80	1.31
a5	−9.10	−41.83	−10.93	−12.20	1.27
a6	−36.51	−21.89	−17.01	−16.50	0.51
a7	−33.24	−24.83	−16.44	−14.00	2.44
a8	−15.38	−28.18	−9.83	−11.10	1.27
a9	−10.53	−27.01	−7.47	−7.40	0.07
a10	−13.95	−10.43	−4.37	−4.50	0.13
a11	−14.84	−15.65	−6.18	−6.40	0.22
a12	−10.04	−29.47	−7.94	−5.00	2.94
a13	−7.84	−28.90	−6.86	−7.40	0.54
a14	−9.78	−30.17	−8.02	−9.10	1.08

All data are obtained using PPC charge scheme.

without the existence of large conformational change of the receptor. This study gave very close results for a6 and a7, due that a similar binding pattern was utilized for both of the ligands.

The variations of the electrostatic and VDW energies along the simulations utilizing AMBER03 force field are showed in Supporting Information Figure S3. Converged ensemble averages have also been reached. LIE components are listed in Table 5. Due that a larger internal dielectric constant (two for AMBER03 charge versus one for PPC; a dielectric constant of one for AMBER03 charge gives much worse results with large random errors.) was utilized when solving the PB equation, the magnitude of the electrostatic energies are much larger than those from PPC. The fitted parameters α, β, and γ, are 0.11, 0.194, and −0.419 kcal/mol, respectively. The correlation between the predicted and experimental binding affinities is plotted in the right panel of Figure 2. The correlation coefficient is 0.926 and the RMSD is 1.77 kcal/mol, which are slightly worse than those from PPC. Our previous experience with AMBER charge elucidated that the trajectory under AMBER charge was not stable enough to well maintain the binding pattern of ligand molecule.[12] Although systematic error can be eliminated in LIE, it still adds random errors to energy components.

Ligand-BACE system

Ligand-BACE system is more challenging than ligand–avidin set, because some of the ligands are very flexible and heavily charged and biological water molecules may play an essential role. The electrostatic and the VDW energies over the simulation time in BACE-inhibitor systems are very stable as showed in Supporting Information Figure S4. The LIE components are listed in Table 6. The fitted α, β, and γ, parameters are −0.019, 0.306, and 13.67, respectively. Such a small α indicates weak contribution of the electrostatic interaction to the binding affinity. Therefore, VDW term is the dominant driving force for the binding process. This observation can also be verified by the correlation between experimental binding affinities and energy components shown in Figures 3a and 3b. The correlation between the binding affinities and the VDW interactions is 0.934, while that between the binding affinities and the electrostatic contributions is only −0.149. The correlations between the experimental and computed binding free

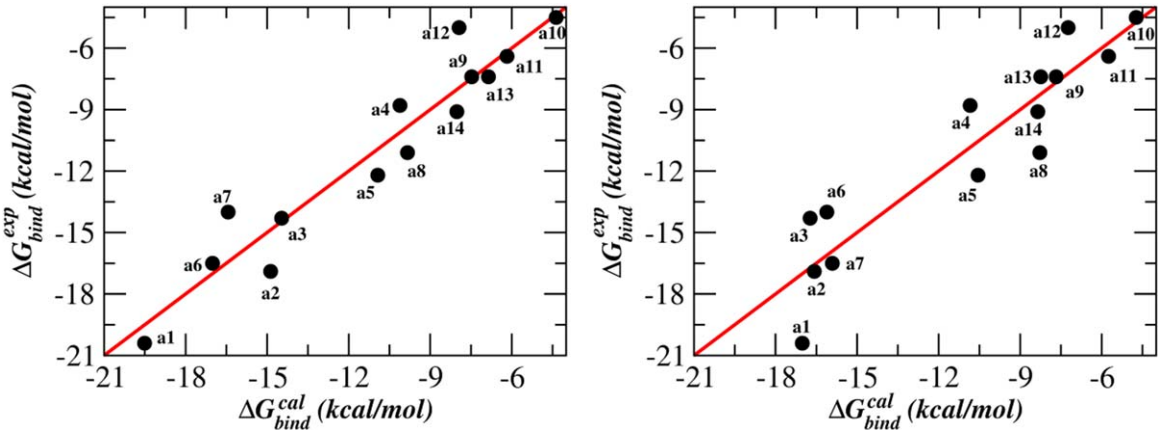


Figure 2. Binding energies calculated by LIE method versus the experimental measurement for avidin-binding systems utilizing (left) PPC and (right) AMBER03 charge.

Table 4. Leave-one-out analysis of the avidin–biotin set under PPC.

Excluded ligand	α	β	γ	ΔG_{bind}^{calc}	ΔG_{bind}^{exp}
a1	0.404	0.266	3.917	−19.07	−20.40
a2	0.407	0.267	4.075	−14.58	−16.90
a3	0.422	0.274	4.379	−14.47	−14.30
a4	0.412	0.309	5.012	−7.01	−8.80
a5	0.425	0.251	3.921	−10.44	−12.20
a6	0.428	0.273	4.414	−17.19	−16.50
a7	0.448	0.277	4.714	−17.05	−14.00
a8	0.424	0.274	4.519	−9.72	−11.10
a9	0.421	0.274	4.341	−7.49	−7.40
a10	0.423	0.279	4.551	−4.26	−4.50
a11	0.423	0.278	4.537	−6.09	−6.40
a12	0.405	0.273	3.787	−8.32	−5.00
a13	0.425	0.275	4.511	−6.77	−7.40
a14	0.427	0.274	4.549	−7.89	−9.10

Table 6. Computed interaction energy components and the computed and experimental binding free energies for the BACE–ligand set.

Ligand	$\langle E_{elec} \rangle$	$\langle E_{vdW} \rangle$	ΔG_{bind}^{calc}	ΔG_{bind}^{exp}	$\lvert \Delta G_{bind}^{calc} - \Delta G_{bind}^{exp} \rvert$
b1	−18.50	−85.58	−12.15	−12.06	0.09
b2	−23.29	−84.67	−11.79	−13.05	1.26
b3	−27.64	−66.86	−6.23	−6.38	0.15
b4	−29.92	−71.46	−7.69	−7.55	0.14
b5	−26.29	−76.53	−9.23	−8.16	1.07
b6	−28.91	−76.16	−9.07	−9.90	0.83
b7	−31.46	−82.96	−11.11	−11.30	0.19
b8	−41.43	−82.30	−10.73	−10.02	0.71
b9	−27.93	−85.94	−12.09	−11.02	1.07
b10	−43.12	−73.26	−7.92	−7.91	0.01
b11	−43.87	−83.36	−11.01	−11.81	0.80
b12	−36.11	−83.67	−11.24	−11.11	0.13

All data are obtained using PPC charge scheme.

energies for all the BACE ligands are plotted in Figure 4 (left panel). The correlation coefficient under PPC is 0.937 and the RMSD is 0.695 kcal/mol. This RMSD is smaller than that in the avidin system. However, it should be kept in mind that the binding free energies of BACE set differ by only 7 kcal/mol, but those of avidin set differ by 15 kcal/mol. Nevertheless, this RMSD is still smaller than a previous study[42] and the correlation coefficient is no worse than a specific LIE model built for BACE.[53] The spearman's rank correlation coefficient is 0.88. Compared with another LIE study of systems in which the binding free energies differ by similar magnitude,[38] this result is acceptable. LOOCV analysis was also applied. The obtained parameters and calculated binding energies are listed in Table 7. The RMSD between theoretical and experimental results is 0.890 kcal/mol and the correlation coefficient is 0.899.

On the whole, the properties like correlation coefficient and spearman's rank correlation of BACE complex are lower than those of ligand–avidin complex. It may due that ligands binding to β-secretase are larger and more flexible, which makes the computation of atomic charges more difficult and the description of electrostatic interaction between molecules less accurate. Besides, entropy is related to the flexibility of the system but the entropy of receptor or ligand itself was not explicitly included in LIE equation. This property is an essential part of the binding free energy. The more flexibility the system has, the more difficulty there will be in describing binding energy using LIE method. For this BACE binding set, we use the number of rotatable bonds to approximately represent the entropy effect. This scheme has been used in some previous studies.[54,55] The number of the rotatable bonds is 29 for b1 but for b3 there are only 17 rotatable bonds. This large difference reminds us that explicitly treatment of the entropy effect in LIE method is necessary. The entropy-included LIE equation can be expressed as

$$\Delta G = \alpha(\langle V_{l-c}^{ele} \rangle - \langle V_{l-f}^{ele} \rangle) + \beta(\langle V_{l-c}^{vdw} \rangle - \langle V_{l-f}^{vdw} \rangle) + \gamma + \delta \cdot \mathrm{Num}_{rot\_bond},$$

where Num_{rot_bond} is the number of rotatable bonds for each ligand. After least-square fitting, the fitted α, β, γ, and δ parameters are 0.019, 0.218, 12.09, and −0.195, respectively. In Table 8, the number of rotatable bonds and calculated binding energies are listed. Comparing with the experimental data, the RMSD and Spearman's rank correlation coefficient are 0.496 and 0.93 kcal/mol, respectively. Therefore, by explicitly considering the entropy effect, an obvious improvement is achieved. A direct visual comparison between the theoretical and target data is plotted in Figure 5 (left panel) and a high correlation coefficient of 0.969 is obtained.

LIE components obtained using AMBER03 force field were listed in Table 9 and their variations over simulation time were showed in Supporting Information Figure S5. Using these converged energy components, the fitted α, β, and γ, were −0.002, 0.317, and 17.71, respectively. This is similar to the PPC scheme that VDW interaction is the major driving factor for ligand-BACE system and electrostatic interaction has almost no contribution for the binding. Direct visual comparison between the calculated and experimental data is given in Figure 4 (right panel). The correlation coefficient is 0.957 and the RMSD is 0.574 kcal/mol, indicating that these two sets matched very well and AMBER03 force field is capable of

Table 5. Computed interaction energy components and the computed and experimental binding free energies for the avidin–biotin set.

Ligand	$\langle E_{elec} \rangle$	$\langle E_{vdW} \rangle$	ΔG_{bind}^{calc}	ΔG_{bind}^{exp}	$\lvert \Delta G_{bind}^{calc} - \Delta G_{bind}^{exp} \rvert$
a1	−87.55	−35.62	−17.02	−20.40	3.38
a2	−84.09	−35.32	−16.57	−16.90	0.33
a3	−86.45	−34.77	−16.73	−14.30	2.43
a4	−17.10	−43.89	−10.84	−8.80	2.04
a5	−17.72	−42.04	−10.55	−12.20	1.65
a6	−84.67	−31.58	−15.91	−16.50	0.59
a7	−86.10	−31.80	−16.11	−14.00	2.11
a8	−15.69	−31.50	−8.28	−11.10	2.82
a9	−11.04	−31.07	−7.68	−7.40	0.28
a10	−14.05	−14.20	−4.73	−4.50	0.23
a11	−13.52	−19.74	−5.75	−6.40	0.65
a12	−9.24	−29.80	−7.23	−5.00	2.23
a13	−16.20	−31.05	−8.25	−7.40	0.85
a14	−15.15	−32.20	−8.35	−9.10	0.75

All data are obtained using AMBER03 force field.

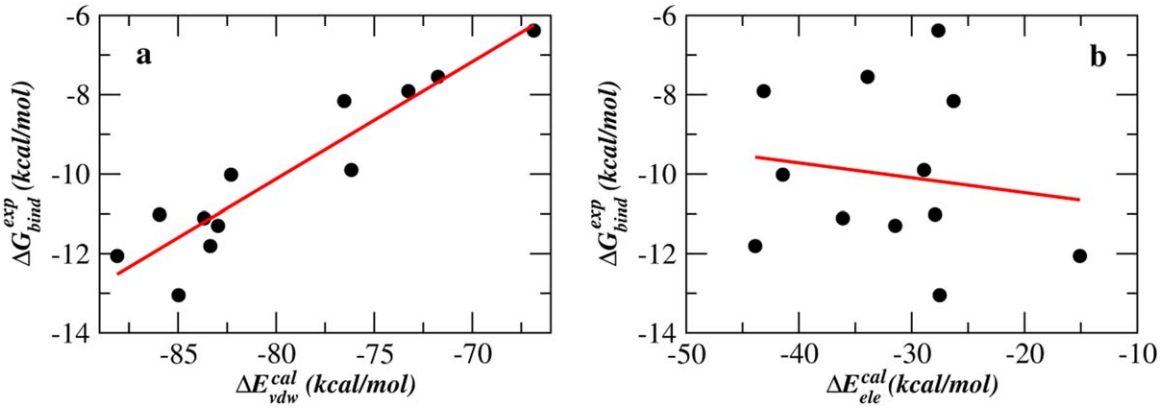


Figure 3. The correlation between the experimental binding energy and individual energy terms for BACE binding systems under PPC. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

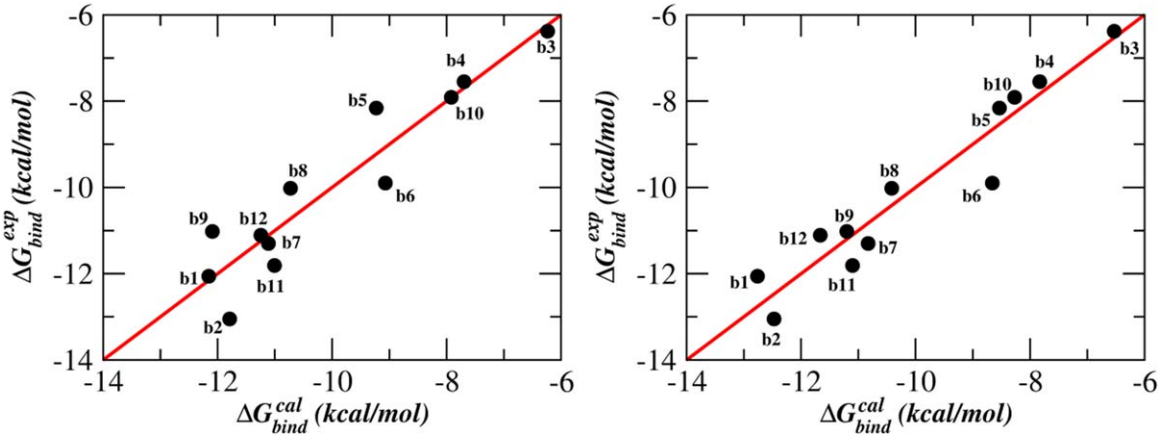


Figure 4. Binding energies calculated by LIE method versus the experimental values for BACE binding systems utilizing (left) PPC and (right) AMBER03 charge. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

giving reliable LIE results for systems of which VDW term is the dominant factor. The entropy-included LIE equation is also used for this AMBER03 force field analysis. Figure 5 (right panel) gives a direct visual comparison between the theoretical and experimental data. The correlation coefficient is 0.970 and the RMSD value is 0.481 kcal/mol. Improvement in not only seen for PPC, but also for AMBER03 force field when entropy effect is considered explicitly.

Table 7. Leave-one-out analysis of the BACE complex under PPC.

Excluded ligand	α	β	γ	ΔG_{bind}^{calc}	ΔG_{bind}^{exp}
b1	−0.021	0.309	13.87	−12.21	−12.06
b2	−0.001	0.289	13.07	−11.41	−13.05
b3	−0.016	0.316	14.55	−6.10	−6.38
b4	−0.019	0.304	13.51	−7.74	−7.55
b5	−0.028	0.298	12.67	−9.41	−8.16
b6	−0.014	0.315	14.60	−8.96	−9.90
b7	−0.018	0.306	13.69	−11.09	−11.30
b8	−0.005	0.315	14.74	−10.96	−10.02
b9	−0.024	0.327	15.06	−12.35	−11.02
b10	−0.018	0.307	13.81	−7.92	−7.91
b11	−0.040	0.294	12.19	−10.58	−11.81
b12	−0.017	0.309	13.97	−11.27	−11.11

Conclusions

Free energy calculation is still a grand challenge in computer simulation. Many efforts have been devoted to pursue a quick and accurate method to predict the free energy change

Table 8. The number of rotatable bonds and the computed and experimental binding free energies for the BACE–ligand set.

Ligand	No. of rotatable bonds	ΔG_{bind}^{calc}	ΔG_{bind}^{exp}	$\lvert\Delta G_{bind}^{calc} - \Delta G_{bind}^{exp}\rvert$
b1	29	−12.56	−12.06	0.49
b2	31	−12.84	−13.05	0.21
b3	17	−6.32	−6.38	0.06
b4	18	−7.62	−7.55	0.07
b5	18	−8.59	−8.16	0.43
b6	20	−8.95	−9.90	0.95
b7	21	−10.68	−11.30	0.63
b8	21	−10.73	−10.02	0.71
b9	21	−11.26	−11.02	0.24
b10	19	−8.40	−7.91	0.49
b11	22	−11.20	−11.81	0.61
b12	22	−11.12	−11.11	0.01

The entropy added LIE equation is used and all data are obtained using PPC charge scheme.

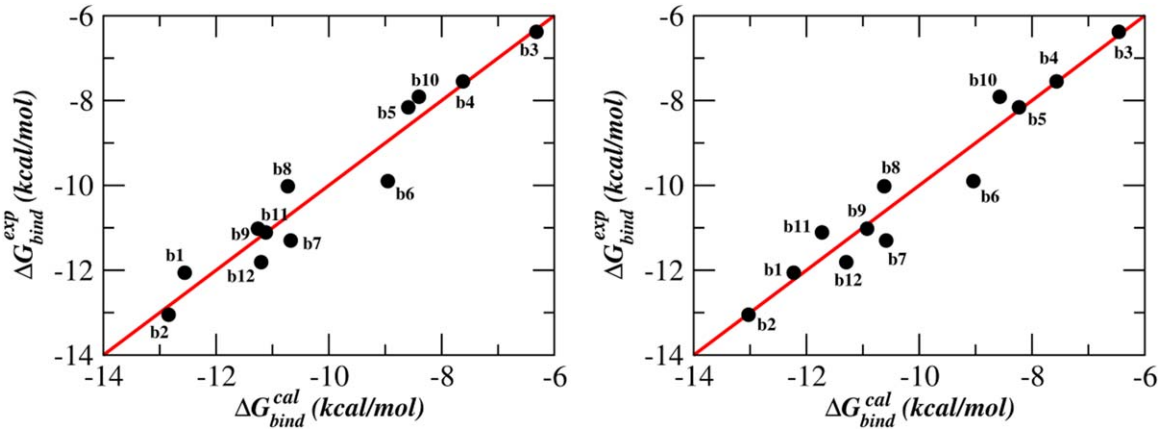


Figure 5. Binding energies calculated by entropy-included LIE method versus the experimental values for BACE binding systems from trajectories utilizing (left) PPC and (right) AMBER03 charge. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Table 9. Computed interaction energy components and the computed and experimental binding free energies for the BACE–ligand set.

Ligand	$\langle E_{elec} \rangle$	$\langle E_{vdW} \rangle$	ΔG_{bind}^{calc}	ΔG_{bind}^{exp}	$\lvert \Delta G_{bind}^{calc} - \Delta G_{bind}^{exp} \rvert$
b1	16.38	−96.27	−12.83	−12.06	0.77
b2	22.02	−94.34	−12.21	−13.05	0.84
b3	−37.71	−76.53	−6.53	−6.38	0.15
b4	−36.94	−80.68	−7.83	−7.55	0.28
b5	−43.64	−82.94	−8.55	−8.16	0.39
b6	−42.31	−83.33	−8.67	−9.90	1.23
b7	−40.24	−90.25	−10.87	−11.30	0.43
b8	−47.02	−88.95	−10.44	−10.02	0.42
b9	−43.42	−91.44	−11.25	−11.02	0.23
b10	−46.12	−82.10	−8.26	−7.91	0.35
b11	−44.17	−91.13	−11.13	−11.81	0.68
b12	−47.01	−92.94	−11.72	−11.11	0.61

All data are obtained using AMBER03 force field.

associated with protein/ligand binding. LIE method is very effective in predicting the binding free energy of a new ligand from the data for a family of similar ligands. This method assumes a linear response of the environment to the variation of the ligand. The accuracy of LIE calculation depends on force field and scale of conformational space been sampled. PPC, which is believed to represent the true electronic distribution, has been proved to be quite successful in the studies of protein structure and dynamics. In this work, the applicability of PPC in LIE method is justified through the calculation of the affinities for series of ligands binding to avidin and β-secretase. Although the result is encouraging, we noticed that other inherent defects of LIE must be fixed in order to achieve further improvement. For ligands without knowledge of its complex structure with the receptor, long time MD simulation might be essential to relax the complex toward the true binding pattern. Differences in entropy change upon ligand binding, which is often neglected in LIE approach, may be essential.

Acknowledgment

The authors thank Supercomputer Center of East China Normal University for CPU time support.

Keywords: polarization · linear interaction energy · force field · charge model · avidin · β-secretase

How to cite this article: X. Jia, J. Zeng, John Z. H. Zhang, Y. Mei. *J. Comput. Chem.* **2014**, *35*, 737–747. DOI: 10.1002/jcc.23547

Additional Supporting Information may be found in the online version of this article.

[1] D. L. Mobley, K. A. Dill, *Structure* **2009**, *17*, 489.
[2] G. J. Rocklin, D. L. Mobley, K. A. Dill, *J. Chem. Theory Comput.* **2013**, *9*, 3072.
[3] D. G. Fedorov, T. Ishida, M. Uebayasi, K. Kitaura, *J. Phys. Chem. A* **2007**, *111*, 2722.
[4] V. Gogonea, L. M. Westerhoff, K. M. Merz, *J. Chem. Phys.* **2000**, *113*, 5604.
[5] M. A. Collins, *J. Chem. Phys.* **2007**, *127*, 024104.
[6] D. W. Zhang, Y. Xiang, J. Z. H. Zhang, *J. Phys. Chem. B* **2003**, *107*, 12039.
[7] C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, *97*, 10269.
[8] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1996**, *118*, 2309.
[9] C. G. Ji, Y. Mei, J. Z. H. Zhang, *Biophys. J.* **2008**, *95*, 1080.
[10] C. G. Ji, J. Z. H. Zhang, *J. Comput. Chem.* **2012**, *33*, 1416.
[11] Y. Mei, Y. L. Li, J. Zeng, J. Z. H. Zhang, *J. Comput. Chem.* **2012**, *33*, 1374.
[12] Y. Tong, Y. Mei, Y. L. Li, C. G. Ji, J. Z. H. Zhang, *J. Am. Chem. Soc.* **2010**, *132*, 5137.
[13] Y. Xiang, L. L. Duan, J. Z. H. Zhang, *J. Chem. Phys.* **2011**, *134*, 205101.
[14] M. M. Francl, C. Carey, L. E. Chirlian, D. M. Gange, *J. Comput. Chem.* **1996**, *17*, 367.
[15] J. S. Tan, S. X. M. Boerrigter, R. P. Scaringe, K. R. Morris, *J. Comput. Chem.* **2009**, *30*, 733.
[16] H. Hu, Z. Y. Lu, W. T. Yang, *J. Chem. Theory Comput.* **2007**, *3*, 1004.
[17] J. Zeng, L. L. Duan, J. Z. H. Zhang, Y. Mei, *J. Comput. Chem.* **2013**, *34*, 847.
[18] A. de Ruiter, S. Boresch, C. Oostenbrink, *J. Comput. Chem.* **2013**, *34*, 1024.
[19] P. Kollman, *Chem. Rev.* **1993**, *93*, 2395.
[20] J. Kastner, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 932.
[21] M. Souaille, B. Roux, *Comput. Phys. Commun.* **2001**, *135*, 40.
[22] T. Simonson, J. Carlsson, D. A. Case, *J. Am. Chem. Soc.* **2004**, *126*, 4167.

[23] I. D. Kuntz, J. M. Blaney, S. J. Oatley, R. Langridge, T. E. Ferrin, *J. Mol. Biol.* **1982**, *161*, 269.
[24] D. H. Williams, J. P. L. Cox, A. J. Doig, M. Gardner, U. Gerhard, P. T. Kaye, A. R. Lal, I. A. Nicholls, C. J. Salter, R. C. Mitchell, *J. Am. Chem. Soc.* **1991**, *113*, 7020.
[25] J. Aqvist, C. Medina, J. E. Samuelsson, *Protein Eng.* **1994**, *7*, 385.
[26] J. Aqvist, J. Marelius, *Combinatorial Chem. High Throughput Screen.* **2001**, *4*, 613.
[27] J. Aqvist, V. B. Luzhkov, B. O. Brandsdal, *Acc. Chem. Res.* **2002**, *35*, 358.
[28] M. Almlof, J. Aqvist, A. O. Smalas, B. O. Brandsdal, *Biophys. J.* **2006**, *90*, 433.
[29] B. O. Brandsdal, J. Aqvist, A. O. Smalas, *Protein Sci.* **2001**, *10*, 1584.
[30] W. Wang, J. Wang, P. A. Kollman, *Proteins* **1999**, *34*, 395.
[31] D. Ben-Amotz, R. Underwood, *Acc. Chem. Res.* **2008**, *41*, 957.
[32] J. Aqvist, T. Hansson, *J. Phys. Chem.* **1996**, *100*, 9512.
[33] D. Huang, A. Caflisch, *J. Med. Chem.* **2004**, *47*, 5791.
[34] D. K. Jones-Hertzog, W. L. Jorgensen, *J. Med. Chem.* **1997**, *40*, 1539.
[35] V. Zoete, O. Michielin, M. Karplus, *J. Comput. Aided Mol. Des.* **2003**, *17*, 861.
[36] R. H. Zhou, R. A. Friesner, A. Ghosh, R. C. Rizzo, W. L. Jorgensen, R. M. Levy, *J. Phys. Chem. B* **2001**, *105*, 10388.
[37] K. B. Ljungberg, J. Marelius, D. Musil, P. Svensson, B. Norden, J. Aqvist, *Eur. J. Pharm. Sci.* **2001**, *12*, 441.
[38] E. Stjernschantz, J. Marelius, C. Medina, M. Jacobsson, N. P. E. Vermeulen, C. Oostenbrink, *J. Chem. Inf. Model.* **2006**, *46*, 1972.
[39] G. Wallin, M. Nervall, J. Carlsson, J. Aqvist, *J. Chem. Theory Comput.* **2009**, *5*, 380.
[40] M. Almlof, B. O. Brandsdal, J. Aqvist, *J. Comput. Chem.* **2004**, *25*, 1242.
[41] J. Aqvist, *J. Comput. Chem.* **1996**, *17*, 1587.
[42] B. A. Tounge, C. H. Reynolds, *J. Med. Chem.* **2003**, *46*, 2074.
[43] M. A. Alam, P. K. Naik, *J. Mol. Graph.* **2009**, *27*, 930.
[44] O. Livnah, E. A. Bayer, M. Wilchek, J. L. Sussman, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 5076.
[45] B. Kuhn, P. A. Kollman, *J. Med. Chem.* **2000**, *43*, 3786.
[46] B. Kuhn, P. A. Kollman, *J. Am. Chem. Soc.* **2000**, *122*, 3909.
[47] J. Wang, R. Dixon, P. A. Kollman, *Proteins* **1999**, *34*, 69.
[48] J. Zeng, X. Y. Jia, J. Z. H. Zhang, Y. Mei, *J. Comput. Chem.* **2013**, *34*, 2677.
[49] L. Hong, G. Koelsch, X. L. Lin, S. L. Wu, S. Terzyan, A. K. Ghosh, X. C. Zhang, J. Tang, *Science* **2000**, *290*, 150.
[50] R. Vassar, B. D. Bennett, S. Babu-Khan, S. Kahn, E. A. Mendiaz, P. Denis, D. B. Teplow, S. Ross, P. Amarante, R. Loeloff, Y. Luo, S. Fisher, L. Fuller, S. Edenson, J. Lile, M. A. Jarosinski, A. L. Biere, E. Curran, T. Burgess, J. C. Louis, F. Collins, J. Treanor, G. Rogers, M. Citron, *Science* **1999**, *286*, 735.
[51] M. S. Malamas, J. Erdei, I. Gunawan, J. Turner, Y. Hu, E. Wagner, K. Fan, R. Chopra, A. Olland, J. Bard, S. Jacobsen, R. L. Magolda, M. Pangalos, A. J. Robichaud, *J. Med. Chem.* **2010**, *53*, 1146.
[52] S. Mishra, A. Caflisch, *Biochemistry* **2011**, *50*, 9328.
[53] B. A. Tounge, R. T. Rajarnani, E. W. Baxter, A. B. Reitz, C. H. Reynolds, *J. Mol. Graph.* **2006**, *24*, 475.
[54] K. Raha, K. M. Merz, *J. Am. Chem. Soc.* **2004**, *126*, 1020.
[55] A. V. Ishchenko, E. I. Shakhnovich, *J. Med. Chem.* **2002**, *45*, 2770.
[56] A. Fiser, A. Sali, *Bioinformatics* **2003**, *19*, 2500.
[57] R. Anandakrishnan, B. Aguilar, A. V. Onufriev, *Nucleic Acids Res.* **2012**, *40*, W537.
[58] P. Cieplak, W. D. Cornell, C. Bayly, P. A. Kollman, *J. Comput. Chem.* **1995**, *16*, 1357.
[59] J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157.
[60] R. Rajamani, C. H. Reynolds, *J. Med. Chem.* **2004**, *47*, 5159.
[61] A. Jakalian, B. L. Bush, D. B. Jack, C. I. Bayly, *J. Comput. Chem.* **2000**, *21*, 132.
[62] A. Jakalian, D. B. Jack, C. I. Bayly, *J. Comput. Chem.* **2002**, *23*, 1623.
[63] W. Rocchia, S. Sridharan, A. Nicholls, E. Alexov, A. Chiabrera, B. Honig, *J. Comput. Chem.* **2002**, *23*, 128.
[64] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian 09, Revision B.01; Gaussian, Inc.: Wallingford, CT, **2010**.
[65] D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. Crowley, R. C. Walker, W. Zhang, K. M. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossvary, K. F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, V Hornak, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, P. A. Kollman, B. P. Roberts, AMBER 11; University of California: San Francisco, **2010**.
[66] R. W. Pastor, B. R. Brooks, A. Szabo, *Mol. Phys.* **1988**, *65*, 1409.
[67] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684.
[68] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089.
[69] B. R. Miller, T. D. McGee, J. M. Swails, N. Homeyer, H. Gohlke, A. E. Roitberg, *J. Chem. Theory Comput.* **2012**, *8*, 3314.
[70] H. Gohlke, C. Kiel, D. A. Case, *J. Mol. Biol.* **2003**, *330*, 891.
[71] L. Li, C. Li, Z. Zhang, E. Alexov, *J. Chem. Theory Comput.* **2013**, *9*, 2126.
[72] C. N. Schutz, A. Warshel, *Proteins* **2001**, *44*, 400.
[73] A. Warshel, P. K. Sharma, M. Kato, W. W. Parson, *Biochim. Biophys. Acta (BBA) - Proteins and Proteomics* **2006**, *1764*, 1647.
[74] R. R. Picard, R. D. Cook, *J. Am. Stat. Assoc.* **1984**, *79*, 575.

Received: 15 September 2013
Revised: 15 November 2013
Accepted: 5 January 2014
Published online on 5 February 2014