

Evaluation on performance of MM/PBSA in nucleic acid-protein systems

Yuan-Qiang Chen(陈远强), Yan-Jing Sheng(盛艳静), Hong-Ming Ding(丁泓铭), and Yu-Qiang Ma(马余强)
Citation: Chin. Phys. B, 2022, 31 (**4**): 048701. DOI: 10.1088/1674-1056/ac3a5c
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Biophysical model for high-throughput tumor and epithelial cell co-culture in complex biochemical microenvironments

Guoqiang Li(李国强), Yanping Liu(刘艳平), Jingru Yao(姚静如), Kena Song(宋克纳), Gao Wang(王高), Lianjie Zhou(周连杰), Guo Chen(陈果), and Liyu Liu(刘雳宇)
Chin. Phys. B, 2022, 31 (**2**): 028703. DOI: 10.1088/1674-1056/ac381c

Tunable inhibition of *β*-amyloid peptides by fast green molecules

Tiantian Yang(杨甜甜), Tianxiang Yu(俞天翔), Wenhui Zhao(赵文辉), and Dongdong Lin(林冬冬)
Chin. Phys. B, 2021, 30 (**8**): 088701. DOI: 10.1088/1674-1056/ac0690

Modeling hydrogen exchange of proteins by a multiscale method

Wentao Zhu(祝文涛), Wenfei Li(李文飞), and Wei Wang(王炜)
Chin. Phys. B, 2021, 30 (**7**): 078701. DOI: 10.1088/1674-1056/abe377

Coarse-grained simulations on interactions between spectrins and phase-separated lipid bilayers

Xuegui Lin(林雪桂), Xiaojie Chen(陈晓洁), and Qing Liang(梁清)
Chin. Phys. B, 2021, 30 (**6**): 068701. DOI: 10.1088/1674-1056/abf129

Enhancements of the Gaussian network model in describing nucleotide residue fluctuations for RNA

Wen-Jing Wang(王文静) and Ji-Guo Su(苏计国)
Chin. Phys. B, 2021, 30 (**5**): 058701. DOI: 10.1088/1674-1056/abe1a5

Evaluation on performance of MM/PBSA in nucleic acid-protein systems

Yuan-Qiang Chen(陈远强)[1], Yan-Jing Sheng(盛艳静)[1], Hong-Ming Ding(丁泓铭)[1,†], and Yu-Qiang Ma(马余强)[2,‡]

[1] *Center for Soft Condensed Matter Physics and Interdisciplinary Research, School of Physical Science and Technology, Soochow University, Suzhou* 215006, *China*

[2] *National Laboratory of Solid State Microstructures and Department of Physics, Collaborative Innovation Center of Advanced Microstructures, Nanjing University, Nanjing* 210093, *China*

(Received 11 October 2021; revised manuscript received 11 November 2021; accepted manuscript online 17 November 2021)

The molecular mechanics/Poisson–Boltzmann surface area (MM/PBSA) method has been widely used in predicting the binding affinity among ligands, proteins, and nucleic acids. However, the accuracy of the predicted binding energy by the standard MM/PBSA is not always good, especially in highly charged systems. In this work, we take the protein–nucleic acid complexes as an example, and showed that the use of screening electrostatic energy (instead of Coulomb electrostatic energy) in molecular mechanics can greatly improve the performance of MM/PBSA. In particular, the Pearson correlation coefficient of dataset II in the modified MM/PBSA (*i.e.*, screening MM/PBSA) is about 0.52, much better than that (< 0.33) in the standard MM/PBSA. Further, we also evaluate the effect of solute dielectric constant and salt concentration on the performance of the screening MM/PBSA. The present study highlights the potential power of the screening MM/PBSA for predicting the binding energy in highly charged bio-systems.

Keywords: molecular mechanics/Poisson–Boltzmann surface area (MM/PBSA), screening electrostatic interaction, protein, nucleic acid, molecular dynamics simulation

PACS: 87.14.E–, 87.14.G–, 87.15.A–, 87.15.kj **DOI:** 10.1088/1674-1056/ac3a5c

1. Introduction

The molecular mechanics/Poisson–Boltzmann surface area (MM/PBSA) and the molecular mechanics/generalized Born surface area (MM/GBSA) methods are two popular approaches to estimate the binding free energy of ligand–biomolecule and biomolecule–biomolecule interactions.[1–4] Compared to the alchemical free energy (AFE) methods like free energy perturbation (FEP) and thermodynamic integration (TI),[5–7] the MM/PBSA and MM/GBSA methods require much lower computational cost; moreover, they can also reproduce and rationalize experimental findings better than the molecular docking.[8–11] Thus, the MM/PBSA and MM/GBSA have been widely used in many studies, especially in bio-systems like protein–ligand binding,[1,2,12–16] protein–protein interactions.[2,17–20]

However, recently more and more studies have shown that the performance of the MM/PBSA or MM/GBSA could be just moderate and even poor in some cases,[1,2,21,22] particularly in highly charged systems. To overcome such problem, the high value of solute dielectric constant in the MM/PBSA and MM/GBSA or the residue-type-dependent dielectric constant in the MM/GBSA was recommended.[2,9,14,17,19,23,24] It is shown the latter one can indeed greatly improve the performance of the MM/GBSA.[19,23,24] Moreover, the enhanced sampling is another possible way, but the improvement is usually limited.[17] Notably, it is still difficult to introduce the residue-type-dependent dielectric constant in the MM/PBSA, thus presently the performance of the MM/PBSA (even if a very large solute dielectric constant was set) is usually worse than that of the MM/GBSA in the highly charged systems. Therefore, it is a challenging problem to improve the performance of the MM/PBSA in such systems.

Recently, we raised a modified approach of MM/PBSA that considers the screening effect of ions on the electrostatic interaction between the proteins. The modified method (*i.e.*, screening MM/PBSA) showed a good performance on the protein–protein interactions,[25] especially in the SARS-CoV-2 RBD-ACE2 interaction.[26] However, it is still unclear how about its performance on other bio-systems like the nucleic acid-involved systems.

Actually, the nucleic acids (including DNA/RNA) are highly negative-charged biomolecules. For example, the linear charge density of the single strand DNA is about 0.17 e/nm while that of the double strand DNA is about 0.34 e/nm. Thus, there are many counter ions that could condensed on the DNA, namely the well-known Manning condensation,[27] which could greatly decrease the effective charges of DNA molecules. As a result, the DNA/RNA may be compacted closely to form specific structures in nature[28] and can be assembled into various three-dimensional (3D) structures in the experiments.[29–31] Moreover, their interaction with proteins is of great importance in bio-systems like transcription, replica-

†Corresponding author. E-mail: dinghm@suda.edu.cn
‡Corresponding author. E-mail: myqiang@nju.edu.cn

© 2022 Chinese Physical Society and IOP Publishing Ltd

tion and recombination.[28] For example, the expression level of genes is regulated by transcription-factor proteins, which can recognize a specific sequence of the DNA in the binding domain.[32]

In this work, we take the protein–nucleic acid interaction as an example, and show that the use of the screening electrostatic energy (instead of the coulomb electrostatic energy) in molecular mechanics can greatly improve the performance of the MM/PBSA in the systems, especially when the proteins are also negatively charged. Moreover, the effect of the solute dielectric constant and the salt concentration on the performance of the screening MM/PBSA is also investigated.

2. Model and methods

2.1. Preparation of the datasets for the protein–nucleic acid complexes

In order to systematically assess the performance of the standard MM/PBSA and screening MM/PBSA for the nucleic acid-protein interaction, two different datasets (*i.e.*, dataset I and dataset II) were provided, which were listed in Table S1 and Table S2. The dataset I was mainly selected from the dataset by Hou *et al.*[21] The dataset I covered a broad range of 8 orders of magnitude (experimental binding affinities between 750 μM and 2.14 pM). The proteins in dataset I were all positively charged (or charge-neutral), thus there may exist the attractively electrostatic interactions between the proteins and the nucleic acids. On the contrary, the proteins in dataset II were all negatively charged, and were downloaded from the PDBbind database.[33,34] Notably, since the nucleic acids were negatively charged, there may exist the repulsively electrostatic interaction between the nucleic acids and the proteins. Thus, the data points were not easily collected in dataset II. Nevertheless, it also covered a relatively broad range of 6 orders of magnitude (experimental binding affinities between 50 μM and 90 pM).

2.2. All-atom molecular dynamics (MD) simulations

Each system (*i.e.*, the protein–nucleic acid complex) was solvated in TIP3P water[35] (the minimum distances from the surfaces of the box to the complex atoms were set to 15 Å) and NaCl to neutralize the systems at salt concentration of 0.15 M. In this work, all-atom MD simulations were carried out by using GROMACS 2019.03 package[36,37] with Amber ff14sb_parmbsc1 force field.[38] The LINCS constraints[39] were used to all bonds involving hydrogen atoms. The particle mesh Ewald method was used when calculating the long-range electrostatic interactions,[40] and the Lennard–Jones (LJ) interactions were cut off at a distance of 1.0 nm. The periodic boundary conditions were adopted in all three directions.

During the simulations, the system was first energy-minimized by the steepest descent method for about 10000 cycles when the heavy atoms of the protein and the nucleic acid were harmonically restrained with spring constant of 1000 kJ·mol^{-1}·nm^{-2}. Then the system was energy-minimized by the steepest descent method for about 10000 cycles with turning off all the constraints.[19] After that, the system was gradually heated from 0 K to 298 K in the NVT ensemble over a period of 100 ps, and then relaxed by 100 ps in the NPT ensemble, where the temperature was controlled at 298 K by the V-rescale thermostat[41] with a time constant of 0.2 ps and the pressure was kept at 1 atm (1 atm = 1.01325×10^5 Pa) by the Parrinello–Rahman barostat[42] with a time constant of 1.0 ps. The protein and the nucleic acid were harmonically restrained with spring constant of 1000 kJ·mol^{-1}·nm^{-2} in the above two stages. Finally, 5-ns free NPT simulations were performed for each system.[17]

2.3. MM/PBSA calculation of the protein–nucleic acid binding energy

In the binding free energy calculation, 100 frames (with an interval of 10 ps in the final 1 ns) of the MD trajectories or the minimized structure were used to calculate the binding energy in each system via the MM/PBSA method. All the MM/PBSA calculations were performed by using the modified shell script gmx_mmpbsa.[43]

In standard MM/PBSA, the binding free energy $\Delta G_{\rm bind}$ is calculated as follows:[1–3]

$$\Delta G_{\rm bind} = \Delta E_{\rm ele} + \Delta E_{\rm vdw} + \Delta G_{\rm PB} + \Delta G_{\rm SA} - T\Delta S,$$

where $\Delta E_{\rm ele}$ and $\Delta E_{\rm vdw}$ are the electrostatic energy and van der Waals energy between the protein and the nucleic acid at the gas-phase, respectively. $\Delta G_{\rm PB}$ and $\Delta G_{\rm SA}$ are the polar contribution and non-polar contribution of the desolvation free energy, respectively. $-T\Delta S$ is the entropy change upon binding, and is not considered here due to the high computational cost and low prediction accuracy.[21,22]

Notably, $\Delta E_{\rm ele}$ usually has the Coulombic form in the standard MM/PBSA:[1–3]

$$\Delta E_{ij}^{\rm ele} = \sum_i \sum_j \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_{\rm in} r_{ij}},$$

where q_i is the charge of atom i in the protein, q_j is the charge of atom j in the nucleic acid, r_{ij} is the distance between atom i and atom j, ε_0 is the dielectric constant in vacuum, and $\varepsilon_{\rm in}$ is the relative dielectric constant of the solute.

While in the screening MM/PBSA, $\Delta E_{\rm ele}$ has the following form:[25,26]

$$\Delta E_{ij}^{\rm ele} = \sum_i \sum_j \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_{\rm in} r_{ij}} {\rm e}^{-r_{ij}/\lambda_{\rm D}},$$

where $\lambda_{\rm D}$ is the Debye screening length and is about 0.8 nm when the salt (NaCl) concentration is 0.15 M.

The experimental binding energy was estimated as $\Delta G = k_{\rm B}T\ln K_{\rm d}$, where $k_{\rm B}$ is the Boltzmann constant, T is the temperature, $K_{\rm d}$ is the equilibrium dissociation constant. Then,

the Pearson correlation coefficient r_p was employed to evaluate the linear correlation between the predicted binding free energies and the experimental ones.

3. Results and discussion

3.1. Comparison of accuracy of binding energy between standard MM/PBSA and screening MM/PBSA

The binding free energy of the protein–nucleic acids in dataset I was first calculated using the standard MM/PBSA methods. In terms of the correlation coefficients, the accuracy of the standard MM/PBSA in dataset I was relatively satisfied — the Pearson correlation coefficient (r_P) was about 0.69, 0.69, 0.67, respectively when the solute dielectric constants (ε_{in}) were set as 2.0, 4.0, and 8.0 (see Figs. 1(a)–1(c) and Figs. S1(a)–S1(c)). Notably, the MM/PBSA calculations based on the minimized structures was nearly the same as the MM/PBSA calculations based on the MD trajectories, which agreed with previous studies.[9,22] Moreover, such results also indicated that here r_P was nearly independent on ε_{in}. Nevertheless, when ε_{in} was small (*e.g.*, 2.0 or 4.0), the predicted binding energy was over thousands of kJ/mol in some cases, which was extremely larger than the experimental value ($\sim$ dozens of kJ/mol), even though we did not consider the entropy term here. On the contrary, when using a relatively larger ε_{in}, the predicted binding energy greatly decreased to tens or hundreds of kJ/mol. However, there existed some positive values in the binding energy (see Fig. 1(c)), meaning that the complex of the protein and the nucleic acid may be dissolved, probably due to the fact that the larger ε_{in} may underestimate the attractively electrostatic interaction between the protein and the nucleic acid.

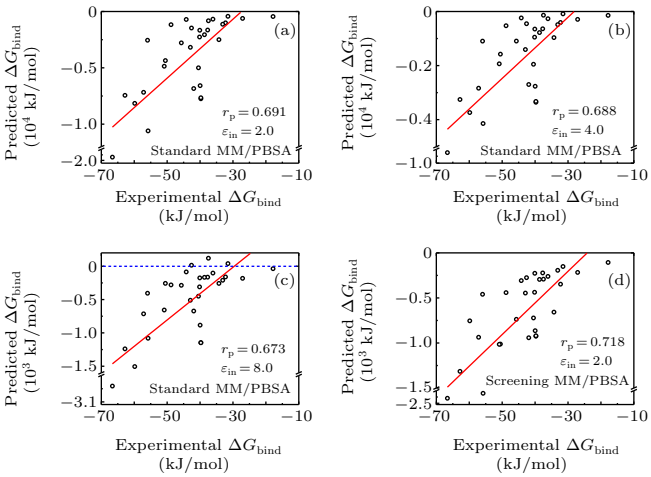


Fig. 1. Pearson correlation coefficient for the dataset I based on the minimized structures in the standard and screening MM/PBSA methods. (a) Standard MM/PBSA under solute dielectric constant 2.0; (b) standard MM/PBSA under solute dielectric constant 4.0; (c) standard MM/PBSA under solute dielectric constant 8.0; (d) screening MM/PBSA under solute dielectric constant 2.0.

The binding energy of the protein–nucleic acid complexes in dataset I was then calculated using the screening MM/PBSA method based on the screening electrostatic energy. As shown in Fig. 1(d) and Fig. S1(d), the Pearson correlation coefficient r_P (0.718–0.730) was slightly better that (0.667–0.695) in the standard MM/PBSA using different ε_{in}. Besides, the MM/PBSA calculation based on the minimized structures ($r_P = 0.718$) was also similar to the MM/PBSA calculations based on the MD trajectories ($r_P = 0.730$). More importantly, the predicted binding energy predicted by screening MM/PBSA was between -100 kJ/mol and -2000 kJ/mol, and no positive values were observed, which was more reasonable than that in the standard MM/PBSA to some extent.

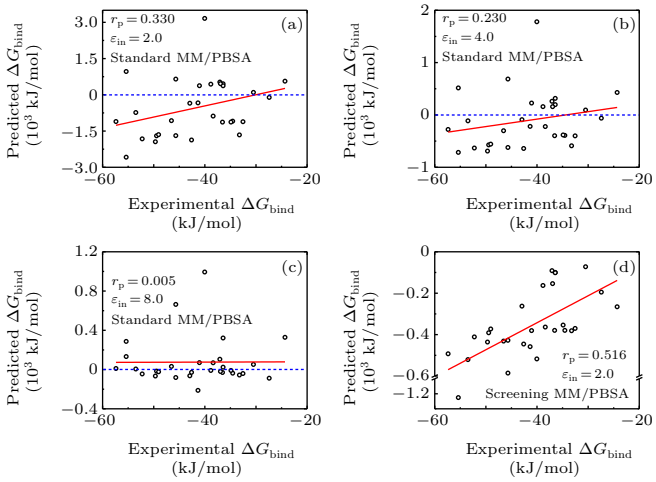


Fig. 2. Pearson correlation coefficient for the dataset II based on the minimized structures in the standard and screening MM/PBSA methods. (a) Standard MM/PBSA with solute dielectric constant 2.0; (b) standard MM/PBSA with solute dielectric constant 4.0; (c) standard MM/PBSA with solute dielectric constant 8.0; (d) screening MM/PBSA with solute dielectric constant 2.0.

The binding energy of the protein–nucleic acid complexes in dataset II was further calculated using the standard MMPBSA and the screening MM/PBSA (Fig. 2 and Fig. S2). In terms of the correlation coefficients, the accuracy of standard MM/PBSA in dataset II was quite poor ($-0.060 < r_P < 0.330$, the average of r_P was about 0.15), where the MM/PBSA calculations under $\varepsilon_{in} = 8.0$ gave the worst prediction and even predicted the opposite trend using the MD trajectories ($r_P = -0.060$). Moreover, there were many positive binding energies predicted by the standard MM/PBSA in dataset II, namely nearly half of the values was positive under $\varepsilon_{in} = 8.0$ and over one third of the values was positive under $\varepsilon_{in} = 2.0$ or 4.0. All these results indicated that the performance of the standard MM/PBSA in dataset II was much weaker than that in dataset I, which is mainly caused by the repulsively electrostatic interactions between the proteins and the nucleic acids carrying the charges of the same sign.

On the contrary, the accuracy of the screening MM/PBSA in dataset II was much better than that of the standard MM/PBSA and was in the moderate range (Fig. 2(d) and Fig. S2(d)), where the Pearson correlation coefficient was 0.516 based on the minimized structures and it was 0.528 based on the MD trajectories. Moreover, the predicted binding

energies were all negative. We analyzed the electrostatic energy in the cases where the binding energy predicted by the standard MM/PBSA was positive. The electrostatic energy was divided into two parts on the basis of the Debye screening length (0.8 nm), namely, the long-range part ($r > 0.8$ nm) and the short-range part ($r < 0.8$ nm). As shown in Table S3, the long-range electrostatic energy was always positive (*i.e.*, repulsive) when using the standard MM/PBSA. More importantly, the repulsively long-range energy was comparable to (and even greater than) the attractively short-range energy. As a result, the total electrostatic energy was small and can be even positive. On the contrary, the long-range electrostatic energy was negative or weakly positive, and the attractively short-range electrostatic energy dominated in most cases when using the screening MM/PBSA. On the basis of above discussion, we can conclude that the usage of Coulomb interaction in the standard MM/PBSA may overestimate the repulsively electrostatic energy (especially in the long range) between the proteins and the nucleic acids carrying the charges of the same sign. On the contrary, the screening interaction corrected the overestimated electrostatic energy between the proteins and the nucleic acids in the long range, thus it gave a relatively better result.

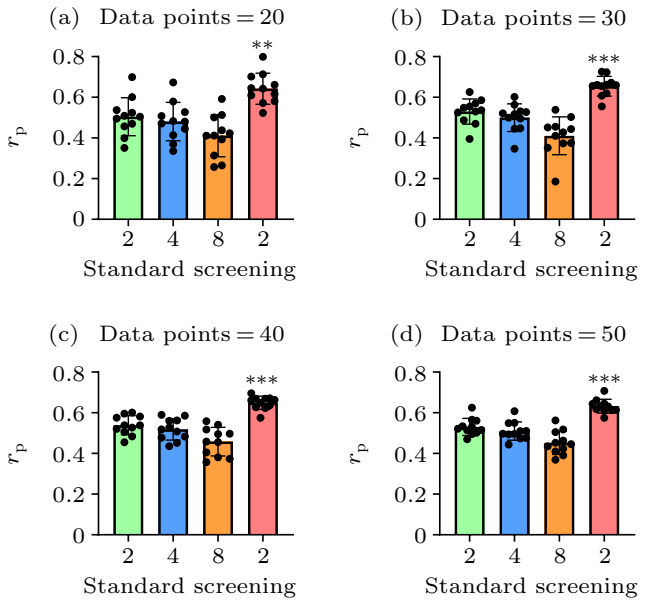


Fig. 3. Overall comparison of the prediction accuracies between the standard and screening MM/PBSA methods using different data points. We randomly selected the data points from dataset I and dataset II for 11 times in each case. The p values were calculated using multiple t tests ($^{**}p < 0.01$, $^{***}p < 0.001$).

Taking dataset I and dataset II together, the screening MM/PBSA method ($r_P = 0.628$) still yields the best performance, following with 0.534 ($\varepsilon_{in} = 2.0$), 0.516 ($\varepsilon_{in} = 4.0$), 0.458 ($\varepsilon_{in} = 8.0$) in the standard MM/PBSA. We also compared the performance of the screening MM/PBSA and the standard MM/PBSA by using different mixtures of dataset I and dataset II. Not surprisingly, the accuracy of the binding energy calculated by the screening MM/PBSA was always the highest, and with the increase of data points, the significant difference between the screening MM/PBSA and the standard MM/PBSA became more obvious (Fig. 3), indicating that the usage of the screening electrostatic energy in molecular mechanics can indeed improve the performance of MM/PBSA in highly charged systems.

3.2. The effect of solute dielectric constant in screening MM/PBSA on accuracy of predicted binding energy

The accuracy of the predicted binding energy is usually related to the solute dielectric constant ε_{in} in the standard MM/PBSA. Here, we investigated the effect of solute dielectric constant on the performance of the screening MM/PBSA. As discussed above, the results based on the minimized structures was similar to that based on the MD trajectories, thus, for the sake of simplicity, we just used the minimized structures for the MM/PBSA calculation in this subsection.

As shown in Table 1, the accuracy of the screening MM/PBSA in dataset I was nearly independent on ε_{in} (r_P was about 0.70) except that the Pearson correlation coefficient was 0.658 when $\varepsilon_{in} = 3.2$. The best performance of the screening MM/PBSA ($r_P = 0.728$) in this case corresponded to $\varepsilon_{in} = 2.4$ or 2.6. While in dataset II, the Pearson correlation coefficient decreased with the increase of the solute dielectric constant, but the difference was not obvious when $\varepsilon_{in} <= 2.0$. With further increase of the solute dielectric constant, the performance of the screening MM/PBSA became worse, particular when $\varepsilon_{in} > 2.6$, indicating that the large ε_{in} may underestimate the electrostatic energy and was harmful for the prediction accuracy in this case. Actually, when the proteins and the nucleic acids carried the opposite charges in dataset I, there should be a strong polarization at the interface of the protein–nucleic acid complex. Thus a large ε_{in} should be used to well describe the polarization but this value should not be very large to avoid overestimating the polarization. While in dataset II, the proteins and the nucleic acids carried the like charges, there should be no obvious polarization at their interface. Thus, a small ε_{in} (*i.e.*, 1.0) should be used. Taking the two datasets together, the screening MM/PBSA showed good performance when $\varepsilon_{in} <= 2.6$ (see the last column in Table 1), where it ($r_P = 0.628$) showed the best performance under $\varepsilon_{in} = 2.0$.

We also investigated the relationship between the correlation coefficient and the solute dielectric constant in the standard MM/PBSA (Table 2). Similar to that in the screening MM/PBSA, there was also an optimal value of ε_{in} in dataset I while the Pearson correlation coefficient decreased with the increase of the solute dielectric constant in dataset II. However, since such difference in dataset II was much more obvious than that in dataset I, the Pearson correlation coefficient also decreased with the increase of the solute dielectric constant in the whole dataset.

Table 1. Overall prediction accuracies of the screening MM/PBSA under different solute dielectric constants.

ε_{in}	r_P in dataset I	r_P in dataset II	r_P in all datasets
1.0	0.698	0.563	0.623
1.2	0.702	0.557	0.624
1.4	0.705	0.550	0.625
1.6	0.710	0.540	0.626
1.8	0.714	0.530	0.627
2.0	0.718	0.516	0.628
2.2	0.723	0.493	0.627
2.4	0.728	0.462	0.625
2.6	0.728	0.415	0.618
2.8	0.723	0.349	0.602
3.0	0.700	0.255	0.568
3.2	0.658	0.140	0.500

Table 2. Overall prediction accuracies of the standard MM/PBSA under different solute dielectric constants.

ε_{in}	r_P in dataset I	r_P in dataset II	r_P in all datasets
1.0	0.688	0.381	0.540
2.0	0.691	0.330	0.534
3.0	0.692	0.281	0.525
4.0	0.688	0.230	0.516
5.0	0.687	0.176	0.505
6.0	0.684	0.119	0.492
7.0	0.680	0.058	0.477
8.0	0.673	−0.002	0.458

3.3. Effect of salt concentration in screening MM/PBSA on accuracy of predicted binding energy

In real experiment, the conditions for measuring the protein–nucleic acid binding affinity could be very different, where the type of salts could be various and the concentration could be in a broad range. Since the Debye length in the screening electrostatic energy and the polar energy was related to the salt concentration, it is of great importance to investigate the effect of the salt concentration on the accuracy of binding energy predicted by the screening MM/PBSA. Here, for the sake of simplicity, three different salt concentrations were considered (0.05 M, 0.15 M, 0.30 M) and the salt type (*i.e.*, NaCl) was kept the same.

As shown in Fig. 4, when the salt concentration was 0.05 M, most of the predicted binding energy became larger due to the weaker screening effect of the electrostatic interaction, particularly in dataset I. On the contrary, most of the predicted binding energy become smaller due to the stronger screening effect when the salt concentration was 0.30 M. In terms of the correlation coefficients, the accuracy of the screening MM/PBSA at 0.05 M was very similar to that at 0.15 M (*i.e.*, 0.709 *versus* 0.718 in dataset I and 0.519 *versus* 0.516 in dataset II). However, the accuracy of the screening MM/PBSA at 0.30 M became worse than that at 0.15 M, especially in dataset II (0.291 *versus* 0.516), indicating that the high screening effect could be harmful for the prediction, which was similar to the standard MM/PBSA using a very large ε_{in} (*e.g.*, 8.0) or the screening MM/PBSA using a large ε_{in} (*e.g.*, 3.0). As we know, the interaction energy at the interface between the protein and the nucleic acid usually played an important role in total binding energy. The high screening of electrostatic interaction would certainly underestimate the attractively electrostatic energy at the binding interface, particularly in dataset II (the protein was also negatively charged), thereby it would lead the inaccuracy of the screening MM/PBSA.

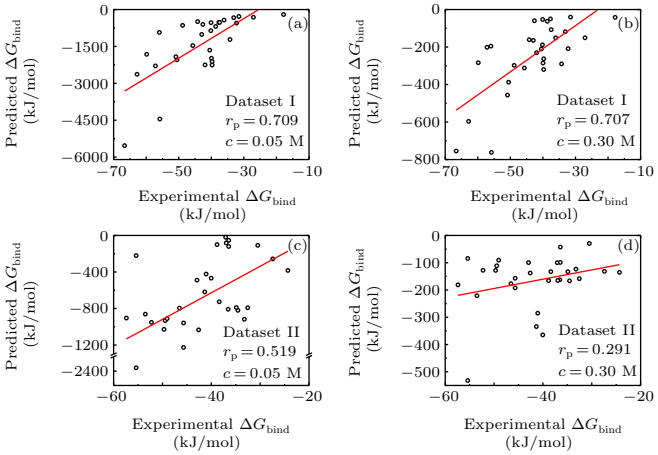


Fig. 4. Pearson correlation coefficient for the dataset I and dataset II based on the minimized structures in the screening MM/PBSA methods. (a) Dataset I under the concentration of 0.05 M; (b) dataset I under the concentration of 0.30 M; (c) dataset II under the concentration of 0.05 M; (d) dataset II under the concentration of 0.30 M.

4. Conclusion

In this study, we have applied a modified approach via the screening electrostatic energy within the framework of the MM/PBSA method. Based on 60 protein–nucleic acid complexes in dataset I and dataset II, the screening MM/PBSA shows an excellent performance in dataset I and a moderate performance in dataset II; while the standard MM/PBSA just yields a good performance in dataset I and a poor performance in dataset II (probably due to the overestimated long-range electrostatic energy). Besides, we also evaluated the performance of the screening MM/PBSA under different solute dielectric constant ε_{in}, and recommend that a small ε_{in} should be used. In general, the screening MM/PBSA can indeed provide reasonable predictions on the binding energy in charged bio-systems.

Acknowledgment

Project supported by the National Natural Science Foundation of China (Grant Nos. 11874045 and 11774147).

References

[1] Wang E, Sun H, Wang J, Wang Z, Liu H, Zhang J Z H and Hou T 2019 *Chem. Rev.* **119** 9478
[2] Genheden S and Ryde U 2015 *Expert. Opin. Drug. Dis.* **10** 449
[3] Wang C, Greene D, Xiao L, Qi R and Luo R 2018 *Front. Mol. Biosci.* **4** 87

[4] Homeyer N and Gohlke H 2012 *Mol. Inform.* **31** 114
[5] Williamsnoonan B J, Yuriev E and Chalmers D K 2017 *J. Med. Chem.* **61** 638
[6] Chodera J D, Mobley D L, Shirts M R, Dixon R W, Branson K and Pande V S 2011 *Curr. Opin. Struc. Biol.* **21** 150
[7] Pohorille A, Jarzynski C and Chipot C 2010 *J. Phys. Chem. B* **114** 10235
[8] Pu C, Yan G, Shi J and Li R 2017 *Med. Chem. Comm.* **8** 1452
[9] Sun H, Li Y, Shen M, Tian S, Xu L, Pan P, Guan Y and Hou T 2014 *Phys. Chem. Chem. Phys.* **16** 22035
[10] Hou T, Wang J, Li Y and Wang W 2011 *J. Comput. Chem.* **32** 866
[11] Ahinko M, Niinivehmas S, Jokinen E M and Pentikainen O T 2019 *Chem. Biol. Drug. Des.* **93** 522
[12] Hou T, Wang J, Li Y and Wang W 2011 *J. Chem. Inf. Model.* **51** 69
[13] Liu J, He X and Zhang J Z H 2013 *J. Chem. Inf. Model.* **53** 1306
[14] Wang E, Liu H, Wang J, Weng G, Sun H, Wang Z, Kang Y and Hou T 2020 *J. Chem. Inf. Model.* **60** 5353
[15] Mikulskis P, Genheden S and Ryde U 2014 *J. Chem. Inf. Model.* **54** 2794
[16] Greenidge P, Kramer C, Mozziconacci J and Wolf R M 2013 *J. Chem. Inf. Model.* **53** 201
[17] Wang E, Weng G, Sun H, Du H, Zhu F, Chen F, Wang Z and Hou T 2019 *Phys. Chem. Chem. Phys.* **21** 18958
[18] Maffucci I and Contini A 2016 *J. Chem. Inf. Model.* **56** 1692
[19] Liu X, Peng L and Zhang J Z H 2019 *J. Chem. Inf. Model.* **59** 272
[20] Sun Z, Yan Y N, Yang M and Zhang J Z H 2017 *J. Chem. Phys.* **146** 124124
[21] Chen F, Sun H, Wang J, Zhu F, Liu H, Wang Z, Lei T, Li Y and Hou T 2018 *RNA* **24** 1183
[22] Sun H, Li Y, Tian S, Xu L and Hou T 2014 *Phys. Chem. Chem. Phys.* **16** 16719
[23] Ravindranathan K P, Tiradorives J, Jorgensen W L and Guimaraes C R W 2011 *J. Chem. Theory Comput.* **7** 3859
[24] Li Y, Cong Y, Feng G, Zhong S, Zhang J Z H, Sun H and Duan L 2018 *Struct. Dynam.* **5** 064101
[25] Sheng Y J, Yin Y W, Ma Y Q and Ding H M 2021 *J. Chem. Inf. Model.* **61** 2454
[26] Ding H M, Yin Y W, Ni S D, Sheng Y J and Ma Y Q 2021 *Chin. Phys. Lett.* **38** 018701
[27] Manning G S 1969 *J. Chem. Phys.* **51** 924
[28] Liljas A, Liljas L, Piskur J, Linblom G, Nissen P and Kjeldgaard M 2009 *Textbook of Structural Biology* (Singapore: World Scientific Publishing)
[29] Ding H, Li J, Chen N, Hu X, Yang X, Guo L, Li Q, Zuo X, Wang L and Ma Y 2018 *ACS Central. Sci.* **4** 1344
[30] Douglas S M, Dietz H, Liedl T, Hogberg B, Graf F and Shih W M 2009 *Nature* **459** 414
[31] Hu Q, Li H, Wang L, Gu H and Fan C 2018 *Chem. Rev.* **119** 6459
[32] Orphanides G and Reinberg D 2002 *Cell* **108** 439
[33] Liu Z, Su M, Han L, Liu J, Yang Q, Li Y and Wang R 2017 *Acc. Chem. Res.* **50** 302
[34] Wang R, Fang X, Lu Y and Wang S 2004 *J. Med. Chem.* **47** 2977
[35] Jorgensen W L and Madura J D 1983 *J. Am. Chem. Soc.* **105** 1407
[36] Der Spoel D V, Lindahl E, Hess B, Groenhof G, Mark A E and Berendsen H J C 2005 *J. Comput. Chem.* **26** 1701
[37] Abraham M J, Murtola T, Schulz R, Pall S, Smith J C, Hess B and Lindahl E 2015 *SoftwareX* **1–2** 19
[38] Maier J, Martinez C, Kasavajhala K, Wickstrom L, Hauser K and Simmerling C 2015 *J. Chem. Theory Comput.* **11** 3696
[39] Hess B, Bekker H, Berendsen H J C and Fraaije J G E M 1997 *J. Comput. Chem.* **18** 1463
[40] Essmann U, Perera L, Berkowitz M L, Darden T, Lee H and Pedersen L G 1995 *J. Chem. Phys.* **103** 8577
[41] Bussi G, Donadio D and Parrinello M 2007 *J. Chem. Phys.* **126** 014101
[42] Parrinello M and Rahman A 1981 *J. Appl. Phys.* **52** 7182
[43] https://jerkwin.github.io/gmxtool