

Molecular Dynamics Simulation of Protein Crystal with Polarized Protein-Specific Force Field

Yongxiu Li,^{†,‡} John Z. H. Zhang,^{†,¶} and Ye Mei^{*,†,¶}

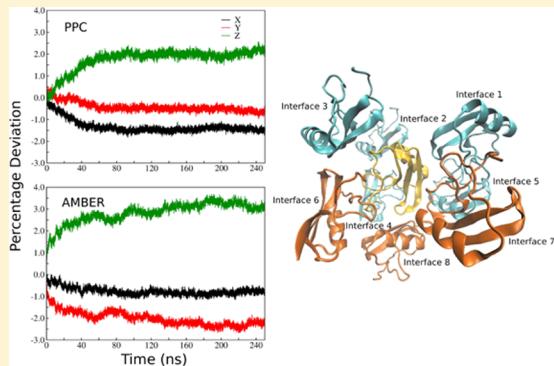
[†]State Key Laboratory of Precision Spectroscopy, Department of Physics and Institute of Theoretical and Computational Science, East China Normal University, Shanghai 200062, China

[‡]School of Chemistry and Chemical Engineering, Anshun University, Anshun 561000, China

[¶]NYU-ECNU Center for Computational Chemistry at NYU Shanghai, Shanghai 200062, China

Supporting Information

ABSTRACT: Two 250 ns molecular simulations have been carried out to study the structure and dynamics of crystal toxin protein II from the scorpion *Androctonus australis* Hector employing the polarized protein-specific charge (PPC), as well as the standard AMBER99SB force field, to investigate the electrostatic polarization on the simulated crystal stability. Results show that under PPC, the monomers in unit cell as well as the lattice in supercell are more stable with smaller root-mean-square deviations and more accurate lattice atomic fluctuations compared with the crystallographic B-factors than under AMBER99SB force field. Most of the interactions at interfaces in the X-ray structure are quite well-preserved, underscoring the important effect of polarization on maintaining the crystal stability. However, the results also show that the hydrogen bond between Asp53 and Gln37 and the cation–π interaction between Arg56 and His64 are not stable, indicating that further optimization of force field, especially the van der Waals interaction parameters, is desired.



INTRODUCTION

X-ray crystallography is one of the major experimental means to determine protein structure. However, determination of atomic positions for protein is a high-dimensional mathematical problem. It may resort to molecular dynamics simulation to optimize the conformations. Proteins in crystal lattice pack very closely to each other as a crowded environment. Without a thick layer of water molecules screening the strong Coulomb interaction, the heavily charged proteins generate heterogeneous electrostatic environment leading to strong polarization effect. Directional hydrogen bond interaction and salt bridges may play important roles in determining the crystal lattice and protein structure. However, contemporary protein force fields are mainly developed for simulations in aqueous solution or in membrane. Their applicability in simulating a crystal system is not guaranteed.

Many studies have been dedicated to this issue. Cerutti et al.¹ investigated the stability of the toxin II protein lattice using various combinations of force fields (AMBER99SB,² AMBER03,³ OPLS,⁴ CHARMM22,^{5,6}) and water models (TIP3P,⁷ SPC/E,⁸ TIP3P-Ewald,⁹ and TIP4P-2005¹⁰). They found that AMBER99SB force field outperformed other force fields in maintaining the lattice structure and producing the most realistic atomic fluctuations. The choice of water model had a minor effect on the structural properties of the monomers or the lattice as a whole. However, none of the force fields gave correct distances between the centers of mass (COM) of the

monomers. Excessive formation of hydrogen bonds or salt bridges between polar groups and the loss of hydrophobic interaction occurred during simulation. All force fields they used were fine-grained with explicit representation of all the hydrogen atoms and were additive for nonbonded interactions with effective mean-field partial charges. Many critical features of the crystal energy landscape (such as aspherical features of the electron density inherent to lone pairs and π electrons, and perturbations such as conformational flexibility or transfer between vapor, liquid, and crystalline environments) were neglected by fixed atomic charge force fields such as AMBER, CHARMM, or OPLS-AA. CL&P force field¹¹ is an evolution of OPLS-AA force field, in which the electron density computed at MP2/cc-pVTZ(-f) level is used to calculate the point CHarges from Electrostatic Potentials using a Grid-based method (CHelpG). The point charges are placed at the nuclear centers of the atoms. With an overall charge-rescaling factor of 0.8, CL&P force field provided an improved description of the organic ionic plastic crystal, which was in line with the known experimental data.¹² Svärd et al.¹³ simulated seven crystal structures of small aromatic monomolecular compounds using several molecular mechanics (MM) force fields combined with different methods to assign atomic point charges. Their

Received: April 23, 2014

Revised: October 5, 2014

Published: October 6, 2014

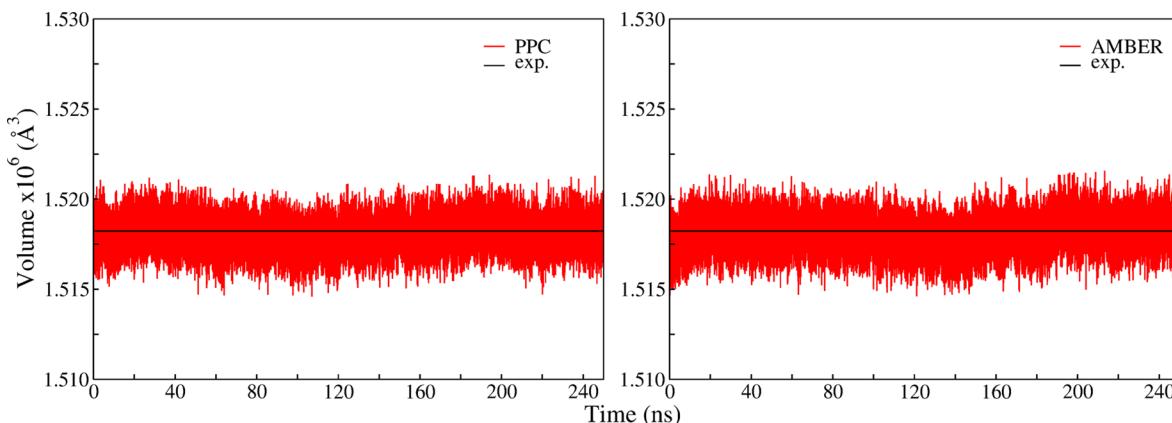


Figure 1. Supercell volume in the simulations employing PPC and AMBER99SB force field.

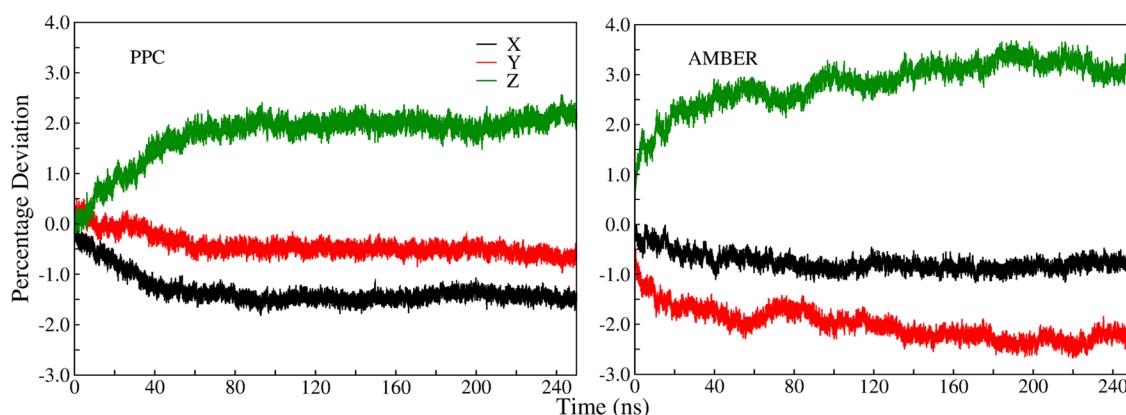


Figure 2. Supercell dimensions in the simulations employing PPC and AMBER99SB force field.

performance was evaluated on identifying the absolute crystal lattice energies to check whether the experimental structure was a minimum on the potential energy hypersurface. The results were highly dependent on the charge schemes, particularly Drieding. Most of the tested methods were unable to predict absolute lattice energies reliably. Therefore, incorporating polarization effect is needed in the study of protein crystals. Schnieders et al.¹⁴ carried out molecular dynamics (MD) simulations of protein crystals of different resolutions with the Atomic Multipole Optimized Energetics for Biomolecular Applications (AMOEBA) force field.¹⁵ This force field has been very successful in the study of structure, thermodynamic stability, and organic crystal solubility.¹⁶ However, AMOEBA uses high-level multipoles to describe the electrostatic interaction, which is much more expensive than those utilizing point charges only.

The polarized protein-specific charge (PPC)¹⁷ has been used successfully in describing electrostatic interaction in protein-related systems.^{18–22} Unlike the mean-field charge in AMBER force field, PPC is derived from *in situ* quantum mechanical calculation of protein using the molecular fractionation with conjugate caps (MFCC) approach^{17,23,24} and Poisson–Boltzmann (PB) solvation model.^{25,26} Atomic charges depend on their specific conformations and orientations in the electric field. Even the same kind of amino acids have different atomic charges. Hence, PPC provides a more reliable description of charge distribution at a fixed conformation. PPC has shown important properties for the study of protein–ligand binding^{18,19,27} and protein folding.²¹ However, PPC has not been

used in a simulation of protein crystal yet. Employing only the fixed and prepolarized zeroth order multipole, it is much faster than AMOEBA and other polarizable force fields, which makes it a promising scheme for the simulation of protein crystals. It is necessary to examine the performance of PPC in a simulation of protein crystal by comparing with other force fields with mean-field charge schemes.

The structure of toxin protein II from the scorpion *Androctonus australis* Hector has been determined by X-ray at a high resolution of 0.96 Å [Protein Data Bank (PDB) accession codes 1AHO²⁸ and 1PTX²⁹]. It is chemically stable in 0.2 M ammonium acetate at pH 6.8 and crystallizes in an orthorhombic $P2_12_12_1$ space group by slow evaporation at 4 °C.³⁰ The molecular dynamics simulation of this protein crystal is feasible due to the mild experimental conditions. In this article, two simulations of the toxin II protein crystal were carried out using PPC and AMBER99SB force field, respectively. The results showed that PPC was superior to AMBER99SB force field in maintaining the stability of this protein crystal.

RESULTS AND DISCUSSION

Cell Dimensions. The trajectories analyzed below are the final 250 ns production in NPT ensemble. The average volume deviations from the experimental measurement under PPC and AMBER99SB force field are all about –0.01% (see Figure 1). However, examinations of each dimension show that the constructed supercell still has some anisotropic deformation, especially under AMBER99SB force field (see Figure 2). The Z

dimension tends to expand, while the X dimension shrinks. The Y dimension under PPC is very stable. But under AMBER99SB, the contraction in the Y dimension is significant. The expansion and contraction of the supercell reach plateaus in 60 ns under PPC and 100 ns under AMBER99SB force field.

Analysis of Protein Structures. Atomic root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF) are computed by two different ways labeled as “ASU” and “Lattice” respectively. The “ASU RMSD” for M ASUs with N atoms in each ASU is computed as³¹

$$\text{ASU RMSD} = \sqrt{\frac{1}{M} \sum_{j=1}^M \left[\frac{1}{N} \sum_{i=1}^N |\mathbf{r}_{i,j} - \mathbf{r}_{i,j}^*|^2 \right]} \quad (1)$$

where $\mathbf{r}_{i,j}$ and $\mathbf{r}_{i,j}^*$ are the coordinates of atom i in ASU j in the simulation and X-ray. The global rotational and translational motions of the backbone heavy atoms in each monomer against the crystal 1AHO structure are removed using the Kabsch algorithm³² prior to the ASU RMSD calculation. As shown in Table 1, the internal degree-of-freedom (DoF) of backbone

Table 1. Average ASU and Lattice Backbone RMSD During the 250 ns NPT Simulations

RMSD	PPC	AMBER99SB
ASU	0.58	0.64
lattice	1.11	1.28

atoms measured by ASU RMSD remains very close to that in the crystallographic structure. This is much smaller than the usual RMSD for protein after a moderate timescale simulation in aqueous solution because the neighbors of protein units in the crystallographic condition impede its thermal motion. The ASU RMSD under PPC is even smaller than that under AMBER99SB force field. The time evolutions of the “ASU RMSDs” are shown in Figure 3. It shows that the deformation of each ASU is very fast. Over 90% of deformations take place in the first 20 ns. The ASU RMSD under AMBER99SB force field has a much broader distribution than that under PPC. The deformation comes from the backbone atoms and the side chains almost equally, which can be read from the difference of the RMSDs for the backbone atoms and for all the heavy atoms shown in Figure S1 of the Supporting Information.

The distance deviations for α -carbons in each monomer are plotted in Figure 4. The distance deviations are the difference between the average distances obtained over the last 150 ns simulation and the corresponding distances in the X-ray structure. Because the overall deformations of ASUs are very small, the distance deviations are mainly distributed in a narrow range from -0.5 \AA to 1.0 \AA (see Figure 5). It shows that the overall distribution under PPC shifts toward longer distances, and that under AMBER99SB force field, shifts to the right side a little bit further. This observation is consistent with the fact that PPC gives stronger intraprotein interaction than AMBER charge through strengthened Coulomb interaction such as a hydrogen bond.

Some patterns in the deviations appear to occur under both force fields. Two random coil regions in this protein (residues 7–10 and residues 52–56) are drawn closer to the α region (residues 19–29). One random coil region (residues 10–18) moves away from nearly most residues. Both force fields produce very little deviation in the juxtaposition of the helix (residues 19–29) with the β -strand (residues 45–51). The

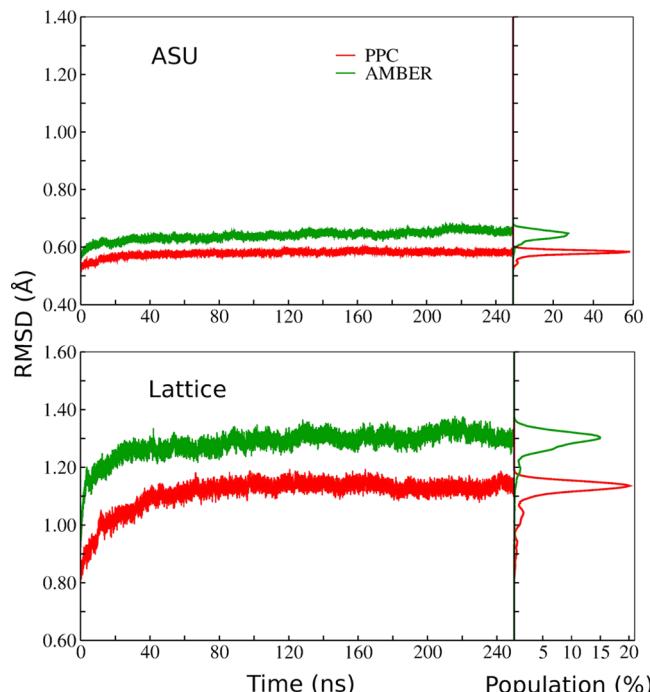


Figure 3. Backbone root-mean-square deviations for ASU and lattice in the simulations. Histograms of the RMSD are given in the right panel.

interaction of these two secondary structures is reinforced by two disulfide bonds. Under both force fields, Ala39 and Gly43 move away from most of the other residues, but Ser40, Pro41, and Tyr42 stand still. Hence, two β -hairpins flip outward but the T-stacking contact between Tyr42 and Tyr5 is preserved.

Nonetheless, there are many different patterns between these two force fields. Under PPC, Asp9 moves closer to nearly all other parts of protein, while under AMBER99SB force field Asp9 moves closer to residues 15–25 but away from residues 26–32 and residues 49–56. Tyr49 moves away from most other residues under AMBER99SB force field, while it takes nearly the same position as in the crystal under PPC. As shown in Figure 6, PPC maintains the correct orientations of the side chains in Tyr5, Tyr42, Tyr47, and Tyr49, whereas AMBER99SB force field causes much larger displacement of Tyr49 toward Tyr47. Therefore, PPC outperforms AMBER99SB force field in maintaining the structures of ASUs.

We use “Lattice RMSD” to measure the arrangement of protein monomers in the unit cell. Lattice RMSD is computed using the same formula as the ASU RMSD, except that these unit cells are superimposed using optimal quaternion alignments against the crystal lattice structure. Before measuring the Lattice RMSD, the unit cell in each snapshot was rescaled to the same size as the original one, and all atomic positions were rescaled accordingly. In contrast to ASU RMSDs, Lattice RMSDs ascend more quickly and have much larger amplitudes (see the lower panel in Figure 3). It indicates that in spite of the small deformations for the internal coordinates of ASUs, the relative position and orientation of ASUs in the lattice show larger deviations, especially under AMBER99SB force field. Besides, the difference between the Lattice RMSDs calculated for the backbone atoms and for all the heavy atoms is small relative to their absolute value as shown in Figure S1 of the Supporting Information, indicating translational and rotational motion of each ASU as a whole in the lattice.

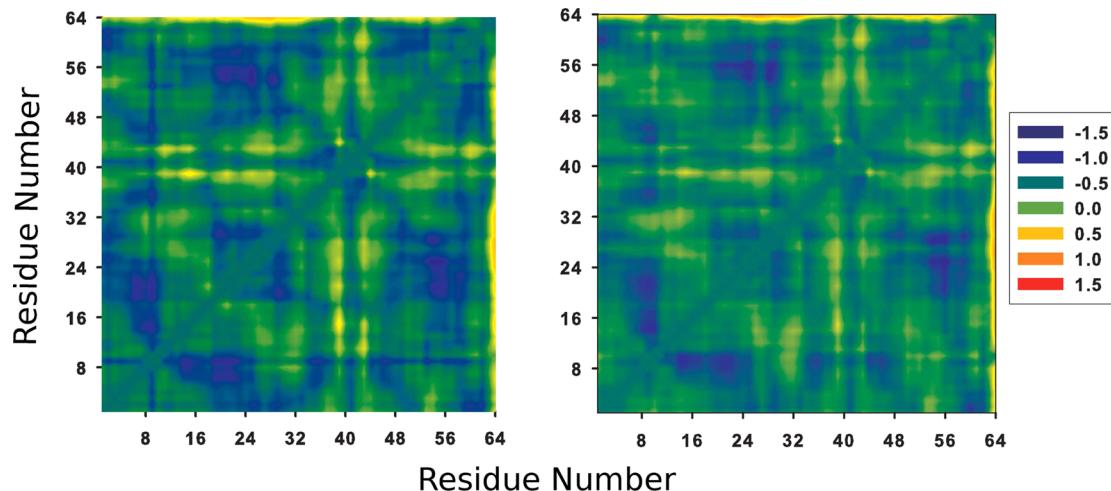


Figure 4. Deviations of the distance between α -carbon atoms from the structure over the last 150 ns simulations and the experimental structure, which is averaged over 108 distances. Same color scheme as in Cerutti's paper¹ was adopted.

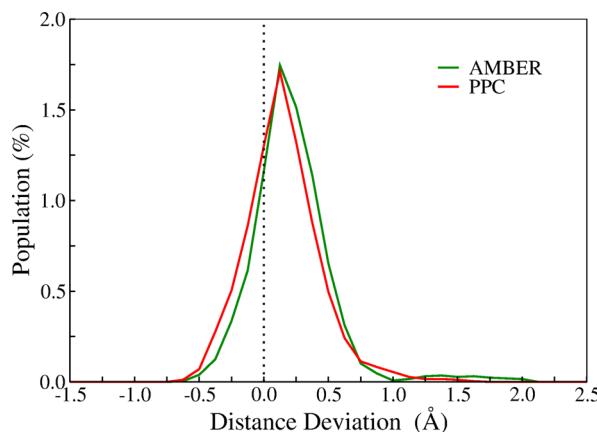


Figure 5. Simulated distribution of distance deviation for α -carbon atoms.

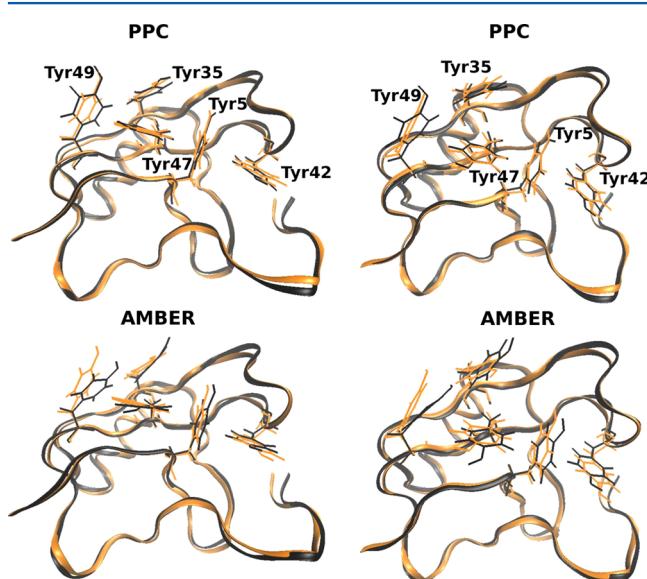


Figure 6. Distortion of the β -hairpin loop induces the loss of interaction in Tyr residues. The X-ray structure is shown in orange, and the last structures in the simulations are shown in black.

Crystallographic “B factor” is a good metric characterizing the thermal motion of atoms near its average position.³³ The atomic RMS fluctuation (F) can be computed using the formula

$$F = \sqrt{\frac{3}{8\pi^2}B} \quad (2)$$

As a crystal system, it is necessary to use ASU RMSF and Lattice RMSF to describe the movements of atoms within an ASU and that of proteins within a lattice. When computing ASU RMSF, each trajectory is broken into 108 monomer trajectories and these monomers are superimposed using optimal quaternion alignments against the average monomer structure. Lattice RMSF is computed in the similar way as ASU RMSF but considering each lattice as a whole. As shown in Figure 7, without the Lattice disorder the ASU RMSFs are very small under both force fields, indicating that all the protein monomers are rigid and stable. This observation is consistent with the small values of ASU RMSD. Peaks in the plot correspond to the random coil region and the chain terminus. However, the unit cells are not rigid. Lattice RMSFs under both force fields are larger than ASU RMSFs. They are in better agreement with the experimental value. Compared to AMBER99SB force field, PPC performs even better (the RMSD of α -carbons Lattice RMSF is 0.11 Å under PPC and 0.15 Å under AMBER99SB force field).

Thaw of Monomers in Supercell. The deformation of supercell dimensions and Lattice RMSDs described above show that monomers in supercell shift from their initial positions. In a supercell, each monomer has eight interfaces. The lattice deformations can also be reflected by the variations in the distances between the COMs of the monomers across each interface. Shown in Figure 8 are the average distances between the COMs of monomers along eight interfaces. Both increases and decreases in the distance are observed under both force fields, reflecting a nonuniform distortion of the lattice. It is quite clear that both simulations compress interface 3 and 7 but stretch interface 4 and 6. The variation of the distance between COMs of monomers may show that there are some gain and loss of interaction at these interfaces. Even so, the lattice has less distortion under PPC than under AMBER99SB force field.

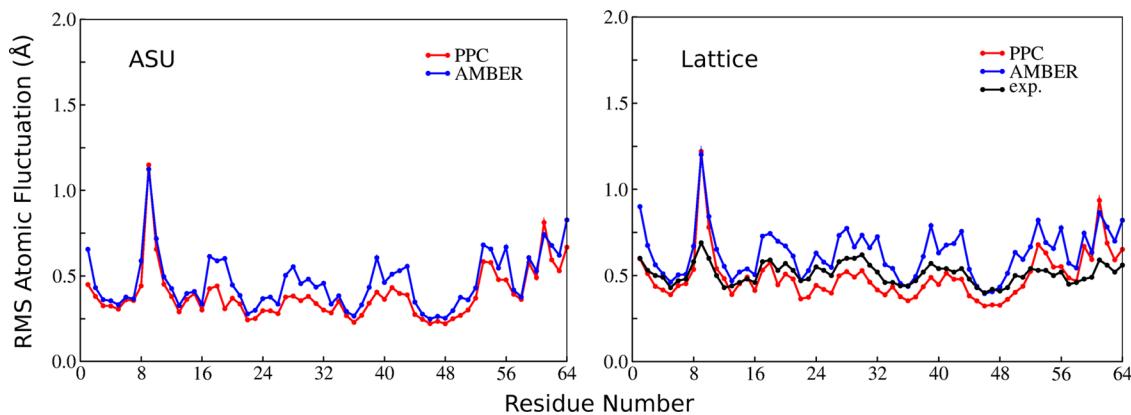


Figure 7. (Left) ASU RMSF and (right) Lattice RMSF for α -carbons over the last 150 ns simulation.

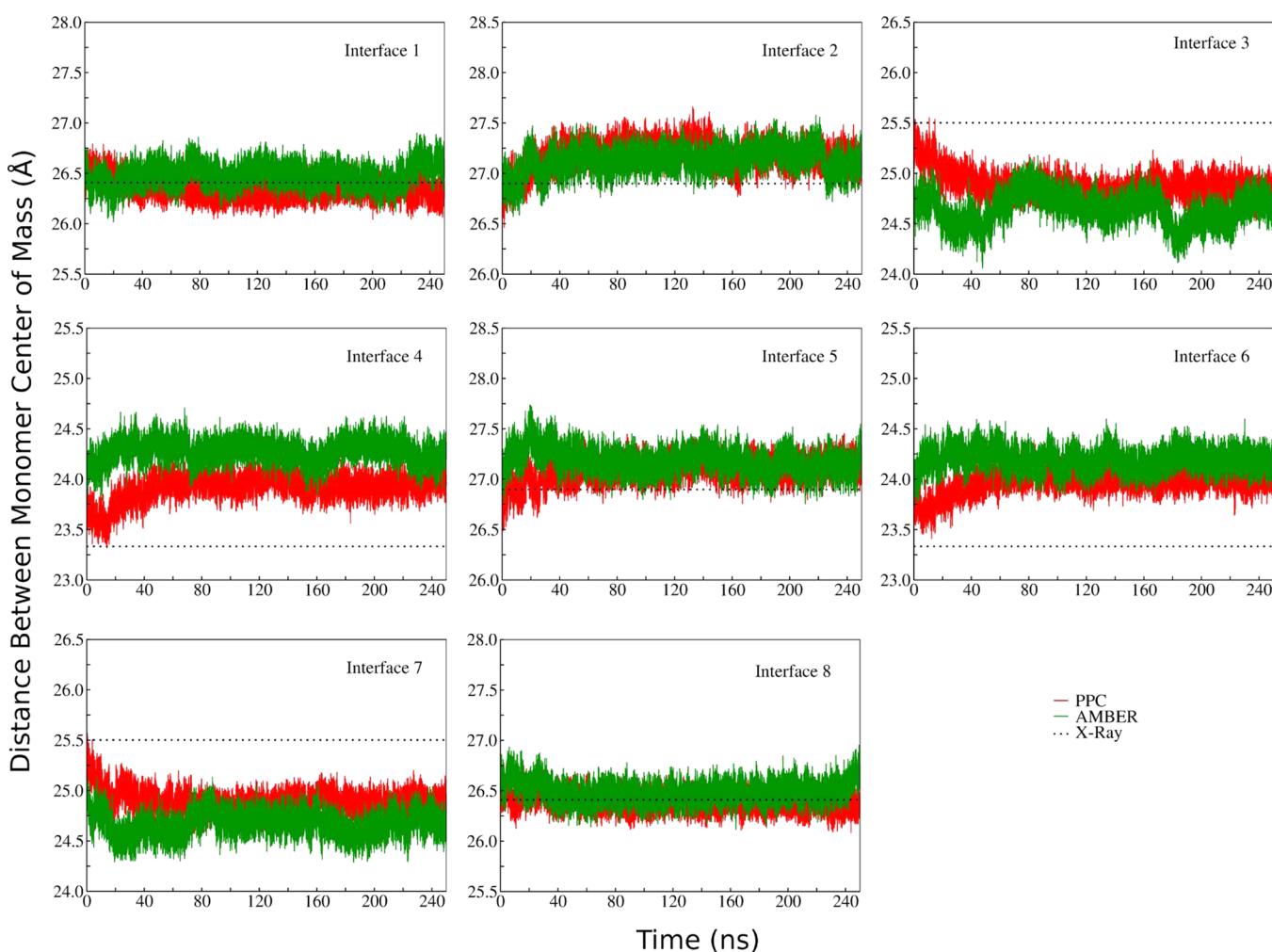


Figure 8. Deformation of the lattice measured by the average distances between protein monomer centers of mass across eight interfaces from simulations using AMBER99SB force field (green) and PPC (red). Experimental data are also indicated (black dot line).

The protein crystal is maintained by protein–protein interactions. In order to evaluate these interactions, the charge difference, electrostatic potential difference surface, and residue–residue contacts at each interface under PPC and AMBER99SB force field are analyzed. The charge differences presented in Figure S2 of the Supporting Information show that there are some obvious differences between PPC and AMBER99SB charge. These differences can also be presented by electrostatic potential difference surface in Figure S3 of the

Supporting Information. The largest deviations refer to the atoms at eight interfaces, which show the polarization effect can be better described by PPC for the interprotein interactions.

Residue–residue contacts are maintained by diverse favorable interactions, such as hydrogen bonds, salt bridges, and hydrophobic and cation– π interactions. All special interactions at eight interfaces in X-ray structure are presented in Table 2, by a distance criterion of 3.5 Å between the relevant heavy atoms for hydrogen bonds, 4.5 Å for salt bridges, 5.0 Å

Table 2. Interaction between Protein Monomers at Each Contact for X-ray Structure

contact type	interface	atoms in monomer 1	atoms in monomer 2
cation- π	1	Tyr35 ring	Arg62 CZ
cation- π	2	Lys28 CE	Trp38 ring
hydrophobic	3	His54 ring	Tyr47 ring
HB, backbone	3	His54 O	Tyr49 OH
HB, side chain	3	Asp53 OD2	Gln37 NE2
cation- π	4	His64 ring	Arg56 CZ
HB, backbone	4	His64 OXT	Thr57 N
HB, backbone	4	Arg62 NH1	Asp53 O
cation- π	5	Trp38 ring	Lys28 CE
cation- π	6	Arg56 CZ	His64 ring
HB, backbone	6	Asp53 O	Arg62 NH1
HB, backbone	6	Thr57 N	His64 OXT
salt bridge	6	Lys2 NZ	His64 OXT
hydrophobic	7	Tyr47 ring	His54 ring
HB, backbone	7	Tyr49 OH	His54 O
HB, side chain	7	Gln37 NE2	Asp53 OD2
cation- π	8	Arg62 CZ	Tyr35 ring

between centroids of rings for $\pi-\pi$ interaction and between cation and centroid of ring for cation- π interaction.^{28,29} The time evolution of these special hydrogen bond interactions during the simulations under PPC and AMBER99SB force field are shown in Figure S4 of the Supporting Information. It shows that two hydrogen bonds between the hydroxyl group in Try49 and the backbone atoms in His54 across interface 3 and 7 under both force fields are well maintained as in the X-ray structure. Two hydrogen bonds between the carbonyl group in His64 and the backbone of Thr57 and two hydrogen bonds between backbone in Asp53 and the side chain of Arg62 across interface 4 and 6 are perfectly maintained under PPC but fluctuate from the X-ray structure under AMBER99SB force field. Hydrogen bonds between the side chains of Asp53 and Gln37 across interface 3 and 7 are completely lost under both force fields. The salt bridge between Lys2 and His64 across interface 6 is broken at the beginning but reformed in the final stage under both force fields. The cation- π interaction between Trp38 and Lys28 across interface 2 and 5 under both force fields are overestimated as shown in Figure S5 in the Supporting Information. The cation- π interactions between Tyr35 and Arg62 across interface 1 and 8 and the hydrophobic interactions between His54 and Tyr47 across interface 3 and 7 are well-maintained under both force fields. However, neither of these two force fields can reproduce the cation- π interaction between His64 and Arg56 across interface 4 and 6.

The specific contacts in the last frame during both simulation are shown in Figures 9 and 10. PPC can maintain most of these contacts except the hydrogen bond between Gln37 and Asp53. More contacts are lost under AMBER99SB force field for some monomers or even all the monomers. In order to understand the gain or loss of these interactions, four kinds of interactions at each interface for the last frame under PPC and AMBER99SB force field are also analyzed. As mentioned above, the β -hairpin is twisted under AMBER99SB force field, which causes some alternations of interactions for residues from Ser33 to Lys50 with other residues across the interfaces. The alternations of inter-residue interactions at each interface are shown in Figures S6 and S7 of the Supporting Information, and the conformations are shown in Figure S8 to Figure S11 of the Supporting Information. It can be indicated from these figures

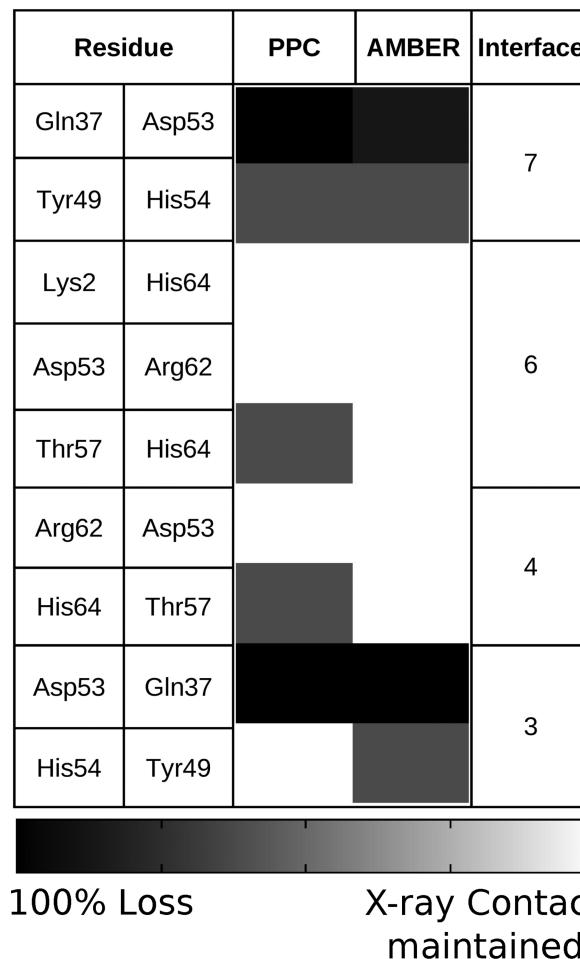


Figure 9. Evolutions of hydrogen bond interaction and salt bridge identified in the 1AHO crystal structure for the last frame in the 250 ns simulations under PPC and AMBER99SB force field. Black square indicates that a contact is completely lost, while white square indicates that a contact is well maintained. The residues involved in each contact are listed.

that some new hydrogen bonds, hydrophobic interactions, and cation- π interaction formed, accompanied by the loss of some original interactions found in the X-ray structure. Only a few of these new interactions exist at the end of each simulation and others show heterogeneities. However, PPC can maintain nearly all the original contacts with fewer new interactions formed.

Since residues are very flexible at the N-terminus, many interactions about Arg62 and His64 are changed under both force fields, especially AMBER99SB. Across interface 1 and interface 8, excess flexibility of Tyr49 under AMBER99SB force field (shown in Figure 4) weakens the hydrogen bond with Arg62 and introduces new interaction with His64. All the cation- π interactions across interface 2 and 5 under PPC and AMBER99SB force fields are maintained. The side chain of Arg18 rotates toward Asp8, forming a salt bridge between them. Another salt bridge is formed between His64 and Lys2 across interface 4 under both force fields. Across interface 6, the side chain of Arg56 and His64 flip together under both force fields. Under AMBER99SB force field, the cation- π interaction between Arg56 and His64 is completely broken. The cation- π interactions under PPC is not completely broken, which weaken the hydrogen bond between His64 and Thr57. The

Residue		PPC	AMBER	Interface
Arg62	Tyr35			8
Tyr47	His54			7
Arg56	His64			6
Trp38	Lys28			5
His64	Arg56			4
His54	Tyr47			3
Lys28	Trp38			2
Tyr35	Arg62			1

Figure 10. Evolutions of hydrophobic and cation–π interaction identified in the 1AHO crystal structure for the last frame in the 250 ns simulation under PPC and AMBER99SB force field. The same color scheme as in Figure 9 is adopted.

interactions between the side chains of Asp53 and Gln37 across interfaces 3 and 7 are lost under both force fields due to the reorientation of side chain of Asp53. There are many alternations of hydrogen bonds, hydrophobic interactions, and cation–π interaction under AMBER99SB force field across interfaces 3 and 7. In accordance with Figure 7, Arg56 has a higher RMSF under AMBER99SB, which implies that this residue is more flexible. The side chain of Arg56 in one monomer reorients to another monomer and forms hydrogen bonds with Ser33, Glu32, and Asp3 and cation–π interaction with the benzene ring of Tyr49, none of which should be there. Under both force fields, Asp9 across interfaces 3 and 7 is very flexible with a large RMSF and forms a new salt bridge with Lys50.

CONCLUSION

In this work, we have studied the crystal of the toxin protein II from the scorpion *Androctonus australis* Hector using molecular dynamics simulation under PPC and AMBER99SB force field. The inclusion of the specific polarization effect in the atomic charge makes it more suitable for the simulation of toxin II protein crystal. Observations in these two parallel simulations can be summarized as following. (1) ASUs are stable under PPC and AMBER99SB force field with low ASU backbone RMSD. Because PPC can give more specific intraprotein interaction through Coulomb interaction such as hydrogen bonds, the ASU structures and lattice thermal fluctuations under PPC are more in line with the experiment. (2) Although lattice deformations measured by the changes of the distances between COMs of monomers can be observed under both force fields, the distortion under PPC is smaller than that under AMBER99SB force field. (3) 10 out of 17 of the hydrogen

bonds, salt bridges, cation–π, and hydrophobic interactions in X-ray structure are maintained under PPC, while nine of them can be preserved under AMBER99SB force field. (4) The deformation of the lattice is accompanied by the breaking of some original interactions and the formation of some new interactions. Under AMBER99SB force field, most of the polar residues, for example Asp, Lys, Arg, His, and Tyr, cannot be modeled well on account of the missing of the polarization effect. Some new interactions across the monomer interfaces appear to saturate the dangling side chains.

It can be inferred from this study that the implementation of electrostatic polarization effect is important for the simulation of toxin II protein crystal. However, energy terms in force fields are highly correlated. There is no clear decomposition scheme of potential energy into various bonded (bond, angle, dihedral) and nonbonded (Coulomb, van der Waals) interactions. Merely improving Coulomb interaction energy cannot always guarantee the improvement of reliability. Corresponding refitting of the van der Waals is not unnecessary, which has also been suggested by Kramer et al.³⁴ Further refinement of other parameters for the force field will be carried out in the future. Atomic charges were fixed in the whole simulation. However, conformation change has been observed. Atomic charges should be updated correspondingly. On-the-fly charge fitting is possible,²¹ but this may significantly increase the computational expense.

THEORETICAL METHODS

Nomenclature. (1) Crystallographic asymmetric unit (ASU): a protein monomer with 64 residues. (2) Lattice: a unit cell made up with four ASUs. (3) Supercell: a $3 \times 3 \times 3$ unit-cell stack containing 108 ASUs.

Construction of the Simulation Cell. The crystal structure of PDB entry 1AHO²⁸ was used to construct the unit cell by applying the symmetry operations of the $P2_12_12_1$ space group. The cell dimension is $45.9 \times 40.7 \times 30.1 \text{ \AA}^3$. The crystallographic asymmetric unit (ASU) is stabilized by four disulfide bridges between cysteine residues and T-stacking interactions between aromatic rings of tyrosine residues. Each unit cell consists of four symmetry-related monomers. Two monomers form a “contact” when their heavy atoms separated from each other by a distance smaller than 5 Å. For this orthorhombic $P2_12_12_1$ space group, each monomer has eight interfaces.

The preprocessing of the protein structure to generate the supercell is similar to that in the study of Cerutti.¹ The initial coordinates were downloaded from PDB (entry 1AHO). For simplicity, only the first position was selected for each atom with multiple locations. A few atoms, such as the distal atoms in the side chain of Lys30 and Lys50 as well as part of the carboxyl group in Asp9, were missing in the original crystallographic structure. These atoms were reconstructed with the tLEaP module of AMBER11 package.³⁵ The reason to choose the 1AHO structure over 1PTX was that a larger number of water molecules in the 1AHO structure had been refined by the SnB program.³² The final structure of 1AHO consists of 925 protein atoms and 129 water molecules, among which 43 water molecules are only partially occupied. Of the 106 water molecules introduced in the 1PTX structure during refinement, 97 occupy single site and others have ambiguous locations. Although all of these crystallographic water molecules were included in constructing the initial structure for the simulation, numerous water molecules were still missing. Therefore, extra

water molecules are needed to completely hydrate the unit cell. However, this might induce many steric clashes among atoms. To relieve this tension, a few steps of energy minimization process and some picoseconds of dynamics were carried out after adding water molecules.

Periodic boundary condition is a valid approximation for simulation of crystal lattices. However, the limitation from the rigid lattice and symmetric long-range ordering may introduce some artifacts of the periodic model.¹ In this work, a supercell was constructed as the minimal periodic element with 27 individual unit cells in a $3 \times 3 \times 3$ arrangement to suppress any artifacts from the periodicity of our model. The central unit cell was generated with UnitCell module of the AMBER11 package. Seven acetate ions, three ammonium ions, and a rough number of water molecules were added sequentially using "AddToBox" module to mimic the experimental buffer conditions. The new added ions and water molecules were placed at least 2.8 Å from a solute atom and a pre-existing solvent atom or ion. These acetate and ammonium ions were modeled with the general AMBER force field (GAFF).³⁶ TIP3P water model⁷ was used to parametrize the solvent. A three-step energy minimization was carried out to relax this unit cell. In the first step, the buffer ions and water molecules were energy minimized while proteins heavy atoms positions were tightly fixed with a 1000 kcal/(mol Å²) restraint applied. Then, the protein components were energy minimized while solvent positions were tightly restrained to their new positions. Finally, all the components were fully optimized with no restraints applied. After this three-step energy minimization, "PropPDB" module of AMBER11 package was used to replicate this unit cell to produce the supercell. A same three-step energy minimization, a 25 ps heating to 287 K with a 1000 kcal/(mol Å²) restraint applied on all protein atoms, several-step restrained equilibration and a 10 ns NTP MD simulation were performed to relax the whole system. Equilibration approaches were carried out in the same way as in Molecular Dynamics Simulations below, but the simulation time was only 1.39 ns. If the volume of the system was smaller than the experimental measurement, more water molecules were added into the lattice. These procedures were iterated until the volume of supercell reached an equilibrium value with an error no larger than 0.3% of the experimental value. The simulation supercell was finally determined with 108 protein monomers, 189 acetate ions, 81 ammonium ions, and 19953 water molecules under AMBER99SB force field and 20628 water molecules under PPC. The supercell dimensions are 137.7 × 122.1 × 90.3 Å³.

PPC Charge Calculation. Recently, a new charge scheme, termed delta restrained electrostatic potential (dRESP) charge, has been developed,³⁷ which was employed here to fit the atomic charges to the electrostatic potential for each residue. This scheme has been proved numerically more stable than the original RESP charge fitting method.³⁸ The electronic structures of the residues were calculated one-by-one employing the molecular fractionation with conjugate caps approach.¹⁷ The supercell is considered as a central unit cell plus 26 duplicates with the same charge distribution. Therefore, we only fitted the atomic charges of the proteins in the central unit cell (four ASUs not just one) and then copied the charges to other image cells. During the quantum mechanical calculation of the electronic structure for each residue, other residues (also including those in the 26 image unit cells), buffer ions, and water molecules were taken as background charges to generate an electric field. After five iterations, the atomic charges

converged [the root-mean-square deviation (RMSD) of the atomic charges between the last two cycles was below $10^{-4}\epsilon$]. Finally, the atomic charges were averaged among the four ASUs in each lattice to remove the accidental difference. The PPCs were used in the subsequent MD simulation and remained unchanged. All quantum mechanical calculations were performed at the B3LYP/6-31G* level with Gaussian 09.³⁹ However, the polarization effect from the atomic charges outside this $3 \times 3 \times 3$ supercell has been truncated during the PPC fitting. A more rigorous scheme with periodic boundary conditions can be employed to remove this truncation error.⁴⁰

Molecular Dynamics Simulations. The initial structure for the MD simulation was obtained as described in Construction of the Simulation Cell. Two MD simulations under AMBER99SB force field and PPC were performed with AMBER11.³⁵ When performing simulation under PPC, the standard charges of AMBER99SB force field were simply replaced by PPCs while the rest of the AMBER parameters were kept intact. In MD simulation, the van der Waals interaction was truncated at 9.0 Å, and the Coulomb interaction was calculated by the particle mesh Ewald method with a 9.0 Å cutoff in real space. To ensure the stability of the entire system before the production dynamics, the minimization, the heating, and the restrained equilibration were carried out in the same way as in the work of Cerutti et al.^{1,41,42} A three-step minimization was also performed to relax the whole system in a mild way as described above. The system was heated up to 287 K for 500 ps with a 1000 kcal/(mol Å²) restraint applied on all protein atoms. The restraint was then reduced to 900 kcal/(mol Å²) in a 500 ps simulation at the same temperature. In order to allow a comprehensive sampling in the phase space for the buffer ions and water molecules in the supercell, a 10 ns simulation at 500 K was conducted with a 800 kcal/(mol Å²) restraint applied on all protein atoms, followed by a 40 ns simulation at the same temperature with the restraint reduced from 700 kcal/(mol Å²) to 400 kcal/(mol Å²). With the temperature reduced to 287 K and the restraint kept unchanged, a 5 ns simulation was conducted. Restrained equilibration at 287 K was extended for 7.5 ns with the restraint reduced from 300 kcal/(mol Å²) to 0 kcal/(mol Å²). In all these steps, a constant volume constraint was applied. An isothermal-isobaric production dynamics at 287 K was performed for 250 ns with a 2 fs time step. This system temperature was maintained using Andersen temperature coupling scheme, in which the velocities were randomized to a distribution corresponding to 287 K every 2 ps. Volume rescaling was anisotropic depending on all three elements of the virial tensor perpendicular to each face of supercell.

ASSOCIATED CONTENT

Supporting Information

Backbone and heavy atoms ASU RMSD and Lattice RMSD, charge difference, electrostatic potential difference surface, variation in distance between the relevant heavy atoms or ring for hydrogen bonds, salt bridge, π–π and cation–π, alternations and conformations of each kind of interacting contacts. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: ymei@phy.ecnu.edu.cn.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China (Grants 10974054, 20933002, and 21173082) and the Shanghai PuJiang Program (09PJ1404000). We thank Supercomputer Center of East China Normal University for CPU time support.

REFERENCES

- (1) Cerutti, D. S.; Freddolino, P. L.; Duke, R. E.; Case, D. A. Simulations of a Protein Crystal with a High Resolution X-ray Structure: Evaluation of Force Fields and Water Models. *J. Phys. Chem. B* **2010**, *114*, 12811–12824.
- (2) 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.
- (3) 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.
- (4) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- (5) Mackerell, A. D., Jr.; Feig, M.; Brooks, C. L., III Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. *J. Comput. Chem.* **2004**, *25*, 1400–1415.
- (6) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; 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.
- (7) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (8) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
- (9) Price, D. J.; Brooks, C. L., III A Modified TIP3P Water Potential for Simulation With Ewald Summation. *J. Chem. Phys.* **2004**, *121*, 10096–10103.
- (10) Abascal, J. L. F.; Vega, C. A General Purpose Model for the Condensed Phases of Water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, 234505.
- (11) Canongia Lopes, J. N.; Pádua, A. A. H. CL&P: A Generic and Systematic Force Field for Ionic Liquids Modeling. *Theor. Chem. Acc.* **2012**, *131*, 1129.
- (12) Carignano, M. A. Structure and Dynamics of $[PF_6]^-[P_{1,2,2,4}]$ from Molecular Dynamics Simulations. *J. Phys. Chem. B* **2013**, *117*, 15176–15183.
- (13) Svärd, M.; Rasmussen, A. C. Force Fields and Point Charges for Crystal Structure Modeling. *Ind. Eng. Chem. Res.* **2009**, *48*, 2899–2912.
- (14) Schnieders, M. J.; Fenn, T. D.; Pande, V. S. Polarizable Atomic Multipole X-Ray Refinement: Particle Mesh Ewald Electrostatics for Macromolecular Crystals. *J. Chem. Theory Comput.* **2011**, *7*, 1141–1156.
- (15) 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.; et al. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114*, 2549–2564.
- (16) Schnieders, M. J.; Baltrusaitis, J.; Shi, Y.; Chattree, G.; Zheng, L.; Yang, W.; Ren, P. The Structure, Thermodynamics, and Solubility of Organic Crystals from Simulation with a Polarizable Force Field. *J. Chem. Theory Comput.* **2012**, *8*, 1721–1736.
- (17) Ji, C.; Mei, Y.; Zhang, J. Z. H. Developing Polarized Protein-Specific Charges for Protein Dynamics: MD Free Energy Calculation of pK_a Shifts for Asp²⁶/Asp²⁰ in Thioredoxin. *Biophys. J.* **2008**, *95*, 1080–1088.
- (18) Mei, Y.; Li, Y. L.; Zeng, J.; Zhang, J. Z. H. Electrostatic Polarization Is Critical for the Strong Binding in Streptavidin-Biotin System. *J. Comput. Chem.* **2012**, *33*, 1374–1382.
- (19) Ji, C. G.; Zhang, J. Z. H. Effect of Interprotein Polarization on Protein-Protein Binding Energy. *J. Comput. Chem.* **2012**, *33*, 1416–1420.
- (20) Wei, C.; Tung, D.; Yip, Y. M.; Mei, Y.; Zhang, D. Communication: The Electrostatic Polarization Is Essential to Differentiate the Helical Propensity in Polyalanine Mutants. *J. Chem. Phys.* **2011**, *134*, 171101.
- (21) Duan, L. L.; Mei, Y.; Zhang, D.; Zhang, Q. G.; Zhang, J. Z. H. Folding of a Helix at Room Temperature Is Critically Aided by Electrostatic Polarization of Intraprotein Hydrogen Bonds. *J. Am. Chem. Soc.* **2010**, *132*, 11159–11164.
- (22) Jia, X.; Zeng, J.; Zhang, J. Z. H.; Mei, Y. Accessing the Applicability of Polarized Protein-Specific Charge in Linear Interaction Energy Analysis. *J. Comput. Chem.* **2014**, *35*, 737–747.
- (23) Mei, Y.; Ji, C.; Zhang, J. Z. H. A New Quantum Method for Electrostatic Solvation Energy of Protein. *J. Chem. Phys.* **2006**, *125*, 094906.
- (24) Mei, Y.; Zhang, D. W.; Zhang, J. Z. H. New Method for Direct Linear-Scaling Calculation of Electron Density of Proteins. *J. Phys. Chem. A* **2005**, *109*, 2–5.
- (25) Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Honig, B.; Ringnalda, M.; Goddard, W. A., III Accurate First Principles Calculation of Molecular Charge Distributions and Solvation Energies from Ab Initio Quantum Mechanics and Continuum Dielectric Theory. *J. Am. Chem. Soc.* **1994**, *116*, 11875–1188.
- (26) Honig, B.; Sharp, K.; Yang, A. S. Macroscopic Models of Aqueous Solutions: Biological and Chemical Applications. *J. Phys. Chem.* **1993**, *97*, 1101–1109.
- (27) Mei, Y.; Wei, C. Y.; Yip, Y. M.; Ho, C. Y.; Zhang, J. Z. H.; Zhang, D. W. Folding and Thermodynamic Studies of Trp-cage Based on Polarized Force Field. *Theor. Chem. Acc.* **2012**, *131*, 1168.
- (28) Smith, G. D.; Blessing, R. H.; Ealick, S. E.; Fontecilla-Camps, J. C.; Hauptman, H. A.; Housset, D.; Langs, D. A.; Miller, R. Ab Initio Structure Determination and Refinement of a Scorpion Protein Toxin. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **1997**, *53*, 551–557.
- (29) Housset, D.; Habersetzer-Rochat, C.; Astier, J. P.; Fontecilla-Camps, J. C. Crystal Structure of Toxin II From the Scorpion Androctonus Australis Hector Refined at 1.3 Å Resolution. *J. Mol. Biol.* **1994**, *238*, 88–103.
- (30) Fontecilla-Camps, J. C.; Habersetzer-Rochat, C.; Rochat, H. Orthorhombic Crystals and Three-Dimensional Structure of the Potent Toxin II From the Scorpion Androctonus Australis Hector. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 7443–7447.
- (31) Janowski, P. A.; Cerutti, D. S.; Holton, J.; Case, D. A. Peptide Crystal Simulations Reveal Hidden Dynamics. *J. Am. Chem. Soc.* **2013**, *135*, 7938–7948.
- (32) Miller, R.; Gallo, S. M.; Khalak, H. G.; Weeks, C. M. SnB: Crystal Structure Determination Via Shake-and-Bake. *J. Appl. Crystallogr.* **1994**, *27*, 613–621.
- (33) Allen, T. W.; Andersen, O. S.; Roux, B. On the Importance of Atomic Fluctuations, Protein Flexibility, and Solvent in Ion Permeation. *J. Gen. Physiol.* **2004**, *124*, 679–690.
- (34) Kramer, C.; Gedeck, P.; Meuwly, M. Multipole-Based Force Fields from ab Initio Interaction Energies and the Need for Jointly Refitting All Intermolecular Parameters. *J. Chem. Theory Comput.* **2013**, *9*, 1499–1511.
- (35) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R.; Zhang, W.; Merz, K. M. et al. AMBER 11, University of California: San Francisco, 2010.

- (36) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
- (37) Zeng, J.; Duan, L.; Zhang, J. Z. H.; Mei, Y. A Numerically Stable Restrained Electrostatic Potential Charge Fitting Method. *J. Comput. Chem.* **2013**, *34*, 847–853.
- (38) 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.
- (39) 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*, revision B.01, Gaussian, Inc.: Wallingford, CT, 2010.
- (40) Holden, Z. C.; Richard, R. M.; Herbert, J. M. Periodic Boundary Conditions for QM/MM Calculations: Ewald Summation for Extended Gaussian Basis Sets. *J. Chem. Phys.* **2013**, *139*, 244108.
- (41) Cerutti, D. S.; Trong, I. L.; Stenkamp, R. E.; Lybrand, T. P. Dynamics of the Streptavidin-Biotin Complex in Solution and in Its Crystal Lattice: Distinct Behavior Revealed by Molecular Simulations. *J. Phys. Chem. B* **2009**, *113*, 6971–6985.
- (42) Cerutti, D. S.; Le Trong, I.; Stenkamp, R. E.; Lybrand, T. P. Simulations of a Protein Crystal: Explicit Treatment of Crystallization Conditions Links Theory and Experiment in the Streptavidin-Biotin Complex. *Biochemistry* **2008**, *47*, 12065–12077.