Chemical Physics Letters 706 (2018) 303–307

Contents lists available at ScienceDirect

Chemical Physics Letters

journal homepage: www.elsevier.com/locate/cplett

Research paper

Polarization effect within a protein crystal: A molecular dynamics simulation study

Tingting Zhu [a], Chao Wu [a], Jianing Song [b], Jeffrey R. Reimers [a,c], Yongle Li [a,*]

[a] Department of Physics, International Center for Quantum and Molecular Structures, and Shanghai Key Laboratory of High Temperature Superconductors, Shanghai University, 99 Shangda Road, Shanghai 200444, China
[b] NYU-ECNU Center for Computational Chemistry, NYU Shanghai, Shanghai, China
[c] School of Mathematical and Physical Sciences, The University of Technology Sydney, Sydney, NSW 2007, Australia

ARTICLE INFO

Article history:
Received 25 May 2018
In final form 11 June 2018
Available online 15 June 2018

Keywords:
Protein crystal
MD simulations
Polarized protein-specific charge

ABSTRACT

Systematical molecular dynamics with different force fields are performed to simulate the structure and dynamics of crystal of hen egg-white lysozyme, including AMBER and three versions of protein-specific charge, PPC. The electrostatic polarization within the crystal is studied with the comparison among four 250 ns trajectories under them. Results show that under appropriate parameterized PPC, the protein can be stable during simulations, indicated by both smaller root-mean-square deviation and closer crystallographic B-factors to the experimental values. This work also shows how the selection of dielectric constant affects the results of utilizing PPC.

1. Introduction

X-ray crystallography is the major method to resolve the structure of proteins [1]. Although it needs crystals of protein, which may be different from the natural state of it, its ability to achieve very high resolution is indispensible. In the protein database, most resolved structures obtained using this method.

MD simulation is a powerful tool to reveal both the structure and dynamics of molecules in the condensed phase [2–4]. In recent years MD has attained extraordinarily success, and with state-of-the-art super computers, it can simulate huge systems and with the time scale of several milliseconds. Recently MD simulation began to be applied to simulating protein crystals [5,6].

Different from the MD simulations in water, MD simulation of crystals leads application to highly inhomogeneous systems in which the proteins are crowded together within crystal lattice. Since the parameters of proteins are obtained from the calculations which use proteins solvated in water, the MD simulation of protein crystals remains challenging [7]. Only in recent several years have there been some such successful MD simulations [7–10]. But as mentioned above, the force field parameters may not perfect for simulating crystalized phase, as the results still deviate from experimental data. Previously, only a polarizable force field, AMOEBA, succeeded for the MD simulation of protein crystals [11]. But the time-consuming feature hindered its widespread usage. The unsuccessfully MD simulation also partly stems from the problem of combination between polarizable or polarized parameters of protein with un-polarized parameters of water [12]. In a recent work, the polarized force field, known as the polarized protein-specific charge (PPC) [13], succeeded in predicting correct structures and dynamic behaviors in solution phase, yet suffered a similar deviation from experiment data [9].

From published works, the non-negligible deviations from the experiments include the distribution of water on the surface of proteins, and missing hydrogen bonds (H-bonds). From recent work, some major H-bonds within crystalized protein can be destroyed during MD simulations even with the polarized force field PPC, in which the electrostatic polarization was considered to be good in liquid phase simulations [9]. It was proposed that the deviation stemmed from the insufficiency of simulating non-electrostatic parts of force field existed, such as van der Waals (vdW) effect without further validation. However, considering that vdW affects the system only in large space scale, while the backbone H-bond of the protein in a crystal is a local property, the parameters related to electrostatic effects such as partial charge must play non-trivial role. So in this work, we systematically investigated the parameters generated by the PPC, by adjusting the major input parameter, the dielectric constant, ε, which affects the property of the solvent simulated during quantum chemical calculations in preparing the PPC. We hope our investigation can help to reveal the effect of the electrostatic interaction in crystals,

* Corresponding author.
E-mail address: yongleli@shu.edu.cn (Y. Li).

https://doi.org/10.1016/j.cplett.2018.06.018
0009-2614/

and help the preparation of the PPC for the simulation of protein crystals.

The hen egg-white lysozyme (HEWL) is one of the most thoroughly investigated proteins [14,15], and one of its X-ray crystallographic structure is with ultra-high resolution of 0.65 Å [16]. Recently, J. Reimers et al. investigated its crystal refinement by quantum chemistry [17]. Such comprehensive structural information provided us with an ideal testing system. Thus in this work, we have benchmarked MD simulation using this highest resolution of HEWL crystal and evaluated the simulated polarization effect with different values of ε used in the PPC.

2. Methods

2.1. Construction of the model system

The system is based on the experimental X-ray diffraction structure, with PDB entry 2VB1, and with space group P1 (shown in Fig. 1). The cell constants are $27.07 \times 31.25 \times 33.76$ Å^3, and the three angles: $87.98 \times 108.00 \times 112.11°$. There are four disulfide bonds between cysteine residues stabilize the crystallographic asymmetric unit (ASU). In each unit cell, there is only one single-chain protein. Proteins between different cells are separated by unstructured water molecules, which cannot be recognized experimentally.

For preprocessing the protein structure we adopted the method proposed by Cerutti [6], so only a brief summary is provided here, and further details can be found in previous works of MD simulations of protein crystals. From the downloaded experimental structure (entry 2VB1), we retained all the structural solvent molecules, including structural water, which are within 3 Å from protein, acetate acid anion (ACT), 1,2-ethanediol (EDO), and nitrate anion (NO3). We used the first position for each atom with multiple locations for simplicity. Missing atoms were added by the tLeap module in the AMBER16 [18] package. There were 1001 heavy protein atoms, 151 oxygen atoms belonging to water, 7 atoms in 1 ACT molecule, 30 atoms in 3 EDO molecules, and 36 atoms in 9 NO3 molecules.

To use periodic boundary condition (PBC) and reduce the error induced by the rigid lattice and symmetric long-range ordering, we constructed a super cell in a $3 \times 3 \times 3$ arrangement to avoid the deficiency of using PBC in our MD simulation. The central unit cell was created by the UnitCell module in the AMBER16 package. After constructing the super cell, we added extra water molecules using another module, AddToBox, to simulate the condensation occurring in experimentally.

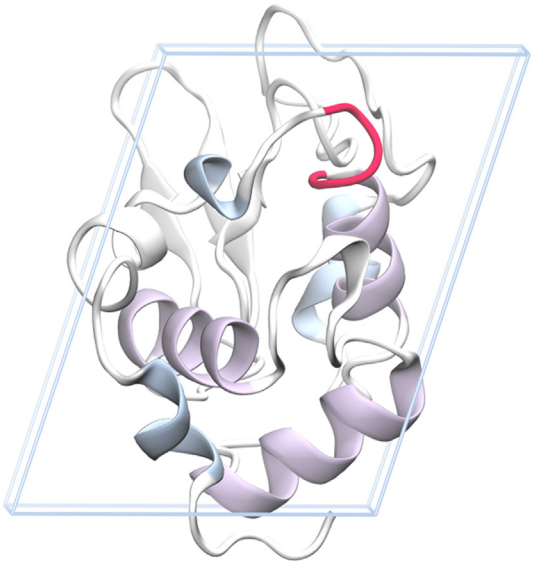

Fig. 1. The single cell model of the HEWL crystal used in this work. The loop with the largest B-factor is shown in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

For the protein we used the popular AMBER99SB force field [19] as the control group. For the PPC, we adopted the methodology in the work of Mei et al.,[9] but used 3 different values of ε for comparison: the default value $\varepsilon = 80$ denoted as PPC1, $\varepsilon = 10$ as PPC2, and $\varepsilon = 20$ as PPC3. The consideration is the protein environment is considered with approximate ε with value 10. On the other hand, since $\varepsilon = 20$ lies in between the two extreme cases of pure water and pure protein, this choice may mimic the protein crystal with water molecules.

The force field parameters for solvent molecules other than water were taken from the generalized AMBER force field (GAFF) [20], and SPC/E [21] was used for water. After optimization of the single cell, the PropPDB module in the AMBER16 package was used for constructing the super cell. The final model contained 52,920 atoms belonging to 27 single-chain proteins, 27 ACT, 81 EDO and 243 NO3 molecules.

To obtaining a proper number of water, under each different force field we benchmarked several different numbers of water molecules, each with a 100 ns MD simulation. The optimum number of water was defined to be that which gave the lowest deviation in volume of the system compared with the experimental value. Then the systems with the optimum number of water were used for production runs of 250 ns. Finally, we obtained four 250 ns trajectories for further analysis, with the optimum number of water molecules as shown in Table 1.

2.2. Molecular dynamics simulations

First we relaxed the entire model by 1000 steps of steepest-descent followed by another 1000 steps of conjugate gradient to minimize the water molecules while experimentally resolved protein atoms were fixed by a force constant 1000 kcal·mol^{-1}·Å^{-2}). Then the system was then heated to 800 K within 200 ps and kept at equilibrium for 10 ns, to fully relax the water molecules. After this stage, the temperature was reduced to 292 K within 1 ns to mimic its growth temperature. Then we performed an 18 ns simulation at the same temperature with the restraint reduced stepwise from 1000 down to 10 kcal·mol^{-1}·Å^{-2}, with 2 ns interval for each step change of the value of the force constant: 1000, 800, 600, 400, 200, 100, 50, 20, and 10 kcal·mol^{-1}·Å^{-2}. Throughout all these steps, the systems were maintained in NVT ensemble, and the PBC was applied. Then 100 ns production runs in NPT ensemble without any restraint were taken for the systems with different number of water. After the optimum number of water for systems under

Table 1
Comparison of the optimum number of water molecules and number of H-bonds from MD simulations under different force fields. The experimental value is also shown. The different versions of the polarized protein-specific charge, PPC1-3, are explained in the text.

Method	N_{WAT}[a]	ε[b]	$N_{H\text{-}bond}$[c]
AMBER	6453	N/A	54 ± 1
PPC1	6372	80	50 ± 1
PPC2	6642	10	75 ± 1
PPC3	6615	20	75 ± 1
Expt.	N/A	N/A	79

[a] Optimum number of water molecules giving smallest relative error of crystal volume to the experimental value.
[b] Dielectric constant used for obtaining PPC.
[c] Number of H-bonds during MD simulations.