BL-5A, AR-NW12A/2018G613

Anisotropic Distribution of Ammonium Sulfate Ions in Protein Crystallization

Tyuji Hoshino[1,*] Mariko Kitahara[1] Satoshi Fudo[1] Tomoki Yoneda[1] and Michiyoshi Nukaga[2]
[1] Graduate School of Pharmaceutical Sciences, Chiba University,
1-8-1 Inohana, Chuo-ku, Chiba 260-8675, Japan
[2] Faculty of Pharmaceutical Sciences, Josai International University,
Gumyo 1, Togane-shi Chiba 283-8555, Japan

1 Introduction

X-ray crystal analysis is one of the most popular approaches to determine the protein structure. The growth of single crystal is a critical step in the X-ray crystallography. Ammonium sulfate (AS) is one of the most popular precipitants to induce crystallization of protein molecules. AS is, however, not always effective for all kinds of proteins in crystal growth. Namely, some proteins are readily crystallized by utilizing AS as a precipitant, but other proteins are not. The reason for this difference has been discussed in literature, but isn't still well-understood. Like AS, polyethylene glycols (PEGs) are frequently utilized in protein crystallization. Crystal growth of some kinds of proteins is effectively induced by PEGs, instead of AS. The cause for the dependency of protein crystallization on the precipitants has been barely studied in the atomistic level so far.

In our previous study, crystals of a single kind of protein were grown with three different precipitants, including AS and PEG. Molecular dynamics (MD) simulations were carried out for the proteins in the presence of precipitants. The simulations showed that ammonium and sulfate ions (AS ions) played roles not only of decreasing the protein solubility but also of restricting the contact sites on the protein surface to other protein molecules, which was reflected in the molecular packing of the crystal structure. Since molecular packing of protein crystals is highly linked to space groups, the relationship between space groups and precipitants was also investigated for several kinds of proteins by surveying more than a thousand of crystal structures.

In this work, we performed X-ray structure analysis by growing single crystals for four kinds of proteins; human carbonic anhydrase II (CAII), horse myoglobin (Mb), hen egg white lysozyme (HEWL), and human serum albumin (HSA). The former two proteins, CAII and Mb, are known to be easily crystallized by AS. On the other hand, the latter ones, HEWL and HSA, are known to be not. No structure has been deposited in the protein data bank (PDB) for HEWL crystallized with AS due to its difficulty in crystallization. To our knowledge, only photo image and crystallographic data of HEWL crystals by AS are available. Since the crystal growth of HSA by AS is also difficult, no PDB entry of HSA crystallized with AS is available. A description of the HSA crystal was seen in a report on the study about the cross-linking of the protein crystals. It should be noted that a large number of crystal structures are available in PDB both for HEWL and HSA, but all of them were obtained by growing crystals with precipitants other than AS. We also performed MD simulations for the four proteins in the calculation models with and without AS ions. The AS distribution around the protein was deduced from the simulation trajectory. In addition, the electrostatic potentials were drawn for the space surrounding the proteins. To clarify the reason for readiness for and difficulty in crystal growth by AS, these computational results were compared among the proteins.

2 Experiment

Protein single crystals were grown for CAII, Mb, HEWL, and HSA at the AS concentrations of 2.7, 3.2, 1.2, and 1.2 M, respectively. The max resolutions for the respective proteins were 1.32, 1.32, 1.13, and 3.90 Å. The major precipitant reported in the literature for crystallization of CAII and Mb is AS, and the space groups are $P12_11$ for CAII and $P6$ or $P12_11$ for Mb. Therefore, the crystals obtained in this work reproduced the previous reports. The space group of the HEWL crystal grown by AS was $P1$, which was different from the most popular space group of HEWL crystals, $P4_32_12$. The max resolution for HEWL in this study, 1.13 Å, was better than the average max resolution of the crystal structures of HEWL deposited in PDB, 1.73 Å. The space group of the HSA crystal by AS was $C121$, and the max resolution was 2.92 Å. The major space group of HSA crystals deposited in PDB is $C121$, and the average max resolution is 2.60 Å. In terms of the resolution, the quality of the HSA crystal grown by AS was inferior to the crystals obtained by other precipitants.

Calculation models for CAII, Mb, HEWL, and HSA were built from the crystal structures obtained above. MD simulations were performed for 200 ns in the presence of AS and also for 200 ns in the absence of AS. The simulations with AS were executed in triplicate for reproducibility. The concentrations of AS ions in the calculation models were set to be equal to the initial droplet in the experimental setup for protein crystallization, and other minor ingredients were omitted from the models. For reference, the simulations without protein were also performed in the presence of the same concentrations of AS as the respective models.

3 Results and Discussion

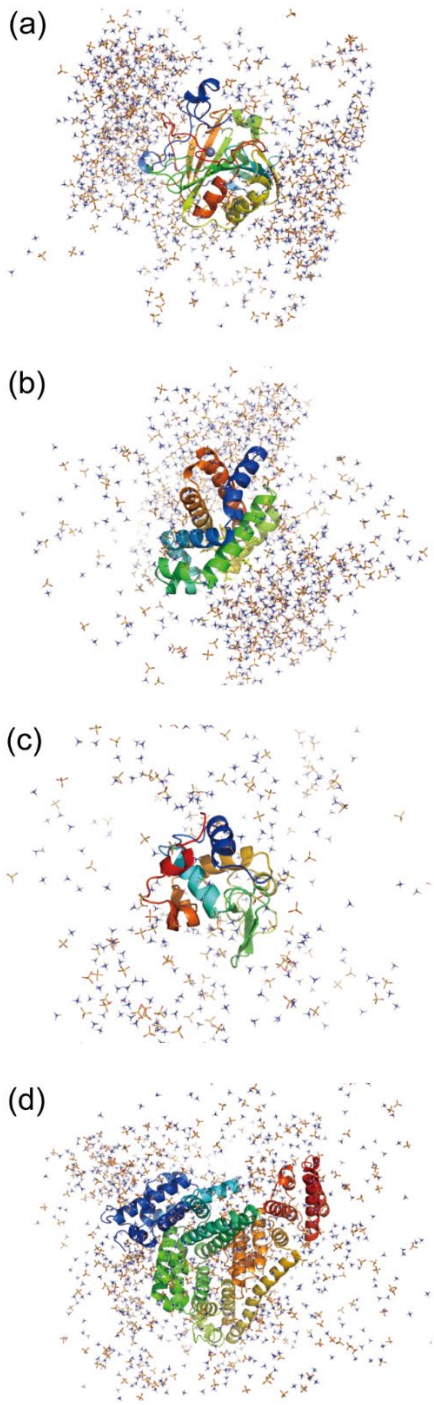

Figure 1. Representative structures of the simulation trajectory for (a) CAII, (b) Mb, (c) HEWL, and (d) HSA. Protein molecules are represented by the cartoon, changing colors from blue to red as the residue goes from N- to C-terminal sides. N, S, and O atoms of AS ions are colored blue, yellow, and red, respectively. Sodium and chloride ions and water molecules are not shown for clarity.

The average structure of the protein was obtained from the simulation trajectory for the last 20 ns. A snapshot structure that was the closest to the average one was extracted from the trajectory and defined as the representative structure (Figure 1). The representative structures demonstrate a keen difference in the AS distributions between the easily crystalized proteins by AS, CAII and Mb, and the not easily crystallized ones, HEWL and HSA. That is, the distribution of AS ions was markedly anisotropic for CAII and Mb. On the other hand, the distribution was isotropic for HEWL and HSA.

In CAII, AS ions were localized at two regions in the solvent, and the regions made contacts with the protein from both sides on an axis across the protein. Since the distributions of ammonium ions and sulfate ones were almost identical to each other, electrical polarity caused by AS ions was zero everywhere around CAII. AS distributions were compatible among the three simulations, and particularly the AS sparse areas around the protein are well matched among the three. In Mb, AS ions were localized at two regions, in which Mb was sandwiched between the two regions as in CAII. The AS dense areas were common among the three simulations. Electrical polarity due to AS ions was almost zero everywhere because of the identical distributions of ammonium and sulfate ions. In HEWL, AS ions were not localized in all of the three simulations, and instead, AS ions sparsely distributed in the solvent. The gatherings of AS ions were temporarily observed at local areas, while the positions were different among the simulations. In HSA, AS ions were not localized as well as HEWL. Small gatherings of AS ions were observed at a few regions near the protein surface. When no protein was included in the calculation model, AS ions were randomly distributed in solvent. The conformational changes of the tertiary structures were hardly observed for all the proteins, while noticeable increases in atom fluctuation were occasionally seen at the loop domain for HEWL.

The motions of protein and precipitant molecules were monitored all through the 200 ns simulation, and the snapshot structures at 0, 50, 100 ns were extracted from simulation trajectory both for the models with and without AS. In every simulation without AS, the protein molecule substantially moved in the solvent, and the direction of the principal molecular axis was ceaselessly altered. Since the molecular weights of Mb and HEWL are small compared to those of CAII and HSA, the displacements of the center of mass for Mb and HEWL were much larger than those of CAII and HSA. In the presence of AS ions, the motion of the protein molecule was reduced especially for CAII and Mb. In the case of CAII with AS, not only the center of mass but also the direction of the molecular axis did not easily alter. The initial setup of AS positions, in which AS ions were randomly placed in the solvent, was reflected in the AS distribution at 0 ns. At 50 ns, AS ions were relocated not to cover all over the protein surface but to make contacts with CAII partially. The AS distribution was not uniform, and instead, the localization of AS ions was observed. The localized distribution became remarkable at 100 ns. In the simulation of Mb with AS, while AS distribution at 0 ns was random, AS ions were arranged to form a localized distribution at 50 ns. AS dense areas appeared at two regions around the protein, in which two AS dense areas and the protein were almost aligned in a straight line. The positions of AS dense areas were

scarcely changed at 100 ns, which suggested that the distribution was stable. In HEWL with AS, AS ions were randomly distributed during the simulation. The random distribution of AS ions suggested that the surrounding space of HEWL was unfavorable for the gathering of ions. In HSA with AS, the AS dispersed distribution was kept through the simulation for 100 ns.

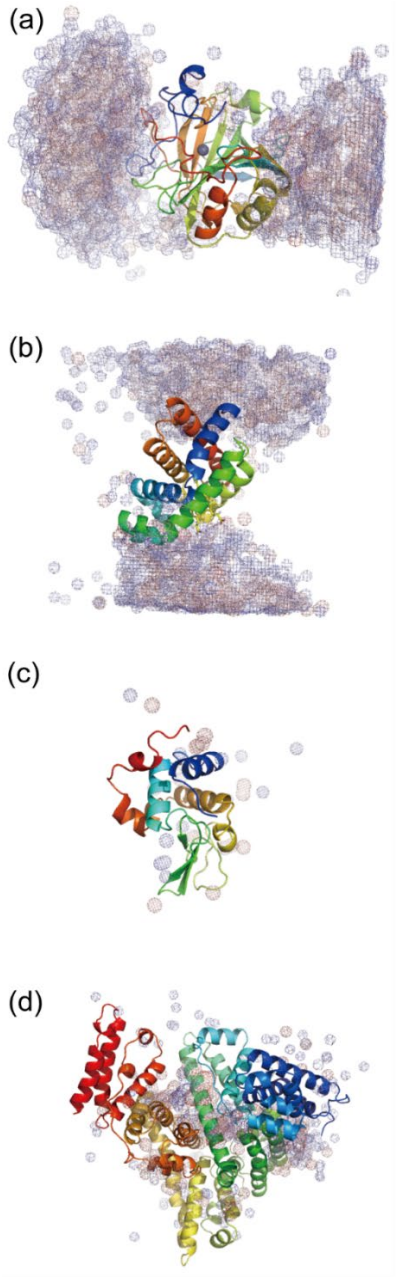

Figure 2. Distributions of AS ions in the simulations for (a) CAII, (b) Mb, (c) HEWL, and (d) HSA. The distributions of AS ions are depicted in mesh, according to the rate of the presence of ions for the last 20 ns. The population of ammonium ions is colored from light to deep with the increase of ion density in blue and that of sulfate ions is colored in red.

In order to explain the localized distribution of AS ions, the iso-surface of electrostatic potential around the protein was depicted (Figure 3). In CAII and Mb, the positive and negative areas of electrostatic potential were clearly separated, in which both areas had spherical shapes and were almost equal to each other in volume. The AS gatherings observed in the simulations were compatible with the iso-surface of electrostatic potential. AS ions were likely to be located at the regions where the absolute value of electrostatic potential was high. That is, the surroundings of the protein were polarized, and the electric field lines were crowded at these polar regions. In contrast, a marked difference was observed in volume between the positive and negative areas in HEWL and HSA. In HEWL, the whole protein was completely covered with the positive potential surface. AS ions were not inclined to gather at specific areas on the protein surface. In HSA, the electrostatic potential was greatly biased to the negative. Most of AS ions were not strongly attracted to specific local areas, while there was a small region where AS ions gathered due to the positive electrostatic potential (Figure 3d). Consequently, the electrostatic potential is responsible for the localized distribution of AS ions, which leads to the stabilization of the mixture of AS ions and the protein in cases of CAII and Mb. It is known that the protein crystals contain a large amount of solvent. The protein atoms occupy only about 43% of the unit cell in volume on average. The distributions of precipitant agents and the packing of protein molecules were suggested to be complementary to each other. Therefore, AS ions will stabilize the protein crystals by staying at the space not occupied by the protein atoms.

It will be expected from the findings of this study that an introduction of amino acid mutation to alter the electrostatic potential around a protein can enhance crystallization even if the experimental condition for the crystal growth of the protein with AS is not well-established. The optimal zone of precipitant concentration is sometimes narrow for crystal growth with AS. The electrostatic potential will be utilized as a guide to how to modify recombinant proteins suitable for the crystal growth with AS. For example, when a protein crystal is obtained with a precipitant other than AS, but the resolution of X-ray diffraction is not satisfactorily high, a better quality of protein crystals is occasionally required. Modeling of the protein structure from the low-resolution X-ray diffraction, modification of the model by introducing amino acid mutation, and calculation of the electrostatic potential for the modified model will give a good suggestion on the adequate modification of the protein to promote the crystallization with AS. The electrostatic potential significantly depends on the pH of solution. Multiple calculations of the electrostatic potentials for the models changing the protonation states of the titratable residues will provide the information on the appropriate pH condition. Hence, the analysis of the electrostatic potential enables us to take the rational strategy in experiments.

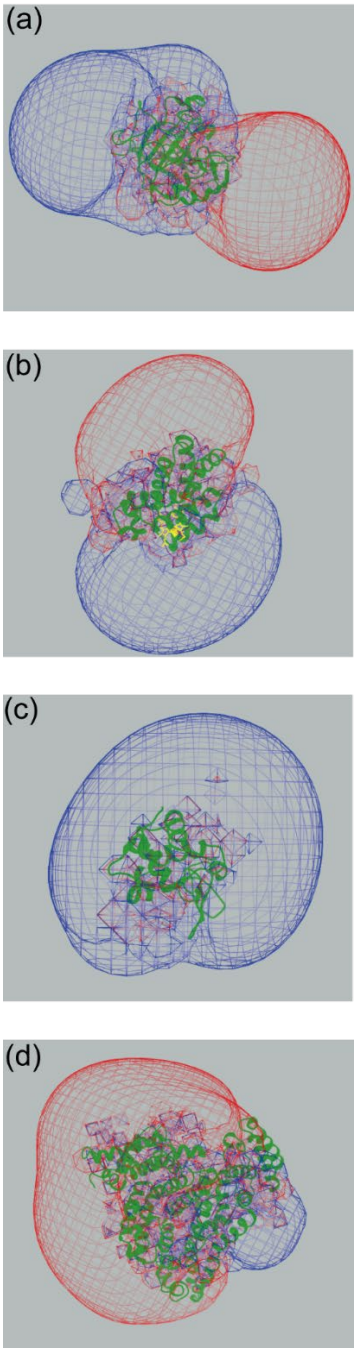

Figure 3. Positive and negative iso-surfaces of electrostatic potential around the proteins for (a) CAII, (b) Mb, (c) HEWL, and (d) HSA. The positive and negative iso-surfaces are shown in the blue and red mesh representation, respectively. The contouring values of iso-surfaces are (a) +0.43 and -0.14 for CAII, (b) +0.39 and -0.13 for Mb, (c) +1.75 and -0.004 for HEWL, and (d) +0.43 and -1.96 for HSA in unit of kT/e. Protein molecules are represented by green cartoons. Every electrostatic potential was obtained from the simulation in the absence of AS.

In summary, protein crystals were grown for CAII and Mb, which are easily crystallized by AS, and also for HEWL and HSA, which are not. X-ray diffractions were acquired from the respective protein crystals, and the structures were determined using the molecular replacement method. Based on the determined protein structure, 200 ns MD simulations were performed in triplicate with the models including AS ions at the same concentration as the experimental condition in the crystal growth for every protein. A 200 ns simulation without AS ions was also performed for comparison. The presence of AS ions reduced the motion of the proteins during the simulations, especially for CAII and Mb. The distribution of AS ions was not random but highly anisotropic around the protein with the ions localized at two areas. The localized distribution was caused by the electrostatic potential around the protein, and AS ions gathered at the regions where the absolute value of the electrostatic potential was high. In CAII and Mb, the positive and negative areas of electrostatic potential were almost equally separated, and both areas had a spherical shape. In contrast, either one of the positive or negative areas was dominant in HEWL and HSA. Therefore, the shape of the iso-surface of the electrostatic potential is highly responsible for the readiness in the crystal growth with AS. If electrostatic potential around a protein can be controlled by adjusting pH or introducing amino acid mutation, the quality of crystallization by AS will be improved even for the proteins which are not easily crystallized by AS.

Acknowledgement

Calculations were performed at Research Center for Computational Science, Okazaki, Japan and at Information Technology Center of the University of Tokyo.

* hoshino@chiba-u.jp