BL-17A/2018G613

Electrostatic Potentials around the Proteins Crystallized by Ammonium Sulfate Preferably

Tyuji Hoshino[1,*] Yan Guo[1] Liang Qu[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 crystallography is the primary experimental method for determining protein structure. The crystal growth of protein molecules is, however, a critical step for this method. In general, a variety of parameters has influence on the molecular interaction of proteins and is related to the protein crystal growth. Hence, the optimization of the crystallization condition is a time-consuming process. In particular, the selection of precipitant agent is a key parameter for obtaining a good quality of protein crystals. Ammonium sulfate (AS) and polyethylene glycol (PEG) are the most popular precipitants to induce protein crystallization. Some proteins are easily crystallized by AS, while others are not. Instead, some proteins are crystallized PEGs, preferably. The reason why the preference of precipitant is different among proteins is not still clear in spite of many previous studies on the function of precipitants.

The protein solubility is affected by the ionic strength of the solution and the net charge of the protein. The surface tension at the water-protein interface and the activity of waters in the hydration layer are related to the solubility. According to the well-accepted mechanism, Hofmeister series of ions for protein hydration, the strongly hydrated anions such as sulfate and phosphate compete for water molecules in the second hydration layer, which leads to the decrease of the water activity in the first layer for solvating the protein. The strongly hydrated anions further increase the surface tension of the solvent, making the protein to minimize its solvent accessible surface area. The decrease of water activity and the increase of surface tension eventually result in the reduction of the protein solubility by the salting-out effect. Another effect of salts on proteins is the protein-ion interaction. The electric charge of protein is an essential factor related to the solubility. The strongly hydrated ions form strong protein-ion interaction, which have significant influence on the hydration of protein. AS is one of the strong salting-out agents and widely used in crystallization due to its high solubility, low heat of solution, and low denaturation of protein structure. Hence, the protein crystal growth by AS is attributed to hydration and dehydration effects involving competition for waters. Although the salt-out mechanism is broadly accepted, the understanding of the reason for the preference of precipitants is still limited.

In our previous study [1], a kind of protein was crystallized with three different precipitants, including AS and PEG. Computational analysis showed that AS ions were anisotropically distributed around the protein, which was reflected in the molecular packing of the crystals. Because the molecular packing of protein crystal is linked to space groups, the relationship between space groups and precipitants was investigated for several kinds of proteins. Furthermore, we performed X-ray structure analysis by growing single crystals for four kinds of proteins [2]. Two of them are known to be easily crystallized by AS, while the other two are not. Molecular dynamics simulations of the four proteins with AS ions indicated that the distribution of AS ions in the proteins easily crystallized by AS was not random but highly anisotropic around the protein with the ions localized in 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. The findings suggested that the shape of the iso-surface of the electrostatic potential was responsible for the easiness of the crystal growth with AS.

In this work, we picked up 100 crystal structures that had been crystallized by high concentrations of AS. The electrostatic potentials were drawn for the space surrounding the proteins. The shapes of electrostatic potentials were contrasted to those of 20 proteins that had been crystallized by PEGs. In addition, the electrostatic potentials were examined for the proteins that had been crystallized by low concentrations of AS and other precipitant agents. The difference in the shape of electrostatic potential was quantitatively characterized by an index representing the polarity. This study aims to clarify the reason for the preference for precipitants in protein crystal growth.

2 Experiment

The crystal structures analyzed in this work were downloaded from protein data bank (PDB) under the following criteria. (1) Crystal was grown in the *apo*-form without ligand or inhibitors. (2) As for 100 structures, proteins were crystallized by AS at high concentrations over 2.0 M. (3) As for 10 structures, proteins were crystallized by AS at low concentrations below 2.0 M. For comparison, 20 crystal structures obtained by utilizing PEGs were picked up. Additionally, 10 structures of the

proteins crystallized by other kinds of salts or agents were included. Totally, 140 crystal structures were surveyed in this work.

The electrostatic potential around every protein was depicted for all the downloaded crystal structures in the following manner. In preparation, the missing residues in the crystal structure were added by the homology modeling with Modeller 9.2. The pKa calculation was performed by Propka 3.1, in which the pKa values of the titratable residues were predicted by electrostatic continuum theory. The protonation state of each titratable residue was selected so that the net charge of the whole-body of the protein became zero at a certain pH, i.e., iso-electric point. The hydrogen atoms were generated by leap module of AmberTools16. The electrostatic potential was obtained by solving the Poisson-Boltzmann equation with the electrostatic continuum method by Delphi 5.1, and the potential map was visualized using Chimera 1.12 or PyMOL 1.7.

Single crystal of chitosanase was grown by utilizing AS. An *Escherichia coli* strain, Rosetta (DE3) pLysS, was transformed with pET50b vector (Novagen) containing the gene coding N-terminal side-truncated chitosanase. Chitosanase was expressed as a 6×His-fused Nus-tag-conjugated form due to the use of pET50b vector. After the pre-culture in LB medium at 37°C, the protein was expressed at 28°C overnight by the induction with 0.2 mM isopropyl-β-thiogalactopyranoside (IPTG) at an OD600 value of 0.6. The conjugated protein was purified a Co metal-affinity chromatography, followed by the cleavage of 6×His-fused Nus-tag by human rhinovirus 3C (HRV-3C) protease. The cleaved Nus-tag, HRV-3C protease, and uncleaved conjugated protein were removed by Ni-nitrilotriacetic acid (NTA) agarose resin. The protein was further purified by gel filtration with a running buffer of 10 mM Tris-HCl at pH 8.0 and 150 mM NaCl. Finally, the protein was concentrated to 5.4 mg/mL.

The single crystal of chitosanase was grown by the sitting drop vapor diffusion method at 18°C. For setting up a droplet for crystallization, 1.0 µL of protein solution was mixed with 1.0 µL of the precipitant solution consisting of 100 mM sodium citrate at pH 5.0 and 3.4 M AS. The mixture was placed on the well plate with the reservoir filled with 300 µL of precipitant solution. The crystals were grown within two weeks without any seeding technique. The crystals were cryo-protected by brief immersion in a solution containing 25% (v/v) glycerol, followed by freezing in liquid nitrogen.

X-ray diffraction data were collected at 100 K on BL-17A beamline of photon-factory (PF, Tsukuba, Japan). The diffraction data were indexed, scaled, and merged with XDS. Intensities were converted into structure factors, and 5% of the reflections were flagged for R_{free} calculations. The protein structure was determined by the molecular replacement with MolRep in CCP4 software package using a structure obtained by the PEG-grown crystal (PDB code: 1v5c) as a search model. Structure refinement and model building were carried out using Refmac and Phenix. Molecular structure and electron density map were visualized by COOT.

3 Results and Discussion

The electrostatic potentials were depicted for all the 140 proteins with determining the protonation state of every titratable residue at a pH value in which the net charge of the protein was zero. Figure 1 shows the electrostatic potentials at the surrounding space of the typical 20 proteins that are preferably crystallized by high concentrations of AS. All the 20 proteins display a common shape in electrostatic potential. The positive and negative areas are almost equally separated, and each area has a spherical lobe form. The protein molecule is positioned at the boundary region of the two areas, that is, each spherical lobe stems from the protein and expanded in the opposite direction to each other. It should be noted that the positive and negative areas distinctly split even at their boundary, and the contact region is limited to the protein. The electrostatic potentials around the 10 proteins that are preferably crystallized by PEGs are shown for comparison. The separation between the positive and negative areas is unclear and the electrostatic potential at the boundary region is not neat for the proteins crystallized by PEGs. The positive and negative areas are interlaced with each other.

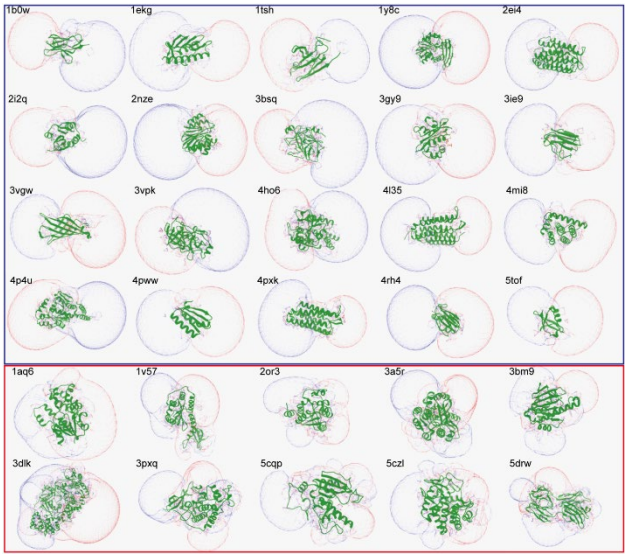

Figure 1: Positive and negative iso-surfaces of electrostatic potential around the selected 30 proteins. Out of 30, 20 proteins were crystallized by high concentrations of AS, and 10 were by PEGs. The AS-crystallized 20 proteins are illustrated in the blue frame and the PEG-crystallized 10 are in the red one. The positive and negative iso-surfaces are shown in the blue and red mesh representation, respectively. The contouring values of iso-surfaces are +0.5 and -0.5 in the unit of kT/e. Protein molecules are represented by green cartoons.

The electrostatic potentials for all the 140 proteins are shown in Figure 2. The initial 100 proteins have been crystallized by AS at the concentrations above 2.0 M. As described for the 20 typical proteins in Figure 1, the positive and negative areas are distinctly separated, and each of the areas has a spherical lobe form. In other words,

the positive and negative areas of the electrostatic potential clearly split, in which both areas are round in shape and almost equal to each other in volume. The next 10 proteins have been crystallized by AS at the concentrations below 2.0 M. The positive and negative areas are clearly separated for some proteins such as 1skz and 1toh. At the same time, the separation is not clear for a few other proteins as 5go6 and 5wko. For the 20 proteins that have been crystallized by PEGs, the boundary between the positive and negative areas is not clear, with some exceptions like 1v5c. In 1ao6 and 1y62, the positive and negative areas are separated, but the boundary is not distinguishable. The additional 6, 2, and 2 proteins have been crystallized by sodium formate, glycerol, and others, respectively. The proteins crystallized by sodium formate show the distinct separation of the positive and negative areas as seen in the proteins crystallized by AS. Therefore, it can be concluded that the positive and negative areas neatly split in the electrostatic potentials for the proteins preferably crystallized by the strongly hydrated ion-containing precipitants such as AS and sodium formate.

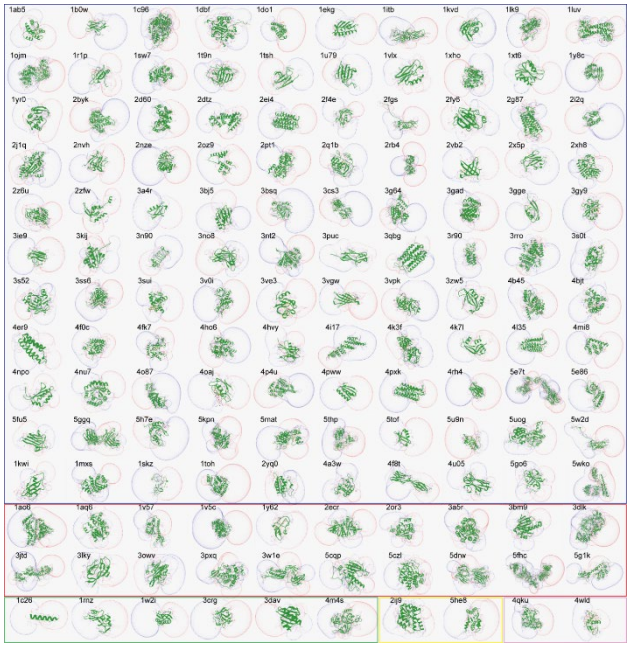

Figure 2: Electrostatic potentials for the 140 proteins. The contouring values of the positive and negative iso-surfaces are +0.5 and -0.5 in the unit of kT/e, respectively. The positive and negative iso-surfaces are shown in the blue and red mesh representation, and the protein molecule is represented by the green cartoon. The AS-crystallized 110 proteins are displayed in the blue frame, and the PEG-crystallized 20 are in the red one. Six proteins in the green frame were crystallized by sodium formate, and two proteins in yellow were by glycerol. The rest two proteins in pink were crystallized by ethylene glycol and sodium chloride.

Among 20 crystal structures crystallized by PEGs, the electrostatic potential of 1v5c exceptionally displayed a clear separation of negative and positive areas. Since the electrostatic potential had the shape commonly observed for the AS-crystallized proteins, the crystal growth of the protein will be possible by AS, too. The PDB entry: 1v5c is the crystal structure of the inactive form of chitosanase, which is an enzyme to hydrolyze the partly acetylated chitosan. The inactive form of chitosanase was crystallized by the precipitant of 20% (w/v) PEG4000, including 0.4M AS at pH3.7 in a previous report. The space group of 1v5c was I222, and the max resolution was 2.00 Å. In the present work, we tried the crystallization of chitosanase by AS. A single crystal of chitosanase was grown under the condition of 3.4 M AS at pH5.0. No PEGs or organic ingredient was included in the precipitant solution. The space group of the AS-grown crystal was $P2_12_12_1$, and the max resolution was 1.75 Å. Two chitosanase molecules were included in the asymmetric unit, unlike 1v5c. The structures of the two chitosanase molecules were almost equal to each other, as shown in the drawing of their superimposition (Figure 3a). The root-mean-square deviation (RMSD) between two molecules was 0.131 Å. The structure of the AS-grown chitosanase is almost the same as the PEG-grown one. The crystal structure of one chitosanase molecule in the asymmetric unit was superimposed on the 1v5c structure (Figure 3b). The RMSD between them was 0.136 Å.

One chitosanase molecule in the AS-grown crystal made contact with six surrounding molecules in the crystal packing. The electrostatic potential of the AS-grown chitosanase was drawn with the surrounding molecules (Figure 3c). The positive and negative spherical lobes of the electrostatic potential were located at the space unoccupied by other molecules in the crystal packing. Other molecules were likely to be positioned at the intermediate between two spherical lobes. In other words, the molecules in the crystal made contacts among them in the areas where the absolute value of the electrostatic potential was low. According to a simulation study of AS ions in the protein crystal growth, the distribution of AS ions around a protein was anisotropic if the protein was easily crystallized by AS solution as a precipitant. Instead, AS ions are likely to gather at the local areas of the protein surface. The distribution of AS ions around the protein coincided with the shape of the electrostatic potential. Therefore, AS ions were considered to be attracted to the areas where the absolute value of the electrostatic potential was high, that is, the space unoccupied by chitosanase molecules in crystal packing. The electrostatic potential of the PEG-grown chitosanase was also drawn with the surrounding molecules for comparison (Figure 3d). The locations of the surrounding molecules seem to correlate with the electrostatic potential. The surrounding molecules, with one exception, tended to make contact with the center molecule at the intermediate areas between two spherical lobes where the absolute value of the electrostatic potential was low. Furthermore, no molecule was observed in the positive area of the electrostatic potential. Since the PEG-grown chitosanase was reported to be crystallized with including a low concentration of AS, the mixed AS ions might assist the protein growth with gathering at the positive area even for the PEG-grown chitosanase.

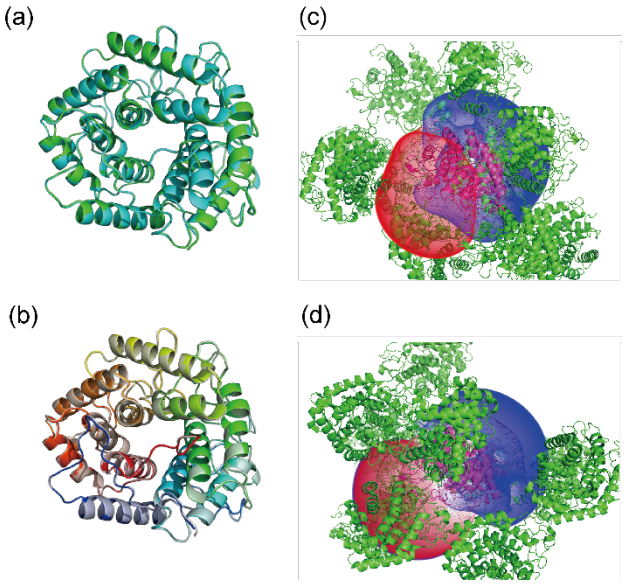

Figure 3: (a) Structures of the two chitosanase molecules in an asymmetric unit of the AS-grown crystal. One molecule colored green is superimposed on the other colored cyan. (b) Superimposition of the AS-crystallized chitosanase onto the PEG-crystallized one. Protein molecules are represented by the cartoon, changing colors from blue to red as the residue goes from N- to C-terminal sides for the AS-crystallized one. The PEG-crystallized molecule is colored grey. Water molecules are not shown for clarity. (c) Electrostatic potentials for the AS-crystallized chitosanase with the surrounding molecules. A molecule of interest is colored cyan, and the surrounding ones are green. The positive and negative iso-surfaces are drawn in blue and red, with their contouring values at +0.5 and -0.5 in the unit of kT/e, respectively. (d) Electrostatic potentials for the PEG-crystallized chitosanase with the surrounding molecules. The coloring and drawings are the same as (c).

Acknowledgement

Calculations were performed at the Research Center for Computational Science, Okazaki, Japan and at the Information Technology Center of the University of Tokyo. X-ray diffractions were acquired at the Photon Factory, Tsukuba, Japan (proposal no. 2018G613). A part of this work was supported by a grant for Scientific Research C from the Japan Society for the Promotion of Science.

References

[1] Fudo, S., Qi, F., Nukaga, M., Hoshino *Cryst. Growth Des.* **17**, 534-542 (2017).

[2] Kitahara, M., Fudo, S., Yoneda, T., Nukaga, M., Hoshino, T. *Cryst. Growth Des.* **19**, 11, 6004-6010 (2019).

* hoshino@chiba-u.jp