Cell-free protein crystallization for structural analysis

In nature, some proteins spontaneously crystallize in living cells. Such crystals are known to have biological functions such as protein storage, virus protection, heterogeneous catalysis, and immune system activation [1,2]. Since the structure of polyhedra, one of the in-cell protein crystals, was determined in 2007 [3], in-cell protein crystallization (ICPC) has attracted attention as a next-generation structural biology tool because it does not require multistep purification processes or large-scale crystallization screening. Several ICPC methods have been developed, including high-throughput screening and the optimization of the cell culture process. However, significant issues remain to be resolved in obtaining various protein crystals in sufficient amounts and quality for structure determination by ICPC because crystals are often formed incidentally in the cells. Therefore, several technical challenges must be overcome to apply this method to protein structural analysis. If a new ICPC method can be established, it will be expected to become a more accessible structural analysis technique.

Cell-free protein synthesis (CFPS), a protein preparation technique used in synthetic biology, is very effective for rapidly screening protein synthesis [4]. However, it has been considered unsuitable for structural biology efforts that require large amounts of protein, such as crystallization. Here, we report the development of cell-free protein crystallization (CFPC), a direct protein crystallization method using CFPS [5]. We have (1) established small-scale and rapid crystallization using CFPS and (2) manipulated crystallization by adding chemical reagents. The polyhedra crystal (PhC), produced in insect cells by infection with cytoplasmic polyhedrosis virus (CPV), is one of the most studied in-cell protein crystals. The most crucial advantages of CFPC are that the reaction scale and time can be minimized and that various reagents can be added during the reaction.

The crystallization of polyhedrin monomers (PhMs) was performed using the Wheat Germ Protein Synthesis kit (WEPRO7240 Expression Kit) because these extracts have been identified as having the highest protein expression activity among eukaryotic systems. Translation reactions were performed by the bilayer method. A 20 μL reaction mixture containing 10 μL of WEPRO7240 and 10 μL of mRNA solution was placed in a 1.5 mL microtube, overlaid with 200 μL of SUB-AMIX SGC solution, and incubated at 20°C for 24 h (Fig. 1(a)). The reaction mixture was centrifuged, and a white precipitate was collected (Fig. 1(b)). The crystalline precipitate was observed under an optical microscope (Fig. 1(c)). The crystals prepared from the CFPC reaction (PhC_CF) have the same cubic morphology as PhC synthesized in insect cells (PhC_IC). The average size of PhC_CF (580 nm) measured by scanning electron microscopy (SEM) is about one-fifth that of PhC_IC (2700 nm) (Fig. 1(d)).

To collect diffraction data of nanosized PhC_CF isolated from the reaction mixture, a micro-X-ray beam at SPring-8 **BL32XU** equipped with Serial Synchrotron Rotation Crystallography (SS-ROX) was used for diffraction. PhC_CF was refined with a resolution of 1.80 Å and has a space group (I23) and lattice parameters identical to those of PhC_IC (Fig. 2). The significant difference between PhC_CF and PhC_IC is the absence of the electron density of nucleotide triphosphate (NTP) bound to the monomer interface observed in PhC_IC. The result indicates that NTP binding is not essential for the crystallization of PhMs. After the elucidation of the translation time, it was found that the CFPC reaction of PhC produces nanocrystals of sufficient quality to obtain a high-resolution structure in only 6 h. This reaction time is markedly shorter than the incubation time (>3 days) required to obtain equivalent high-quality crystals using insect cells. By the dialysis method for CFPC, the crystal that gives a resolution of 1.95 Å was obtained with only 20 μL of the reaction mixture.

We applied this method to the structure determination of crystalline inclusion protein A (CipA) with the addition of chemical reagents to the reaction solution to suppress twin crystal formation. CipA, a hydrophobic protein

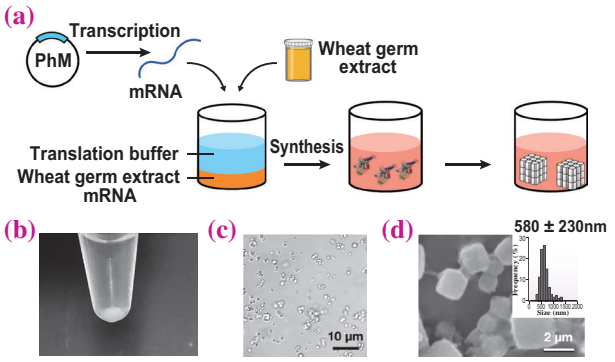

Fig. 1. **(a)** Schematic illustration of cell-free protein crystallization (CFPC) of polyhedrin monomer (PhM) using the Wheat Germ Protein Synthesis kit. **(b)** Photograph of the tube after CFPC. **(c)** Differential interference contrast (DIC) image of PhC_CF. **(d)** Scanning electron micrograph of PhC_CF. Size distribution of PhC_CF determined from the SEM image.

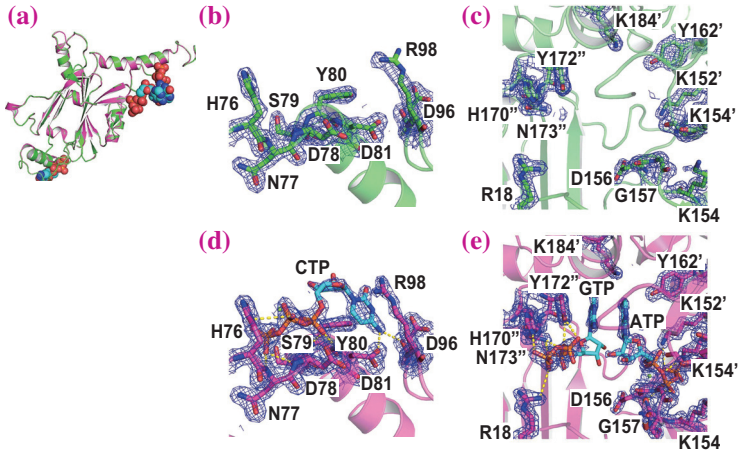

Fig. 2. Crystal structures of PhC_CF20°C/24h and PhC_IC (PDB ID: 5gqm). **(a)** Superimposed structures of PhC_CF 20°C/24h (green) and PhC_IC (magenta). **(b,c)** Close-up views of (b) CTP and (c) ATP/GTP binding sites in PhC_CF20°C/24h, **(d,e)** Close-up views of (d) CTP and (e) ATP/GTP binding sites in PhC_IC. The selected $2|Fo| - |Fc|$ electron density maps at 1.0σ are shown in blue. Hydrogen bonds are indicated with yellow dotted lines.

of 104 amino acid residues, spontaneously forms crystalline particles in *Photorhabdus luminescens*, an entomopathogenic bacterium. When CipA was expressed in CFPC by the dialysis method, a white precipitate appeared in the solution mixture after 24 h. The structural analysis of CipA_CF was attempted with 1.61 Å resolution data obtained by the small wedge method at SPring-8 BL32XU. However, the structure could not be determined because the twin fraction was high. Therefore, we attempted to overcome this problem by adding a reagent that inhibits twinning to the reaction mixture of CFPC. X-ray diffraction experiments of the crystal show that CipA_CF crystallized in the presence of 3 v/v% 1,4-dioxane had a markedly reduced twinning fraction of 0.10 with a resolution of 2.11 Å. The structure was determined by molecular replacement using the search model created by AlphaFold2 (Fig. 3). The monomer structure of CipA is a typical oligonucleotide/oligosaccharide-binding (OB) fold, consisting of an α-helix and a β-strand. In the lattice structure, the four α-helices from each monomer form a four-helix bundle and exist as a tetramer. This tetramer is considered to be the basic unit of crystal growth. The interactions between the edges of each tetramer form the crystal lattice, which is further stabilized by embedding the neighboring monomer's N-terminal arm in the cleft that forms between the tetramer-tetramer interface.

We have established a CFPC method to rapidly obtain protein crystals in microliter volumes within a few hours without the need for purification and crystallization procedures. This technology, which integrates in-cell and *in vitro* crystallization, significantly expands the tools available for high throughput protein structure determination, particularly in unstable, low-yield, or substrate-binding proteins, which are difficult to analyze by conventional methods.

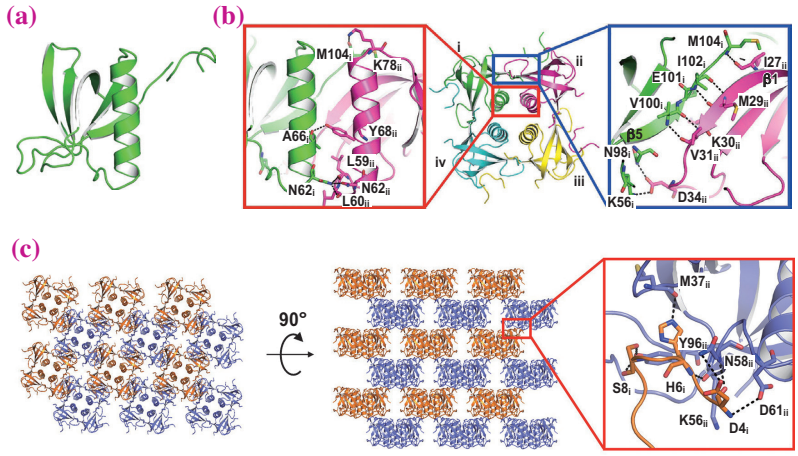

Fig. 3. Crystal structure of CipAC-CF with 1,4-dioxane. Structures of **(a)** monomer and **(b)** tetramer. **(a)** The CipA monomer consists of the N-terminal arm followed by three β-strands (β1, β2, and β3), an α-helix, and two β-strands (β4 and β5). **(b)** Interactions between monomers (i, ii, iii, and iv) in the tetramer. **(c)** Lattice structure and interactions between tetramers. Hydrogen bonds are indicated with black dotted lines.

Satoshi Abe* and Takafumi Ueno

Tokyo Institute of Technology

*Email: saabe@bio.titech.ac.jp

References

[1] M. Kojima *et al.*: Biomater. Sci. **10** (2022) 354.
[2] R. Schonherr *et al.*: Biol. Chem. **399** (2018) 751.
[3] F. Coulibaly *et al.*: Nature **446** (2007) 97.
[4] M. Harbers: FEBS Lett. **588** (2014) 2762.
[5] *S. Abe, J. Tanaka, M. Kojima, S. Kanamaru, K. Hirata, K. Yamashita, A. Kobayashi and T. Ueno: Sci. Rep.* ***12*** *(2022) 16031.*