Crystal Structure Determination using a Protein Fusion that Forms a Highly Porous Lattice

Protein crystallization requires numerous screenings, which hinders structural studies. To overcome this, an R1EN-fusion method was developed, which consists of preparing an R1EN-protein fusion and assembling the same crystal lattice as R1EN, consequently obtaining the crystal structure of the fusion guest. R1EN-ubiquitin fusions were prepared and crystals with the same lattice under the same conditions as the original R1EN were successfully obtained. The crystals diffracted to 1.7–2.4 Å resolution, and the ubiquitin structure could be determined by molecular replacement using R1EN. The R1EN-fusion method eliminates the need for crystallization screening and is useful for *de novo* protein structure determination.

A crystallization process requires setting a large number of conditions for screening and requires a large amount of high-purity protein, hindering crystal structure analysis. Accordingly, it is extremely difficult to develop an easy and versatile protein crystallization technique. In 1982, Seeman suggested a strategy of protein crystallization in which a crystal lattice made of DNA origami is used as a scaffold and a DNA-binding protein is arranged therein, consequently solving the crystal structure [1]; however, a long-stretched DNA itself is too flexible to retain the crystal lattice. Instead of DNA, using a protein that assembles a highly porous crystal lattice is a simple way to incorporate the protein of interest to the crystalline framework just by constructing a protein fusion. A previous study [2] focused on R1EN, which forms a hollow honeycomb crystal lattice with an inner diameter of ~110 Å (**Fig. 1**) [3], and demonstrated that R1EN-ubiquitin fusion protein (R1EN-Ub) could be assembled with the same honeycomb crystalline framework as R1EN, thus determining the X-ray structure of ubiquitin.

At first, a series of R1EN-Ub with various linker lengths were prepared and crystallized by the micro-seeding technique, resulting in the same hexagonal crystals under the same condition as the R1EN. X-ray diffraction data with 1.7–2.4 Å resolution were collected at AR-NW12A (PF) and BL44XU (SPring-8). The phases were obtained by molecular replacement using R1EN (PDB:2EI9) and ubiquitin (1UBQ). In $R1EN_{223}$-Ub (linker length: 3aa), the electron density of ubiquitin in the central cavity was observed, while in the long (≥11aa) linker constructs, the electron density of ubiquitin could not be observed. There are two contact faces between ubiquitin and symmetric ubiquitin [**Fig. 2(a)**], one of which (contact-A) seems to be too close to neighboring ubiquitin. The steric collision in contact-A was not fixed even when the space group was converted to P1. The R1EN moiety was almost the same as the previous structure. Interestingly, the electron density of Met5-Pro9, which was disordered in the previous R1EN structure, was observed. There are four hydrogen bonds between ubiquitin Lys33 and Met5-Pro9, and

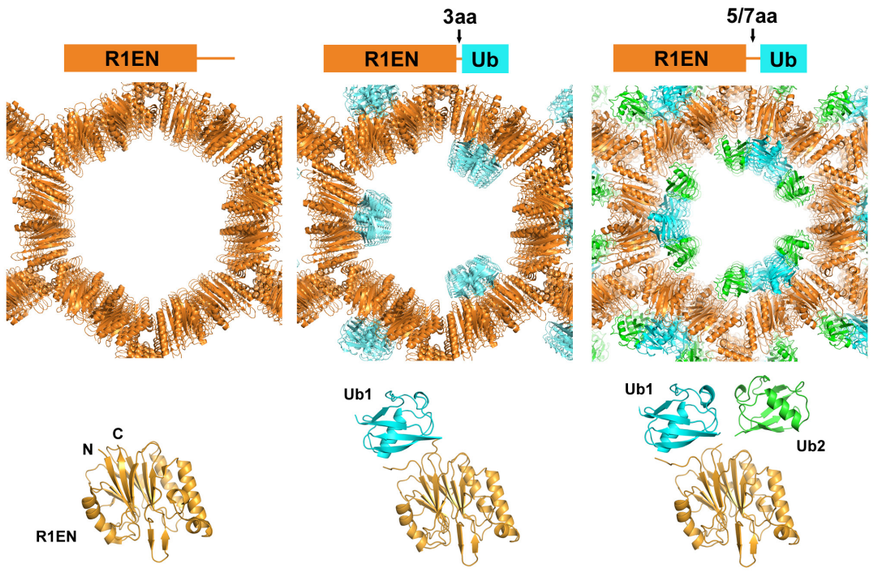

Figure 1: Crystal lattices of R1EN (left), $R1EN_{223}$-Ub (center), and $R1EN_{225/227}$-Ub (right).

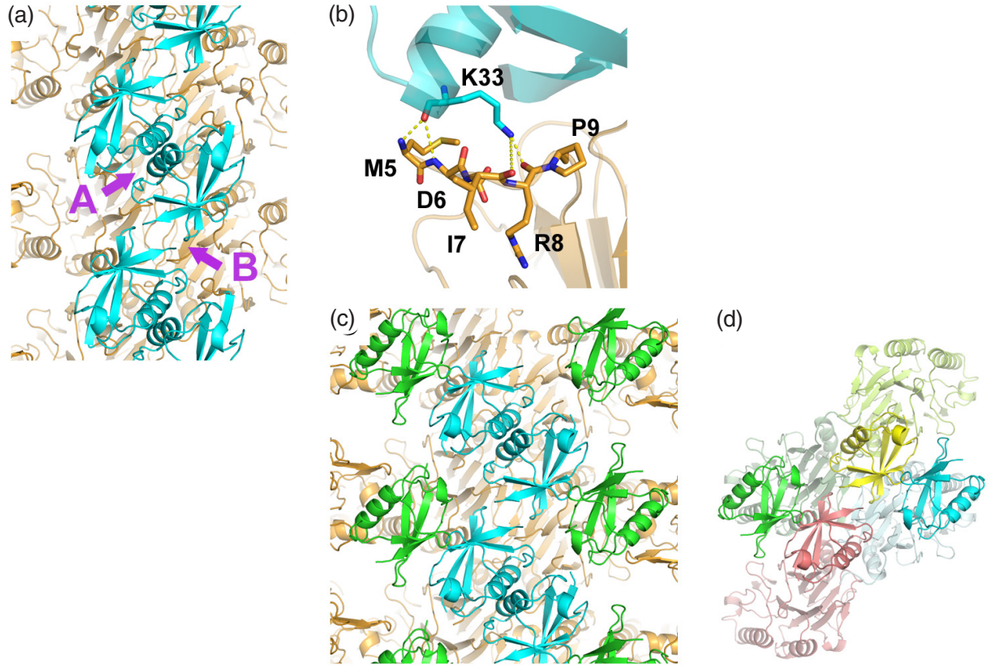

Figure 2: Interaction of ubiquitin moiety. **(a)** Two contact faces of Ub1. Steric collision in contact-A was observed. **(b)** Interaction between the N-terminal region of R1EN and Ub1. **(c)** Symmetry-related contacts of Ub1 and Ub2 in $R1EN_{225/227}$-Ub. **(d)** The proposed model of tetrameric R1EN-Ub unit in $R1EN_{225/227}$-Ub crystal. For clarity, R1EN-Ubs are colored yellow, green, cyan, and red, respectively.

presumably the Met5-Pro9 region interacted with ubiquitin, stabilizing each other [**Fig. 2(b)**].

Next, an attempt was made to solve the ubiquitin structure by the phases calculated from molecular replacement of R1EN alone. Phase improvement and auto-building were performed, and the resulting ubiquitin model contained 65 out of 72 residues of the main chain, including 36 of the correctly modelled side chains.

In $R1EN_{225}$-Ub (linker length: 5aa) and $R1EN_{227}$-Ub (linker length: 7aa), another ubiquitin subunit was observed [**Fig. 1, 2(c)**]. For clarity, the ubiquitin observed in $R1EN_{223}$-Ub is referred to as Ub1, and the second ubiquitin as Ub2. Since the N-terminus of Ub1 and Ub2 are close to each other, Ub2 is considered to be Ub1's multi-conformer. The distance between the Met1 in Ub2 and the C-terminus R1EN is 9.4 Å, which is too far for the $R1EN_{223}$-Ub to place the ubiquitin at the Ub2 position. Ub2 has no definite interaction with cis-R1EN, however, the Gln31-Gly35 region of Ub2 interacts with symmetry-related R1EN by water-mediated interactions. Moreover, Lys6 in Ub2 interacts with Asn60 and Gln62 in symmetry-related Ub1. Collectively, the Ub1-Ub1 interface at α1 (contact-A) is too close, while other Ub1-Ub1 (contact-B) and Ub2-Ub1 interfaces are quite reasonable [**Fig. 2(c)**]. These observations suggest that in the crystalline lattice, four R1EN-Ubs compose one unit: two Ub1s in contact-B relationships and the other two are placed in Ub2 [**Fig. 2(d)**]. The four R1EN-Ubs conformation is able to complete along the c-axis arrangement and compose one side of the hexagon, however, each side of the hexagon is independent in assembling the honeycomb structure, resulting in the overlapped electron density which seems to cause crushing of Ub1 at α1. Similarly, in $R1EN_{223}$-Ub, two Ub1s may be in contact-B and other two subunits are disordered.

The refined ubiquitin structures in all fusion proteins are almost identical to the reported structure (1UBQ) with the main chain Cα positional root mean square deviation range of 0.46–0.55 Å, indicating that the structure determined by the R1EN-fusion method is quite reliable. In conclusion, it is demonstrated that highly porous protein crystal could be used as a novel protein crystallization and structure determination tool by the fusion technique.

REFERENCES

[1] N. C. Seeman, *J. Theor. Biol.* **99**, 237 (1982).
[2] N. Maita, *J. Am. Chem. Soc.* **140**, 13546 (2018).
[3] N. Maita, H. Aoyagi, M. Osanai, M. Shirakawa and H. Fujiwara, *Nucl. Acids Res.* **35**, 3918 (2007).

BEAMLINE

AR-NW12A

N. Maita (Tokushima Univ.)