Biomaterials Science

Engineering of an in-cell protein crystal for fastening a metastable conformation of a target miniprotein

Journal:	*Biomaterials Science*
Manuscript ID	BM-ART-10-2022-001759.R1
Article Type:	Paper
Date Submitted by the Author:	14-Dec-2022

Complete List of Authors:	Kojima, Mariko; Tokyo Institute of Technology, Life Science and Technology; Tokyo Institute of Technology Abe, Satoshi; Tokyo Institute of Technology, Department of Biomolecular Engineering Furuta, Tadaomi; Tokyo Institute of Technology, School of Life Science and Technology Tran, Duy; Ton Duc Thang University, Department for Management of Science and Technology Development; Tokyo Institute of Technology - Ookayama Campus, Hirata, Kunio; RIKEN/SPring8-Center Yamashita, Keitaro; RIKEN/SPring8-Center Hishikawa, Yuki; Tokyo Institute of Technology Kitao, Akio; Tokyo Institute of Technology Ueno, Takafumi; Tokyo Institute of Technology,

ARTICLE

Engineering of an in-cell protein crystal for fastening a metastable conformation of a target miniprotein

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Mariko Kojima,[a] Satoshi Abe,[a] Tadaomi Furuta,[a] Duy Phuoc Tran,[a] Kunio Hirata,[b] Keitaro Yamashita,[b, d] Yuki Hishikawa,[a] Akio Kitao,[a] and Takafumi Ueno*[a,c]

Protein crystals can be utilized as porous scaffolds to capture exogenous molecules. Immobilization of target proteins using protein crystals is expected to facilitate X-ray structure analysis of proteins that are difficult to crystallize. One of the advantages of scaffold-assisted structure determination is the analysis of metastable structures that are not observed in solution. However, efforts to fix target proteins within pores of scaffold protein crystals have been limited due to the lack of strategies to control protein-protein interactions formed in the crystals. In this study, we analyze the metastable structure of the miniprotein, CLN025, which forms a β-hairpin structure in solution, using polyhedra crystal (PhC), an in-cell protein crystal. CLN025 is successfully fixed within the PhC scaffold by replacing the original loop region. X-ray crystal structure analysis and molecular dynamics (MD) simulation reveals that CLN025 is fixed as a helical structure in a metastable state by non-covalent interactions in the scaffold crystal. These results indicate that modulation of intermolecular interactions can trap various protein conformations in the engineered PhC and provides a new strategy for scaffold-assisted structure determination.

Introduction

Protein crystals have attracted attention in biomaterial science because their framework structures allow them to capture exogenous molecules.[1, 2] Protein crystals can serve as porous scaffold materials by retaining solvent channels that range from 30% to 65% of the volume of an entire crystal.[3] Various foreign molecules, such as organic compounds, metal complexes, metal nanoparticles, and proteins, can be captured within pores of protein crystals by soaking or co-crystallization methods.[4-6] Immobilizing target molecules at well-defined positions within protein crystals has led to successful X-ray diffraction structural characterization of various exogenous molecules with different sizes.[7-15]

One of the challenges of scaffold-assisted structure determination using protein crystals is to analyze target molecules with metastable structures that are not observed in solution.[6, 16] The interaction between the molecules and the scaffold crystals is expected to allow a target molecule to remain in a metastable state, such as the intermediate states of cofactors and the transition states in reactions of organic molecules.[17-20] For example, a biphenyl moiety immobilized in threonyl-tRNA synthetase as a non-canonical amino acid was trapped in its transition state by steric hindrance of the surrounding amino acids.[18] This indicates that the engineering of scaffold protein crystals can be harnessed to fix the metastable conformation of peptides and proteins. Characterization of such metastable conformations is important in context of understanding the folding process and the dynamics in protein/peptides. Although computational studies have identified various metastable states, several states have not been experimentally validated. A few attempts have been made to determine the structure of the flexible domains in the scaffold protein crystals.[21-23] However, determining the metastable structure of the target protein in the scaffold crystals remains challenging because of the lack of methodology to fix the state by controlling non-covalent interactions.

In-cell protein crystals are suitable candidates for development as molecular scaffolds for structure determination.[24-26] It has been reported that some proteins crystalize spontaneously in living cells when they are expressed at high concentrations.[27, 28] The rapid development of X-ray diffraction data collection and analysis, such as serial femtosecond and multiple-crystal synchrotron crystallography, enables the structure determination of various in-cell protein crystals of less than 10 μm.[26, 28-38] The high porosity and stability of in-cell protein crystals permit accumulation of foreign synthetic molecules and proteins.[24, 39-43] These advantageous properties will promote collection of diffraction data for the target molecules in the crystals. While the structures of metal complexes immobilized in the pores have been determined,[44, 45] the structures of exogenous proteins and peptides have not yet been reported because procedures for fixation of the

[a] School of Life Science and Technology, Tokyo Institute of Technology, Nagatsuta-cho 4259, Midori-ku, Yokohama 226-8501, Japan
[b] SR Life Science Instrumentation Unit, RIKEN/SPring-8 Center, 1-1-1, Kouto, Sayo-cho, Sayo-gun, Hyogo 679-5148, Japan
[c] International Research Frontiers Initiative (IRFI), Tokyo Institute of Technology, Nagatsuta-cho 4259, Midori-ku, Yokohama 226-8501, Japan
[d] MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, UK
†Electronic supplementary information (ESI) available. See DOI: 10.1039/x0xx00000x

proteins and peptides have not been established.[24-26] Therefore, methods for retention of metastable structures of target proteins using in-cell protein crystals for their structural analysis are needed.

Here, we report the X-ray crystallographic analysis of CLN025, a miniprotein consisting of 10 amino acids,[46] by fusing it with cypovirus polyhedra crystal (PhC), an in-cell protein crystal of the polyhedrin monomers (PhM).[29] PhC is a porous protein crystal that enables the immobilization of a wide range of sizes of exogenous molecules, including organic molecules and enzymes.[44, 47, 48] CLN025 is a well-known protein that spontaneously folds into a β-hairpin structure in solution through intramolecular interactions. The MD simulation of CLN025 has suggested the folding pathway from the unfolded structure with the elongated distance between N- and C-termini to the β-hairpin structure, which is the most stable state.[49] However, the intermediate structures in which N- and C-termini are fixed in apart have not been determined experimentally. If the structures are experimentally observed as metastable structures, the conformational change of CLN025 can be understood in detail. Here, we use the term metastable to refer to a stable structure for a protein subdomain that can be observed by a constrained environment in protein crystal. We designed mutant PhMs fused to CLN025 by deleting amino acid residues at the interface in PhC because it forms a space for immobilizing target molecules while retaining the original crystal lattice structure.[44, 47, 48] We succeeded in the in-cell crystallization and high-resolution structure determination of PhC mutants fused with CLN025 (Fig. 1). The observed structure of CLN025 contained the partial helix and the unusual dihedral angles (ϕ, ψ) of D74, that is different from the previously reported β-hairpin turns of CLN025. MD simulations reveal that the structure of CLN025 in PhC obtained by X-ray structure analysis is the metastable structure retained by the

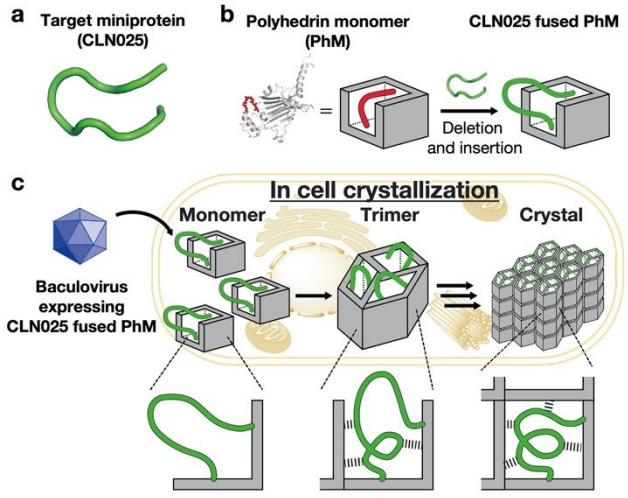

Fig. 1 In-cell crystal engineering for fixing a metastable structure of a target miniprotein. (a) Structure of the target miniprotein, CLN025 (PDB ID: 5AWL). (b) Structure of polyhedrin monomer, PhM (PDB ID: 5GQM) and construction of CLN025 fused PhM. (c) Schematic representation showing the in-cell crystallization of CLN025 fused PhM and structure determination to visualize the real-space structure of a metastable conformer of CLN025. The red loop indicates a selected region which was replaced by CLN025 (green loop).

Intermolecular interactions at the interface with scaffold proteins in the crystal. These results clarified the necessary factors for a target miniprotein to be locked in the metastable state which would not have been predicted by calculations within scaffold protein crystals, suggesting that designing interactions between the target and scaffold proteins in the crystal enables characterization of dynamic structures of various proteins and peptides. Furthermore, the success of this in-cell protein crystal method suggests that it has potential for development as a powerful tool for new high-throughput structural analysis methods.

Results

Crystal design

To determine the metastable structure of CLN025 in PhC, we designed fusion proteins of CLN025 with the polyhedrin monomer (PhM). CLN025 comprises ten amino acids (YYDPETGTWY) and forms a β-hairpin structure in solution by intramolecular interactions (Fig. 2a).[46] The crystal structure of a single CLN025 molecule (CLN025^{5AWL}, PDB ID: 5AWL) shows that the distance of N/Y1–N/Y10 is 4.4 Å (Fig. 2a). We attempted to control the N–N distance in the fused PhM to change the original folding of CLN025. We selected loop 1 (L1) from E70 to N77 in the wild-type PhM (WT-PhM) to fuse the CLN025 fragment into PhC because L1, which has a high *B* factor, is not expected to inhibit the crystallization of PhC (Fig. 2b). We designed three PhM mutants, **1-PhM**, **2-PhM**, and **3-PhM**, in which the E70-N77 (Site 1), the R72-H76 (Site 2), and the E73-G74 (Site 3), of PhM, were replaced with CLN025 (Fig. 2c and Table 1). The distances of N/E70–N/N77, N/R72–N/H76, and N/E73–N/G74 of WT-PhM are 8.6 Å, 7.1 Å, and 3.5 Å, respectively (Fig. 2c). We also introduced a EAAK linker on both sides at Site 1 (**1-PhM**) named **4-PhM** (Fig. 2c and Table 1) by deleting an alanine residue from the rigid helical linker (EAAAK) to fix CLN025 into the restricted space of scaffold protein crystals.[50] Four CLN025 fused PhCs, **1-PhC**, **2-PhC**, **3-PhC**, and **4-PhC**, were successfully obtained in the *Spodoptera frugiperda* cell line IPLB-Sf21-AE (Sf21) (Fig. S1). We obtained crystals with a size range of 0.3-2.5 μm (Fig. S2). SDS-PAGE indicates that each crystal consists of the corresponding protein (Fig. S3). The MALDI-TOF MS spectra of the crystals were assigned to each mutant PhC (Fig. S4). These results show that CLN025 was successfully fused within PhCs without preventing

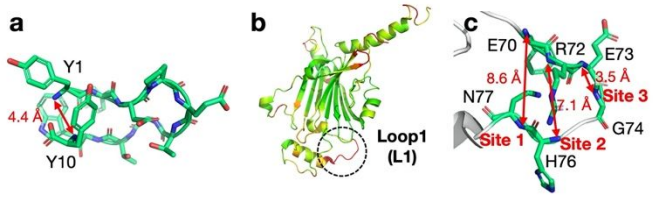

Fig. 2 Molecular structures of target and scaffold proteins. (a) Crystal structure of CLN025^{5AWL}. (b) Crystal structure of WT-PhM (PDB ID: 5GQM). (c) Three sites in the L1 region fused with CLN025 and the N–N distances. N and O atoms in (a) and (c) are colored in blue and red, respectively. The cartoon model in (b) is colored by the *B* factor with values ranging from 6Å^2 (green) to 15 Å^2 (red).

Table 1. Amino acid sequences of L1 regions of the designed mutant crystals of WT-PhM, **1-PhM**, **2-PhM**, **3-PhM**, **4-PhM**, and **2-PhM^{D74R}** and summary of crystal structure analysis.

Mutant	Sequence[a]	Insert structure	Resolution (Å)	PDB ID
	51 61 71 81 91			
WT-PhM	-INKNSQYKVG ISETFKALKE YRKGQHNDSY DEYEVNQSIY YPNGGDARKF-	-	-	5GQM
1-PhM	-INKNSQYKVG ISETFKALK**Y YDPETGTWY**D SYDEYEVNQS IYYPNGGDAR-	disordered	1.55	-
2-PhM	-INKNSQYKVG ISETFKALKE Y**YYDPETGTW Y**NDSYDEYEV NQSIYYPNGG-	ordered	1.75	7WYR
3-PhM	-INKNSQYKVG ISETFKALKE YR**YYDPETGT WY**QHNDSYDE YEVNQSIYYP-	disordered	1.70	-
4-PhM	-INKNSQYKVG ISETFKALKE AAK**YYDPETG TWY**EAAKDSY DEYEVNQSIY-	disordered	1.85	-
2-PhM^{D74R}	-INKNSQYKVG ISETFKALKE Y**YYRPETGTW Y**NDSYDEYEV NQSIYYPNGG-	disordered	1.70	-

[a] The CLN025 sequences are underlined.

crystallization.

Structural determination of the CLN025 fused PhCs

To determine the crystal structures of the mutant PhCs, the microcrystals were diffracted using a micro-X-ray beam of the SPring-8 BL32XU beamline equipped with multiple-crystal synchrotron crystallography (MSX) and ZOO, which is an automatic data processing system.[51] The crystal structures of **1-PhC**, **2-PhC**, **3-PhC**, and **4-PhC** were determined at 1.55-1.85 Å resolution (Table S1 and Fig. S5a-d). All structures have the same space group (*I*23) and lattice constants comparable to WT-PhC. The root-mean-square deviations of the Cα atoms (Cα-RMSD) of all the mutant PhCs from WT-PhC were less than 0.3 Å. The electron density corresponding to the CLN025 fragment in **2-PhC** (**CLN025$^{2\text{-PhC}}$**) was observed between Y71 and N82 (Fig. 3a-c). The full-length **CLN025$^{2\text{-PhC}}$** was determined based on the $2|F_o|-|F_c|$ and $|F_o|-|F_c|$ electron density maps except for the side chains of W80(W9) (ESI). The structure of **2-PhM** indicates that **CLN025$^{2\text{-PhC}}$** fused in **2-PhM-i** contacts two other **2-PhMs**; **2-PhM-ii** within a trimer, and **2-PhM-iv** in the neighboring trimer (Fig. 3a and b). Although only Y70(Y1) and Y79(Y10) of the CLN025 fragments in **1-PhC** were observed, all the fragments in **3-PhC**, and **4-PhC** could not be determined due to a lack of the corresponding electron density (Fig. S6a-c).

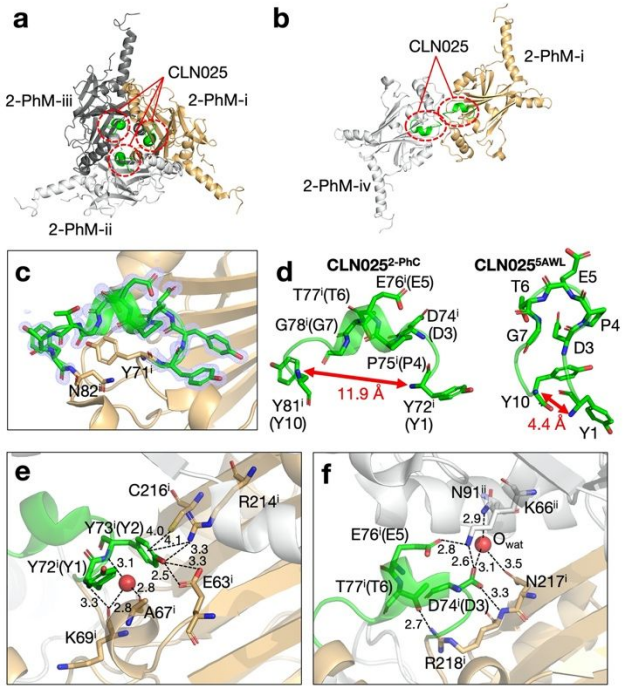

Fig. 3 Crystal structure of **2-PhC** (PDB ID: 7WYR). (a) Structure of the **2PhM** trimer (**2-PhT**). The monomers are designated **2-PhM-i**, **2-PhM-ii,** and **2-PhM-iii**. **CLN025$^{2\text{-PhC}}$s** are located at the interface of **2-PhM-i/2-PhM-ii**, **2-PhM-ii/2-PhM-iii**, and **2-PhM-iii/2-PhM-i**. (b) Structure of two **2-PhMs**, **2-PhM-i** and **2-PhM-iv**, in contact at the interface of two **2-PhTs**. (c) Close-up view of the **CLN025$^{2\text{-PhC}}$** fusion site of **2-PhM-i**. (d) Structure comparison of **CLN025$^{2\text{-PhC}}$** and CLN025^{5AWL}. (e,f) Non-covalent interactions between **CLN025$^{2\text{-PhC}}$** and the surrounding residues at the positions of (e) Y72^{i}(Y1) and Y73^{i}(Y2), and (f) D74^{i}(D3)–T77^{i}(T6). **CLN025$^{2\text{-PhC}}$** is colored green. The monomers interacting with **CLN025$^{2\text{-PhC}}$** are displayed as white cartoon models. The selected $2|Fo|-|Fc|$ electron density maps at 1.0σ are shown in blue. An oxygen atom of a water molecule (O_{wat}) is displayed as a red sphere. The red arrows in (d) show the distance of N/Y72^{i}(Y1)–N/Y81^{i}(Y10) in **CLN025$^{2\text{-PhC}}$** and N/Y1–N/Y10 in CLN025^{5AWL}, respectively. The hydrogen bonds are shown as black dashed lines.

Structure of CLN025 in 2-PhC

CLN025$^{2\text{-PhC}}$ folds differently from the β-hairpin structure of CLN025^{5AWL} (Fig. 3d).[46] The Cα-RMSD value of **CLN025$^{2\text{-PhC}}$** from CLN025^{5AWL} is 6.2 Å. The distance of N/Y72^{i}(Y1)–N/Y81^{i}(Y10) of **CLN025$^{2\text{-PhC}}$** (11.9 Å) is longer than that of CLN025^{5AWL} (4.4 Å). The structural change suggests fixation of an alternate state of CLN025 in **2-PhC** from the original stable structure. The Ramachandran plot of Y73^{i}(Y2)–W80^{i}(W9) of **CLN025$^{2\text{-PhC}}$** reveals that the dihedral angles (ϕ, ψ) of P75^{i}(P4), E76^{i}(E5), T77^{i}(T6), and G78^{i}(G7) are (-67.6, -22.5), (-61.8, -16.7), (-86.2, -8.8), and (-100.4, 5.9), respectively, which are in the helical region (right-handed) (Fig. S7a).[52] Since the dihedral angles of P75^{i}(P4), E76^{i}(E5), and T77^{i}(T6) in **CLN025$^{2\text{-PhC}}$** are located in the same region as CLN025^{5AWL} (Fig. S7b), the formation of the helical structure is attributed to a change in the dihedral angle of G78^{i}(G7) from CLN025^{5AWL}. The dihedral angles of adjacent D74^{i}(D3) (59.5, 51.0) and T79^{i}(T8) (-42.4, -59.1) in **CLN025$^{2\text{-PhC}}$** are significantly changed from D3 (-75.6, 133.6) and T8 (-104.0, 153.3) in CLN025^{5AWL} (Fig. S7). Interestingly, the results show that D74^{i}(D3) is fixed in the left-handed helix conformation in the crystal.[53] Generally, aspartic acid and asparagine are more likely to adopt a left-handed helix conformation than other amino acids except for glycine.[54] One of the reasons was

suggested that an intra-residue interaction between the side and main chain carbonyl groups stabilizes the left-handed helix conformation. However, such interaction was not observed in **CLN025$^{2\text{-PhC}}$**. It indicates that the left-handed helix conformation of D74^{i}(D3) of **CLN025$^{2\text{-PhC}}$** is stabilized by restricted environment in **2-PhC**.

The conformation of **CLN025$^{2\text{-PhC}}$** is fixed by the non-covalent interactions in **2-PhC** (Fig. 3e, f, S8, and Table S2). At the N-terminus of **CLN025$^{2\text{-PhC}}$**, Y73^{i}(Y2) retains its SH---π interaction with Sγ/C216^{i} (4.0 Å), and its NH---π interaction with Nη2/R214^{i} (4.1 Å) (Fig. 3e).[55] In addition, Oη/Y73^{i}(Y2), and N/Y73^{i}(Y2) participate in a hydrogen-bonding network with R214^{i}, E63^{i}, A67^{i}, K69^{i}, and water (Fig. 3e).[56] At the C-terminus of **CLN025$^{2\text{-PhC}}$**, Oη/Y81^{i}(Y10) forms a hydrogen bond with Oδ2/D86^{iv} (2.3 Å) (Fig. S8). The main chain of W80^{i}(W9) forms a hydrogen bond of O/W80^{i}(W9) –N/D83^{iv} (3.3 Å) (Fig. S8). At the middle region of **CLN025$^{2\text{-PhC}}$**, the hydrogen bond at O/P75^{i}(P4)–N/G78^{i}(G7) (3.0 Å) and the salt bridges at Oδ1/D74^{i}(D3)–Nζ/K66^{ii} and Oε1/E76^{i}(E5)-Nζ/K66^{ii} (2.6 Å and 2.8 Å, respectively) are formed, which retains the 3^{10} helix from P75^{i} to G78^{i} (Fig. 3f and S9).[57] Oδ1/D74^{i}(D3), Oδ2/D74^{i}(D3), and Oγ1/T77^{i}(T6) participate in a hydrogen bonding network with N91^{ii}, N217^{i}, R218^{i}, and water molecules (Fig. 3f).[56] To clarify the effect of the non-covalent interactions of D74^{i}(D3) in **CLN025$^{2\text{-PhC}}$**, the D74R mutant (**2-PhC^{D74R}**) was crystallized (Table 1, ESI). While the crystal structure of **2-PhC^{D74R}** was determined with a resolution of 1.70 Å, the CLN025 fragment of **2-PhC^{D74R}** (**D74R-CLN025**) was not observed due to a lack of electron density corresponding to the Y72(Y1)–Y81(Y10) (Fig. S5e, S6d, and Table S1). Moreover, the conformation of the main chain of Y71 of **2-PhC^{D74R}** is altered from that of **2-PhC** (Fig. S10). Structural changes were also observed on K66^{ii}, R103^{iv}, K104^{iv}, and R214^{i} surrounding **D74R-CLN025** (Fig. S10). This result suggests that the D74^{i}(D3) residue is crucial to fix the conformation of **CLN025$^{2\text{-PhC}}$**.

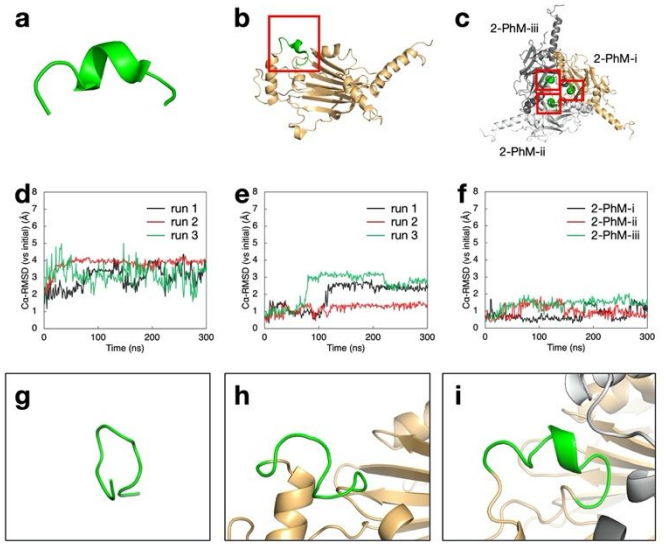

Fig. 4 Analysis of conformations of **CLN025$^{2\text{-PhC}}$** in MD simulations. (a-c) The initial structures in MD simulations of (a) the **CLN025$^{2\text{-PhC}}$** fragment (isolated), (b) **2-PhM** (monomer) and (c) **2-PhT** (trimer). (d-f) Time courses of the Cα-RMSD values of **CLN025$^{2\text{-PhC}}$** in (d) the isolated system, (e) **2-PhM** and (f) **2-PhT** from the initial structures. (g-i) the structures at 300 ns in MD simulations of (g) the isolated system, (h) **2-PhM** and (h) **2-PhT**. Red squares in (b) and (c) show the positions of the **CLN025$^{2\text{-PhC}}$** fragments. Three runs (run 1, 2, and 3) were performed for the isolated system and **2-PhM**, and their data are colored with black, red, and green, respectively (d-e). Data from each monomer, **2-PhM-i**, **-ii**, and **-iii**, in **2-PhT** are colored in black, red, and green, respectively (f).

All-atom MD simulation

All-atom MD simulations using AMBER were performed to investigate the fixation of **CLN025$^{2\text{-PhC}}$** by the interactions with surrounding residues in **2-PhC**. Three systems of the isolated **CLN025$^{2\text{-PhC}}$** fragment, the monomer **2-PhM**, and the trimer of **2-PhM** (**2-PhT**) were subjected to 300 ns MD simulations (Fig. 4, S11, and S12, ESI). The Cα-RMSD value of **CLN025$^{2\text{-PhC}}$** from the initial structure in the isolated system is higher than the corresponding value for **2-PhM** and **2-PhT** (Fig. 4d-f). For the isolated system, the **CLN025$^{2\text{-PhC}}$** fragment did not converge to specific conformation in runs 1 and 3 within the 300 ns simulation (Fig. S11a and S13), while the fragment in run 2 changed to a β-hairpin-like structure formed by intramolecular hydrogen bonds of O/D74^{i}(D3)–N/G78^{i}(G7), O/D74^{i}(D3)–N/T79^{i}(T8), O/P75^{i}(P4)–N/G78^{i}(G7) (Fig. 4g, S11a, S12a, and S12b). On the other hand, the dihedral angles (ϕ and ψ) of D74^{i}(D3) in Ramachandran plot were changed to a β-sheet region from the left-handed helix in all runs (Fig. S13). Moreover, the angles of P75^{i}(P4), E76^{i}(E5), T77^{i}(T6), T79^{i}(T8), and W80^{i}(W9) in the Ramachandran plot also fluctuate between the β-sheet and the helix in runs 1 and 3 because of the conformational flexibility, while the initial region in the Ramachandran plot was maintained in run 2 (Fig. S13). These observations suggest that the molecular environments of **2-PhM** and **2-PhT** play an important role in fixing **CLN025$^{2\text{-PhC}}$**.

The simulations of **2-PhM** and **2-PhT** indicate restraint

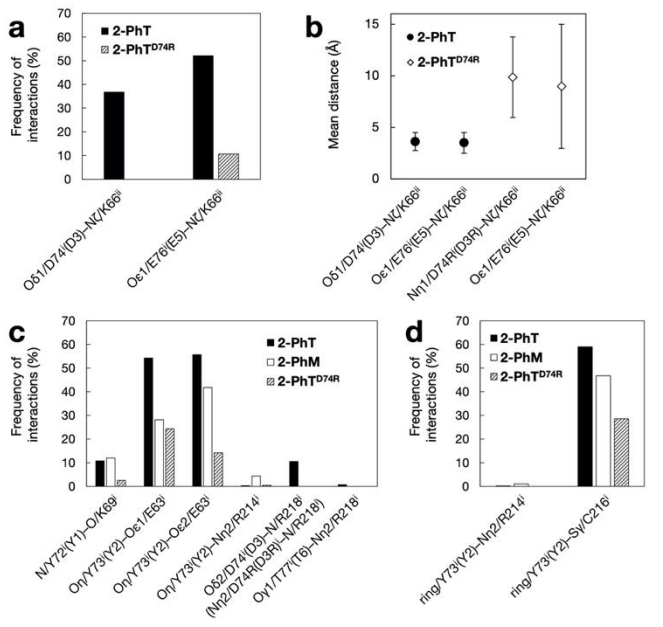

Fig. 5 Analysis of non-covalent interactions stabilizing **CLN025$^{2\text{-PhC}}$** in **2-PhC** in MD simulations. (a) The frequency of salt bridge formation at Oδ1/D74^{i}(D3) (Nη1/D74R^{i})–Nζ/K66^{ii} and Oε1/E76^{i}(E5)–Nζ/K66^{ii} in **2-PhT** and **2-PhT^{D74R}**. (b) The mean distance of the salt bridges at Oδ1/D74^{i}(D3) (Nη1/D74R^{i})–Nζ/K66^{ii} and Oε1/E76^{i}(E5)–Nζ/K66^{ii} in **2-PhT** and **2-PhT^{D74R}**. The frequency of intramolecular (c) hydrogen bonds and (d) NH---π and

SH---π interactions in **2-PhT**, **2-PhM**, and **2-PhT^{D74R}**. The frequencies indicate the percentages of structures in which non-covalent interactions were formed in each MD trajectory (regarding trimers, data for all three **CLN025$^{2\text{-PhC}}$** were included). The cut-off distances of non-covalent interactions are 4.0 Å for salt bridges, 3.5 Å for hydrogen bonds, and 4.2 Å for SH---π and NH---π interactions.[55-57] The cut-off angle for salt bridges and hydrogen bonds is 135°.

behavior of **CLN025$^{2\text{-PhC}}$** by interacting with the other regions in the scaffold proteins. The Cα-RMSD values of **CLN025$^{2\text{-PhC}}$** in **2-PhM** from initial structure vary in the range of 1-3Å (Fig. 4e). For **2-PhM**, runs 1 and 3 demonstrate a structural change of **CLN025$^{2\text{-PhC}}$** to generate a loop structure with conversion of the dihedral angles of D74^{i}(D3) from helix to β-sheet (Fig. 4h, S11b, and S14). The dihedral angles of E76^{i}(E5) and W80^{i}(W9) were retained in the initial state over the simulation in all runs, in contrast to their flexibility in the isolated system (Fig. S14). For **2-PhT**, the Cα-RMSD values of **CLN025$^{2\text{-PhC}}$** were retained at less than 2 Å (Fig. 4f) and the structures maintained their initial conformations in all runs (Fig. 4i, S11c, and S15). The high structural conservation of **CLN025$^{2\text{-PhC}}$** in **2-PhT** suggests that intermolecular interactions within the trimer trigger the fixation of the helical structure of **CLN025$^{2\text{-PhC}}$**.

Based on the crystal structure of **2-PhC**, two intermolecular salt bridges, Oδ1/D74^{i}(D3)–K66^{ii} and Oε1/E76^{i}(E5)–K66^{ii}, are expected to be essential to fix **CLN025$^{2\text{-PhC}}$** in **2-PhT** (Fig. 3f). The frequencies of the salt bridges of Oδ1/D74^{i}(D3)–Nζ/K66^{ii} and Oε1/E76^{i}(E5)–Nζ/K66^{ii} are 36.8% and 52.1%, respectively (Fig. 5a). The mean distances of Oδ1/D74^{i}(D3)–Nζ/K66^{ii} and Oε1/E76^{i}(E5)–Nζ/K66^{ii} during 300 ns simulation are 3.62±0.88 Å and 3.50±1.00 Å, respectively (Fig. 5b and S17).[57] These results suggest that stable bridge geometry is maintained during the entire run time. The slightly higher probability value of the salt bridge of Oε1/E76^{i}(E5)–Nζ/K66^{ii} suggests that E76^{i}(E5)–K66^{ii} mainly contributes to stabilizing the extended structure of **CLN025$^{2\text{-PhC}}$**, which is reinforced by D74^{i}(D3). Comparing the frequency of the other eight intramolecular interactions of **CLN025$^{2\text{-PhC}}$** between **2-PhM** and **2-PhT** shows that the hydrogen bonds of Y73^{i}(Y2)–E63^{i} and the SH---π interaction of Y73^{i}(Y2)–C216^{i} in **2-PhT** exhibit a higher probability value than in **2-PhM** (Fig. 5c and d). Therefore, these interactions are not enough to fix the conformation of CLN025 but enhance the stability of the partially helical structure of **CLN025$^{2\text{-PhC}}$** by the salt bridges at Oδ1/D74i(D3)–Nζ/K66^{ii} and Oε1/E76^{i}(E5)–Nζ/K66^{ii} at the interface of the trimers.

To evaluate the effect of the salt bridge at D74^{i}(D3) for structural fixation of **CLN025$^{2\text{-PhC}}$** in **2-PhC**, an MD simulation of **2-PhT^{D74R}** was performed. The Cα-RMSD value of the **D74R-CLN025** fragment in a monomer of **2-PhT^{D74R}** undergoes more significant fluctuation than in **2-PhT** (Fig. S18). The lack of a salt bridge of D74^{i}(D3)–K66^{ii} and electrostatic repulsion between D74R^{i}(D3R) and K66^{ii} has the effect of decreasing the frequency of other intermolecular and intramolecular interactions (Fig. 5a, c, d, and S19). The mean distances of Nη1/D74R^{i}(D3)–Nζ/K66^{ii} and Oε1/E76^{i}(E5)–Nζ/K66^{ii} during 300 ns simulation are increased to 9.87±3.90 Å and 8.98±6.02 Å, respectively (Fig. 5b). Therefore, D74^{i}(D3) plays a crucial role in fixing **CLN025$^{2\text{-PhC}}$** into the helical structure, consistent with the experimental results.

Discussion

The **2-PhC** structure shows that **CLN025$^{2\text{-PhC}}$** retains a unique extended structure with a partial helical conformation at P75^{i}(P4)–G78^{i}(G7), which is quite different from the hairpin structure of CLN025^{5AWL} (Fig. 3a).[46] It was reported that folding of CLN025^{5AWL} is triggered by the formation of a turn from P4 to G7 followed by formation of hydrogen bonds involving D3 and other residues, and finally stabilized by hydrophobic interactions at the N terminus and the C terminus.[49, 58] In **CLN025$^{2\text{-PhC}}$**, the helix at P75^{i}(P4)–G78^{i}(G7) is stabilized by the hydrogen bond of O/P75^{i}(P4)–N/G78^{i}(G7) (Fig. S9 and Table S3). Additional hydrogen bonds are essential for the β-hairpin formation of CLN025 as suggested by MD simulations.[49] The hydrogen bond of O/P75^{i}(P4)–N/G78^{i}(G7), which is not retained in CLN025^{5AWL}, is a specific interaction formed in the metastable structure predicted from the MD simulation of the CLN025 fragment. The MD simulation using only the **CLN025$^{2\text{-PhC}}$** fragment structure composed of 10 residues includes a fold of the protein into a β-hairpin structure by the hydrogen bonds at O/D74^{i}(D3)–N/G78^{i}(G7) and O/D74^{i}(D3)–N/T79^{i}(T8) spontaneously formed (Fig. S12a and b). The formation of these hydrogen bonds is consistent with the folding pathway of CLN025^{5AWL} revealed by a previous MD simulation study.[49] Therefore, **CLN025$^{2\text{-PhC}}$** appears to be a metastable structure, in which D74^{i}(D3) is prevented from interacting with other residues promoting β-hairpin formation in **2-PhC**.

The metastable structure of **CLN025$^{2\text{-PhC}}$** in **2-PhC** is maintained by intramolecular interactions of **2-PhM**, and intermolecular interactions at the interfaces of **2-PhM-i/2-PhM-ii** and **2-PhM-i/2-PhM-iv** (Fig. 3). A previous study has reported that CLN025 is used as a bending segment of the *de novo* protein, AF.p17.[59] AF.p17 co-crystallizes with a CH2-CH3 domain (CH2-CH3) of the human immunoglobulin G, indicating that Y33(Y1)–W41(W9) (which corresponds to CLN025) forms a β-hairpin structure with hydrogen bonds at N/D35(D3)–O/T40(T8) and Oγ1/D35(D35)–N/G39(G7) (Fig. S20a, b, and Table S3, PDB ID: 3WKN).[59] The CLN025 fragment in AF.p17 is fixed by an intramolecular interaction at Y33(Y1) and Y34(Y2) and intermolecular interactions at Y33(Y2), D35(D3), and E37(E5), while retaining the β-hairpin structure (Fig. S20c, d, and Table S3). Y34(Y2) forms intramolecular hydrogen bonds at N/Y34(Y2)–O/W18 and O/Y34(Y2)–Oγ1/T17 and includes β-sheet structure from E27 to T32 (Fig. S20c and Table S3). These results indicate that the **2-PhM** fusion site makes **CLN025$^{2\text{-PhC}}$** an extended conformation, inhibiting the formation of hydrogen bonds that allow the fragment to form a β-hairpin conformation.

The intermolecular salt bridges fix the helical structure of **CLN025$^{2\text{-PhC}}$** at D74^{i}(D3)–K66^{ii} and E76^{i}(E5)–K66^{ii} (Fig. 3f). The MD simulation for **2-PhT** revealed that these salt bridges retain the helix structure from P75^{i}(P4) to G78^{i}(G7) of three **CLN025$^{2\text{-PhC}}$s** in **2-PhT** over the entire run time (Fig. 5a and b). Two salt bridges act as an anchor to guide formation of subsequent intramolecular hydrogen bonds and an SH---π interaction of Y73^{i}(Y2)–E63^{i} and Y73^{i}(Y2)–C216^{i}, respectively (Fig. 5c and d). The low electron density observed corresponding to **D74R-**

CLN025 in **2-PhC^{D74R}** indicates that a missing salt bridge at Oδ1/D74^{i}(D3)–Nζ/K66^{ii} destabilizes the helical structure of **CLN025$^{2\text{-PhC}}$** (Fig. S6d). The MD simulation of **2-PhT^{D74R}** indicates that the electrostatic repulsion around D74R^{i} influences the multi-conformation of CLN025 fragment (Fig. S19). On the other hand, in the CLN025 fragment of AF.p17, D35(D3) forms an intermolecular salt bridge which contributes to stabilization of the β-hairpin conformation (Fig. S20d).[59] This result suggests that the exact arrangement of salt bridges in the crystal controls the conformation of the CLN025 fragment. Thus, the lack of electron density of CLN025 in **1**, **3**, **4-PhC**, and **2-PhC^{D74R}** could be due to failure of D74^{i}(D3) and K66^{ii} to adopt configurations which induce formation of the intermolecular salt bridges.

Conclusion

In conclusion, the metastable structure of CLN025 was successfully fastened using non-covalent interactions in a protein crystal scaffold. The crystal structure analysis reveals that CLN025, which has a hairpin structure in solution, takes on an extended structure including a helix in **2-PhC**. The MD simulation and mutation experiments of **2-PhC** indicate that intermolecular salt bridges in **2-PhT** are essential for maintaining the extended structure of CLN025, which is further stabilized by intramolecular interactions in **2-PhM**. Our findings will lead to new methods to trap and analyze intermediate or transition state conformations of proteins, which have been studied only by transient observations, by reconstructing the environment that induces various intermolecular interactions in protein crystals. Furthermore, this method can be expanded to enhance structure determination and stabilization of metastable proteins, such as amyloid-β oligomers and intrinsically disordered proteins. An additional advantage of in-cell protein crystals, such as PhC, is that they can be adapted to construct pipelines for automated crystallization and structure determination.[45, 60] Thus, engineering of in-cell protein crystals is expected to open up new areas of structural biology and biomaterials science.

Author Contributions

K.M., S.A., and T.U. designed the research. K.M., T. F., D.P.T., K.H., K.Y., and Y.H. carried out the experiments. All authors analyzed the data and discussed the results. K.M. S.A., and T.U. wrote the main manuscript text. All authors reviewed the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Nos. JP19H02830, JP20K21244, and Grant-in-Aid for Scientific Research on Innovative Areas “Molecular Engines” (JP18H05421) to T.U. and JP18K05140 to S.A., and Adaptable and Seamless Technology Transfer Program through Target-driven R&D (JPMJTR20U1) from the Japan Science and Technology Agency to T.U. Synchrotron radiation experiments were conducted under the approval of 2019A2561, 2019B2561, and 2020A2541 at SPring-8. This work was supported by SUNBOR Grant from Suntory Foundation for Life Sciences to M.K. This research was partially supported by Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)) from AMED under Grant Number JP21am0101070 (support number 1854).

Notes and references

1. A. L. Margolin and M. A. Navia, *Angew. Chem. Int. Ed.*, 2001, **40**, 2204-2222.
2. A. R. Ward and C. D. Snow, *Curr. Opin. Struct. Biol.*, 2020, **60**, 85-92.
3. L. Z. Vilenchik, J. P. Griffith, N. St Clair, M. A. Navia and A. L. Margolin, *J. Am. Chem. Soc.*, 1998, **120**, 4290-4294.
4. S. Abe, B. Maity and T. Ueno, *Curr. Opin. Chem. Biol.*, 2018, **43**, 68-76.
5. L. F. Hartje and C. D. Snow, *Wiley Interdiscip. Rev. Nanomed.*, 2019, **11**, e1547.
6. M. Kojima, S. Abe and T. Ueno, *Biomater. Sci.*, 2022, **10**, 354-367.
7. T. Koshiyama, N. Kawaba, T. Hikage, M. Shirai, Y. Miura, C. Y. Huang, K. Tanaka, Y. Watanabe and T. Ueno, *Bioconjugate Chem.*, 2010, **21**, 264-269.
8. T. R. Huber, E. C. McPherson, C. E. Keating and C. D. Snow, *Bioconjugate Chem.*, 2018, **29**, 17-22.
9. T. Matsumoto, R. Nakashima, A. Yamano and K. Nishino, *Biochem. Biophys. Res. Commun.*, 2019, **518**, 402-408.
10. H. Tabe, S. Abe, T. Hikage, S. Kitagawa and T. Ueno, *Chem. Asian J.*, 2014, **9**, 1373-1378.
11. H. Tabe, K. Fujita, S. Abe, M. Tsujimoto, T. Kuchimaru, S. Kizaka-Kondoh, M. Takano, S. Kitagawa and T. Ueno, *Inorg. Chem.*, 2015, **54**, 215-220.
12. S. I. Mann, T. Heinisch, T. R. Ward and A. S. Borovik, *Chem. Commun.*, 2018, **54**, 4413-4416.
13. P. Ecsedi, G. Gogl, H. Hof, B. Kiss, V. Harmat and L. Nyitray, *Structure*, 2020, **28**, 943-953.
14. P. Ernst, A. Pluckthun and P. R. E. Mittl, *Sci. Rep.*, 2019, **9**, 15199.
15. N. Maita, *J. Am. Chem. Soc.*, 2018, **140**, 13546-13549.
16. S. Koide, *Curr. Opin. Struct. Biol.*, 2009, **19**, 449-457.
17. C. Cavazza, C. Bochot, P. Rousselot-Pailley, P. Carpentier, M. V. Cherrier, L. Martin, C. Marchi-Delapierre, J. C. Fontecilla-Camps and S. Menage, *Nat. Chem.*, 2010, **2**, 1069-1076.
18. A. D. Pearson, J. H. Mills, Y. F. Song, F. Nasertorabi, G. W. Han, D. Baker, R. C. Stevens and P. G. Schultz, *Science*, 2015, **347**, 863-867.
19. I. Schlichting, J. Berendzen, K. Chu, A. M. Stock, S. A. Maves, D. E. Benson, B. M. Sweet, D. Ringe, G. A. Petsko and S. G. Sligar, *Science*, 2000, **287**, 1615-1622.
20. A. Volbeda, J. S. Cabodevilla, C. Darnault, O. Gigarel, T. H. L. Han, O. Renoux, O. Hamelin, S. Agnier-de-Choudens, P. Amara and J. C. Fontecilla-Camps, *ACS Chem. Biol.*, 2018, **13**, 1209-1217.

21. K. Takano, S. Endo, A. Mukaiyama, H. Chon, H. Matsumura, Y. Koga and S. Kanaya, *FEBS J.*, 2006, **273**, 150-158.
22. R. Matsuoka, A. Shimada, Y. Komuro, Y. Sugita and D. Kohda, *Protein Sci.*, 2016, **25**, 754-768.
23. S. Bala, S. Shinya, A. Srivastava, M. Ishikawa, A. Shimada, N. Kobayashi, C. Kojima, F. Tama, O. Miyashita and D. Kohda, *Biochim. Biophys. Acta. Gen. Subj.*, 2020, **1864**, 129418.
24. Y. Baskaran, K. C. Ang, P. V. Anekal, W. L. Chan, J. M. Grimes, E. Manser and R. C. Robinson, *Nat. Commun.*, 2015, **6**, 8681.
25. B. Heater, Z. F. Yang, M. Lee and M. K. Chan, *J. Am. Chem. Soc.*, 2020, **142**, 9879-9883.
26. N. Nagaratnam, Y. Y. Tang, S. Botha, J. Saul, C. F. Li, H. Hu, S. Zaare, M. Hunter, D. Lowry, U. Weierstall, N. Zatsepin, J. C. H. Spence, J. Qiu, J. LaBaer, P. Fromme and J. M. Martin-Garcia, *Acta Crystallogr. Sect. F:Struct. Biol. Commun.*, 2020, **76**, 278-289.
27. J. P. K. Doye and W. C. K. Poon, *Curr. Opin. Colloid Interface Sci.*, 2006, **11**, 40-46.
28. R. Schonherr, J. M. Rudolph and L. Redecke, *Biol. Chem.*, 2018, **399**, 751-772.
29. F. Coulibaly, E. Chiu, K. Ikeda, S. Gutmann, P. W. Haebel, C. Schulze-Briese, H. Mori and P. Metcalf, *Nature*, 2007, **446**, 97-101.
30. F. Coulibaly, E. Chiu, S. Gutmann, C. Rajendran, P. W. Haebel, K. Ikeda, H. Mori, V. K. Ward, C. Schulze-Briese and P. Metcalf, *Proc. Natl. Acad. Sci. U.S.A.*, 2009, **106**, 22205-22210.
31. L. Redecke, K. Nass, D. P. DePonte, T. A. White, D. Rehders, A. Barty, F. Stellato, M. N. Liang, T. R. M. Barends, S. Boutet, G. J. Williams, M. Messerschmidt, M. M. Seibert, A. Aquila, D. Arnlund, S. Bajt, T. Barth, M. J. Bogan, C. Caleman, T. C. Chao, R. B. Doak, H. Fleckenstein, M. Frank, R. Fromme, L. Galli, I. Grotjohann, M. S. Hunter, L. C. Johansson, S. Kassemeyer, G. Katona, R. A. Kirian, R. Koopmann, C. Kupitz, L. Lomb, A. V. Martin, S. Mogk, R. Neutze, R. L. Shoeman, J. Steinbrener, N. Timneanu, D. J. Wang, U. Weierstall, N. A. Zatsepin, J. C. H. Spence, P. Fromme, I. Schlichting, M. Duszenko, C. Betzel and H. N. Chapman, *Science*, 2013, **339**, 227-230.
32. D. Axford, X. Y. Ji, D. I. Stuart and G. Sutton, *Acta Crystallogr. Sect. D: Struct. Biol.*, 2014, **70**, 1435-1441.
33. C. Gati, G. Bourenkov, M. Klinge, D. Rehders, F. Stellato, D. Oberthur, O. Yefanov, B. P. Sommer, S. Mogk, M. Duszenko, C. Betzel, T. R. Schneider, H. N. Chapman and L. Redecke, *IUCrJ*, 2014, **1**, 87-94.
34. M. R. Sawaya, D. Cascio, M. Gingery, J. Rodriguez, L. Goldschmidt, J. P. Colletier, M. M. Messerschmidt, S. Boutet, J. E. Koglin, G. J. Williams, A. S. Brewster, K. Nass, J. Hattne, S. Botha, R. B. Doak, R. L. Shoeman, D. P. DePonte, H. W. Park, B. A. Federici, N. K. Sauter, I. Schlichting and D. S. Eisenberg, *Proc. Natl. Acad. Sci. U.S.A.*, 2014, **111**, 12769-12774.
35. E. Chiu, M. Hijnen, R. D. Bunker, M. Boudes, C. Rajendran, K. Aizel, V. Olieric, C. Schulze-Briese, W. Mitsuhashi, V. Young, V. K. Ward, M. Bergoin, P. Metcalf and F. Coulibaly, *Proc. Natl. Acad. Sci. U.S.A.*, 2015, **112**, 3973-3978.
36. H. M. Ginn, M. Messerschmidt, X. Y. Ji, H. W. Zhang, D. Axford, R. J. Gildea, G. Winter, A. S. Brewster, J. Hattne, A. Wagner, J. M. Grimes, G. Evans, N. K. Sauter, G. Sutton and D. I. Stuart, *Nat. Commun.*, 2015, **6**, 6435.
37. J. P. Colletier, M. R. Sawaya, M. Gingery, J. A. Rodriguez, D. Cascio, A. S. Brewster, T. Michels-Clark, R. H. Hice, N. Coquelle, S. Boutet, G. J. Williams, M. Messerschmidt, D. P. DePonte, R. G. Sierra, H. Laksmono, J. E. Koglin, M. S. Hunter, H. W. Park, M. Uervirojnangkoorn, D. K. Bideshi, A. T. Brunger, B. A. Federici, N. K. Sauter and D. S. Eisenberg, *Nature*, 2016, **539**, 43-47.
38. C. Gati, D. Oberthuer, O. Yefanov, R. D. Bunker, F. Stellato, E. Chiu, S. M. Yeh, A. Aquila, S. Basu, R. Bean, K. R. Beyerlein, S. Botha, S. Boutetg, D. P. DePonte, R. B. Doak, R. Fromme, L. Galli, I. Grotjohann, D. R. James, C. Kupitz, L. Lomb, M. Messerschmidt, K. Nass, K. Rendek, R. L. Shoeman, D. J. Wang, U. Weierstall, T. A. White, G. J. Williams, N. A. Zatsepin, P. Fromme, J. C. H. Spence, K. N. Goldie, J. A. Jehle, P. Metcalf, A. Barty and H. N. Chapman, *Proc. Natl. Acad. Sci. U.S.A.*, 2017, **114**, 2247-2252.
39. M. S. Nair, M. M. Lee, A. Bonnegarde-Bernard, J. A. Wallace, D. H. Dean, M. C. Ostrowski, R. W. Burry, P. N. Boyaka and M. K. Chan, *PloS One*, 2015, **10**, e0127669.
40. H. Tabe, T. Shimoi, K. Fujita, S. Abe, H. Ijiri, M. Tsujimoto, T. Kuchimaru, S. Kizaka-Kondo, H. Mori, S. Kitagawa and T. Ueno, *Chem. Lett.*, 2015, **44**, 342-344.
41. H. Tabe, T. Shimoi, M. Boudes, S. Abe, F. Coulibaly, S. Kitagawa, H. Mori and T. Ueno, *Chem. Commun.*, 2016, **52**, 4545-4548.
42. B. S. Heater, M. M. Lee and M. K. Chan, *Sci. Rep.*, 2018, **8**, 12783.
43. Q. Sun, S. W. Cheng, K. Cheung, M. M. Lee and M. K. Chan, *Crystals*, 2019, **9**, 287.
44. S. Abe, K. Atsumi, K. Yamashita, K. Hirata, H. Mori and T. Ueno, *Phys. Chem. Chem. Phys.*, 2018, **20**, 2986-2989.
45. M. Boudes, D. Garriga, A. Fryga, T. Caradoc-Davies and F. Coulibaly, *Acta Crystallogr. Sect. D: Struct. Biol.*, 2016, **72**, 576-585.
46. S. Honda, T. Akiba, Y. S. Kato, Y. Sawada, M. Sekijima, M. Ishimura, A. Ooishi, H. Watanabe, T. Odahara and K. Harata, *J. Am. Chem. Soc.*, 2008, **130**, 15327-15331.
47. S. Abe, H. Tabe, H. Ijiri, K. Yamashita, K. Hirata, K. Atsumi, T. Shimoi, M. Akai, H. Mori, S. Kitagawa and T. Ueno, *ACS Nano*, 2017, **11**, 2410-2419.
48. T. K. Nguyen, S. Abe, M. Kasamatsu, B. Maity, K. Yamashita, K. Hirata, M. Kojima and T. Ueno, *ACS Appl. Nano Mater.*, 2021, **4**, 1672-1681.
49. M. Sobieraj and P. Setny, *J. Chem. Theory Comput.*, 2022, **18**, 1936-1944.
50. X. Y. Chen, J. L. Zaro and W. C. Shen, *Adv. Drug Deliv. Rev.*, 2013, **65**, 1357-1369.
51. K. Hirata, K. Yamashita, G. Ueno, Y. Kawano, K. Hasegawa, T. Kumasaka and M. Yamamoto, *Acta Crystallogr. Sect. D: Struct. Biol.*, 2019, **75**, 138-150.
52. J. K. Sun and A. J. Doig, *Protein Sci.*, 1998, **7**, 2374-2383.
53. M. Novotny and G. J. Kleywegt, *J. Mol. Biol.*, 2005, **347**, 231-241.
54. C. M. Deane, F. H. Allen, R. Taylor and T. L. Blundell, *Protein Eng.*, 1999, **12**, 1025-1028.
55. M. A. AlDamen and M. Sinnokrot, *J. Struct. Chem.*, 2014, **55**, 53-60.
56. E. N. Baker and R. E. Hubbard, *Prog. Biophys. Mol. Biol.*, 1984, **44**, 97-179.
57. D. J. Barlow and J. M. Thornton, *J. Mol. Biol.*, 1983, **168**, 867-885.

58. K. A. McKiernan, B. E. Husic and V. S. Pande, *J. Chem. Phys.*, 2017, **147**, 104107.
59. H. Watanabe and S. Honda, *Chem. Biol.*, 2015, **22**, 1165-1173.
60. Y. Y. Tang, J. Saul, N. Nagaratnam, J. M. Martin-Garcia, P. Fromme, J. Qiu and J. LaBaer, *Sci. Rep.*, 2020, **10**, 13323.