Multiple binding Modes of a moderate Ice-binding Protein from a polar Microalga

Hidemasa Kondo [1ab], Kenji Mochizuki [1cd], Maddalena Bayer-Giraldi [*e]

1: The authors contributed equally to the study

a Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Sapporo 062-8517, Japan

b Graduate School of Life Science, Hokkaido University, Sapporo 060-0810, Japan

c Institute for Fiber Engineering, Shinshu University, Ueda, Nagano 386-8567, Japan

d Faculty of Textile Science and Technology, Shinshu University, Ueda, Nagano 386-8567, Japan

e Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Department of Glaciology (Geosciences), Am Alten Hafen 26, 27568 Bremerhaven, Germany

* Corresponding author: maddalena.bayer@awi.de

ABSTRACT

Ice-binding proteins (IBPs) produced by cold-tolerant organisms interact with ice and strongly control crystal growth. The molecular basis for the different magnitude of activities displayed by various IBPs (moderate and hyperactive) has not yet been clarified. Previous studies questioned whether the moderate activity of some IBPs relies on their weaker binding modus to the ice surface, compared to hyperactive IBPs, rather than relying on binding only to selected faces of the ice crystal. We present the structure of one moderate IBP from the sea-ice diatom *Fragilariopsis cylindrus* (*fc*IBP) as determined by X-ray crystallography and investigate the protein's binding modes to the growing ice-water interface using molecular dynamics simulations. The structure of the *fc*IBP is the IBP-1 fold, defined by a discontinuous β-solenoid delimitated by three faces (*A*, *B* and *C*-faces) and braced by an α-helix. The *fc*IBP structure shows capping loops on both N- and C-terminal parts of the solenoid. We show that the protein adsorbs on both the prism and the basal faces of ice crystals, confirming experimental results. The *fc*IBP binds irreversibly to the prism face using the loop between the *B* and the *C*-faces, involving also the *B*-face in water immobilization despite its irregular structure. The α-helix attaches the protein to the basal face with a partly reversible modus. Our results suggest that *fc*IBP has a loser attachment to ice and that this weaker binding modus is the basis to explain the moderate activity of the *fc*IBP.

Keywords: Ice-binding proteins; Antifreeze proteins; DUF3494; IBP-1 fold; molecular dynamics simulation.

INTRODUCTION

Ice-binding proteins (IBPs) are defined by their ability to attach to ice and influence its growth [1, 2]. IBPs lower the freezing point of a solution, affect the ice growth kinetics during the crystallization process and inhibit ice-grain boundary migration (recrystallization) in polycrystalline ice [3, 4]. A variety of different IBPs, also called *antifreeze proteins* (AFPs), has

been found in several polar and cold-tolerant organisms. One common classification of IBPs is based on the protein's effectiveness in causing a thermal hysteresis (TH), i.e., a separation of the freezing point below the melting point of a solution. At identical protein concentration, moderate IBPs induce a TH of less than 1°C, whereas hyperactive IBPs cause a much stronger freezing point depression [3]. The mechanisms underlying this difference in activity are currently under dispute. Some studies suggest that hyperactivity is related to the ability of IBPs to bind the basal face of ice crystals and suppress growth along the *c*-axis [5, 6]. Other studies indicate that the TH activity shown by IBPs is related less to the crystallographic face bound by the proteins, but rather to the strength of the binding of IBPs to the ice crystal surface [7, 8].

The different IBP families show an amazing diversity of structures, including α-helix, globular polypeptides with mixed folds, polyproline type-II coils and β-solenoids [3]. The ice-binding site (IBS), often determined by point mutagenesis, is described as a broad, flat, somewhat hydrophobic surface. IBPs exhibit various ice-binding mechanisms driven by hydrogen bonding [9], hydrophobic interaction [10-13], and anchored clathrate motif [14-16].

The IBPs from the sea-ice diatom *Fragilariopsis cylindrus* (*fc*IBP), a dominant species within polar sea-ice microbial assemblages, belong to an IBP family very common among psychrophilic microorganisms [17]. This IBP family is characterized by the "domain of unknown function" (DUF) 3494, as the domain is called in the Pfam database. The DUF3494-IBP family represents today the most widespread of the known IBP families and can be found in bacteria [18-20], diatoms [17, 21, 22], yeast and other fungi [23-27], among others. Studies on one *fc*IBP isoform, *fc*IBP11, revealed that *fc*IBP11 binds to the prism and basal faces of ice crystals and stops growth along the *c*-axis despite its moderate TH activity [8, 28]

The structures of DUF3494-IBPs known until now are a β-solenoid, characterized by a discontinuous β-helix with a triangular cross-section defined by the *A*-, *B*- and *C*-faces. An α-helix runs along the *A*-face, parallel to the longitudinal axis of the β-helix. This fold, typical for DUF3494-IBPs, has been called IBP-1 fold [3]. The capping regions masking the hydrophobic core of the proteins can be more or less extended in the individual proteins. Until now, eight IBP-1 folds have been determined by X-ray crystallography. The solved structures belong to sequences from Antarctic bacteria [19, 29-32], a snow mold fungus [7, 33] and an Arctic yeast [34]. Despite the broad distribution of DUF3494-IBPs among polar diatoms, only the structure of the *Chaetoceros neogracile* IBP, estimated by 3D modelling, has been reported [35, 36].

Point mutation, structural analyses of the topography at the protein surface and docking studies have been examined to identify the IBS [7, 19, 29-36]. In all these cases, results suggest that the protein *B*-face is involved in ice-binding, despite its lack of structural regularity. Furthermore, some studies mention a possible relevance of the *C*-face and of the loop adjacent to the *B*-face [7, 29, 30]. However, although the computational docking studies give an

insight about the structural matching between IBPs and ice surfaces, we must consider that IBPs bind to a growing ice-water interface rather than to an ice crystal face alone [37].

In the following, we present the structure of *fc*IBP11 determined by X-ray crystallography and its ice-binding mode inferred by molecular dynamics simulation. This is the first case of a DUF3494-IBP from diatom cells solved by crystallography and the first molecular dynamics simulation of binding of a DUF3494-IBP to the growing ice surface. We investigate whether *fc*IBP11 binds to both the primary prism and basal surfaces and suggest where its IBS is.

MATERIAL AND METHODS

Crystallization and structure analysis of *fc*IBP11

*fcIBP*11 (GenBank Acc Nr DR026070), an isoform of *fc*IBP with molecular weight of 26 kDa and of moderate TH activity, was recombinantly expressed (EMBL Heidelberg, Germany) as explained elsewhere (Bayer-Giraldi et al 2011). This isoform was chosen based on previous studies, which demonstrated its relevance for *F. cylindrus* in cold response [17, 28]. The protein was lyophilized for storage, then dissolved in cold water and dialyzed against 100 mM Tris-HCl pH 8.2 for buffer exchange and desalting. Prior to crystallization, the dialysate was concentrated to 20 mg/ml using Amicon Ultra-4 centrifugal filter units (Merck KGaA, Germany). The crystallization condition was screened using Crystal Screen, Crystal Screen 2, Index, PEG/Ion, PEG/Ion 2 (Hampton Research, CA, USA), and Wizard Classic 1 and 2 (Molecular Dimensions, UK). By using nanoliter dispenser mosquito (TTP Labtech, UK) 0.1 µL of protein solution was mixed with the same volume of reservoir solution in a 96-well sitting-drop plate [38], then incubated at 20 and 4°C.

The diffraction data from the *fc*IBP11 crystal were collected at the beamline BL17A in Photon Factory, KEK, Japan [39], using the synchrotron radiation of 0.9800 Å. The crystal was soaked into a crystallization solution containing 30% glycerol, which was used as a cryoprotectant, and then mounted on a nylon loop, followed by flash cooling to 95 K by a cryocooling device. Diffraction images were processed by program XDS [40], and CCP4 program suite [41]. The crystal structure of *fc*IBP11 was determined by a molecular replacement method using the program Phenix [42] applying the coordinate of *Typhula isikariensis* AFP6 (PDB ID: 3VN3) as the search model. The molecular model of *fc*IBP11 was build and manually corrected using Coot [43] and further refined using REFMAC5 [44]. The structure of *fc*IBP11 was superposed against other known DUF3494-IBP structures by utilizing secondary-structure matching (SSM) option implemented in Coot. Root mean square distance (RMSD) between equivalent Cα atoms was employed for assessing structural similarities.

Models

*fcIBP*11 was modeled with full atomistic detail using CHARMM27 [45, 46] (CHARMM22 [47] plus CMAP [48] for proteins). The experimentally obtained crystal structure was used, in which N-terminal 2 residues (S and T) and *C*-terminal 11 residues (TRRGLRGLQVA) were not included in the structure model due to the indistinct electron density map. Water was

represented by the TIP4P/2005 model [49], which provides a more realistic description of the bulk liquid density and the hydration thermodynamics of simple molecules [50-52]. The intermolecular interactions were truncated at 0.85 nm. The Lennard-Jones parameters for cross-interactions were obtained using the Lorentz−Berthelot combination rules: $\epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}}$ and $\sigma_{ij} = (\sigma_{ii} + \sigma_{jj})/2$. The long-range Coulombic interactions were evaluated with the particle-mesh Ewald algorithm and the dispersion corrections were implemented in the evaluation of the energy and pressure.

Molecular Dynamics Simulations

Molecular dynamics (MD) simulations were carried out using GROMACS 2016.4 [53, 54] integrating the equations of motion with the leapfrog algorithm using a time step of 2 fs. The temperature T and pressure P were controlled with the Berendsen algorithm [55]. Periodic boundary conditions were applied in the three directions.

*fcIBP*11 in bulk water

In order to investigate the hydration shell structure around the protein, we performed MD simulations for *fc*IBP11 dissolved in 20,000 water molecules with 5 sodium ions. The energy minimization using the steepest descent method is followed by a 10 ns *NPT*-MD run at 300 K. Then, the production *NPT*-MD run of 20 ns was performed at 250 K for the conformational sampling. We also performed the same simulation without the protein.

Adsorption of *fc*IBP11 on the growing ice surface

We performed non-equilibrium *NVT*-MD simulations with cells containing slabs of vapor, liquid and ice [10]. The size of the simulation box for the system in which the primary prism face was exposed to liquid water was set to 9.06 × 13.00 × 8.88 nm^3 and that for the basal face was 9.06 × 8.63 × 13.00 nm^3. The simulation cells consisted of one *fc*IBP11 molecule, 5 sodium molecules, 20,000 free liquid water molecules and two restrained ice layers. These two layers of proton disordered ice Ih (1920 and 1760 molecules for the systems exposing the primary prism and basal planes, respectively) were generated with the program GenIce [56]. The 20,000 free water molecules were placed on one side of the ice layers to let ice grow in a single direction. The oxygen atoms of the molecules in these two layers of ice were harmonically restrained at their original positions with a force constant of 1000 kJ mol^{-1} nm^{-2}. The *fc*IBP11 was initially placed 1.0~1.5 nm above the ice surfaces in three different ways as *A*-, *B*- and *C*-faces of the protein faced towards the ice surface. We first performed an energy minimization using the steepest descent method, followed by a 400 ps *NVT* equilibration run at 300 K with freezing the protein coordination. Then, the production *NVT*-MD run was evolved for 600 ns at 248 K, which is 2.5 K lower than the freezing temperature of ice Ih in this model (250.5 K) [57]. We gave three different momenta to the equilibrated configurations, so that 9 independent trajectories were generated for each system. The name of the trajectory indicates the exposed ice plane (P for prism face, or B for basal face), the face of *fc*IBP11 (*A* or *B* or *C*) that initially faced toward the ice surface and the given initial momentums (1 or 2 or 3). We also performed the same simulations without the protein in order to investigate the influence of the protein on the ice growth. In the protein bound trajectories, we determined the amino acids that possibly interact directly with the ice face by checking which residues are aligned with a crystal face.

Root mean square displacement

To assess the adsorption of *fc*IB11 we computed the time evolution of root mean square displacement (RMSD) for αC atoms of *fc*IBP11 using the equation of $\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{r}_i - \boldsymbol{r}_i^0\right)^2}$, where $\boldsymbol{r}_i$ and $\boldsymbol{r}_i^0$ are the coordination vectors of *i*th αC atom at time *t* and the initial state (t = 0), respectively. In the RMSD calculations, the highly flexible regions of *fc*IBP11 (with residue numbers 3-13, 102-120, 239-246) were excluded.

RESULTS AND DISCUSSION

1. Structural Analyses

1.1 Crystal Structure of *fc*IBP11

Crystals of *fc*IBP11 were obtained under 0.1M Tris-HCl pH 8.5 and 2.0 M ammonium dihydrogen phosphate at 4°C and grown in a plate-like shape. Diffraction data at 1.4 Å resolution were collected at the synchrotron beamline. The crystal belongs to orthogonal space group $P222_1$ with unit cell parameters of a=36.05, b=47.59, and c=134.52 Å, containing one molecule in an asymmetric unit. A clear solution for *fc*IBP11 structure was provided by molecular replacement calculation, then corrected manually and applied for the crystallographic refinement. N-terminal 2 residues and C-terminal 11 residues were not included in the structure model due to the indistinct electron density map, which implies the disordered conformation in these regions. At the late stage of the structure refinement, water molecules were introduced to the model by inspecting the electron density map. The final *fc*IBP11 structure contains 244 residues out of 257, and 279 water molecules with *R* factor of 0.137 and Free *R* factor [58] of 0.162. The quality of the main-chain conformations was validated by Ramachandran-plot calculated by MolProbity [59], showing that most residues fell into the favored and allowed region. The statistics for data collection and refinement was summarized in Table 1. The coordinates are deposited in the Protein Data Bank (PDB) under ID 6A8K.

The crystal structure of *fc*IBP11 exhibits an IBP-1 fold characteristic for DUF3494-IBPs (Figure 1). The structure is dominated by a distinct β-solenoid [60], which is composed of a helical structure of parallel β-sheets. The β-solenoid of *fc*IBP11 is folded into a right-handed helix with a triangular cross-section, which forms three side faces (the *A*-, *B*- and *C*-faces) made up of eight- or six-stranded β-sheets with 2-5 residues each. The β-solenoid domain with 182 residues is composed of 56 N-terminal residues, from Val20 to Pro75, and 126 C-terminal residues, from Gly124 to Ala239. The N-terminal part forms a helical coil (β1) toward the end of the solenoid, followed by a capping loop structure. The other end of the solenoid is also covered by a capping loop, followed by six helical coils (β2-β6) toward the one-third middle of the solenoid. Accordingly, the N- and C-terminal units are stacked together with a "head-to-tail" manner to locate the N- and C-terminal residues (Val20 and Ala239, respectively) at an adjoining position. A long α-helix with 20 residues (Gly82-Ala101) is situated in the middle of the polypeptide chain, lying parallel to the β-solenoid

along the *A*-face, and connecting each end of the solenoid. This topological arrangement as a discontinuous solenoid braced by a parallel α-helix has been uniquely observed for the 3D structure of DUF3494-IBPs.

Considering the faces of the solenoid, the *B*-face is formed by regularly aligned β-strands and displays the flattest molecular surface of the three faces of the solenoid. The *C*-face also constitutes a flat surface, but two strands (β2 and β3) near the end of the solenoid elongate to form a small bulge. The N-terminal region of the polypeptide chain (Ala3-Asp18), which lies antiparallel to the α-helix and covers the *A*-face of the solenoid, forms an extended helical structure, showing a close similarity to polyproline II helix followed by a short 3_{10} helix.

The similarity between the overall structure of *fc*IBP11 and those of other DUF3494-IBPs was assessed by superposition. The RMSD between corresponding Cα atoms in the β-solenoid domain and the long α-helix ranges from 0.72 Å to 1.41 Å, showing close similarity with other microbial IBPs. This is reflecting the overall sequence similarity (29–49%) typical for the residues that constitute the hydrophobic core of the molecule.

Some differences among the IBP-1 folds can be seen when considering the capping structures (Figure 2). The local conformation of the N-terminal capping loop of the DUF3494-IBPs can be classified into four groups, with eukaryotic IBP-1 folds (*Typhula ishikariensis Tis*IBP and *Leucosporidium* sp. *Le*IBP) constituting a separate group [30, 32]. The N-terminal loop structure of *fc*IBP11 shows close similarity with the eukaryotic group (Figure 2A) and forms an antiparallel strand composed of 14 residues (Asp53 – Thr66). In the known DUF3494-IBPs structures the C-terminal edge of the solenoid is less covered with distinct loop segment. However, *fc*IBP11 possesses a unique loop segment of 12 residues, from Gly106 to Thr117, which is inserted between the long α-helix and the β-solenoid and covers the C-terminal unit of the solenoid (Figure 2B). In many β-solenoid proteins, the loop or helical components are situated at both ends of the solenoid and prevent the exposure of the hydrophobic core of the molecule by covering the solenoid. The capping loop also builds hydrogen bonds with the β-sheets at the edge of the solenoid in order to avoid the aggregation of different molecules by intermolecular hydrogen bonds.

1.2 Putative IBS inferred from structural analysis

The preceding papers about the structure analysis of DUF3494-IBPs reported that the flat *B*-face of the solenoid is involved in binding of the protein to ice [19, 29, 33, 34]. It was estimated that the IBS is constituted by approximately 30 residues, aligned on the β-sheet of the *B*-face, on the adjacent loop region and on the *C*-face [7, 29, 30]. The residues on the putative IBS mainly have short side chains, but are poorly conserved among the homologous IBPs. *fcIBP*11 displays a less flat surface on its corresponding face (Figure 3A), which reflects a lower contents of Gly, Ala, Ser and Thr on the *B*-face. The content of these residues is 56% (14 residues out of 25) for *fc*IBP11, whereas for the hyperactive *Tis*IBP8, for example, it is 80% (20 residues out of 25). In addition, residues with a long side chain such as Lys24, 165, 209 and 213, and Glu193 form a small bulge on the *B*-face of *fc*IBP11. Lys213 and Asn195, which are located at the center region of the putative IBS, adopt multiple conformations in

their side chain. This structural variety of side chains implies that the IBS of *fc*IBP11 possesses a less regular and a less restricted structure compared with other DUF3494-IBPs.

Furthermore, our analyses show little regularity among the 50 bound waters on the putative IBS of the *fc*IBP11 *B*-face (Figure 3B). Previous reports of crystal structure of DUF3494-IBPs identified regularly aligned water molecules, which occupy the shallow grooves on their IBS [7, 29, 33]. In contrast, the putative IBS on the *B*-face of *fc*IBP11 exhibits less regularly aligned waters and distinct surface grooves. No bound waters are found at the center of the putative IBS around Lys213 and Asn195, which adopt multiple conformations in their side chain. This also seems to reflect the high relief surface and the less restricted property of the IBS of *fc*IBP11. Therefore, in order to identify further putative IBS of *fc*IBP11, we proceeded with MD simulations, which allow analyzing the binding process of the protein to the prism and basal face of ice in a dynamic situation.

2. MD simulations

2.1. Hydration shell structure in bulk water

Some IBPs indirectly bind to ice through ordered "ice-like" or "clathrate-like" hydration water molecules, which are formed beside regularly spaced residues on the ice-binding site of IBP. Although the three-dimensional structure of these bound water molecules is distinct from the ice, it extensively matches the spacing of water molecules in ice lattice and forms hydrogen bonds with them [14-16, 61].

To assess whether such an ordered hydration array exists around *fc*IBP11 in solution, we first identify immobile (or solid-like) water molecules based on its translational mobility [15, 62, 63]. The value δ_i^2 is defined by $\delta_i^2 = \langle\{\{\boldsymbol{r}_i(t+\Delta t) - \boldsymbol{r}_i(t)\}^2\rangle$, where $\boldsymbol{r}_i$ is the coordination vector of oxygen of *i*th water molecule, Δt is set to 100 ps and $\langle \dots \rangle$ is the average over 1 ns. We compute the distribution of δ_i^2 for bulk liquid water molecules at 250 K and found that the population in the region of $\delta_i^2 < 0.06$ nm^2 is negligibly small (Figure 4A). Thus, we define immobile water molecules if its δ_i^2 is smaller than 0.06 nm^2. The distribution for the water molecules in a *fc*IBP11 solution (green line in Figure 4A) shows that the population of immobile water molecules significantly increases upon addition of *fc*IBP11 (380 molecules on average), indicating that *fc*IBP11 slows down the translational displacement of water molecules in the hydration shell.

Snapshots in Figure 4B show that *fc*IBP11 is fully covered by immobile water molecules. However, we do not find any clue for ordered hydration layers in the vicinity of *fc*IBP11, presumably because *fc*IBP11 lacks the structural regularity on the surface. These results are consistent with the analyses examined in the section 1.2.

2.2. Adsorption on Primary Prism Surface

We observe the adsorption of *fc*IBP11 on the growing primary prism surface in 2 of 9 trajectories (P_b1 and P_b3, Figure 5). The movies for these trajectories clearly demonstrate that the *fc*IBP11 tightly binds to the ice surface and halts its growth after wandering on the ice-water interface in the first 200 ns (SI movie P_b1 and P_b3). In these trajectories, the RMSDs with regard to the initial conformation become almost constant after 200 ns due to the adsorption on the ice surface (Figure 6A). Then, the adsorption immediately halts the

growth of ice, while the amount of ice monotonically increases without *fc*IBP11 (Figure 7A). In the trajectory of P_b1, the *fc*IBP11 rotates 90 degrees in the first 200 ns and binds to the ice surface through the loop region between the *B*- and the *C*-faces (Figure 8A). The residues of Thr50, Val30, Gln233, Ile215, Thr216 and Ala198 align with the *x*-axis (perpendicular to the *c*-axis of the ice lattice). Three residues of Thr50, Val30 and Gln233 directly bind to the well-defined ice lattice consisting of only 6-member rings, while there are non-ice-like immobile water molecules between the other residues (Ile215, Thr216 and Ala198) and the ice surface. Such immobile water molecules are also formed under the whole *B*-face (see the dashed circle in Figure 8B). In the trajectory of P_b3, we also observe that the *fc*IBP11 binds to ice through the loop region between the *B*- and the *C*-faces, although the loop region aligns with the *z*-axis in this trajectory. Figure 8D shows the residues of Val30, Thr145, Ala171, Ala198, Ile215, Thr216 and Gln233 directly bind to the ice lattice. The immobile water molecules without ice structure are also observed below the *B*-face (the dashed circle in Figure 8C). These results indicate that the loop region between the *B*- and the *C*-faces of *fc*IBP11 directly binds to the ice lattice, but the flat *B*-face may also contribute the adsorption, in consistent with the point mutation experiment for the same IBP family [33]. Although there are at least two different binding modes on the primary prism surface, we do not observe the transformation between them and each binding is irreversible in the computational time scale.

2.2. Adsorption on Basal Surface

We observe that the *fc*IBP11 binds to the basal ice surface in 2 of 9 independent trajectories (see SI movies B_a1 and B_a3). The adsorption resulting in the halt of ice growth is supported by the RMSD (Figure 6B) and the amount of immobile water molecules (Figure 7B). The ice-binding occurs through the α-helix on the *A*-face in both trajectories. The residues directly binding to the ice lattice are Ser79, Thr83, Thr87, Ser90, Thr94, Asp98 and Ala101 (Figure 9A and 9C). A small portion of the α-helix is buried into the ice lattice in the trajectory B_a1, while the α-helix in the trajectory B_a3 is parallel to the ice surface (Figures 9A and 9C). Because the direction of the α-helix on the *x*-*y* plane is almost identical (Figures 9B and 9D), we recognize these binding modes are the same. In the end of trajectory B_a3, *fc*IBP11 is released from the ice surface, indicating a reversible ice-binding.

3. Binding mode and TH activity of *fc*IBP11

We show that *fc*IBP11 can bind to ice with different modes, depending on the crystallographic face considered. We observe, within the computational time scale, irreversible attachment of *fc*IBP11 to the prism face of the ice crystal and partly reversible binding of *fc*IBP11 to the basal face, suggesting that the adsorption on the basal face is not strong.

Our results can be considered in the frame of the discussion about the mechanisms underlying differences in TH activity displayed by IBPs. The basis for hyperactivity of IBPs has often been correlated with the adsorption of the proteins on the basal face of ice crystals and the suppression of its growth [5, 6]. Other works have focused on the binding strength of IBPs to ice, rather than to the affinity for specific crystallographic faces. For example, results from fluorescence microscopy [64, 65] and MD simulations [10, 16] show that the hyperactive protein

from the insect *Tenebrio molitor* (*Tm*AFP) binds irreversibly to ice whereas the moderate fish antifreeze glycoprotein (AFGP) 8 binds reversibly, suggesting a loser binding mode for the moderate protein. Previous publications on various DUF3494-IBPs [7, 8, 19, 29], conclude that the binding energy of the proteins to ice crystal surface plays an important role in IBP hyperactivity. Furthermore, it has been shown that the moderate *fc*IBP11 can attach, at least in some amount, to both the prism and the basal face of ice crystals and suppress growth of the basal face, resulting in a growing pattern of the ice crystal usually ascribed to the presence of hyperactive IBPs [8]. Our results match experimental evidence of affinity of *fc*IBP11 also for the basal face of ice crystals and indicate that the moderate TH activity of *fc*IBP11 is possibly related to a partly reversible, lose attachment of the proteins to the basal face of ice. Anyhow, also the lose binding mode of *fc*IBP11 to the basal face enables its growth suppression as experimentally shown before, and therefore must be strong or fast enough to stop ice growth.

CONCLUSIONS

In this study, we show for the first time the structure of a diatom DUF3494-IBP and determine that the conformation of *fc*IBP11 groups with that of other eukaryotic IBPs with IBP-1 fold. *fc*IBP11 binds to both the primary prism and basal surfaces, consistent with the experimental results. The binding site of *fc*IBP11 to the primary prism surface is the loop region between the *B*- and *C*-faces and *fc*IBP11 has at least two different adsorption alignments. The *B*-face may also contribute to ice, although it shows a less regular and a less restricted structure compared with other DUF3494-IBPs as indicated by structural analysis on this face. The binding to the basal surface occurs through the α-helix parallel to the *A*-face and we further observe partial detachment of the protein from the ice surface. We here show selected putative binding options of *fc*IBP11 to ice, not considering statistical significance. However, we suggest that this binding mode explains the kinetics of ice growth in the presence of *fc*IBP11, a protein with moderate TH activity but causing basal face growth inhibition. Further studies are required to determine which factors, e.g. hydrogen bonding or hydrophobic interaction, predominantly contribute to the adsorption and to clarify the binding kinetics of *fc*IBP11 to stop crystal growth along the *c*-axis.

Conflicts of interest

There are no conflicts to declare.

Aknowledgements

We acknowledge support from the Japan Society for the Promotion of Science (Kakenhi 18K19060), from the Deutsche Forschungsgemeinschaft SPP1158 (German Research Association Special Program 1158, Grant BA 3694/2-1) and from the Innovation Fund from the Alfred Wegener Institute (Project MIAA). K. M. thanks Prof. Masakazu Matsumoto for insightful discussions. A part of calculations was performed at the Research Center for Computational Science in Okazaki, Japan.

Table 1 Data collection and refinement statistics for *fc*IBP11

Data collection	
Space group	$P222_1$
Unit-cell parameters (*a*, *b*, *c*), (Å)	36.05, 47.59, 134.52
Number of molecules in asymmetric unit	1
Beam line	Photon Factory BL-17A
Wavelength (Å)	0.9800
Resolution range (Å)	47.6–1.4 Å
R_{merge} [*,†]	0.057 (0.434)
Observed reflections	571,099
Independent reflections	46,394
Completeness (%) [*]	99.7 (98.7)
Multiplicity [*]	12.3 (11.6)
$\langle I/\sigma(I)\rangle$ [*]	24.1 (6.3)
Refinement	
R factor [*,‡]	0.137 (0.192)
Free *R* factor [*,‡,§]	0.162 (0.200)
R.M.S bond lengths (Å)	0.029
R.M.S bond angles (°)	2.478
Residues	244
Number of non-hydrogen atoms	
Protein	1758
Solvent	279
Ramachandran plot (%) [¶]	
Residues in favored regions	97
Residues in allowed regions	3
Residues in outliner regions	0
Average B factor (Å^2)	15.0

[*] Values in parentheses are for the highest resolution shell (1.48–1.4 Å for data collection and 1.44–1.4 Å for refinement)

[†] $R_{merge} = \Sigma\Sigma_j|\langle I(h)\rangle - I(h)_j| / \Sigma\Sigma_j\langle I(h)\rangle$, where $\langle I(h)\rangle$ is the mean intensity of a set of equivalent reflections.

[‡] R factor $= \Sigma||F_{obs}(h)| - |F_{calc}(h)||/ \Sigma |F_{obs}(h)|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively.

[§] A randomly chosen 5.0% of the data were used to calculate the free *R* factor [58].

[¶] Statistics were obtained from MolProbity [59].

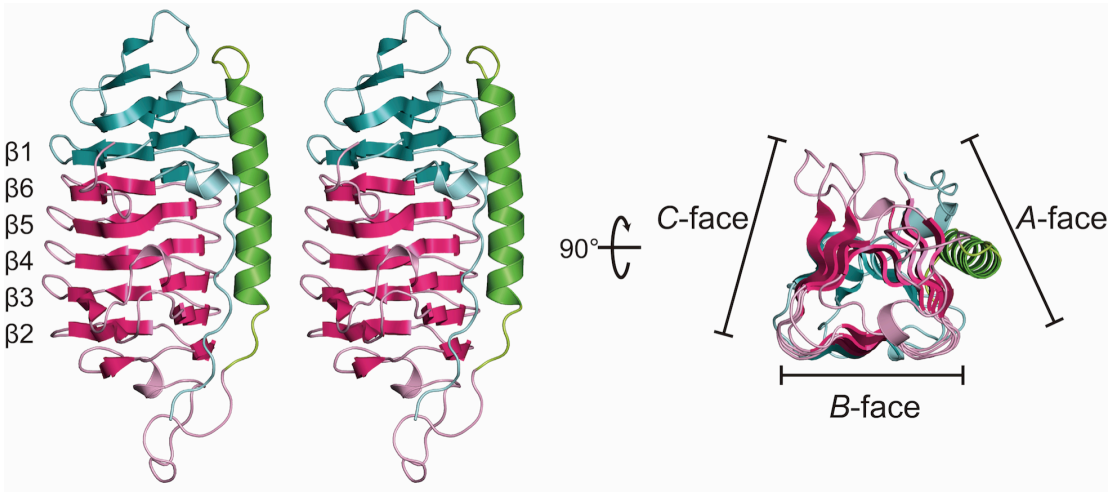

Figure 1. The crystal structure of *fc*IBP11 drawn with schematic illustrations. *fcIBP*11 is composed of a discontinuous β-solenoid braced by an α-helix. The major coils of the solenoid are indicated as β1-6. The N- and C-terminal parts of the β-solenoid are indicated in cyan and red, respectively. The long α-helix situated along β-solenoid is colored green. The β-solenoid of *fc*IBP11 folds into a right-handed helix with a triangular cross-sections, which displays *A*-, *B*- and *C*-faces on its molecular surface. The illustrations in Figure 1, 2 and 3 were prepared by PyMOL [66].

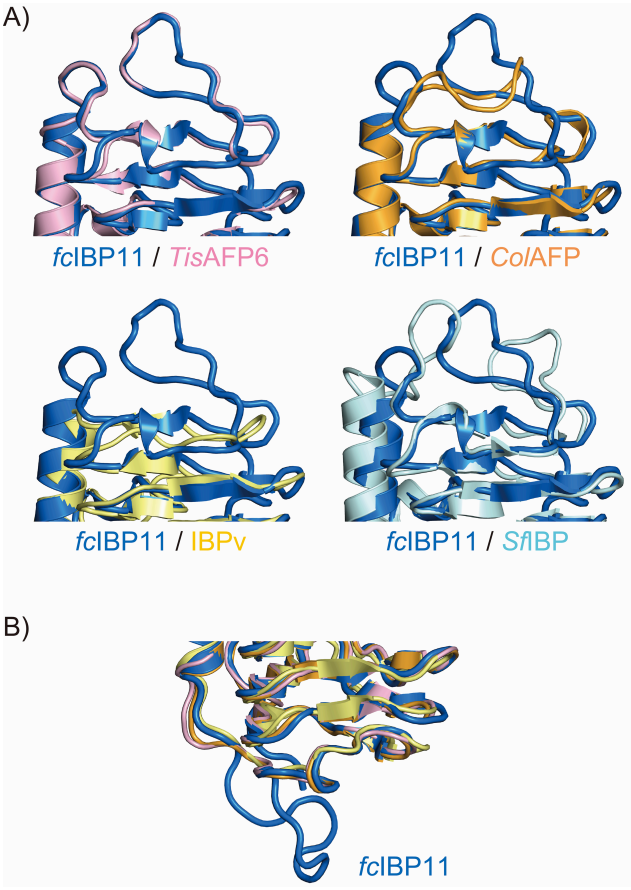

Figure 2. Close-up views of the capping loops of *fc*IBP11 and other microbial IBPs. (A) Pairwise superpositions of *fc*IBP11 N-terminal loop (blue) with *Tis*AFP6 (PDB ID: 3VN3, pink), *Col*AFP (PDB ID: 3WP9, orange), IBPv (PDB ID: 5UYT, yellow), and *Sf*IBP (PDB ID: 6EIO, cyan). (B) Close-up view of the C-terminal capping loop.

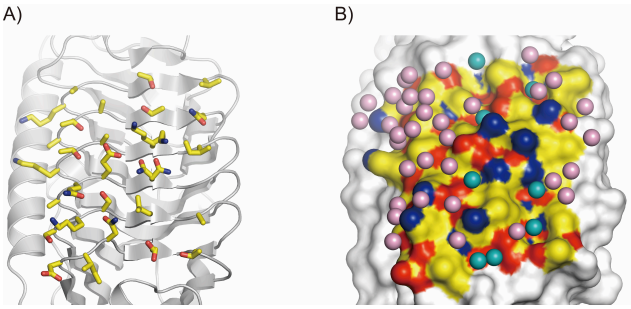

Figure 3. Surface residues and bound waters located in putative ice-binding site (IBS) of *fc*IBP11.

(A) The side chains of IBS residues are shown as sticks. C, O and N atoms are colored yellow, red and blue, respectively. (B) Water molecules located within 5Å from the IBS are shown as balls. Waters located on the concave surface of the IBS are colored cyan. The molecular surface of *fc*IBP11 is also represented and the IBS is drawn with same colors as (A).

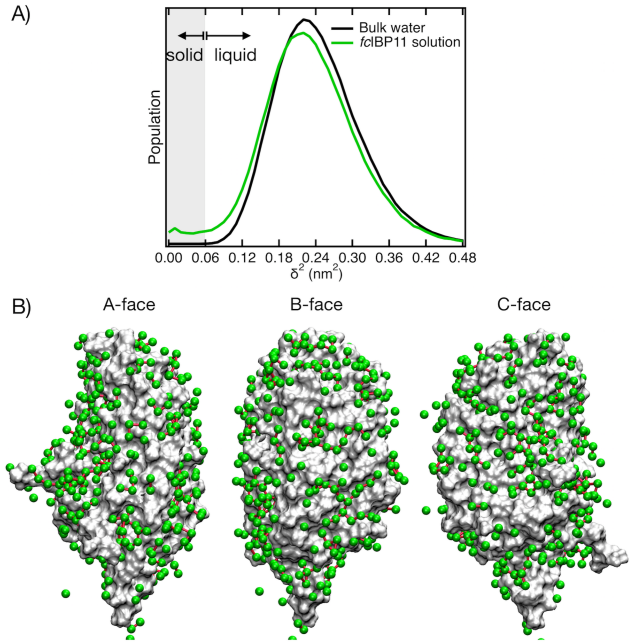

Figure 4. (A) Distributions of the δ_i^2 of water molecules in bulk liquid (black line) and *fc*IBP11 solution (green line) at 250 K. The shade region indicates immobile water molecules. (B) A typical structure of immobile molecules in the *fc*IBP11 solution from three different angles. The oxygen atoms of immobile water molecules with $\delta_i^2 < 0.06$ nm^2 are described by green spheres and two immobile water molecules are connected by a red line when their oxygen-oxygen distance is smaller than 0.35 nm.

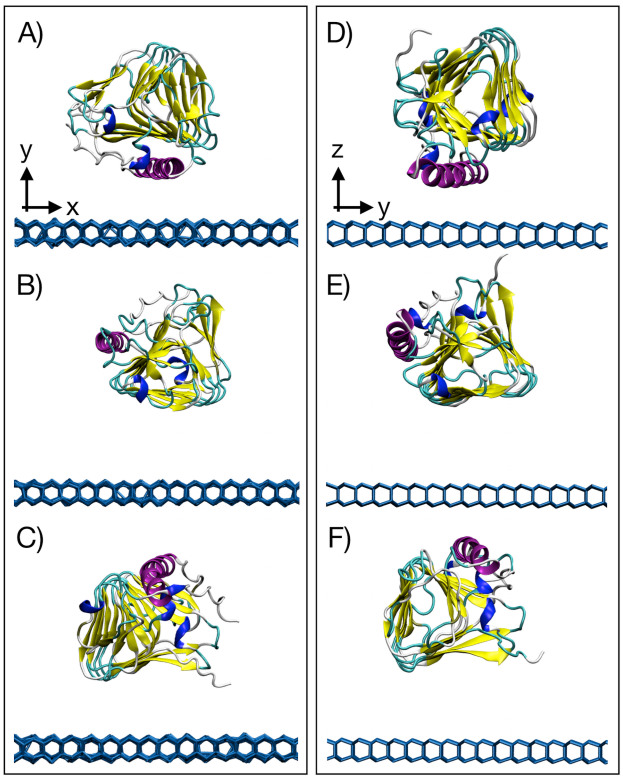

Figure 5. Initial configurations for the MD simulations in which the ice surface exposes (A-C) primary prism and (D-F) basal surfaces. These correspond to the trajectories named (A) P_a1, P_a2 and P_a3; (B) P_b1, P_b2 and P_b3; (C) P_c1, P_c2 and P_c3; (D) B_a1, B_a2 and B_a3; (E) B_b1, B_b2 and B_b3; (F) B_c1, B_c2 and B_c3, by giving three different initial

momenta. The restrained two ice layers are shown by dark blue lines, while the other water molecules are not shown in the figures. The *z*-axis corresponds to the *c*-axis of the ice lattice.

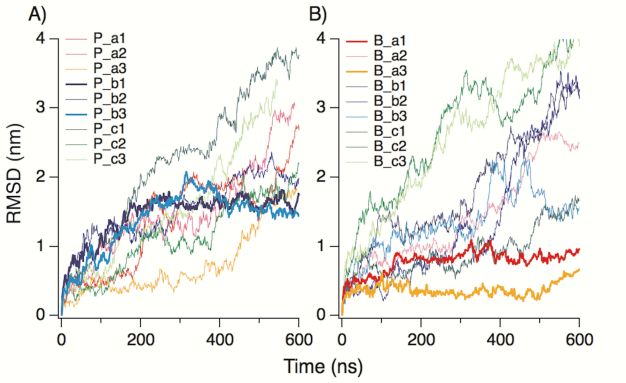

Figure 6. RMSD of *fc*IBP11 on the (A) primary prism and (B) basal planes. The adsorption of *fc*IBP11 is observed in the trajectories named P_b1 and P_b3 in the panel (A), and B_a1 and B_a3 in the panel (B).

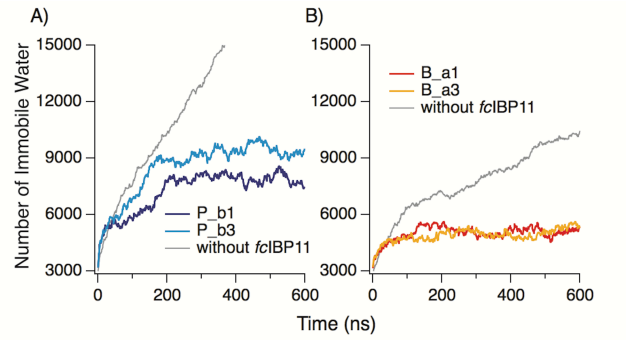

Figure 7. Time evolution of immobile water molecules in the system exposing (A) primary prism and (B) basal surfaces. The plotted are the trajectories in which the adsorption of *fc*IBP11 is observed and the trajectories separately computed without *fc*IBP11. The plot without the protein is the average obtained from three independent trajectories.

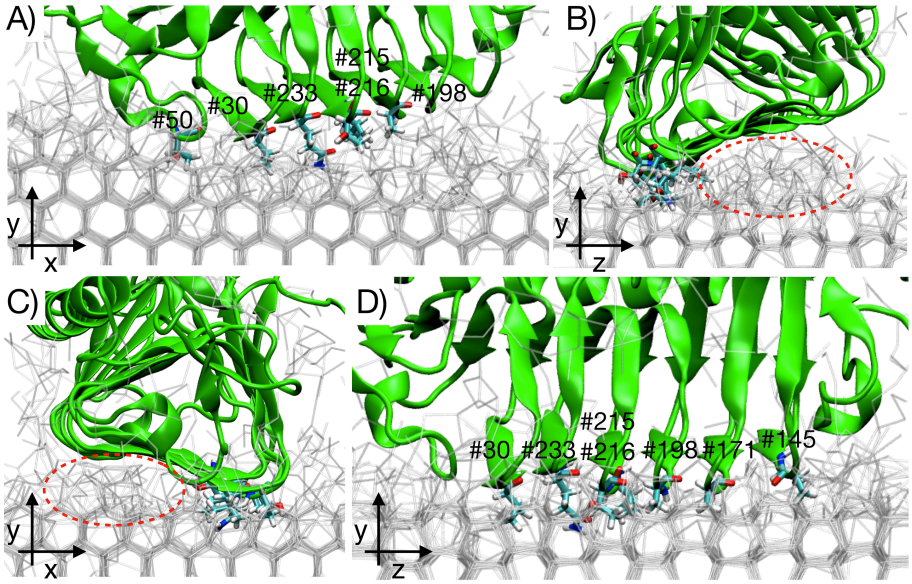

Figure 8. Snapshots of the adsorbed *fc*IBP11 on the primary prism surface, obtained from the last flame (600 ns) of trajectory (A, B) P_b1 and (C, D) P_b3. The immobile water molecules are shown by gray lines. The residues in the loop between *B*- and *C*-faces which directly bind to the ice lattice are shown by sticks with the ID. The red dashed circles indicate the immobile water molecules without ice structure below the *B*-face.

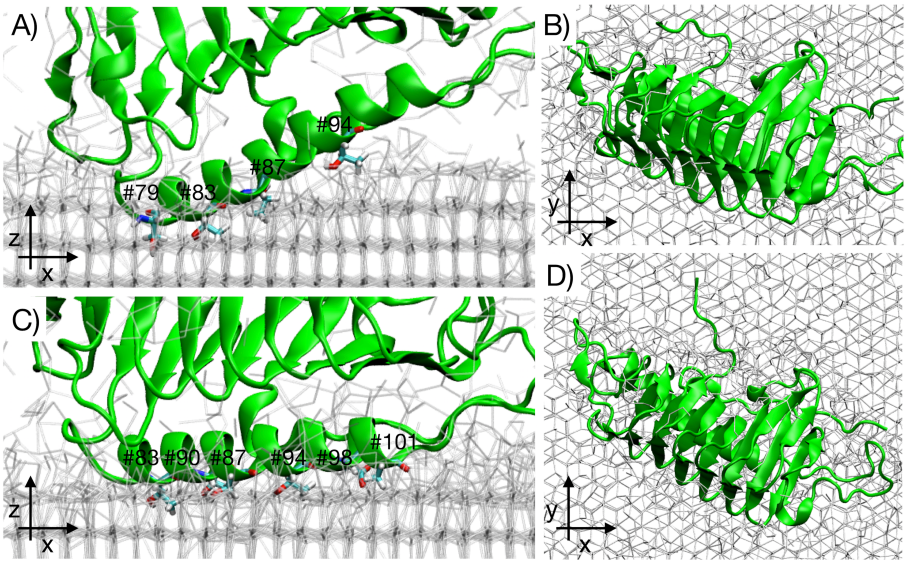

Figure 9. Snapshots of the adsorbed *fc*IBP11 on the basal surface, obtained from the last flame (600 ns) of trajectory (A, B) B_a1 and (C, D) B_a3. The z-axis corresponds to the *c*-axis of ice lattice. The immobile water molecules are shown by gray lines. The residues in the alpha-helix which directly bind to the ice lattice are shown by sticks with the ID.

BIBLIOGRAPHY

1. C. A. Knight, *Nature*, 2000, 406, 249-250.
2. J. A. Raymond and A. L. DeVries, *Proceedings of the National Academy of Sciences, USA*, 1977, 74, 2589-2593.
3. M. Bar Dolev, I. Braslavsky and P. L. Davies, *Annual Review of Biochemistry*, 2016, 85, 23.21-23.28.
4. S. Venketesh and C. Dayananda, *Critical Reviews in Biotechnology*, 2008, 28, 57-82.
5. A. J. Scotter, C. B. Marshall, L. A. Graham, J. A. Gilbert, C. P. Garnham and P. L. Davies, *Cryobiology*, 2006, 53, 229-239.
6. N. Pertaya, C. B. Marshall, Y. Celik, P. L. Davies and I. Braslavsky, *Biophysical Journal*, 2008, 95, 333-341.
7. J. Cheng, Y. Hanada, A. Miura, S. Tsuda and H. Kondo, *Biochemical Journal*, 2016, 473, 4011-4026.
8. M. Bayer-Giraldi, G. Sazaki, N. Ken, S. Kipfstuhl, D. A. Vorontsov and Y. Furukawa, *Proceedings of the National Academy of Sciences, USA*, 2018, 115, 7479-7484.
9. P. M. Naullage, L. Lupi and V. Molinero, *The Journal of Physical Chemistry C*, 2017, 121, 26949-26957.
10. K. Mochizuki and V. Molinero, *J. Am. Chem. Soc.*, 2018, 140, 4803-4811.
11. H. Nada and Y. Furukawa, *J. Phys. Chem. B*, 2008, 112, 7111-7119.
12. D. S. Yang, W. C. Hon, S. Bubanko, Y. Xue, J. Seetharaman, C. L. Hew and F. Sicheri, *Biophysical Journal*, 1998, 74, 2142-2151.
13. E. I. Howard, M. P. Blakeley, M. Haertlein, I. P. Haertlein, A. Mitschler, S. J. Fisher, A. C. Siah, A. G. Salvay, A. Popov, C. M. Dieckmann, T. Petrova and A. Podjarny, *Journal of Molecular Recognition*, 2011, 24, 724-732.
14. C. P. Garnham, R. L. Campbell and P. L. Davies, *Proc. Natl. Acad. Sci. U.S.A.*, 2011, 108, 7363-7367.
15. M. J. Kuiper, C. J. Morton, S. E. Abraham and A. Gray-Weale, *Elife*, 2015, 4, e05142.
16. A. Hudait, N. Odendahl, Y. Qiu, F. Paesani and V. Molinero, *J. Am. Chem. Soc.*, 2018, 140, 4905-4912.
17. M. Bayer-Giraldi, C. Uhlig, U. John, T. Mock and K. Valentin, *Environmental Microbiology*, 2010, 12, 1041-1062.
18. J. A. Raymond, C. Fritsen and K. Shen, *FEMS Microbiology Ecology*, 2007, 61, 214-221.
19. H. Do, S.-J. Kim, H. J. Kim and J. H. Lee, *Acta Crystallographica Section D: Biological Crystallography*, 2014, D70, 1061-1073.
20. M. Mangiagalli, M. Bar Dolev, P. Tedesco, A. Natalello, A. Kaleda, S. Brocca, D. de Pascale, S. Pucciarelli, C. Miceli, I. Braslavsky and M. Lotti, *FEBS Journal*, 2017, 284, 163-177.
21. I. G. Gwak, W. s. Jung, H. J. Kim, S.-H. Kang and E. Jin, *Marine Biotechnology*, 2010, 12, 630-639.
22. M. G. Janech, A. Krell, T. Mock, J.-S. Kang and J. A. Raymond, *Journal of Phycology*, 2006, 42, 410-416.
23. J. A. Raymond and M. G. Janech, *Cryobiology*, 2009, 58, 151-156.
24. N. Xiao, K. Suzuki, Y. Nishimiya, H. Kondo, A. Miura, S. Tsuda and T. Hoshino, *FEBS Journal*, 2010, 277, 394-403.
25. J. K. Lee, K. S. Park, S. Park, P. Hyun, Y. H. Song, S.-H. Kang and H. J. Kim, *Cryobiology*, 2010, 60, 222-228.
26. N. H. F. Hashim, I. Bharudin, D. L. S. Nguong, S. Higa, F. D. A. Bakar, S. Nathan, A. Rabu, H. Kawahara, R. M. Illias, N. Najimudin, N. M. Mahadi and A. M. Murad, *Extremophiles*, 2013, 17, 63-73.

27. T. Hoshino, M. Kiriaki, S. Ohgiya, M. Fujiwara, H. Kondo, Y. Nishimiya, I. Yumoto and S. Tsuda, *Canadian Journal of Botany/Revue Canadienne de Botanique*, 2003, 81, 1175-1181.
28. M. Bayer-Giraldi, I. Weikusat, H. Besir and G. Dieckmann, *Cryobiology* 2011, 63, 2010-2019.
29. Y. Hanada, Y. Nishimiya, A. Miura, S. Tsuda and H. Kondo, *FEBS Journal*, 2014, 281, 3576-3590.
30. M. Mangiagalli, G. Sarusi, A. Kaleda, M. Bar Dolev, V. Nardone, V. F. Vena, I. Braslavsky, M. Lotti and M. Nardini, *FEBS Journal*, 2018, 245, 2901-2914.
31. C. Wang, S. Pakhomova, M. E. Newcomer, B. C. Christner and B.-H. Luo, *PLoS ONE*, 2017, 12, e0187169.
32. T. D. R. Vance, L. A. Graham and P. L. Davies, *FEBS Journal*, 2018, 285, 1511-1527.
33. H. Kondo, Y. Hanada, H. Sugimoto, T. Hoshino, C. P. Garnham, P. L. Davies and S. Tsuda, *Proceedings of the National Academy of Sciences, USA*, 2012, 109, 9360-9365.
34. J. H. Lee, A. K. Park, H. Do, K. S. Park, S. H. Moh, Y. M. Chi and H. J. Kim, *Journal of Biological Chemistry*, 2012, 287, 11460-11468.
35. Y. Gwak, W. Jung, Y. Lee, J. S. Kim, C. G. Kim, J.-H. Ju, C. Song, J.-K. Hyun and E. Jin, *FASEB J.*, 2014, 28, 4924-4935.
36. M. Kim, Y. Gwak, W. Jung and E. Jin, *Marine Drugs*, 2017, 15.
37. H. Nada and Y. Furukawa, *Polym. J.*, 2012, 44, 690-698.
38. A. McPherson and J. A. Gavira, *Acta Crystallographica Section F: Structural biology communications*, 2014, 70, 2-20.
39. N. Igarashi, K. Ikuta, T. Miyoshi, N. Matsugaki, Y. Yamada, M. S. Yousef and S. Wakatsuki, *Journal of synchrotron radiation*, 2008, 15, 292-295.
40. W. Kabsch, *Acta Crystallographica Section D: Biological Crystallography*, 2010, 66, 125-132.
41. N. Collaborative Computational Project, *Acta Crystallographica Section D: Biological Crystallography*, 1994, 50, 760-763.
42. P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart, *Acta Crystallographica Section D: Biological Crystallography*, 2010, 66, 213-221.
43. P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta Crystallographica Section D: Biological Crystallography*, 2010, 66, 486-501.
44. G. N. Murshudov, A. A. Vagin and E. J. Dodson, *Acta Crystallographica Section D: Biological Crystallography*, 1997, 53, 240-255.
45. S. E. Feller and A. D. MacKerell, *J. Phys. Chem. B*, 2000, 104, 7510-7515.
46. J. B. Klauda, B. R. Brooks, A. D. MacKerell, R. M. Venable and R. W. Pastor, *J. Phys. Chem. B*, 2005, 109, 5300-5311.
47. A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, 102, 3586-3616.
48. A. D. MacKerell Jr., M. Feig and C. L. Brooks III, *J. Comput. Chem.*, 2004, 25, 1400-1415.
49. J. L. F. Abascal and C. Vega, *J. Chem. Phys.*, 2005, 123, 234505-234512.

50. H. S. Ashbaugh, N. J. Collett, H. W. Hatch and J. A. Staton, *J. Chem. Phys.*, 2010, 132, 124504.
51. S. R. Pattenaude, B. M. Rankin, K. Mochizuki and D. Ben-Amotz, *PCCP*, 2016, 18, 24937-24943.
52. K. Mochizuki, S. R. Pattenaude and D. Ben-Amotz, *J. Am. Chem. Soc.*, 2016, 138, 9045-9048.
53. B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *Journal of Chemical Theory and Computation*, 2008, 4, 435-447.
54. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, 1-2, 19-25.
55. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, 81, 3684-3690.
56. M. Matsumoto, T. Yagasaki and H. Tanaka, *J. Comput. Chem.*, 2017, 39, 61-64.
57. R. García Fernández, J. L. F. Abascal and C. Vega, *J. Chem. Phys.*, 2006, 124, 144506.
58. A. T. Brunger, *Nature*, 1992, 355, 472-475.
59. C. J. Williams, J. J. Headd, N. W. Moriarty, M. G. Prisant, L. L. Videau, L. N. Deis, V. Verma, D. A. Keedy, B. J. Hintze, V. B. Chen, S. Jain, S. M. Lewis, W. B. Arendall, 3rd, J. Snoeyink, P. D. Adams, S. C. Lovell, J. S. Richardson and D. C. Richardson, *Protein science : a publication of the Protein Society*, 2018, 27, 293-315.
60. J. Jenkins and R. Pickersgill, *Prog. Biophys. Mol. Biol.*, 2001, 77, 111-175.
61. T. Sun, F.-H. Lin, R. L. Campbell, J. S. Allingham and P. L. Davies, *Science*, 2014, 343, 795-798.
62. T. Yagasaki, M. Matsumoto and H. Tanaka, *J. Phys. Chem. B*, 2018, acs.jpcb.7b10356.
63. J. Vatamanu and P. G. Kusalik, *J. Chem. Phys.*, 2007, 126, 124703.
64. S. Zepeda, S. Nakatsubo and Y. Furukawa, *Review of Scientific Instruments*, 2009, 80, 115102.
65. Y. Celik, R. Drori, N. Pertaya-Braun, A. Altan, T. Barton, M. Bar-Dolev, A. Groisman, P. L. Davies and I. Braslavsky, *Proceedings of the National Academy of Sciences, USA*, 2013, 110, 1309-1314.
66. W. L. DeLano, *CCP4 Newsletter On Protein Crystallography*, 2002, 40, 82-92.