

Type II Ice-Binding Proteins Isolated from an Arctic Microalga Are Similar to Adhesin-Like Proteins and Increase Freezing Tolerance in Transgenic Plants

Sung Mi Cho¹, Sanghee Kim², Hojin Cho^{1,3}, Hyungseok Lee^{1,3}, Jun Hyuck Lee^{1,3}, Horim Lee⁴, Hyun Park^{1,3,5}, Seunghyun Kang¹, Han-Gu Choi² and Jungeun Lee^{1,3,*}

¹Unit of Polar Genomics, Korea Polar Research Institute (KOPRI), Yeosu-gu, Incheon 21990, Republic of Korea

²Division of Polar Life Sciences, Korea Polar Research Institute (KOPRI), Yeosu-gu, Incheon 21990, Republic of Korea

³Department of Polar Science, University of Science and Technology, Daejeon 34113, Republic of Korea

⁴Department of Biotechnology, Duksung Women's University, Seoul 01369, Republic of Korea

⁵Division of Biotechnology, College of Life Sciences and Biotechnology, Korea University, Seoul 02841, Republic of Korea

*Corresponding author: E-mail, jelee@kopri.re.kr; Fax, +82(32)7605576.

(Received December 4, 2018; Accepted August 11, 2019)

Microalgal ice-binding proteins (IBPs) in the polar region are poorly understood at the genome-wide level, although they are important for cold adaptation. Through the transcriptome study with the Arctic green alga *Chloromonas* sp. KNF0032, we identified six *Chloromonas* IBP genes (CmIBPs), homologous with the previously reported IBPs from Antarctic snow alga CCMP681 and Antarctic *Chloromonas* sp. They were organized with multiple exon/intron structures and low-temperature-responsive *cis*-elements in their promoters and abundantly expressed at low temperature. The biological functions of three representative CmIBPs (CmIBP1, CmIBP2 and CmIBP3) were tested using *in vitro* analysis and transgenic plant system. CmIBP1 had the most effective ice recrystallization inhibition (IRI) activities in both *in vitro* and transgenic plants, and CmIBP2 and CmIBP3 had followed. All transgenic plants grown under nonacclimated condition were freezing tolerant, and especially 35S::CmIBP1 plants were most effective. After cold acclimation, only 35S::CmIBP2 plants showed slightly increased freezing tolerance. Structurally, the CmIBPs were predicted to have β -solenoid forms with parallel β -sheets and repeated TXT motifs. The repeated TXT structure of CmIBPs appears similar to the AidA domain-containing adhesin-like proteins from methanogens. We have shown that the AidA domain has IRI activity as CmIBPs and phylogenetic analysis also supported that the AidA domains are monophyletic with ice-binding domain of CmIBPs, and these results suggest that CmIBPs are a type of modified adhesins.

Keywords: *Chloromonas* • Ice-binding protein • Freezing tolerance • Ice recrystallization inhibition • Adhesin-like protein.

Introduction

Psychrophilic microalgae are species that tolerate and grow at 0–4°C and have optimal growth rates at temperatures <15°C (Eddy 1960, Morita 1975, Morgan-Kiss et al. 2006), enabling

their survival in snowfields, glaciers and polar freshwater lakes. Under temperatures near the freezing point, the growth of ice crystals causes severe physical damage to cells due to repeated freeze–thaw cycles of water molecules surrounding the cells and the accompanying dehydration (Davies 2014, Bar Dolev et al. 2016).

To protect cells from such damage, psychrophilic microalgae produce ice-binding proteins (IBPs) that bind to the plane of an ice crystal to inhibit its growth (Raymond et al. 2009, Raymond and Morgan-Kiss 2013, Raymond 2014, Raymond and Morgan-Kiss 2017). IBPs have been found in diverse psychrophilic and psychrotolerant organisms, and they are often referred to as antifreeze proteins (AFPs), antifreeze glycoproteins (AFGPs) or ice recrystallization-inhibiting protein (IRIPs) depending on their major activities (Davies 2014, Bar Dolev et al. 2016). Thermal hysteresis (TH) activity, mainly described in studies on AFPs or AFGPs of Antarctic fish (Gonzalez-Aguero et al. 2013) and insects (Graether and Sykes 2004, Duman 2015, Schrodinger 2015), lowers the freezing point below the melting point so that ice crystals form at lower temperatures (Davies 2014, Bar Dolev et al. 2016). Ice recrystallization inhibition (IRI) activity, usually seen in the IRIPs of freeze-tolerant plants (Verbruggen et al. 2009, Middleton et al. 2012, Bredow et al. 2017) and green algae (Raymond et al. 2009, Raymond and Morgan-Kiss 2013, Raymond 2014, Raymond and Morgan-Kiss 2017), inhibits growth of ice crystals at temperatures <0°C, which results in small extra- or intracellular ice crystals (Davies 2014, Bar Dolev et al. 2016).

At present, two major types of IBPs have been identified from microalgae in the polar region based on sequences differences (Raymond and Morgan-Kiss 2013). Type I IBPs have been found in three Antarctic green algae: *Chlamydomonas raudensis* UWO241 (Raymond and Morgan-Kiss 2013), *Chloromonas brevispina* (Raymond 2014) and *Chlamydomonas* ICE-MDV (Raymond and Morgan-Kiss 2017). X-ray crystallographic studies have shown that type I IBPs contain Domain of Unknown Function 3494 (DUF3494), which forms a β -solenoid fold

alongside an α -helix. It has been hypothesized that type I IBPs spread via horizontal gene transfer (HGT) events from bacteria and fungi to proteobacteria, diatoms and copepods (Sorhannus 2011). Unlike type I IBPs, type II IBPs only have been reported from Antarctic snow alga CCMP681 and Antarctic *Chloromonas* sp. (Raymond et al. 2009, Jung et al. 2016). These proteins have found to be more effective in IRI activity than TH activity, and the structural prediction showed that they possess a β -solenoid structure with a parallel plane of ice-binding sites consisting of repeated Thr-rich residues (Raymond et al. 2009, Jung et al. 2016). Although type II IBP proteins have not been found in protein databases, the superficial structure of β -solenoid structures were often found in other IBPs that originated from overwintering plants (Middleton et al. 2012), freeze-resistant insects (Graether and Sykes 2004) or Antarctic marine bacteria, *Marinomonas primoryensis* (Guo et al. 2017).

In this study, we identified type II IBPs from the psychrophilic green alga *Chloromonas* sp. KNF0032, which was isolated from Arctic freshwater. We performed RNA-sequencing (-Seq) analysis and found that this alga progressively expressed multiple IBP genes (*CmlBPs*) with decreasing temperature. Based on in silico predictions, *CmlBP* proteins have a β -solenoid form with a parallel plane consisting of TXT (Thr-X-Thr) motifs, similar to previously identified type II IBPs. To examine the role of *CmlBPs* related to cold adaptation and freezing resistance, we characterized the IRI activity of recombinant *CmlBPs* and evaluated their cryoprotective effect using transgenic *Arabidopsis* plants. Based on the structural similarities and motif-specific features, we found that *CmlBPs* are structurally similar to the AidA domain-containing adhesin proteins present in archaea, bacteria and cyanobacteria and that the AidA domain also retains the IRI activity.

Results

Arctic *Chloromonas* sp. KNF0032

Arctic *Chloromonas* sp. KNF0032 shows a remarkable growth at low temperature compared with other cryophilic strains and has a very high PUFA rate (Jung et al. 2016). To clarify the identity of the strain, the 18S rDNA phylogenetic tree was reconstructed by including four *Chloromonas* strains (Arctic: KNF0012, KNF0030; Antarctic: KSF0057, KSF0063) selected from KCCPM based on phenotype and 18S rDNA sequence information, and other green algae which known to contain type I or type II IBP proteins (Supplementary Fig. S1a). In this tree, *Chloromonas* KNF0032 made a distinct clade with four *Chloromonas* strains of KCCPM AnM0048 and CCMP681. By contrast, *C. brevispina* UTEX: SNO096, *Chlamydomonas* sp. ICE-MDV and UWO241 were dispersed within the *Chloromonas* and *Chlamydomonas* group and did not form a sister relationship. In pairwise distance matrix inferred from 18S rDNA sequences, *Chloromonas* sp. KNF0032 was slightly closer to Arctic strains KNF0012 and KNF0030 (0.003–0.005) than to Antarctic KSF0057, KSF0063, AnM0048 and CCMP681 (0.006–0.007; Supplementary Table S3).

Low-temperature-responsive genes of psychrophilic *Chloromonas* sp. KNF0032

To understand the adaptation mechanisms of the *Chloromonas* sp. KNF0032 strain to low temperature, we investigated the temperature-responsive genes based on the transcriptome profiles at the optimal growth temperature (8°C), low temperature (4°C) and elevated temperatures (12°C and 16°C). Through a computational analysis, 29,631 contigs were assembled de novo, of which 5,153 contigs showing differential gene expression compared with the normal growth temperature (8°C) were identified. To search for the genes involved in low-temperature adaptation, 62 annotated contigs with increased expression at 4°C (4°C vs. 8°C, FDR <0.05, difference >0) were selected. To further narrow the target gene group, we performed hierarchical clustering based on the FPKM values at 4, 8, 12 and 16°C, and finally selected 29 contigs that were specifically expressed at low temperature (Supplementary Table S5). Among these low-temperature-specific expressed contigs, six (~21%) were annotated to IBPs (Fig. 1a).

The genomic fragments of the six contigs were verified by Sanger sequencing. They had 9–12 exons spanning genomic regions of 2.7–4.6 kb in length; therefore, they are referred to as *CmlBP1* to *CmlBP6* in this study (Supplementary Table S6). The FPKM levels of *CmlBP1* to *CmlBP6* increased at low temperatures (Fig. 1b). In particular, *CmlBP1*, *CmlBP2* and *CmlBP3* exhibited remarkably enhanced expression at low temperature (Fig. 1a, b). When the 8°C-acclimated KNF0032 cultures were transferred to 0°C and exposed for 1 week (Fig. 1c), all three *CmlBP* genes were induced more than those at 8°C, and their expression reached the maximum on day 2 (Fig. 1c). Furthermore, *CmlBP1*, *CmlBP2* and *CmlBP3* have known binding motifs (>5) in their promoters for plant transcription factors, such as CBF (Hao et al. 2002), ABI (Gregorio et al. 2014) and WRKY (Ciolkowski et al. 2008), which respond to low temperature and play a key role in abiotic stress signaling in flowering plants (Fig. 2; Supplementary Table S6). These results suggest that *CmlBPs* participate in the adaptive response to freezing temperatures, and that transcription of these genes can be induced by the *cis*-element binding of positive regulators in cold stress signaling pathways, which is presumed to be similar to that of signal transduction in flowering plants.

To confirm that these IBPs were also present in the other *Chloromonas* strains, genomic DNA-PCR was performed on eight additional KCCPM strains using type II IBP-specific primers. As a result, strains in the same clade of KNF0032 had *CmlBP*-homologous genes. We also confirmed that the *CmlBP*-homologous genes were strongly induced at low temperatures (2°C, 7 d; Supplementary Fig. S1).

Thr-rich *CmlBPs* are similar in structure to insect AFP

CmlBP genes were translated into proteins 340–415 amino acids (aa) in length, with expected molecular weights of 32–39 kDa, respectively (Supplementary Fig. S2a). All *CmlBPs* were predicted to have secretory signal peptides (18–31 aa) in the N-terminus according to the SignalP 4.0 server, and *CmlBP3* and

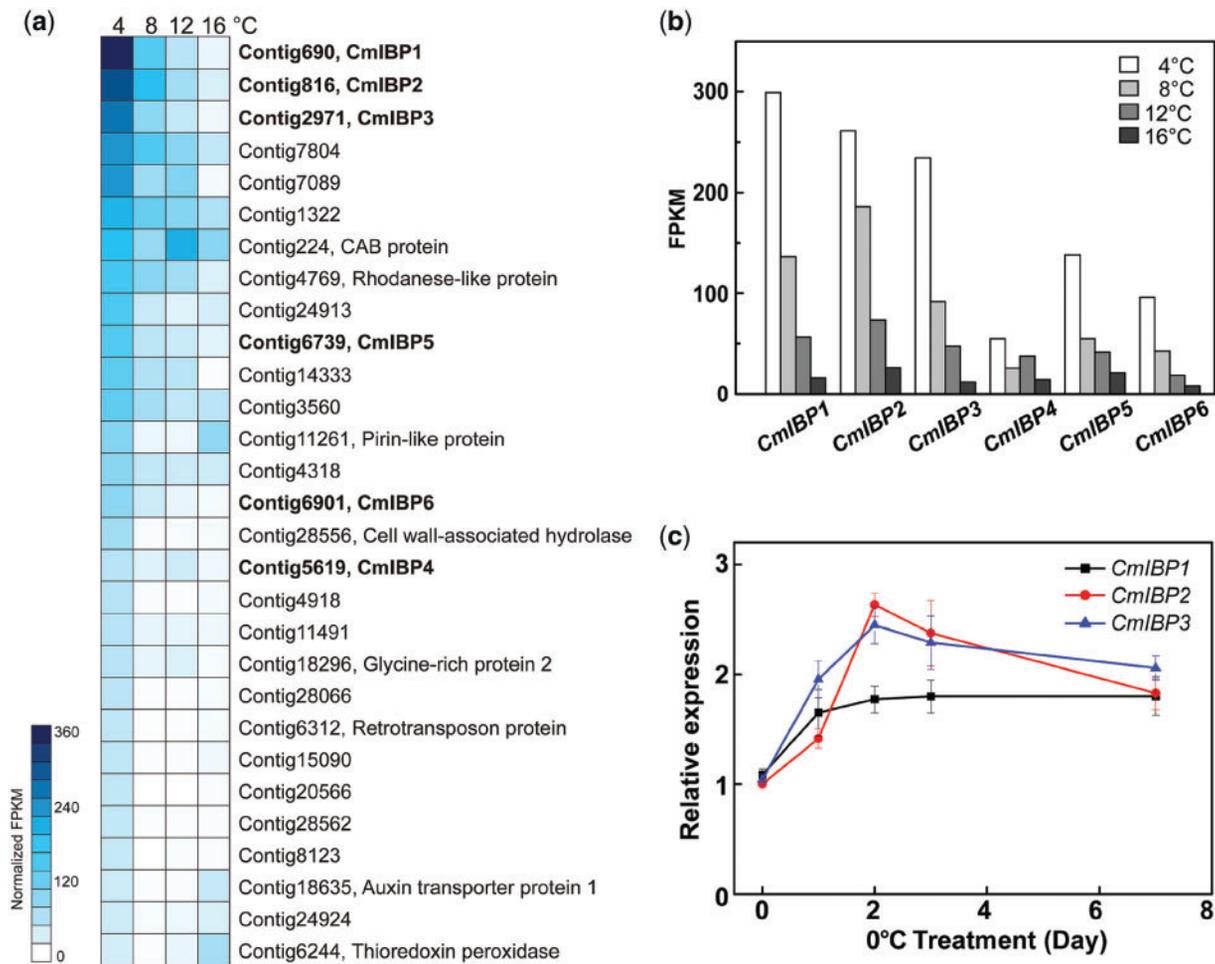
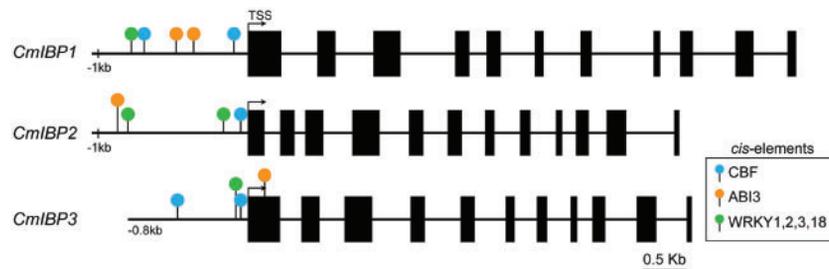


Fig. 1 Transcriptional response of IBPs from the psychrophilic green alga *Chloromonas* sp. KNF0032. (a) Heat map displaying the FPKM levels of 29 annotated genes that increased in response to low temperature. Among those, six genes were identified as CmlBP (*Chloromonas* IBPs). (b) Transcript levels of the CmlBP gene results (CmlBP1–CmlBP6) obtained from RNA-Seq analysis. (c) Transcriptional changes in *CmlBP1*, *CmlBP2* and *CmlBP3* in response to low-temperature treatment quantified by qRT-PCR relative to the expression values of the internal controls contig5016 (methyltransferase-like protein) and contig39 (elongation factor-like protein). Internal control genes were selected and evaluated based on transcriptome data (detailed in the Materials and Methods section). Mean relative expression values are plotted with the standard deviation ($n = 3$).



CmlBP5 were predicted to have an additional transmembrane domain (Supplementary Fig. S2a). A BLAST search of the CmlBP sequences revealed very high similarity with type II

IBPs of *Chlamydomonas* sp. CCMP681 (hereafter, CCMP681 IBPs; Raymond et al. 2009) and to a recently isolated IBP gene of Antarctic *Chloromonas* sp. (hereafter, ChlorolBP;

Jung et al. 2016; Supplementary Fig. S2). The phylogenetic tree clearly showed that the CmIBPs, four CCMP681 IBPs and ChloroIBP formed a clade distinguishable from the previously reported type I IBPs containing the DUF3494 domain from bacteria, fungi, prasinophytes, green algae and diatoms (Fig. 3).

The CmIBPs were composed of 50–53% hydrophobic amino acids (V, I, L, M, F, W, C, G, P and A), but the most abundant amino acids were the hydrophilic amino acid Thr (T; 12–15%). Although CmIBPs had less sequence similarity compared with insect AFP and MpIBP, they both had hydrophilic Thr residues that appeared at a high proportion of 12–15% (Table 1). Notably, Cys residues occupied approximately 5% of CmIBPs, representing a remarkably larger portion compared with MpIBP with no Cys (Table 1). CmIBP1 had seven disulfide bonds, and CmIBP2 and CmIBP3 had six and five disulfide bonds, respectively (Table 1).

Multiple alignments of the CmIBPs and CCMP681 IBPs revealed the presence of seven repeated Thr-X-Thr residues (motifs I–VII in Fig. 4a; Supplementary Fig. S2), previously designated as TXT motifs (Raymond et al. 2009). All TXT motifs were located at regular intervals, but the first and second TXT motifs were slightly apart (40–43 residues) from the third to seventh motifs (25–29 residues). The Thr residues of the TXT motif sequences of CmIBP3, CmIBP4, CmIBP5 and CmIBP6 were fully conserved, but the Thr residues

corresponding to the first or third residue of the CmIBP1 and CmIBP2 TXT motifs were often substituted with Lys(K), Asp(D), Ile(I) or Ser(S) (Fig. 4a; Supplementary Fig. S2).

The tertiary protein structures of the CmIBPs were predicted to have a β -solenoid form (template protein ID: c5gkd, lyase). It was successfully demonstrated that some of the motifs (III–VII) formed a flat plane comprising parallel β -sheets, although the partial sequences of the N-terminus were unpredictable (Fig. 4b). Parallel β -sheets composed of enriched CmIBP TXT motifs were similar to those of insect AFPs, plant IRIPs and an Antarctic bacterial IBP (MpIBP), although they had less sequence similarity and different sizes (Bar Dolev et al. 2016).

Recombinant CmIBP1 shows 5-fold higher IRI activity

Based on the enhanced gene expression of the six CmIBPs under low temperature, we selected three representative proteins, CmIBP1, CmIBP2 and CmIBP3, as targets of further functional studies. To investigate their ice-binding activities, recombinant CmIBP1, CmIBP2 and CmIBP3 were prepared (Supplementary Fig. S3a) and tested for TH and IRI activity. The TH activities of the three recombinant proteins were negligible (TH $<0.1^{\circ}\text{C}$, $0.1\text{ mg}\cdot\text{ml}^{-1}$). However, the IRI activities, a parameter of inhibited ice recrystallization, which thermodynamically favors the formation of smaller ice crystals

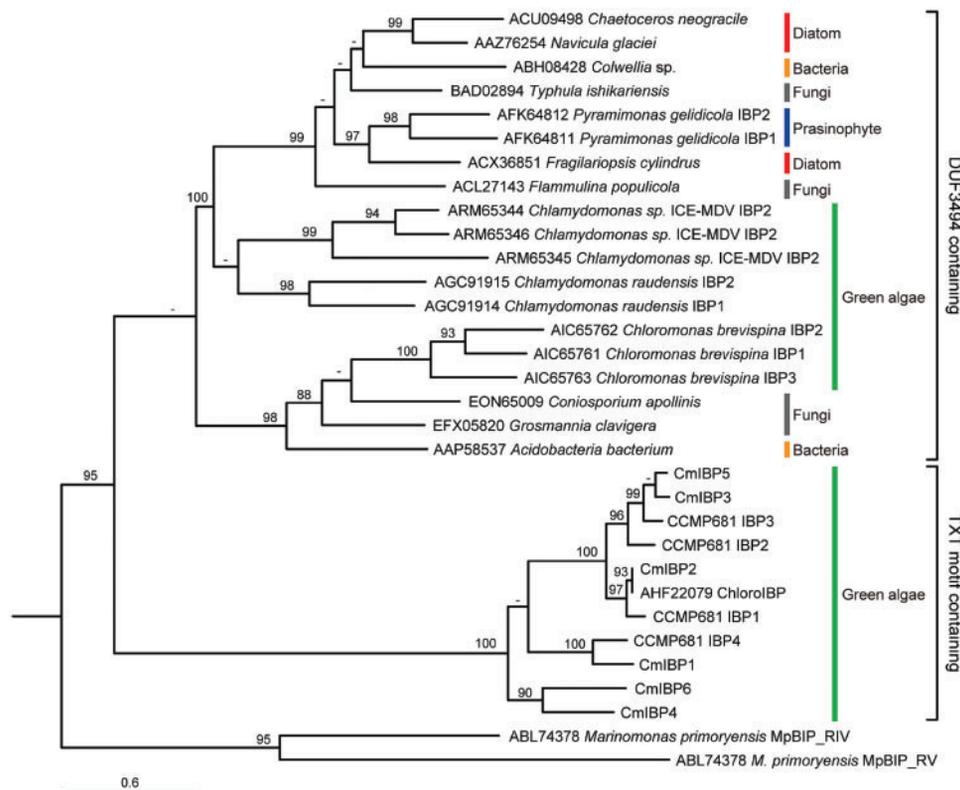


Fig. 3 Maximum-likelihood tree of CmIBPs with previously reported IBP proteins from fungi, bacteria, diatoms, prasinophytes and green algae. The TXT motif-containing IBP group including CmIBPs, CCMP681 IBPs and ChloroIBP formed a monophyletic clade clearly separated from type I IBPs containing the DUF3494 domain. The phylogenetic tree was constructed with PhyML using the WAG+G + F substitution model. Bootstrap analysis was performed with 1,000 replicates and only supporting values >80 are presented on the branches.

Table 1 Comparison of the protein composition of CmlIBPs and other Thr-rich AFPs

	Length (aa)	Thr		Motif	Cys		Disulfide bonds	UniProtKB
		No.	(%)		No.	(%)		
CmlIBP1	341	41	12.0	TXT	17	5.0	7	In this study
CmlIBP2	323	39	12.1	TXT	14	4.3	6	In this study
CmlIBP3	293	39	13.3	TXT	13	4.4	5	In this study
TmAFP	84	17	20.2	TCT	16	19.0	8	O16119
sbwAFP	121	20	16.5	TXT	10	8.3	5	Q9GSA6
MplBP-RIV	326	20	6.1	TXN	0	0	n.a.	A1YIY3

The signal peptide regions were excluded from the calculation.

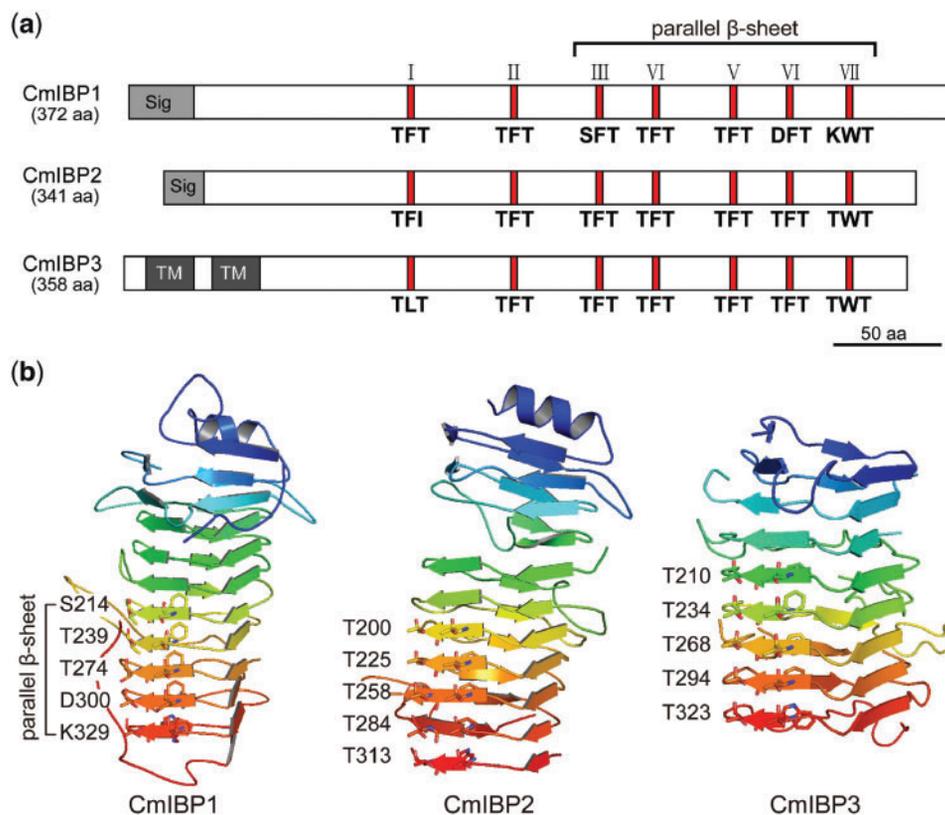


Fig. 4 Prediction of the domains and tertiary structures of the CmlIBP1, CmlIBP2 and CmlIBP3 proteins. (a) Common TXT motifs in CmlIBP1, CmlIBP2 and CmlIBP3 are marked in red and roman numerals (I–VII), and the predicted signal peptides (Sig.) and transmembrane domain are presented as gray boxes. (b) All three proteins were predicted to have a β -solenoid form by the Phyre2 webserver. The third through seventh TXT motifs spaced at regular intervals on protein sequences were arranged on one side of the parallel β -sheets. The aligned sequence regions (CmlIBP1: 150–356 aa, CmlIBP2: 89–337 aa, CmlIBP3: 117–341 aa) to the template (PDB ID: c5gkda) were used for prediction.

(Knight and DeVries 2009), was very high in the CmlIBP recombinant proteins compared with that of the buffer (50 mM Tris-HCl, pH 8.0) and 0.1 mg·ml⁻¹ BSA control (Fig. 5).

As the strength of IRI activity can be compared by how small the ice grain is sustained at a given temperature and duration, the relative ice grains size of the CmlIBPs were calculated from the average sizes of the ice grains against to buffer (Fig. 5). When the recombinant protein samples were treated at -6°C for 30 min, the growth of ice grains was strongly inhibited at 0.2, 0.1 and 0.05 mg·ml⁻¹ CmlIBP1, and no large grains were observed. Under the same treatment conditions at a

concentration of 0.1 mg·ml⁻¹, CmlIBP2 had a moderate IRI effect with delayed ice recrystallization, whereas CmlIBP3 had the least effect, resulting in larger ice grains with polygonal boundaries (Fig. 5a). The relative ice grain size of CmlIBP1 and CmlIBP2 were 5-fold and 2.5-fold stronger than the buffer control, respectively, but the effect of CmlIBP3 was lower (Fig. 5b). In addition, the IRI effect of recombinant CmlIBP1 at 0.2 mg·ml⁻¹ almost fully inhibited the ice recrystallization process (Fig. 5a, b). These results suggest that the IRI strength of the CmlIBPs differs, which may be related to the difference in the affinity of CmlIBPs for ice binding.

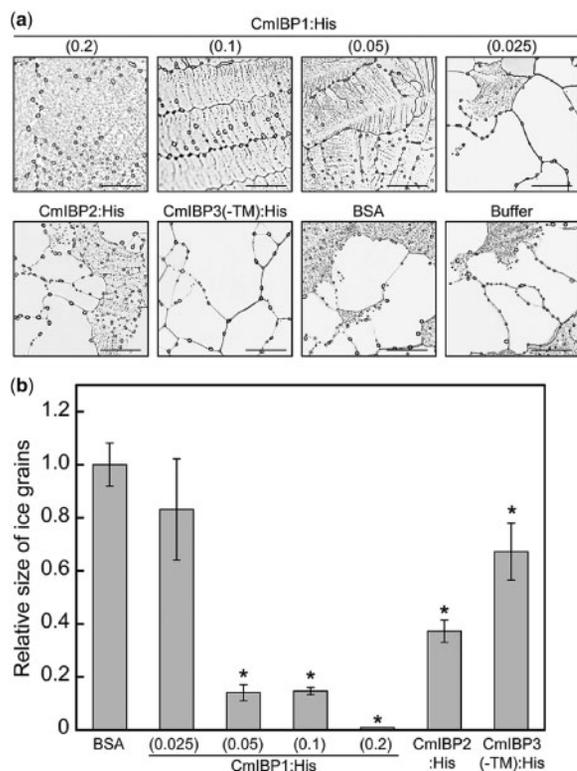


Fig. 5 IRI activity assay of recombinant CmIBPs. (a) Images of ice grains after freezing for 30 min at -6°C in IRI assay. (b) The size of individual ice grains in each sample (3.2 mm^2 in the area) was measured by ImageJ program. Protein concentration was $0.1\text{ mg}\cdot\text{ml}^{-1}$ as a default, and the number of parentheses showed the concentration of CmIBP1. The relative size of ice grains was calculated from the average size of each sample divided by that of the BSA control. The experiments were repeated twice, and the asterisks represent statistical significance ($P < 0.05$) determined by the t -test, compared with the control. The scale bar shows 0.25 mm .

CmIBPs improve freezing tolerance in transgenic *Arabidopsis* plants

The recombinant CmIBPs showed IRI activity thus, we hypothesized that CmIBPs would improve freezing tolerance by protecting cells from freezing damage in a plant expression system. To test this hypothesis, transgenic *Arabidopsis* plants overexpressing *CmIBP1*, *CmIBP2* and *CmIBP3* (hereafter, *35S::CmIBP1*, *35S::CmIBP2* and *35S::CmIBP3*, respectively), were created. No significant phenotypical or developmental differences were observed between transgenic and wild-type plants (Col-0) under the normal growth condition (Fig. 6a). Notably, the RNA transcripts of the transgenes accumulated steadily regardless of cold acclimation, but the protein expression level has differed in each gene (Supplementary Fig. S3). In the case of CmIBP1, the protein expression was less affected by MG132 and cold acclimation treatments, but CmIBP2 and CmIBP3 proteins were not. The intensity of protein band of CmIBP2 appeared to be stronger after cold acclimation, but how the heterologous protein was regulated in plant cells remains unclear. For eliminating the endogenous effect of cold acclimation, nonacclimated (NA) transgenic plants were used for IRI and electrolyte leakage assays.

The IRI experiment with crude protein extracts of *35S::CmIBP* transgenic plants showed dramatic reductions in the size of ice grains or significantly retarded ice crystallization in all tested transgenic plants compared with the wild type (Fig. 6a), indicating that the cell extracts of *35S::CmIBP* plants have IRI activity. The IRI activity was shown in both apoplastic and cytoplasmic regions (Fig. 6b). To investigate whether *35S::CmIBP* transgenic plants have enhanced cell protection ability under freezing conditions, we conducted an electrolyte leakage assay on these plants (Fig. 6c). The level of electrolyte leakage decreased in all transgenic plants grown under NA condition, and especially that of *35S::CmIBP1* plants were significantly lower than others. After cold acclimation, the electrolyte leakage of *35S::CmIBP2* plants further decreased than wild-type plants, but the effects of other transgenic plants were similar. All of the results suggest that the microalgal CmIBPs are secreted into apoplasts and contribute to the increase of freezing tolerance by inhibiting the growth of ice crystals in transgenic plants.

CmIBPs are homologous with the AidA domain-containing adhesin-like proteins, and the AidA domain also exhibits an ice-binding activity

IBPs have been isolated from many eukaryotic microorganisms in cryosphere habitats of polar regions (De Maayer et al. 2014). IBPs harboring the DUF3494 domain, previously defined as type I, have been predominantly found in bacteria, fungi, yeast, sea-ice diatoms, prasinophytes, some green algae and copepods (Uhlir et al. 2015). However, no similar sequence to that of CmIBP has been reported in the NCBI database except type II IBPs isolated from two Antarctic microalgae (Raymond et al. 2009, Jung et al. 2016). In addition, although the DUF3494-containing type I IBP group is presumed to have been obtained from either bacteria or fungi via HGT (Sorhannus 2011, Raymond and Kim 2012), the origin of the type II IBP group (CmIBPs and CCMP681 IBPs) remains unclear.

We detected an AidA domain (Type V secretory pathway, AIDA Adhesin protein, Accession Number: cl25397) in CmIBP2 and CmIBP5 with a Conserved Domain (CD) Search (cutoff: E -value < 0.01) in NCBI. Then, the conserved region of CmIBP2 (positions 173–352) showing homology to the AidA domain was used as a query for position-specific iterated (PSI)-BLAST (cut-off: E -value $< 1\text{-e}5$) for the nonredundant database of NCBI. As a result, 116 orthologous proteins containing an AidA domain were identified (Supplementary Table S4).

From iterative multiple alignments of retrieved protein sequences, we discovered that the conserved motif sequence TXT-X(n)-GGA was repeated at least four times in approximately 14% of AidA domain-containing proteins (16 of 116 proteins). Most (12/16) were adhesin-like proteins derived from Methanobacteriales/Methanomicrobiales (Euryarchaeota). Among the remaining four proteins, three were derived from bacteria (*Bacillus drentensis*, *Syntrophomonas zehnderi* and *Planctomycetes* bacterium) and the other was from the cyanobacterium *Microcystis aeruginosa* (Supplementary Table S4, Fig. S4). Multiple alignments of the sequences of the three CmIBPs and the AidA domains from 16

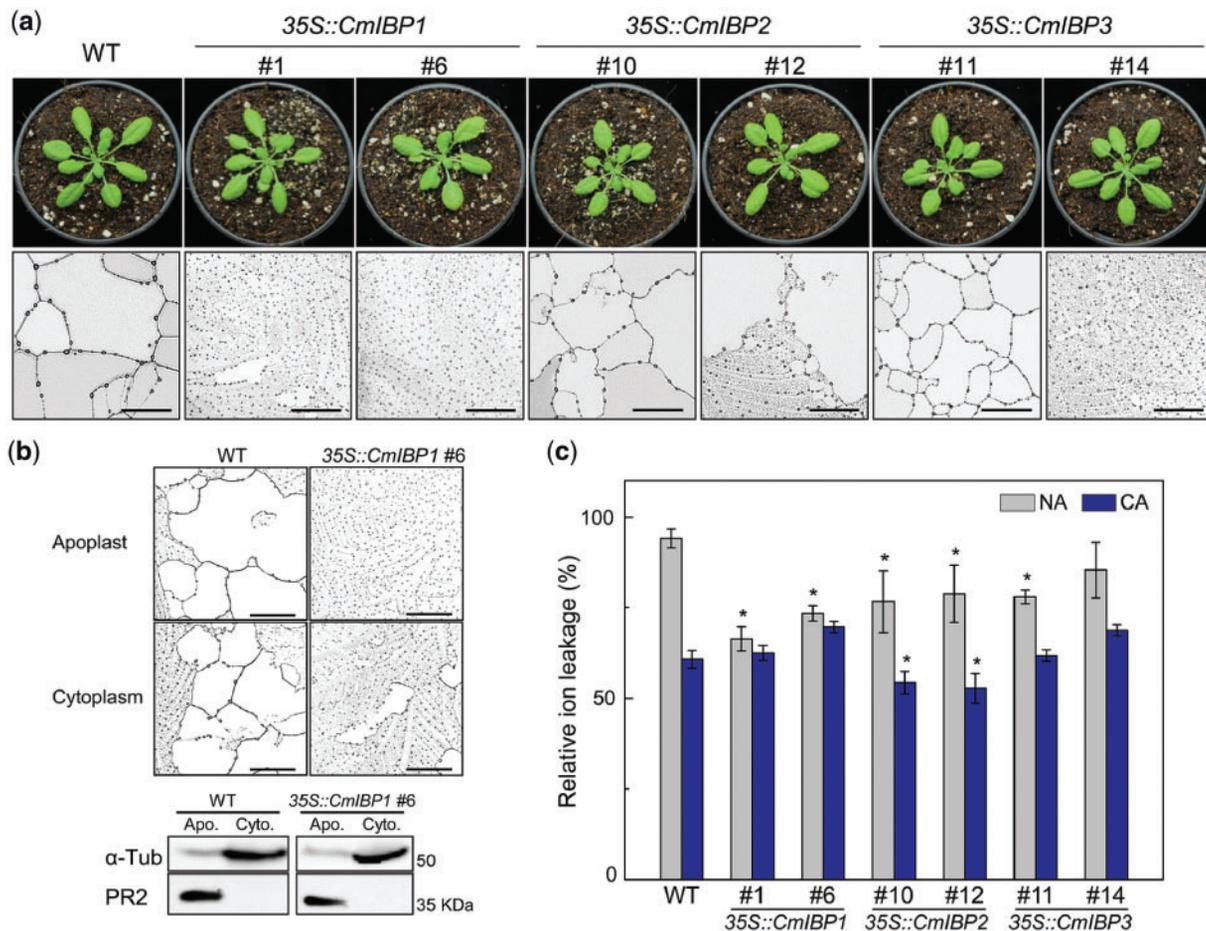


Fig. 6 Transgenic *Arabidopsis* plants overexpressing *CmIBPs* enhance freezing tolerance due to IRI activity. (a) Images of transgenic *Arabidopsis* plants under normal growing conditions at 22°C (upper panel) and the ice grain images of the cell extracts from each plant after freezing for 30 min at −6°C in IRI assay (lower panel). Scale bar shows 0.25 mm. (b) IRI assay and Western blotting results using apoplastic and cytoplasmic protein fractions extracted from wild-type and 35S::CmIBP1 #6 plant leaves. IRI assay was performed as previously described. In Western blotting, anti-PR-2, pathogen-related protein-2, and anti- α -tubulin used as the apoplastic and cytoplasmic marker protein antibody, respectively. (c) Electrolyte leakage measurements after freezing treatment at −6°C of the transgenic plants grown under NA and CA conditions. Relative ion leakages of the samples were compared with that of the wild-type plants grown under NA condition. The experiments were repeated four times and the asterisks represent statistical significance ($P < 0.05$) determined by the t -test compared with the wild-type control.

proteins showed that they shared characteristic repeated TXT sequences at regular intervals (Fig. 7a; Supplementary Fig. S5). A β -solenoid (PDB ID: c5gkda, lyase) was predicted for these proteins by the Phyre2 webserver, and the TXT motifs were laid on the parallel β -sheets forming a flat plane, as were the *CmIBPs* (Figs. 4b, 7a).

The AidA domain has been known to bind to various substrates as a functional domain of bacterial/archaeal adhesin-like protein (Sherlock *et al.* 2004, Fang *et al.* 2005, Ng *et al.* 2016, Poehlein *et al.* 2017). Therefore, we isolated the AidA domain from an archaeal adhesin-like protein of *Methanobacteria paludis* (142–302 in amino acids of WP_013825845) and examined whether this AidA domain also exhibits IRI activity. As shown in Fig. 7b, the IRI activity of the AidA domain was concentration-dependent. At a protein concentration of 0.1 mg·ml^{−1}, the ice grain size was reduced by 40–50% compared with the controls, indicating that the IRI effect of the AidA domain is similar to that of *CmIBP2*.

Taken together, these results suggest that the AidA domain, which has a tertiary structure similar to *CmIBP*, can act as a functional unit that can adhere to the ice surface, such as *CmIBPs*. In addition, the phylogenetic tree (Fig. 7c) perfectly supported that the AidA domains of archaeal, bacterial and cyanobacterial proteins were monophyletic with the conserved motif (III–VII) region of the *CmIBPs*. In particular, some AidA domains of *Methanofollis ethanolicus*, *Methanobacterium subterraneum* and *Methanobacterium formicum* were much closer to the bacterial or cyanobacterial AidA domain-containing proteins (Fig. 7c), suggesting that archaeal AidA domains may have ubiquitously spread in other organisms.

Discussion

Type II IBP was first reported by CCMP681, Antarctic green algae (Raymond *et al.* 2009), and the research provided

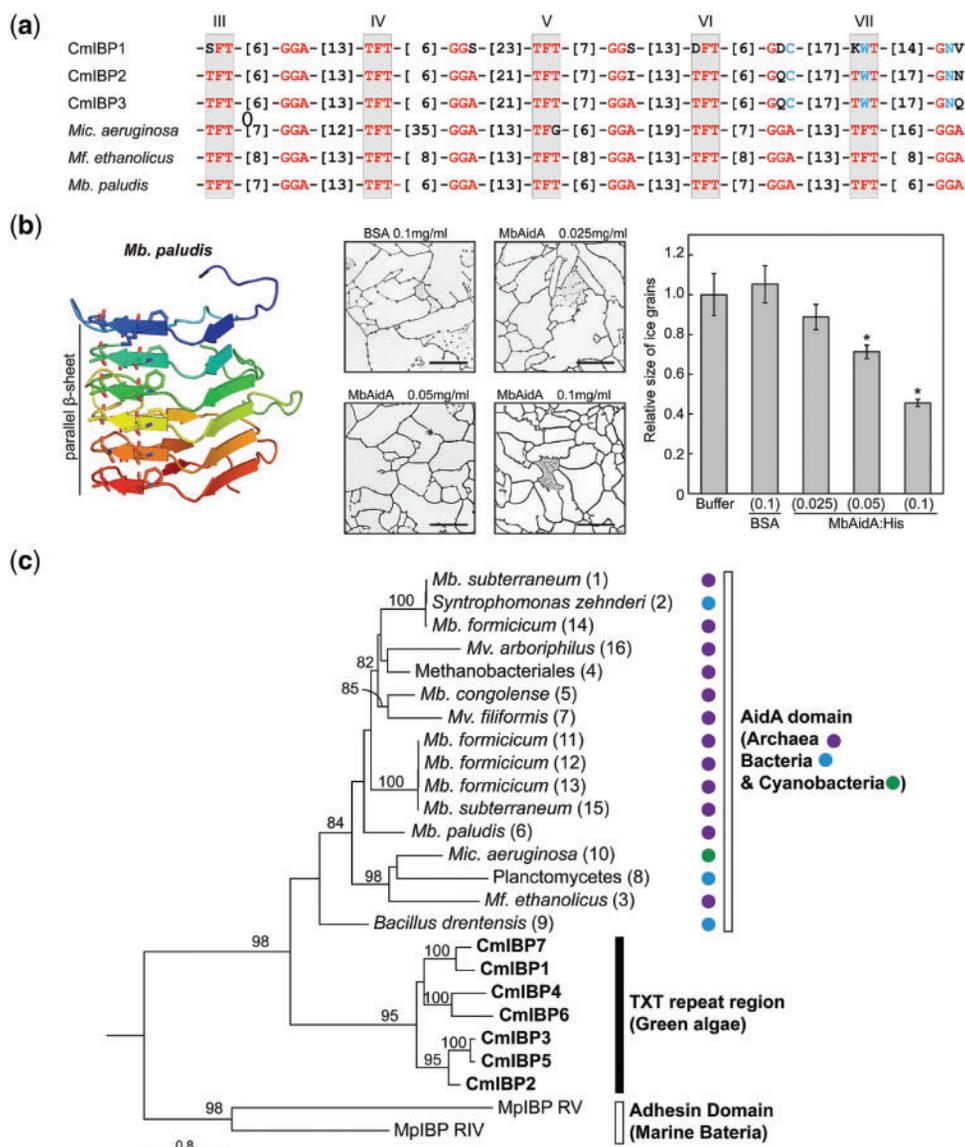


Fig. 7 Structural characteristics shared by CmlBPs and adhesin-like proteins. (a) Representative adhesin-like proteins of *M. ethanolicus* (258–532 aa, WP_067051932), *M. paludis* (142–302 aa, WP_013825845, MbAidA) and *M. aeruginosa* (113–334 aa, WP_103672740) were used to align with CmlBP proteins. Conserved residues are marked in red, and CmlBP-specific residues are in blue. The length between sequences (TXT or GGA) is displayed in parentheses. (b) Protein structure prediction and IRI activity assay of the synthesized MbAidA protein. The predicted protein structure of MbAidA resembled the β -solenoid structure (template protein PDB: d5gkda) of CmlBPs. IRI assay was performed as previously described, and the relative size of ice grain was displayed by compared with the buffer control. The number of parentheses showed the concentration of samples. The experiments were repeated twice, and the asterisks represent statistical significance ($P < 0.05$) determined by the t -test, compared with the control. The scale bar shows 0.25 mm. (c) Maximum-likelihood tree inferred from the AidA domains of 16 adhesin-like proteins (111–156 aa in length) and AidA homologous region of CmlBPs (motifs III–VII, 144–146 aa in length). The number of parentheses indicated an identifier in Supplementary Table S4. The RVI and RV domains of MpiBP (ABL74378) were used as outgroup. The maximum-likelihood tree was constructed using PhyML 3.0 with the WAG+G + F substitution model and the bootstrap analysis was conducted with 1,000 replicates. Supporting values >80 are displayed on the branches.

important biological information about type II IBP. For example, the IRI activity of an algal culture supernatant was confirmed, and four IBP sequences were identified through cDNA library and RACE analysis. And also the IRI activity was confirmed in one of the proteins. Later, Jung et al. (2016) showed the biochemical activity of another homologous gene of type II IBP in Antarctic *Chloromonas* sp. strain isolated and discussed the structural properties of it through modeling.

In this study, through a de novo assembly of RNA-Seq, we demonstrated that type II IBP genes are present in the genome and many of them are regulated in a temperature-dependent manner. We annotated coding sequences of >10 IBPs and analyzed their genic and regulatory regions. We compared the activity of IBP isoforms and found that they differ in IRI activity, and further demonstrated experimentally that they can improve the freezing tolerance when they are expressed

in plants, suggesting their possible role in the agricultural application. Therefore, this is the most comprehensive and advanced study for type II IBPs so far.

The rarity of type II IBP

Green algae are frequently found in icy environments, but the microalgal type I and II IBPs appears to be unevenly distributed in nature. Unlike type I, type II IBPs are found in a particular taxon group consisting of a few strains. In situ expression analyses of Arctic and Antarctic sea-ice communities showed that approximately 90% of the IBP transcripts were grouped with type I IBP sequences and had the DUF3494 domain (Uhlig et al. 2015). In addition, the *Chloromonas* species with type II IBP have not been reported in the Arctic to date. Recently, the research about the bipolar dispersal of red-snow algae using amplicon deep sequencing showed that the ITS2 OTU corresponding to *Stephanospaherinia* clade belonging to KNF0032 was found to be very rare (0.1% of the total reads, 0.066 % of the Antarctic reads; Segawa et al. 2019). As such, their cryosphere biomass ratio is low, so it is considered that they can be isolated as a single strain only through culture enrichment. In addition, previous studies might have excluded them during sampling because they have focused on the red-snow algae which dominated by other snow-algae groups.

In this study, we found that *Chloromonas* strains obtained from KCCPM, a collection of Arctic and Antarctic microalgal strains, form a single clade of and all contain type II IBP in both the Arctic and the Antarctic species (Supplementary Fig. S1, Table S3). Therefore, to define their ecological status and habitat distribution of various microalgae, it is necessary to diversify the sampling environment and collect more data on the polar terrestrial and freshwater algae.

Multiple CmIBPs within the KNF0032 genome and their physiological roles

Multiple IBP genes are found in psychrophilic algae, indicating their biological importance in nature. Studies have predicted at least nine IBP genes in *C. brevispina* (Raymond 2014), *Chlamydomonas raudensis* UWO241 (Raymond and Morgan-Kiss 2013) and *Chlamydomonas* sp. ICE-MDV (Raymond and Morgan-Kiss 2017), and their expression increases at subzero temperatures. We identified at least six CmIBPs genes present as multiple exon–intron structures in the *Chloromonas* sp. KNF0032 genome, with putative cold-responsive *cis*-elements in their promoters.

The CmIBPs differed in their transcriptional response to temperature and each gene had different recrystallization inhibition activities. For example, of the six CmIBPs, CmIBP1, 2 and 3 had markedly higher transcriptional responses to low temperatures. All three proteins showed RI activity, but the extent has differed, and CmIBP1 had the strongest RI activity both in the purified recombinant proteins and in the cell extracts of transgenic plants. Thus, the three CmIBPs act together to improve the overall IRI effects when *Chloromonas* sp. KNF0032 is exposed to low temperature in the natural environment.

Similarly, *Lolium perenne*, a temperate overwintering plant, possesses four IRI proteins and one AFP (Lauersen et al. 2011). Of these, LpIRI2 or LpIRI3 best improve the freezing survival of transgenic *Arabidopsis*. When these two LpIRI genes were co-expressed with LpAFP, the freezing tolerance of transgenic plants is further enhanced (Bredow et al. 2017), implying that these genes have different cellular activities and additive effects under freezing conditions.

It is well known that many extreme organisms have IBP genes in their genome, but there are many things that need to be clarified about the genetic regulatory function of IBPs. Thus, further genetic analysis of the cellular function of the IBPs may be needed to answer whether IBPs are essential for survival for extremophiles. Currently, we are constructing the *Chloromonas* mutants for CmIBP genes. We expect that the phenotypic analysis results can explain whether CmIBPs are a prerequisite for the freezing survival of psychrophilic *Chloromonas* sp.

Structural properties of the CmIBPs related to RI activity

CmIBP1 had stronger IRI effects than CmIBP2 and CmIBP3, although all three had similar sequences and structures. The disulfide bond prediction revealed that CmIBP1 had seven disulfide bonds, and CmIBP2 and CmIBP3 had six and five disulfide bonds, respectively (Table 1). Because the number of disulfide bonds between two Cys residues can affect structural conformation, these results suggest that the protein structural rigidity of CmIBPs may differ. The disulfide bonds in IBPs are important for preserving the hydrophobic core in the protein, while the number of Cys residues and their ability to produce disulfide bonds determines the structural rigidity of the protein (Davies 2014). For example, the TH activity of TmAFP decreases dramatically when the formation of disulfide bonds is inhibited by dithiothreitol (Gauthier et al. 1998). CmIBP1 was predicted to have more disulfide bonds (seven from 17 Cys residues) than CmIBP2 (six from 14 Cys residues) and CmIBP3 (five from 13 Cys residues; Table 1). This observation suggests that the higher number of disulfide bonds in CmIBP1 plays a role in its greater IRI activity. In addition, CmIBP2 and CmIBP3 had a perfectly conserved TXT motif, whereas CmIBP1, with stronger IRI activity, had Asp, Ser or Lys replacing the Thr amino acids in its TXT motifs. In TXT motif, two Thr residues are outward-pointing and X residues, usually Phe(F) and Trp(W) having an aromatic ring, are inward-pointing which has a critical role in maintaining the hydrophobic core structure in IBPs. Most of the water in ice-binding site forms a cage that binds the hydrophobic moiety like the methyl group of Thr, while some water molecules are hydrogen-bonded to nearby hydrophilic groups such as the hydroxyl group of Thr, so that waters enclosed within the structure are temporarily oriented and anchored to fit to the ice-like waters (Bar Dolev et al. 2016). Some studies reported that IRI activity disappears when Thr is replaced with larger or more hydrophobic amino acids such as Leu and Tyr (Graether et al. 2000, Jung et al. 2016). Thus we suggest that the charged amino acids such as Asp(D), Ser(S) or Lys(K) replacing Thr are possible to improve the ice-binding affinity. These differences of CmIBP1

compared with others are presumably applicable to elevate the ice-binding activity for industrial use.

Is CmlBP a specialized adhesin-like protein for binding to ice?

The archaeal adhesin-like proteins, which are present in large numbers in the genomes of both host-associated and nonhost-associated species (Ng et al. 2016, Poehlein et al. 2017), suggest that methanogens in diverse anoxic environments may bind to interacting partners or specific substrates through cell surface-associated functions.

We found structural similarity and a phylogenetic relationship between the ice-binding sites of CmlBPs and the AidA domain, a functional adhesin domain of the archaeal adhesin-like proteins. Our results have shown that the conserved TXT-X(n)-GGA rich region of CmlBPs are similar to the AidA domain through iterative BLAST searches and sequence alignments. Furthermore, we demonstrated that the AidA domain can bind to the ice surfaces and inhibit the ice recrystallization, suggesting that CmlBPs may be one of various types of adhesins.

On the other hand, the recently reported 1.5-MDa RTX adhesin, (MplBP) of the Antarctic marine bacterium *M. primoryensis* demonstrates that this bacterial adhesin is necessary for adaptation to icy environments. To adapt to an ice-covered lake environment, *M. primoryensis* produces MplBPs to attach to both diatom cells and ice surfaces using RIII and RIV domains, respectively, to take up oxygen and nutrients produced by sea-ice diatoms (Guo et al. 2017), suggesting that various types of adhesin proteins can be used as a strategy for organisms to survive in the icy environments.

Taken together, our results suggest that CmlBPs are another example of proteins with adhesin-like domains that bind to ice and inhibit recrystallization. This hypothesis raises additional questions as to how various forms of adhesins have evolved to allow organisms to adapt to diverse environments and whether methanogenic adhesins are involved in cold adaptation of Arctic *Chloromonas* sp.

IRI activity and freezing tolerance phenotype of CmlBP-overexpressing transgenic plants

CmlBP1, CmlBP2 and CmlBP3 have a secretory signal, and CmlBP3 has an additional transmembrane domain at the N-terminus. Therefore, they are presumed to be secreted out of the extracellular spaces or to membrane-anchored to inhibit the growth of ice crystals around the cells. In many microalgae which possess type I or type II IBP showed IRI activity in culture media. In particular, type II IBPs have been detected in culture media of CCMP681 and Antarctic *Chloromonas* sp. (Raymond et al. 2009, Jung et al. 2016). In a recent study, the fusion protein, which combines signal peptides of CCMP681 type II IBP with mCherry fluorescent protein, was successfully secreted out of the cells in *Chlamydomonas reinhardtii* expression system (Molino et al. 2018). However, it was necessary to confirm whether the CmlBPs are secreted out of the cell and function in the apoplastic region in the transgenic plant because ice

nucleation starts in the apoplast because the apoplast has a generally higher freezing point than the cytoplasm. The fact that overexpression of partial LpAFP without secretory signal in *Arabidopsis* does not result in improved freezing tolerance (Bredow et al. 2017) supports the hypothesis that whether ice-binding protein is expressed in an extracellular region or a cytoplasmic region is important regarding freezing tolerance of plants (Bredow and Walker 2017). In our analysis, an IRI activity was observed in both apoplastic and cytoplasmic fractions of transgenic plants, suggesting that CmlBPs exist in both cytoplasmic and apoplastic regions of 35S::CmlBP1 plants. This is probably because the secretory signals of CmlBPs differ from that of plants and secretion efficiency can be lower than that of the plant. More research is needed to see whether these secretory signal peptides interchangeable between higher plants and microalgae. Therefore, it is necessary to confirm the subcellular localization using an experimental method such as a combination of a fluorescence reporter and the secretory signal peptides of CmlBPs.

Our results imply that the heterologous CmlBP proteins can improve the freezing tolerance in the transgenic plant without cold acclimation. In plant cells, the heterologous proteins seem to be regulated at the post-translational level (Supplementary Fig. S3), and thus it should be verified the protein stability, protein (or DNA)-protein interaction, and subcellular localization and so on. By applying the heterologous CmlBPs expression system, however, we suggest that the transgenic plants (and crops) can be cultivated in a wide range of temperatures and elevate resistance to frost. To enhance IBP's capabilities, we need to develop methods for increasing protein stability with higher IRI activity in the cells. For example, optimization of coding sequences had a positive effect in a decrease in plant freezing temperature when insect AFP is expressed in *Arabidopsis* (Huang et al. 2002), and various manipulation methods of protein size to broaden the ice-binding surfaces have been effective in increasing the IBP activity (Bar Dolev et al. 2016). It is also expected that the fusion of the signal peptide, which induces an active extracellular secretion in plants, would be effective. Therefore, it is expected that the transgenic plants with more improved freezing tolerance could be developed by using the optimized genetic engineering methods for CmlBPs.

Materials and Methods

Strain culture conditions and temperature treatments

All KNF and KSF strains used in this study were obtained from the KCCPM (KOPRI Culture Collection), which was originally collected near the Arctic Dasan Research Station, Ny-Alesund, Norway (78°55'N, 11°56'E), and King Sejong Antarctic Station, King George Island, Antarctica (62°14'S, 58°44'W). KCCPM has designated Arctic strains to KNF, and Antarctic strains to KSF. The strains were maintained in Bold's basal medium at 8°C under continuous light (30 $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$). For the temperature response experiment, *Chloromonas* sp. KNF0032 cultures were transferred to growth chambers set to 4, 8, 12 and 16°C at the exponential stage, and grown under the same light conditions for an additional week to assess their response to temperature.

RNA extraction and RNA-sequencing library construction

After the temperature treatment, the algal cells were harvested via centrifugation and homogenized with a mortar and pestle containing liquid nitrogen. Total RNA was extracted using the RNeasy Mini Kit (Qiagen, Hilden, Germany) in conjunction with the RNase free DNase Set (Qiagen) following the manufacturer's protocols as previously described (Cho et al. 2018). Briefly, two biological replicates were prepared. The integrity and concentration of RNA were determined using a Bioanalyzer system (RIN >6; Agilent Technologies, Palo Alto, CA, USA) and a Qubit RNA Broad-range Assay Kit (Life Technologies, Carlsbad, CA, USA), respectively. To construct the RNA-Seq libraries, 1.5 µg of total RNA from each sample was used as input for the TruSeq RNA Sample Prep Kit v2 (Illumina, San Diego, CA, USA). The libraries were quantified using a bioanalyzer and the library quantitative polymerase chain reaction method following the Illumina guidelines. After quantification, they were multiplexed in equal ratios. Sequencing was performed on a MiSeq Sequencer system (Illumina), and a total of 4.98 Gb (28 M paired-end reads) of sequencing data were generated. The sequencing data are available in the NCBI Sequence Read Archive under Accession Number SRP139417.

De novo assembly, annotation and differentially expressed gene analyses

De novo assembly was performed using CLC Genomics Workbench v7.5 software (Qiagen). The reads were filtered by trimming adapter sequences, excluding low-quality sequences (quality score <0.001, ambiguity <2 bp), and removing sequences that were too short (length >50 bp) and duplicates. The resulting reads were assembled using the following parameters (word size = 20, bubble size = 50, length >200 bp). A total of 29,631 assembled contigs (average: 894 bp, N50: 1,246 bp) were subjected to BLASTX searches against the nonredundant protein database with an *E*-value threshold of 1×10^{-3} (Altschul et al. 1990). Gene ontology mapping and annotation (*E*-value < 1×10^{-10}) were performed using the Blast2GO platform (Conesa and Gotz 2008). The expression values were measured in fragments per kilobase of exon model per million mapped reads (FPKM) normalized values at the transcript level. Pairwise comparisons were made with Baggerley's tests in CLC Genomics Workbench, and the differentially expressed genes were identified using a cutoff value (corrected *P*-value of FDR <0.05, difference \neq 0).

Gene structural and in silico promoter analyses

Genomic DNA of the *Chloromonas* sp. KNF0032 strain was prepared using the DNeasy Plant Mini Kit (Qiagen). The coding regions of the six *CmlBP* genes were amplified by PCR using gene-specific primers (Supplementary Table S1), and the sequences were verified by Sanger sequencing. The upstream sequences (~1 kb) of the six *CmlBP* genes were predicted based on the sequence information of the *CmlBP* genomic contigs, which were obtained from partial genome assembly data of KNF0032. The genomic contigs were found by a Blast search using the *CmlBP* sequences as queries, and the sequences found were verified by Sanger sequencing. The confirmed genomic DNA sequences were used to identify the *cis*-regulatory element analysis in TRANSFAC (Matys et al. 2003). The exon–intron boundaries of *CmlBP1*, *CmlBP2* and *CmlBP3* were determined by comparing the sequences of genomic DNA (GenBank ID: MN011069–MN011071) and the coding sequence region (Fig. 2).

The quantitative real-time reverse-transcription PCR analysis

For the quantitative real-time reverse-transcription (qRT-)PCR analysis, *Chloromonas* sp. KNF0032 strain was cultured at 8°C to the exponential growth stage and then cultured in a 0°C chamber for 1, 2, 3 and 7 d. Total RNA was extracted from the harvested cells under the given conditions and purified using the RNeasy Plant Mini Kit (Qiagen). cDNA was synthesized from 2 µg of total RNA. Gene-specific primers were designed according to the sequences of the contigs and are listed in Supplementary Table S1. To select internal control genes, we selected 10 candidate genes based on the average expression value (>30) and coefficient of variation (CV) (<0.1) in all transcriptome data (Supplementary Table S2; Gonzalez-Aguero et al. 2013). An

amplification efficiency test of the 10 candidate genes was performed ($n = 3$) to test expression stability, and the threshold cycle (Ct) values were used with the RefFinder analysis tool (Xie et al. 2012). The RefFinder suggested two reference genes, contig5016 (methyltransferase-like protein) and contig39 (elongation factor-like protein), based on comprehensive comparisons of results from the geNorm, NormFinder, BestKeeper and delta CT methods. The analysis was performed with biological triplicates, and the average expression values were plotted with the standard deviation.

Phylogenetic analyses and structural prediction of proteins

To clarify the phylogenetic position of strains, the 18S rDNA sequences were amplified from five strains of *Chloromonas* sp. using gene-specific primers JO2 and SS17HR (Verbruggen et al. 2009) and those sequences were deposited in NCBI (GenBank ID: MH400028–MH400032). Additional sequences of *Chloromonas*, *Chlamydomonas* and *Chlorella* species were downloaded from NCBI. Multiple alignments performed using MAFFT ver.7 (Katoh et al. 2019), and a phylogenetic analysis was conducted by the Neighbor-Joining method in MEGA7. Bootstrap analyses were conducted by 1,000 replicates and the supporting value (>50) was shown on the branches. The pairwise distance was calculated using the Maximum Composite Likelihood model in MEGA7 (Supplementary Table S3).

The coding sequences of six *CmlBP* genes from *Chloromonas* KNF0032 were amplified from the cDNA which was prepared for RNA-Seq analyses and those sequences were deposited (GenBank ID: MH400035–MH400040) in the NCBI database. To screen the homologous IBP proteins in the remaining four strains of *Chloromonas*, we prepared the cDNA of the samples cultivated at 2°C for a week as described above. RT-PCR was carried out using the gene-specific primers for amplifying the full sequences of CDS (Supplementary Table S1). PCR amplification was performed with following conditions: 98°C for 30 s followed by 28 cycles of 98°C for 10 s, 53°C for 10 s and 72°C for 30 s, with the final extension of 7 min at 72°C. The other *IBP* sequences used in this study were downloaded from the NCBI database.

The AidA domain-containing protein sequences were collected from PSI-BLAST search (*E*-value <0.005) using *CmlBP2* sequence as a query, and then their AidA domains were determined from the CD Search results in NCBI (Supplementary Table S4). Multiple alignments were made using MAFFT ver. 7 with the G-INS-i method (offset value = 0.1; Katoh et al. 2019), and only highly conserved regions were used for the phylogenetic analysis. A maximum-likelihood tree was constructed using PhyML 3.0 in the webserver (Lefort et al. 2017), and the substitution model for both IBP and adhesin phylogenies was selected from the WAG+G+F model based on the Akaike information criterion. Bootstrapping was repeated 1,000 times, and the phylogenetic tree was visualized with FigTree v1.4.3 (Rambaut 2012). Putative signal peptides were predicted with SignalP 4.1 (Petersen et al. 2011), and the domain structure was predicted using SMART (Letunic and Bork 2018). Disulfide bonds were predicted by Dipro in the SCRATCH protein predictor (<http://scratch.proteomics.ics.uci.edu/index.html>, last accessed 26 August 2019). Tertiary structures were predicted with the Phyre2 webserver (Kelley et al. 2015) and the graphical view was manipulated with the PyMol program (Schrodinger 2015).

Cloning and protein purification of recombinant CmlBPs

To construct *CmlBP* recombinant proteins, the coding regions for *CmlBP1*, *CmlBP2* and *CmlBP3*, excluding the signal peptides and transmembrane sequences, were amplified using gene-specific primers (Supplementary Table S1) ligated into the pET22b expression vector (Novagen, Gibbstown, NJ, USA). For the MbAidA protein expression, the codon-optimized sequence of the predicted AidA domain region (142–302 in amino acids) of *M. paludis* (WP_013825845) was synthesized (Bioneer, Daejeon, Korea), amplified using gene-specific primers (Supplementary Table S1), and ligated into pET28a expression vector (Novagen). After verifying the sequences, each construct was transformed into BL21(DE3) cells (Novagen). For protein purification, the cells were grown to an optical density at 600 nm of 0.6, induced for 48 h at 15°C with 0.5 mM isopropyl β-D-1-thiogalactopyranoside, and harvested by centrifugation. Then, the pellets were resuspended in lysis buffer (50 mM Tris-HCl,

300 mM NaCl, 20 mM imidazole, 10% glycerol and 0.1% NP40, pH 8.0) supplemented with cOmplete, EDTA-free protease inhibitor cocktail (Roche, Darmstadt, Germany). The wells were disrupted by ultrasonication (Vibra-Cell; Sonics, Newtown, CT, USA) for 10 min at 35% amplitude on ice, and the lysates were centrifuged at $21,130 \times g$ for 30 min at 4°C. Soluble proteins were purified using a Ni-NTA Agarose column (Qiagen). The proteins were eluted with an elution buffer (50 mM Tris-HCl, 300 mM NaCl and 250 mM imidazole, pH 8.0), and then exchanged with 50 mM Tris-HCl (pH 8.0) using the Amicon Ultracel-10K centrifugal filter (Merck Millipore, Cork, Ireland). The solubility of the purified proteins was verified with a Western blot assay using anti-His6 antibody (#sc-53073; Santa Cruz Biotechnology, Santa Cruz, CA, USA; Supplementary Fig. S3a), and the purified proteins were used immediately for subsequent assays.

Generation of transgenic *Arabidopsis thaliana* plants overexpressing CmlBPs

The open reading frame sequences of the *CmlBPs* were cloned into the pEarleygate203 vector, which has a Myc epitope, a cauliflower mosaic virus (CaMV) 35S promoter, and an octopine synthase gene terminator (Earley et al. 2006), and the verified constructs were transformed into the *Agrobacterium tumefaciens* GV3101 strain. *Agrobacterium*-mediated transformation was conducted using *Arabidopsis thaliana* Col-0 ecotype and the floral-dip method (Zhang et al. 2006). Transformed plants were selected on 0.5 Murashige and Skoog medium containing hygromycin ($50 \mu\text{g}\cdot\text{ml}^{-1}$). Ten independent homozygous lines (T3) were generated for each gene. Data from two independent lines were generated from these lines. The overexpression of heterologous *CmlBP* genes was confirmed by RT-PCR (Supplementary Fig. S3b). NA transgenic plants were cultivated for 3–4 weeks under a 16:8 h (light:dark) cycle with a light intensity of $120 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ and were transferred to 4°C and grown for two more days for the cold-acclimated (CA) treatment and the IRI and ion leakage assays (Bredow et al. 2017).

Protein extraction from *Arabidopsis* transgenic plants

Total crude protein extracts were prepared using 2-week-old transgenic seedlings grown under the NA and CA conditions supplemented with the proteasome inhibitor MG132 (0.1 mM). Harvested seedlings (~30 mg FW) were ground in liquid nitrogen with a pestle, and 200 μl of native protein extraction buffer (10 mM Tris-HCl pH 7.5 and 25 mM NaCl) supplemented with cOmplete EDTA-free protease inhibitor cocktail (Roche, Basel, Swiss) was added. All extraction steps were performed at 4°C or in ice. The samples were vortexed and held on ice for 10 min. After centrifugation at $13,523 \times g$ for 10 min, the supernatants were transferred to new tubes and the total protein concentration was calculated using the DC protein assay (Bio-Rad, Hercules, CA, USA).

For apoplastic protein extraction, 4-week-old plants of wild-type (Col-0) and 35S::*CmlBP1#6* plants were grown under NA condition, and apoplastic proteins were prepared following the protocols previously described with a simple modification (Haslam et al. 2003). Thirty fresh leaves of wild-type and 35S::*CmlBP1#6* plants were harvested, rinsed with 0.05% of Tween20 and rinsed with deionized water (DW) in two times. Washed leaves were completely immersed in 200 ml infiltration buffer (10 mM Tris-HCl pH 7.5 and 25 mM NaCl) supplemented with cOmplete EDTA-free protease inhibitor cocktail, and vacuum-infiltrated in a desiccator at -70 kPa for three times of 5 min. After removing the remaining buffer, the infiltrated leaves rolled with parafilm were placed in a 20-ml syringe and the syringe was placed in a 50-ml tube for centrifugation. Apoplastic proteins were harvested by centrifugation (15 min, $55 \times g$) and cytoplasmic proteins were extracted from the remaining leaves in a way to extract the total proteins described above. Both apoplastic and cytoplasmic proteins were concentrated and their buffer was exchanged to 50 mM Tris-HCl (pH 8.0) using the Amicon Ultracel-10K centrifugal filter (Merck Millipore, Cork, Ireland). Protein concentration was measured by DC protein assay. Total apoplast proteins of wild-type and *CmlBP1#6* transgenic plants (2.18–2.29 g FW) were yielded in $144 \mu\text{g}$ ($4.8 \mu\text{g}$ a leaf) and $126 \mu\text{g}$ ($4.2 \mu\text{g}$ a leaf), respectively. The cytoplasmic protein of wild-type and *CmlBP1#6*

transgenic plants yielded $620 \mu\text{g}$ a leaf and $930 \mu\text{g}$ a leaf, respectively, and thus the apoplast proteins were determined a 0.77% and 0.45% of a total leaf protein, respectively. Prepared proteins were used for Western blotting and IRI assay, right away.

Western blotting assay

Total protein samples (50 μg) were loaded in 12% sodium dodecyl sulfate-polyacrylamide (SDS-PAGE) gels and blotted onto polyvinylidene difluoride (PVDF) membranes according to the manufacturer's protocol. Anti-cMyc (#sc-40; Santa Cruz Biotechnology) was used to detect the CmlBP proteins in transgenic plants, and alkaline phosphatase (AP)-conjugated anti-mouse (#sc-2008; Santa Cruz Biotechnology) antibody was used as a secondary antibody; BCIP/NBT substrates (#sc-24981; Santa Cruz Biotechnology) were used for the colorimetric assay. Apoplastic and cytoplasmic protein samples (50 μg) were loaded in 12% SDS-PAGE gels and blotted on PVDF membranes. Anti- α -tubulin (T5168, Sigma, USA) and anti-PR-2, pathogenesis-related protein-2, (AS12 2366, Agrisera, Sweden) were used to detect the cytoplasm and apoplast protein fractions, respectively. AP-conjugated anti-mouse (#sc-2008; Santa Cruz Biotechnology) antibody was used as a secondary antibody; BCIP/NBT substrates (#sc-24981; Santa Cruz Biotechnology) were used for the colorimetric assay.

IRI assay

Prepared proteins were used for IRI assay without frozen storage. Protein samples and bovine serum albumin as a control were adjusted at a concentration of $0.1 \text{ mg}\cdot\text{ml}^{-1}$. The IRI assay was conducted using the LINKAM THMS600 heating-freezing stage (Linkam Scientific, Tadworth, UK) mounted on a BX2000 light microscope (Olympus, Tokyo, Japan) under the following conditions: -20°C for 5 min, -10°C for 5 min and -6°C for 30 min. The ice grain morphology was assessed after 30 min at -6°C . The size of the ice grains in each sample (3.2 mm^2 in area) was measured as the mean value of the size of ice grains corresponding to 25–75%ile using ImageJ (Schneider et al. 2012). The size of the ice grains in each sample was expressed relative to that of the buffer control. All experiments were repeated three times. Differences were identified with Student's *t*-test ($P < 0.05$).

Electrolyte leakage assay

Transgenic and wild-type (Col-0) plants were cultivated for 4 weeks to analyze electrolyte leakage. The electrolyte leakage assay followed a previous method (Guo et al. 2002) with some modifications. Two mature rosette leaves (seventh and eighth) were cut and placed in glass tubes containing 250 μl of DW. Ten leaves of each plant were prepared and the freezing treatment (-6°C) was performed as follows. The glass tubes were placed in a circulating water bath (EYELA, Tokyo, Japan) at -1°C for 1 h and then ice nuclei were added. After the water started to freeze, the temperature was decreased by $1^\circ\text{C}/\text{h}$, the tubes were removed from the circulating water bath when the temperature was -6°C , and the tubes were thawed at 4°C overnight in the dark. All leaves and DW were transferred to a 50-ml glass tube with an additional 10 ml of DW and agitated overnight at room temperature. Initial conductivity (Ci) and final conductivity (Cf) were measured before and after autoclaving, respectively. The Ci/Cf ratio is presented as a percentage. Relative ion leakage is expressed as a percentage for ion leakage of the wild type under each condition. This experiment was performed four times, and the differences were detected by Student's *t*-test ($P < 0.05$).

Supplementary Data

Supplementary data are available at PCP online.

Funding

This study was supported by the Korea Polar Research Institute (KOPRI) [PE19080, PE19090, PE19180].

Disclosures

The authors have no conflicts of interest to declare.

References

- Atschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* 215: 403–410.
- Bar Dolev, M., Braslavsky, I. and Davies, P.L. (2016) Ice-binding proteins and their function. *Annu. Rev. Biochem.* 85: 515–542.
- Bredow, M., Vanderbeld, B. and Walker, V.K. (2017) Ice-binding proteins confer freezing tolerance in transgenic *Arabidopsis thaliana*. *Plant Biotechnol. J.* 15: 68–81.
- Bredow, M. and Walker, V.K. (2017) Ice-binding proteins in plants. *Front. Plant Sci.* 8: 2153.
- Cho, S.M., Lee, H., Jo, H., Lee, H., Kang, Y., Park, H., et al. (2018) Comparative transcriptome analysis of field- and chamber-grown samples of *Colobanthus quitensis* (Kunth) Bartl, an Antarctic flowering plant. *Sci. Rep.* 8: 11049.
- Ciolkowski, I., Wanke, D., Birkenbihl, R.P. and Somssich, I.E. (2008) Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues into WRKY-domain function. *Plant Mol. Biol.* 68: 81–92.
- Conesa, A. and Gotz, S. (2008) Blast2GO: a comprehensive suite for functional analysis in plant genomics. *Int. J. Plant Genomics* 2008: 619832.
- Davies, P.L. (2014) Ice-binding proteins: a remarkable diversity of structures for stopping and starting ice growth. *Trends Biochem. Sci.* 39: 548–555.
- De Maayer, P., Anderson, D., Cary, C. and Cowan, D.A. (2014) Some like it cold: understanding the survival strategies of psychrophiles. *EMBO Rep.* 15: 508–517.
- Duman, J.G. (2015) Animal ice-binding (antifreeze) proteins and glycolipids: an overview with emphasis on physiological function. *J. Exp. Biol.* 218: 1846–1855.
- Earley, K.W., Haag, J.R., Pontes, O., Opper, K., Juehne, T., Song, K., et al. (2006) Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J.* 45: 616–629.
- Eddy, B. (1960) The use and meaning of the term ‘psychrophilic’. *J. Appl. Microbiol.* 23: 189–190.
- Fang, Y., Ngeleka, M., Middleton, D.M. and Simko, E. (2005) Characterization and immuno-detection of AIDA-I adhesin isolated from porcine *Escherichia coli*. *Vet. Microbiol.* 109: 65–73.
- Gauthier, S.Y., Kay, C.M., Sykes, B.D., Walker, V.K. and Davies, P.L. (1998) Disulfide bond mapping and structural characterization of spruce budworm antifreeze protein. *Eur. J. Biochem.* 258: 445–453.
- Gonzalez-Aguero, M., Garcia-Rojas, M., Di Genova, A., Correa, J., Maass, A., Orellana, A., et al. (2013) Identification of two putative reference genes from grapevine suitable for gene expression analysis in berry and related tissues derived from RNA-Seq data. *BMC Genomics* 14: 878.
- Graether, S.P., Kuiper, M.J., Gagne, S.M., Walker, V.K., Jia, Z., Sykes, B.D., et al. (2000) Beta-helix structure and ice-binding properties of a hyperactive antifreeze protein from an insect. *Nature* 406: 325–328.
- Graether, S.P. and Sykes, B.D. (2004) Cold survival in freeze-intolerant insects: the structure and function of beta-helical antifreeze proteins. *Eur. J. Biochem.* 271: 3285–3296.
- Gregorio, J., Hernández-Bernal, A.F., Cordoba, E. and León, P. (2014) Characterization of evolutionarily conserved motifs involved in activity and regulation of the ABA-INSENSITIVE (ABI) 4 transcription factor. *Mol. Plant* 7: 422–436.
- Guo, S., Stevens, C.A., Vance, T.D.R., Olijve, L.L.C., Graham, L.A., Campbell, R.L., et al. (2017) Structure of a 1.5-MDa adhesin that binds its Antarctic bacterium to diatoms and ice. *Sci. Adv.* 3: e1701440.
- Guo, Y., Xiong, L., Ishitani, M. and Zhu, J.-K. (2002) An *Arabidopsis* mutation in translation elongation factor 2 causes superinduction of CBF/DREB1 transcription factor genes but blocks the induction of their downstream targets under low temperatures. *Proc. Natl. Acad. Sci. USA* 99: 7786–7791.
- Hao, D., Yamasaki, K., Sarai, A. and Ohme-Takagi, M. (2002) Determinants in the sequence specific binding of two plant transcription factors, CBF1 and NtERF2, to the DRE and GCC motifs. *Biochemistry* 41: 4202–4208.
- Haslam, B.R.P., Downie, A.L., Raveton, M., Gallardo, K., Job, D., Pallett, K.E., et al. (2003) The assessment of enriched apoplastic extracts using proteomic approaches. *Ann. Appl. Biol.* 143: 81–91.
- Huang, T., Nicodemus, J., Zarka, D.G., Thomashow, M.F., Wisniewski, M. and Duman, J.G. (2002) Expression of an insect (*Dendroides canadensis*) antifreeze protein in *Arabidopsis thaliana* results in a decrease in plant freezing temperature. *Plant Mol. Biol.* 50: 333–344.
- Jung, W., Campbell, R.L., Gwak, Y., Kim, J.I., Davies, P.L. and Jin, E. (2016) New cysteine-rich ice-binding protein secreted from Antarctic microalga, *Chloromonas* sp. *PLoS One* 11: e0154056.
- Katoh, K., Rozewicki, J. and Yamada, K.D. (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, 20: 1160–1166.
- Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. and Sternberg, M.J.E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10: 845.
- Knight, C.A. and DeVries, A.L. (2009) Ice growth in supercooled solutions of a biological “antifreeze”, AFGP 1-5: an explanation in terms of adsorption rate for the concentration dependence of the freezing point. *Phys. Chem. Chem. Phys.* 11: 5749–5761.
- Lauersen, K.J., Brown, A., Middleton, A., Davies, P.L. and Walker, V.K. (2011) Expression and characterization of an antifreeze protein from the perennial rye grass, *Lolium perenne*. *Cryobiology* 62: 194–201.
- Lefort, V., Longueville, J.-E. and Gascuel, O. (2017) SMS: smart model selection in PhyML. *Mol. Biol. Evol.* 34: 2422–2424.
- Letunic, I. and Bork, P. (2018) 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res.* 46: D493–D496.
- Matys, V., Fricke, E., Geffers, R., Gößling, E., Haubrock, M., Hehl, R., et al. (2003) TRANSFAC[®]: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res.* 31: 374–378.
- Middleton, A.J., Marshall, C.B., Faucher, F., Bar-Dolev, M., Braslavsky, I., Campbell, R.L., et al. (2012) Antifreeze protein from freeze-tolerant grass has a beta-roll fold with an irregularly structured ice-binding site. *J. Mol. Biol.* 416: 713–724.
- Molino, J.V.D., Carvalho, J.C.M. and Mayfield, S.P. (2018) Comparison of secretory signal peptides for heterologous protein expression in microalgae: Expanding the secretion portfolio for *Chlamydomonas reinhardtii*. *Plos One* 13: e0192433.
- Morgan-Kiss, R.M., Prisco, J.C., Pockock, T., Gudynaite-Savitch, L. and Huner, N.P. (2006) Adaptation and acclimation of photosynthetic microorganisms to permanently cold environments. *Microbiol. Mol. Biol. Rev.* 70: 222–252.
- Morita, R.Y. (1975) Psychrophilic bacteria. *Bacteriol. Rev.* 39: 144.
- Ng, F., Kittelmann, S., Patchett, M.L., Attwood, G.T., Janssen, P.H., Rakonjac, J., et al. (2016) An adhesin from hydrogen-utilizing rumen methanogen *Methanobrevibacter ruminantium* M1 binds a broad range of hydrogen-producing microorganisms. *Environ. Microbiol.* 18: 3010–3021.
- Petersen, T.N., Brunak, S., von Heijne, G. and Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8: 785.
- Poehlein, A., Daniel, R. and Seedorf, H. (2017) The draft genome of the non-host-associated *Methanobrevibacter arboriphilus* strain DH1 encodes a large repertoire of adhesin-like proteins. *Archaea* 2017: 1.
- Rambaut, A. (2014) FigTree v1.4.2. Available at: <http://tree.bio.ed.ac.uk/software/figtree/> (last accessed 26 August 2019).

- Raymond, J.A. (2014) The ice-binding proteins of a snow alga, *Chloromonas brevispina*: probable acquisition by horizontal gene transfer. *Extremophiles* 18: 987–994.
- Raymond, J.A., Janech, M.G. and Fritsen, C.H. (2009) Novel ice-binding proteins from a psychrophilic antarctic alga (Chlamydomonadaceae, Chlorophyceae)(1). *J. Phycol.* 45: 130–136.
- Raymond, J.A. and Kim, H.J. (2012) Possible role of horizontal gene transfer in the colonization of sea ice by algae. *PLoS One* 7: e35968.
- Raymond, J.A. and Morgan-Kiss, R. (2013) Separate origins of ice-binding proteins in Antarctic Chlamydomonas species. *PLoS One* 8: e59186.
- Raymond, J.A. and Morgan-Kiss, R. (2017) Multiple ice-binding proteins of probable prokaryotic origin in an Antarctic lake alga, Chlamydomonas sp. ICE-MDV (Chlorophyceae). *J. Phycol.* 53: 848–854.
- Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH image to ImageJ: 25 years of image analysis. *Nat. Methods* 9: 671.
- Schrodinger, L.L.C. (2015) The PyMOL Molecular Graphics System, Version 1.8. Available at: <https://pymol.org/2/> (accessed 26 August 2019).
- Segawa, T., Matsuzaki, R., Takeuchi, N., Akiyoshi, A., Navarro, F., Sugiyama, S. et al. (2019) Bipolar dispersal of red-snow algae. *Nat. Communication* 9: 3094.
- Sherlock, O., Schembri, M.A., Reisner, A. and Klemm, P. (2004) Novel roles for the AIDA adhesin from diarrheagenic *Escherichia coli*: cell aggregation and biofilm formation. *J. Bacteriol.* 186: 8058–8065.
- Sorhannus, U. (2011) Evolution of antifreeze protein genes in the diatom genus *Fragilariopsis*: evidence for horizontal gene transfer, gene duplication and episodic diversifying selection. *Evol. Bioinform. Online* 7: 279–289.
- Uhlig, C., Kilpert, F., Frickenhaus, S., Kegel, J.U., Krell, A., Mock, T., et al. (2015) In situ expression of eukaryotic ice-binding proteins in microbial communities of Arctic and Antarctic sea ice. *ISME J.* 9: 2537–2540.
- Verbruggen, H., Ashworth, M., LoDuca, S.T., Vlaeminck, C., Cocquyt, E., Sauvage, T., et al. (2009) A multi-locus time-calibrated phylogeny of the siphonous green algae. *Mol. Phylogenet. Evol.* 50: 642–653.
- Xie, F., Xiao, P., Chen, D., Xu, L. and Zhang, B. (2012) miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs. *Plant Mol. Biol.* 80: 75–84.
- Zhang, X., Henriques, R., Lin, S.-S., Niu, Q.-W. and Chua, N.-H. (2006) Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nat. Protoc.* 1: 641.