Biochemistry and Chemical Biology
Structural Biology and Molecular Biophysics

Ice nucleation proteins self-assemble into large fibres to trigger freezing at near 0 °C

Thomas Hansen, Jocelyn C. Lee, Naama Reicher, Gil Ovadia, Shuaiqi Guo, Wangbiao Guo, Jun Liu, Ido Braslavsky, Yinon Rudich, Peter L. Davies

Department of Biomedical and Molecular Sciences, Queen's University, Kingston, ON Canada K7L 3N6 • Department of Earth and Planetary Sciences, The Weizmann Institute of Science, Rehovot 7610001, Israel • The Robert H. Smith Faculty of Agriculture, Food and Environment, Institute of Biochemistry, Food Science, and Nutrition, The Hebrew University of Jerusalem, Rehovot 7610001, Israel • Department of Microbial Pathogenesis, Yale University School of Medicine, New Haven, CT 06536

Reviewed Preprint

Revised by authors after peer review.

About eLife's process

Reviewed preprint version 2
November 16, 2023 (this version)

Reviewed preprint version 1
October 5, 2023

Posted to preprint server
August 22, 2023

Sent for peer review
August 22, 2023

Abstract

In nature, frost can form at a few degrees below 0 °C. However, this process requires the assembly of tens of thousands of ice-like water molecules that align together to initiate freezing at these relatively high temperatures. Water ordering on this scale is mediated by the ice nucleation proteins of common environmental bacteria like *Pseudomonas syringae* and *P. borealis*. However, individually, these 100-kDa proteins are too small to organize enough water molecules for frost formation, and it is not known how giant, megadalton-sized multimers, which are crucial for ice nucleation at high sub-zero temperatures, form. The ability of multimers to self-assemble was suggested when the transfer of an ice nucleation protein gene into *Escherichia coli* led to efficient ice nucleation. Here we demonstrate that a positively-charged sub-domain at the C-terminal end of the central beta-solenoid of the ice nucleation protein is crucial for multimerization. Truncation, relocation, or change of the charge of this subdomain caused a catastrophic loss of ice nucleation ability. Cryo-electron tomography of the recombinant *E. coli* showed that the ice nucleation protein multimers form fibres that are ~ 5 nm across and up to 200 nm long. A model of these fibres as an overlapping series of antiparallel dimers can account for all their known properties and suggests a route to making cell-free ice nucleators for biotechnological applications.

eLife assessment

This **valuable** study provides molecular-level insights into the functional mechanism of bacterial ice-nucleating proteins, detailing electrostatic interactions in the domain architecture of multimeric assemblies. The evidence supporting the claims of the authors is **solid**, with results from protein engineering experiments, functional assays, and cryo-electron tomography, while the proposed structural model of protein self-assembly remains hypothetical. The work is of broad interest to researchers in the fields of protein structural biology, biochemistry, and biophysics, with implications in microbial ecology and atmospheric glaciation.

Introduction

Ice crystals grow from ice embryos, which are crystalline aggregates of water molecules that spontaneously form (homogeneous nucleation) in pure H_2O at approximately –38 °C (1). Ice can arise in nature at much warmer temperatures because various surfaces act as stabilizers of ice embryos (heterogeneous nucleation). Only once an ice embryo reaches a critical number of organized water molecules will it become stable enough to spontaneously grow at elevated temperatures, a process called ice nucleation (2). The most active heterogeneous ice nucleators are bacterial ice nucleation proteins (INPs), which can stabilize an ice embryo at temperatures as warm as −2 °C (3). INP-producing bacteria are widespread in the environment where they are responsible for initiating frost (4) and atmospheric precipitation (5). As such, these bacteria play a significant role in the Earth's hydrological cycle and in agricultural productivity.

As described in the literature, INPs are large proteins (up to ~150 kDa) that are thought to form multimers on the surface of the bacteria that express them (6, 7). AlphaFold predictions have provided some insight into the INP monomer structure (**Fig. 1A**) (8). For the INP from *Pseudomonas borealis* (*Pb*INP) AlphaFold predicted a folded domain of ~100 residues at the N terminus followed by a flexible linker of ~50 residues, a repetitive domain composed of 65 16-residue tandem repeats, and a small 41-residue C-terminal capping structure (supported by model confidence metrics, **Fig. S1**). The predicted fold of the repetitive domain agrees with some previous homology-based models in which each 16-residue repeat forms a single coil of a β-solenoid structure (9, 10).

In INP sequences, most coils of the β-solenoid contain putative water-organizing motifs like Thr-Xaa-Thr (TxT) that occupy the same position in each coil to form long parallel arrays, and where Xaa is an inward pointing amino acid residue (**Fig. 1B**). Shorter versions of similar arrays have convergently evolved in insects to form the ice-binding sites of several hyperactive antifreeze proteins (AFPs) (11–13). These arrays are thought to organize sufficient ice-like water molecules on their surface to facilitate AFP adsorption to the ice crystal surface (14). In the much longer INP arrays, which further form multimers, the organizing effect on nearby water molecules is thought to increase to the point where ice embryos can be sufficiently stabilized to cause spontaneous growth of ice at high sub-zero temperatures. Consistent with this idea, we recently demonstrated that interrupting these water-organizing motifs decreased the ice nucleation temperature by the same amount as extensive deletions of the water-organizing coils (8).

Previously we showed that the 12 C-terminal coils lack the water organizing motifs and that deleting these coils resulted in a near total loss of ice nucleation activity (8). Interestingly, the necessity for these C-terminal coils was demonstrated by Green and Warren in 1986 in the first publication of an INP sequence (15) but was not further investigated. While the water-organizing coils (WO-coils) are characterized by their conserved TxT, SxT, and Y motifs, the defining feature of the C-terminal-most coils, other than the lack of these motifs, is that position 12 of the 16-residue coil is typically occupied by arginine. Thus, we refer to these non-WO-coils as R-coils. Since both coil types maintain the same predicted fold but serve different functions, we consider them subdomains of the same β-solenoid (16). In the WO-coils, position 12 is usually occupied by residues of the opposite charge, Asp and Glu. This charge inversion is noteworthy as it has been shown that electrostatic interactions contribute to the formation of INP multimers (17, 18). It has also been shown that INP activity is affected by pH, which is consistent with a role for electrostatic interactions (19). We and others have suggested that INPs may multimerize through salt-bridging of the sidechains in these positions of the coil (8, 20).

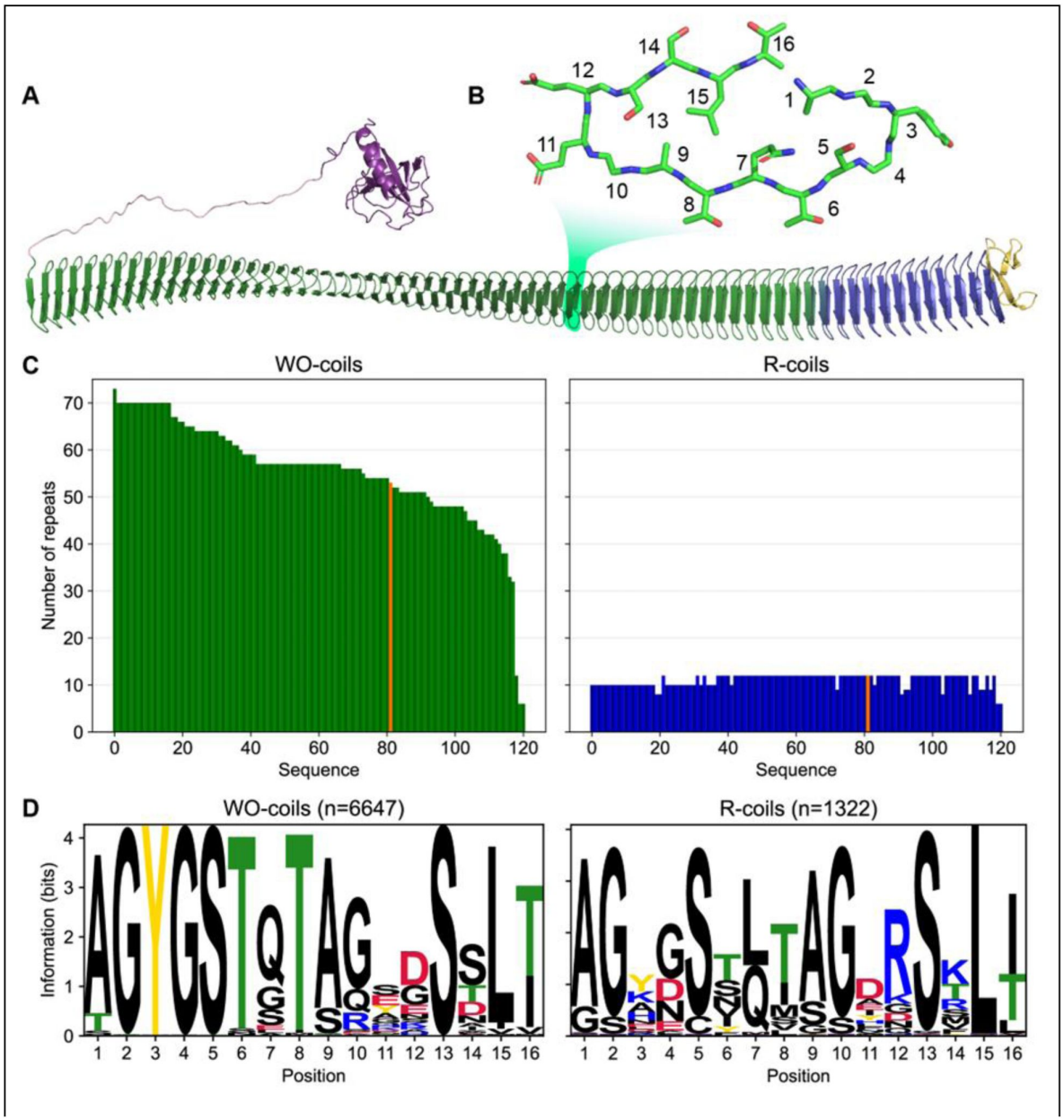

Figure 1.

Classification of unique INP tandem arrays into either WO-coil or R-coil subdomains.

A) AlphaFold 2 model of *Pb*INP coloured by domain. Purple: N-terminal domain, pink: flexible linker, green: water-organizing (WO) coils, blue: arginine-rich (R) coils, yellow: C-terminal cap. The inset shows a cross section through the solenoid coil. **B)** 16-residue tandem repeat forming one coil from the β-solenoid with positions numbered from N-to C-terminus. **C)** The number of repeats in the WO-coils and R-coils for each unique sequence. *Pb*INP is included and indicated in orange. n = 121. **D)** Sequence logos constructed from each 16-residue repeat present in the dataset.

Radiation inactivation analysis suggests a multimer size of 19 MDa (>100 monomers) (6). Computational estimates predict increased activity upon the assembly of up to a 5-MDa (34 INPs) multimer, which is on the same order of magnitude as that determined experimentally (21). The tendency to form such large structures is one of many factors that makes these proteins difficult to work with (3) and may be part of why, despite many attempts, very little is known about them at a molecular level (22). The size of these structures does, however, make them amenable to size-based separation from other proteins (23, 24). INP multimers are also large enough to be visible using negative stain transmission electron microscopy (TEM) on enriched samples, revealing a fibril-like morphology (24, 25).

In nature, these multimers form on the surface of bacteria, anchored to the outer membrane by the N-terminal domain (26). When expressed recombinantly in *E. coli*, INPs have full activity suggesting that multimers are still able to form and that they are the product of self-assembly. Remarkably, INPs with N-terminal truncations are only slightly less active, suggesting that assembly on the cell surface is not mandatory, and it can occur in the cytoplasm whether anchored to the inner surface of the plasma membrane or free in solution (27, 28).

Here, we have studied the role of the R-coils in INPs through a series of mutations and rearrangements. Additionally, using cryo-focused-ion-beam (cryo-FIB) milling and cryo-electron tomography (cryo-ET), we have observed the fibrillar morphology of INP multimers *in situ* within cells recombinantly expressing INPs. The R-coils' length, location, and sequence are critical for INP multimerization and hence INP activity. Although we report results using *Pb*INP, the bioinformatic analysis presented here indicates that these findings are universally applicable to the INP family, including the more commonly studied InaZ from *Pseudomonas syringae*.

Results

Bioinformatic analysis reveals conservation in the number of R-coils across all INPs

A bioinformatic analysis of bacterial INPs was undertaken to identify their variations in size and sequence to understand what is common to all that could guide experiments to probe higher order structure and help develop a collective model of the INP multimer. In *Pb*INP, there are 53 WO-coils and 12 R-coils (**Fig. 1A**), each composed of 16 residues (**Fig. 1B**). To determine whether this ratio of coil types is consistent across known INPs, we analyzed INP and INP-like solenoid sequences in the NCBI's non-redundant protein (nr) database. The long tandem arrays of coils in INPs make them prone to mis-assembly when using short-read DNA sequencing (29) so we opted to limit our dataset to sequences obtained by long-read technologies (Oxford Nanopore and Pacific Biosciences SMRT sequencing). From this bioinformatics study, it is apparent that the number of WO-coils varies considerably from over 70 coils to around 30, with a median length of 58 coils (**Fig. 1C**). In contrast, the length of the R-coil region is much less variable across sequences, with 107 of 120 sequences containing either 10 or 12 R-coils (**Fig. 1C**). The stark difference in length variation between the numbers of WO-coils and the R-coils supports the hypothesis that these two regions have different functions.

The differences observed between the *Pb*INP WO-coil consensus sequence and the R-coil consensus sequence, which include the loss of putative WO motifs in the former and the appearance of basic residues at position 12 in the latter, are consistent across the entire dataset (**Fig. 1D**). This is apparent from the sequence logos comparing them. Also worth noting is the similarity of the sequence logos for WO-coils and R-coils in *Pb*INP (8) with those based on 120 sequences from the database.

Incremental replacement of R-coils with WO-coils severely diminishes ice nucleating activity

Given the remarkable conservation of the R-coil count compared to the variability of the WO-coil numbers, we measured the functional impact of shortening the R-coil region. We designed mutants in which the R-coils were incrementally replaced with WO-coils, shortening the R-coil subdomain from 12 to 10, 8, 6, 4, or 1 coil(s), while retaining the same overall length as wild-type *Pb*INP (**Fig. 2A**). To avoid disrupting any potential interaction between the C-terminal cap structure and the R-coils, one R-coil was left in place to produce the 1 R-coil mutant. As previously described these constructs were tagged with GFP as an internal control for INP production, and its addition had no measured effect on ice nucleation activity (8).

Ice nucleation assays were performed on intact *E. coli* expressing *Pb*INP to assess the activity of the incremental replacement mutants. In theory, replacement of R-coils by WO-coils could result in a gain of function as the water-organizing surface increased in area. However, as the number of R-coils was reduced, the nucleation temperature decreased (**Figs. 2B**, **2C**). Replacing two or four R-coils to leave ten or eight in place resulted in a slight loss of activity compared to the wild-type protein (T_{50} = –9.1 °C and –10.2 °C, respectively, where T_{50} is the temperature at which 50% of droplets have frozen, $p < 0.001$). Reducing the R-coil count to six dramatically decreased the activity (T_{50} = –35.7 °C). The construct with only four R-coils in place showed only the slightest amount of activity, and activity was entirely lost in the construct containing only one R-coil. Evidently, small decreases in the R-coil region length produce disproportionately large decreases in activity. Halving the length of the R-coils by replacing just six coils reduced ice nucleation activity by 26.9 °C, whereas reducing the WO-coil length in half decreased the T_{50} by less than 2 °C (8). Since the R-coils mostly lack the motifs required for water-organizing, we attribute the observed changes in nucleation temperature to changes in INP multimer formation.

The location of the R-coil subdomain is crucial

In addition to its length, we investigated whether the location of the R-coil subdomain is important for ice nucleation activity. We produced constructs where 11 of the 12 R-coils were relocated to either the N-terminal end of the solenoid or the approximate midpoint of the solenoid (**Fig. 3A**). As before, the C-terminal R-coil was left in place adjacent to the cap structure. We also produced an R-coil deletion construct, where the same 11 R-coils were deleted entirely from the protein.

The N-terminal relocation construct displayed markedly lower activity with a T_{50} = –22.4 °C compared to wild-type *Pb*INP, and the midpoint relocation construct displayed almost no activity (T_{50} = –36.1 °C), and was indistinguishable in activity from the construct where the R-coils were deleted (**Fig. 3B**).

Targeted mutations reveal that positively charged residues are important for R-coil function

Having established the importance of R-coil position and length for high activity, we next investigated the features of this subdomain that are required for its activity. Looking at the charge distribution along the solenoid from N terminus to C terminus (**Fig. 4A**), we noted a switch at the start of the R-coils from an abundance of acidic residues to their replacement by basic residues. To probe the significance of this observation, we mutated all basic residues (R/K/H) in the R-coils to match those found in the same repeat positions of the WO-coils (D/G/E for positions 11 and 12, and S for position 14) (**Fig. 1D**). In total, 17 basic residues – 10 Arginine (R), 4 Lysine (K), 3 Histidine (H) – were replaced in the R-coils to generate the RKH replacement mutant. The side chains at these positions are predicted by the AlphaFold model to point outward from the solenoid, so these mutations are unlikely to compromise the stability of the solenoid core.

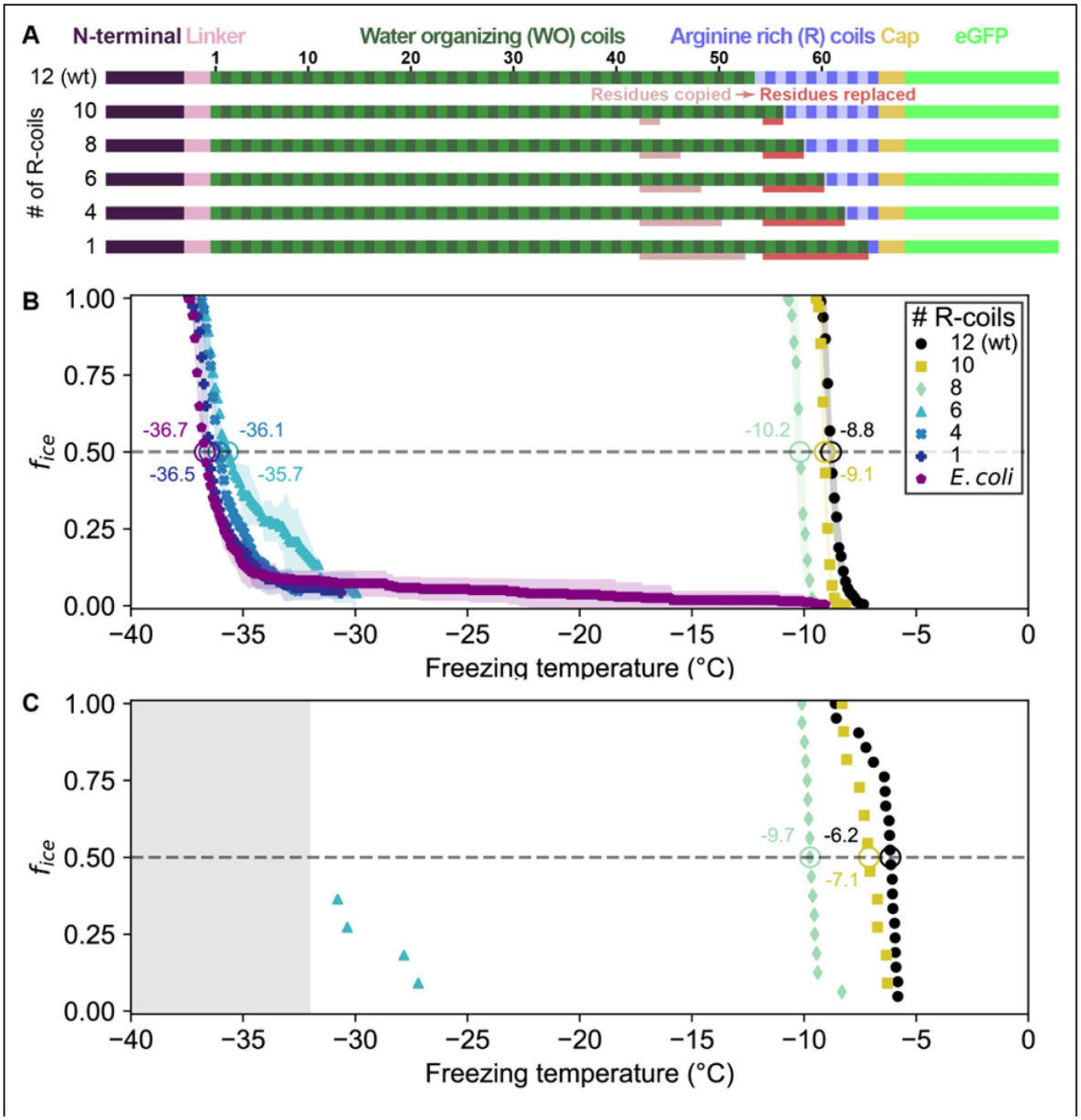

Figure 2.

Ice nucleation activity of mutant INPs in which R-coils are incrementally replaced with WO-coils.

A) Diagram indicating the domain map of *Pb*INP and from it, the design of the INP mutants. Sections of the WO-coils used in the replacement of the R-coils are indicated. wt: wild type. **B**) Ice nucleation temperatures measured by WISDOM for *E. coli* cells expressing *Pb*INP and mutants with different numbers of R-coils. The temperature at which fifty percent of the droplets froze (T_{50}) is indicated with a hollow circle, and its corresponding value written nearby. The shaded region indicates standard deviation. **C)** The same constructs assayed for ice nucleation using a nanoliter osmometer. The grey box indicates temperatures beyond the lower limit of the NLO apparatus for detecting ice nucleation in this experiment.

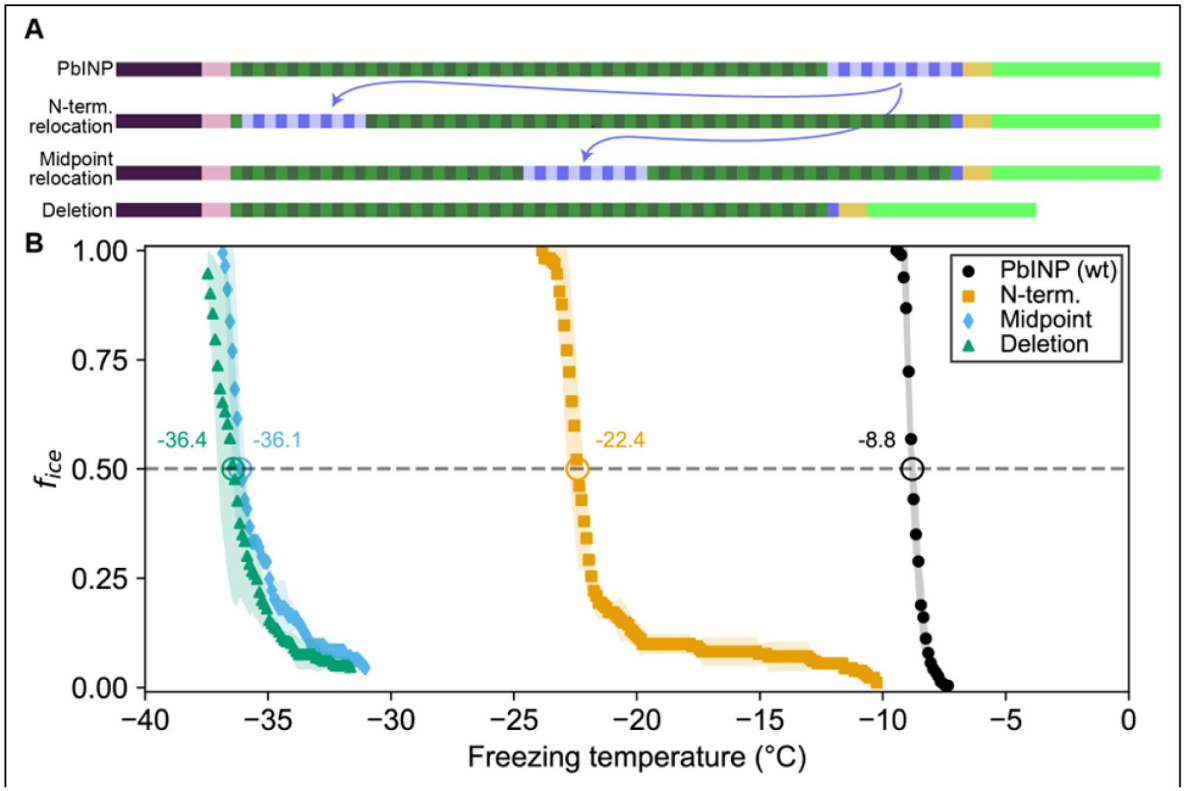

Figure 3.

Ice nucleation activity of mutant INPs with entirely relocated or deleted R-coil subdomain.

A) Diagram indicating the design of the constructs. 11 of the 12 R-coil repeats were either moved within the construct or deleted. **b)** Freezing curves with T_{50} and number of unfrozen droplets indicated where applicable.

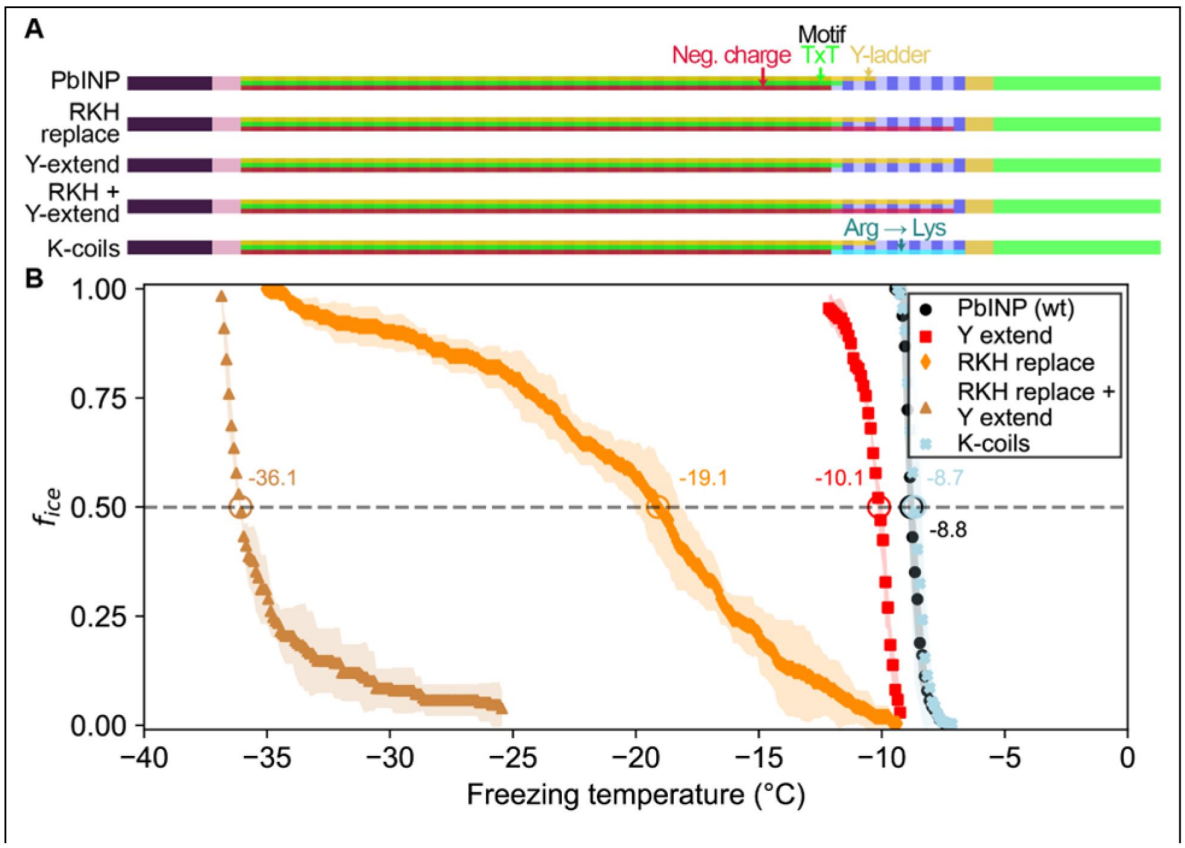

Figure 4.

Site-specific mutagenesis of noteworthy motifs in the R-coil subdomain.

A) Diagram indicating the design of the constructs. Translucent bars indicate continuity of three conserved motifs along the length of the central repetitive domain. Neg. charge: Negative residues present in positions 11, 12, and 14 of the repeat. TxT: Thr-Xaa-Thr motif in positions 6-8 of the repeat. Y-ladder: An entirely conserved Tyr is position 3 of the motif. Three mutants were created in which these motifs were extended into the R-coils. K-coils: All arginines in the R-coils were replaced with lysines. **B)** Freezing curves with T_{50} and number of unfrozen droplets indicated where applicable.

There was a 10.3 °C drop in T_{50} from wild-type activity after RKH replacement (**Fig. 4B**) (T_{50} = –19.1 °C). Although this is a large decrease in activity, it was not as deleterious as the relocation or deletion mutations (**Fig. 3**). The prominent, entirely conserved tyrosine in position 3 of the WO-coils is only present in the first three R-coils and is missing from the following nine coils, making it another candidate for mutation. Upon extending this "tyrosine ladder" through the R-coils (**Fig. 4A**), there was a 1.3-°C loss in activity. However, when combining the RKH replacement with the tyrosine ladder extension, an almost total loss of activity was observed (T_{50} = –36.1 °C on WISDOM) (**Fig. 4B**).

In the final mutated *Pb*INP construct in this series, all arginines in the R-coils were replaced by lysines (K-coils). This mutant nucleated ice formation at essentially the same temperature as the wild type (**Fig. 4b**) (p = 0.89), suggesting that positive charges in these locations are more important than side chain geometry.

Droplet freezing assays show recombinant cell lysate supernatant has ice nucleation activity that is affected by pH

The experiments described above were performed using whole recombinant bacteria rather than extracted INPs. In *E. coli*, the vast majority of the expressed INP is intracellular (27). Indeed, with our GFP-tagged constructs, we observe intense green fluorescence in the cytoplasm. To see how important electrostatic interactions were in the multimerization of *Pb*INP as reflected by its ice nucleation activity, it was necessary to lyse the *E. coli* to change the pH surrounding the INP multimers. After centrifuging the sonicate to remove cell debris and passing the supernatant through a 0.2-µm filter to remove any unbroken cells, the extracts were tested to see how ice nucleation activity is affected by pH between 2.0 to 11.0. The activity of the filtered supernatant was only a few degrees lower than that of whole bacteria (T_{50} = –9.6 °C) (**Fig. 5A**), which agrees with the results of Kassmannhuber *et al.* (28). This indicates that large INP structures are present within the bacterial cytoplasm.

The effect of pH on Snomax activity has been previously reported (18). However, Snomax is comprised of freeze-dried *P. syringae* cells in which the INPs are thought to be membrane bound. Our assays on bacterial lysate tested free, cytoplasmic *Pb*INP complexes, producing a similar trend regarding the effect of pH but with somewhat greater loss of activity on the lower end of the optimal range (**Fig. 5B**). Ice nucleation activity decreased by a few degrees below pH 5.0, and by ~8 °C at pH 2.0. The loss of activity in the alkaline buffers up to pH 11.0 was minimal. Similar to the findings of Chao *et al.* (30), we did not observe a major change in activity (i.e. ΔT_{50} > 10 °C) even at the extremes of pH 2.0 and 11.0, suggesting that the mechanism of ice nucleation is not pH-dependent.

INP activity is remarkably heat resistant

Having access to lysate also provided an opportunity to examine the heat stability of the INP complexes. The filtered lysate was heat-treated to 60, 70, 80, 90, or 99 °C for 10 min in sealed tubes before being chilled and assayed for ice nucleation activity (**Fig. 6A**). The activity of the 60 °C sample (T_{50} = –9.9 °C) was nearly identical to the non-treated wild-type control (T_{50} = –9.6°C), and the 70 °C sample only displayed a minor loss of activity (T_{50} = –10.2 °C). From 80 °C to 99 °C the activity incrementally decreased (T_{50} = –11.3 °C, –12.7 °C, –14.6 °C, respectively), but the activity loss never exceeded 6 °C. Indeed, the heat resistance of the INP complex is remarkable. The C-terminal GFP tag provided an internal control for the effectiveness of heat treatment, as GFP denatures at around 73 °C (31). The green colour of the bacteria was robust at 65 °C and with very few exceptions, gone at 75 °C (**Fig. 6B**). There was no fluorescence at 90°C.

To assess what role the WO-coils play in multimer stability, we also assayed the lysate of a construct from our previous study in which repeats 16-47 (residues 411-923) of the solenoid had been deleted, leaving 32 coils (8). The overall freezing profile remained the same, indicating

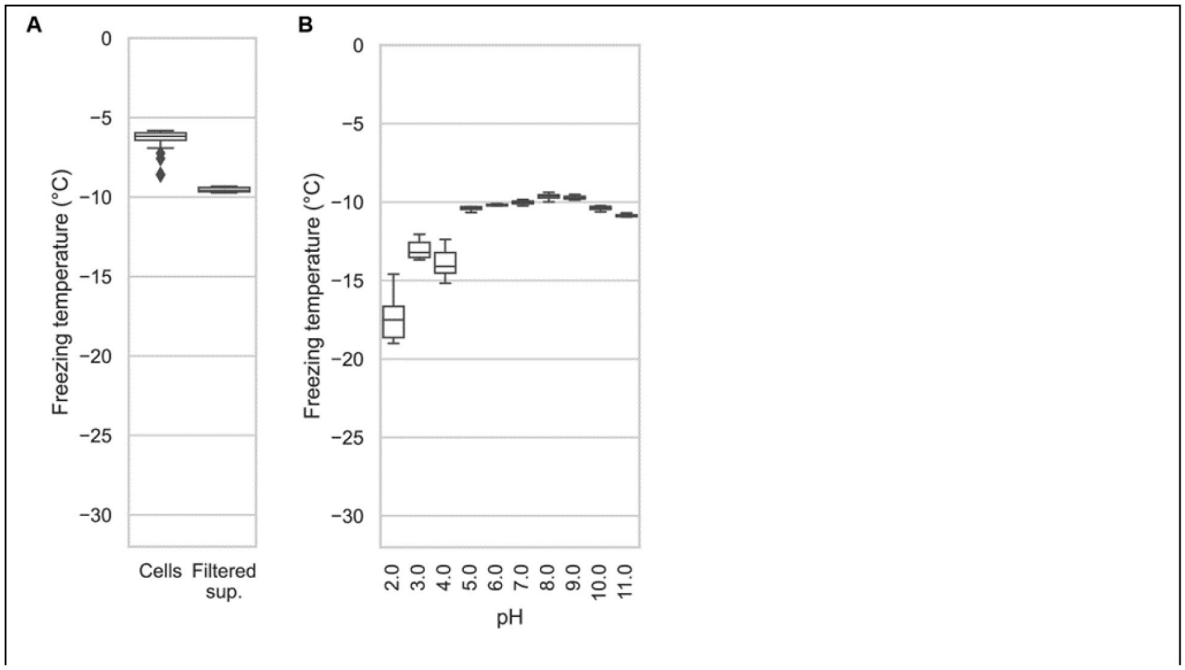

Figure 5.

A) A comparison of the nucleation temperatures of *Pb*INP when assayed using intact *E. coli* cells and when assayed with filtered supernatant. **B)** A box and whisker plot showing the nucleation temperatures of filtered supernatant containing *Pb*INP under different pH conditions. Boxes and bars indicate quartiles, with medians indicated by a centre line. Outliers are indicated by diamonds.

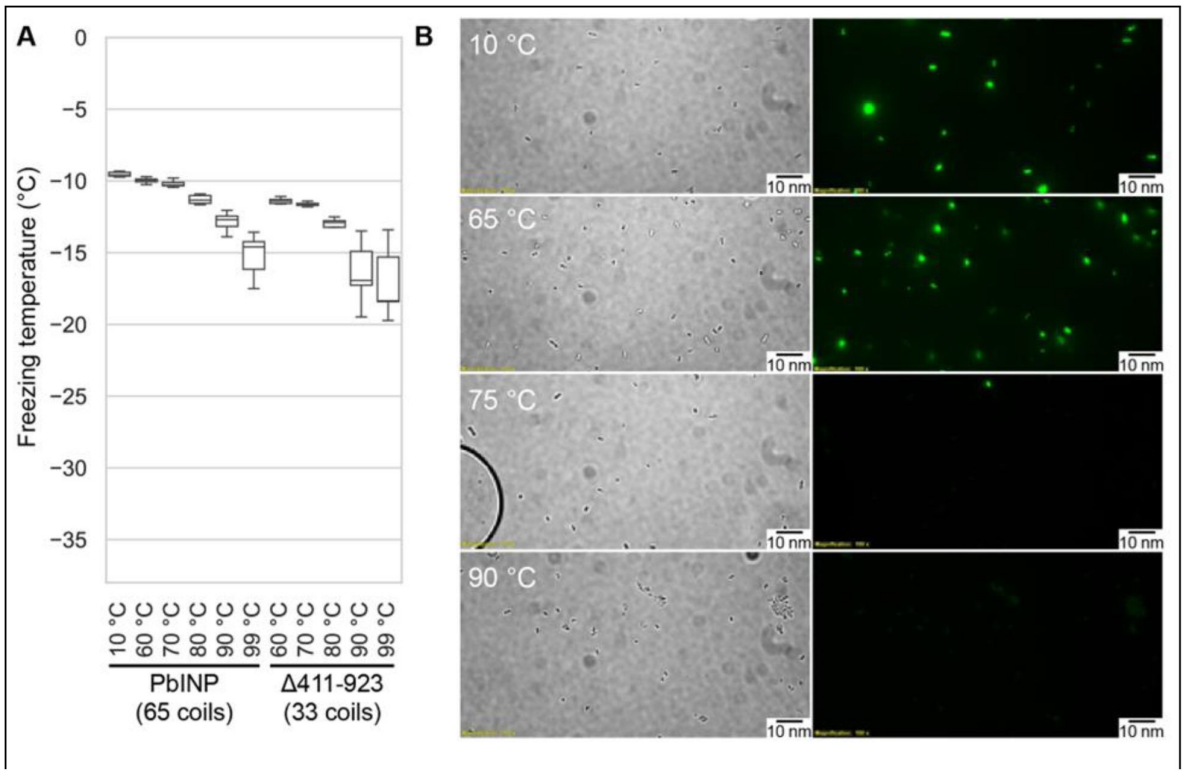

Figure 6.

A) Measured freezing temperatures of heat-treated droplets containing either *Pb*INP or *Pb*INP with the first 32 repeats (counting from the N-terminal end) deleted. Wild-type *Pb*INP freezing without heat treatment (kept at roughly 10 °C) is indicated-on the left. **B)** Fluorescent microscopy images of recombinant *E. coli* cells expressing *Pb*INP tagged with GFP viewed under bright-field (BF) or fluorescent excitatory (GFP) light. Representative images are shown (n = 3). Note: cells that retain their fluorescence after 75 °C treatment are rarely observed.

that this construct is also extremely resistant to heat denaturation, with each temperature sample freezing at slightly lower temperatures than their full-length counterparts (**Fig. 6A**). While the Δ411-923 construct had slightly lower overall activity, its heat resistance was not affected by the truncation, suggesting that the R-coil and C-terminal cap subdomains are mainly responsible for multimer stability.

The β-solenoid of INPs is stabilized by a capping structure at the C terminus, but not at the N terminus

There is a clear C-terminal capping structure in the AlphaFold model (**Fig. 1A**), but a possible N-terminal cap was more nebulous. Most protein solenoids are N-and/or C-terminally capped to help maintain the fold and/or prevent end-to-end associations (32). Looking at the N-terminal sequence, we tested if any part of the extended linker region serves as an N-terminal capping motif. To investigate this, we made a series of incremental N-terminal deletions starting at residues Asp150 (Truncation 1), Gln159 (Truncation 2), and Gln175 (Truncation 3) (**Fig. 7A**).

Truncation 1 lacked most of the N-terminal domain, leaving the last few residues of the unstructured linker. Truncation 2 removed those linker residues so that the putative cap (a single β-strand) was located at the very N-terminal end of the protein. Truncation 3 removed the β-strand along with the rest of the first coil of the solenoid. When tested, there was no difference between the activities of the three truncations and the wild type ($p = 0.82$) (**Fig. 7D**). This result is in line with those from Kassmannhuber *et al.* (28), which showed that deletion of the N-terminal domain does not significantly affect ice nucleation activity.

Previously, we demonstrated that the C-terminal cap is essential for ice nucleation activity (8). Bioinformatic analysis showed a high degree of conservation in the C-terminal cap residues (**Fig. 7B**). Rather than deleting the cap, we made targeted mutations: F1204D, D1208L, and Y1230D, to disrupt the structure predicted by AlphaFold. Residues for mutations were chosen based on the putative key roles of those residues in the AlphaFold model. For an enhanced effect of the mutations hydrophobic residues were replaced with charged ones and vice versa. F1204 sits atop the final R-coil to cover its hydrophobic core. D1208 helps to maintain a tight loop through strategic hydrogen bonds, and Y1230 fills a gap in the surface of the cap (**Fig. 7C**). When comparing these selections to the aligned C-terminal cap sequences, we see that all three residues are highly conserved. The resulting triple mutant displayed greatly reduced activity ($T_{50} = -27.8$°C), which helps validate the AlphaFold-predicted structure of the cap and its importance to the stability of the solenoid it covers.

Cryo-electron tomography reveals INPs multimers form bundled fibres in recombinant cells

The idea that INPs must assemble into larger structures to be effective at ice nucleation has persisted since their discovery (6). In the interim the resolving power of cryo-EM has immensely improved. Here we elected to use cryo-electron tomography to view the INP multimers *in situ* and avoid any perturbation of their superstructure during isolation. *E. coli* cells recombinantly overexpressing INPs were plunge-frozen and milled into ~150-nm thick lamella using cryo-FIB (**Fig. 8A**). Grids containing lamellae were transferred into either a 200-or a 300-kV transmission electron microscope for imaging under cryogenic conditions. Many *E. coli* cells were observed within the low-magnification cryo-TEM overview image of the lamella (**Fig. 8B**). Tilt series were collected near individual *E. coli* cells, and 3-D tomograms were reconstructed to reveal cellular and extracellular features. Strikingly, *E. coli* cells overproducing wild-type INPs appear to be lysed after three days of cold acclimation at 4 °C and contain clusters of fibres in the cytoplasm (**Fig. 8 C, D, E**, tomograms in **Movies S2** and **S3**). Individual fibres are up to a few hundred nanometers in length but only a few nanometers in width. Intriguingly, these fibre

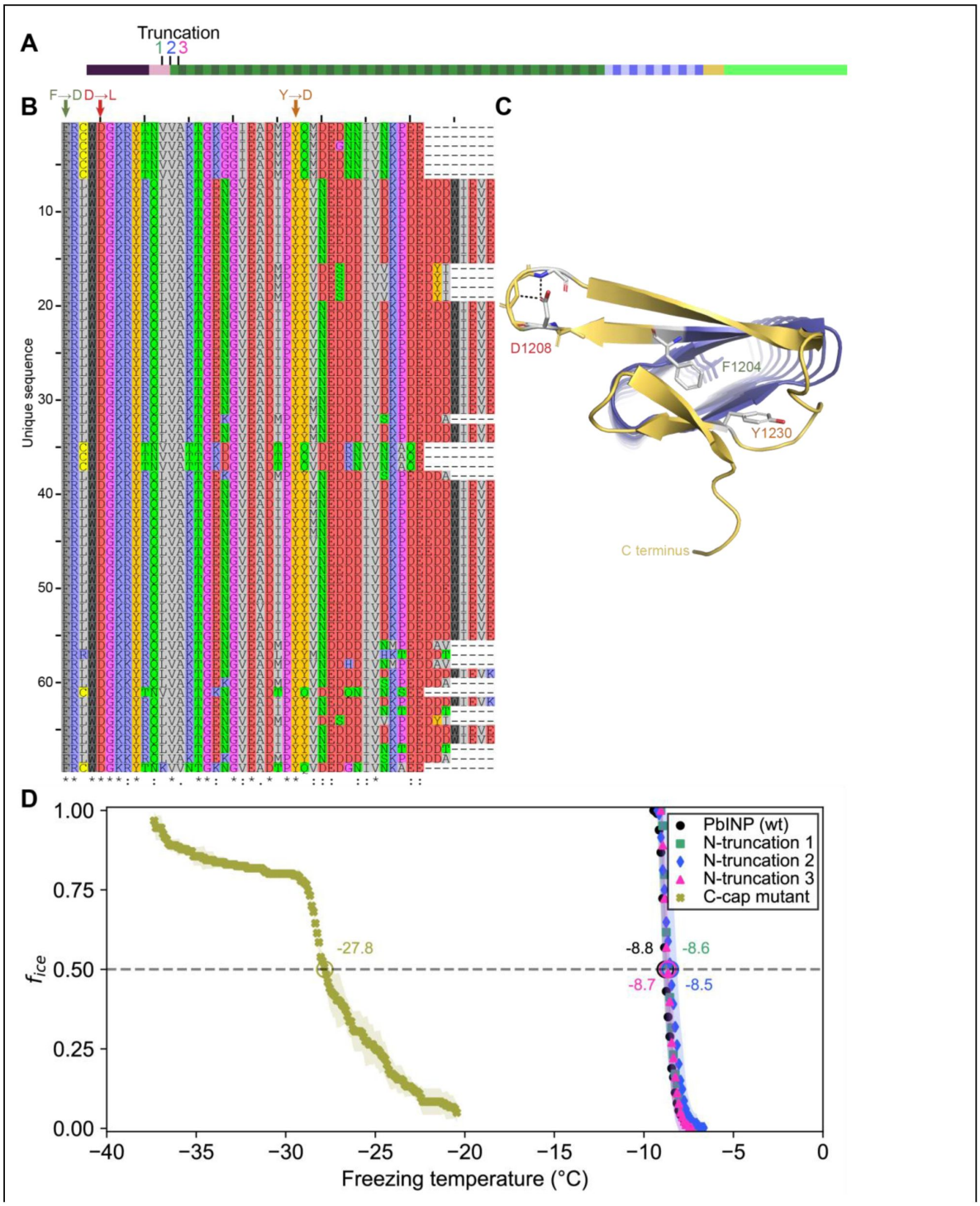

Figure 7.

A) Sites of N-terminal truncations to *Pb*INP, indicating the location of the starting residue in the shortened construct. **B)** Alignment of representative INP C-terminal domains from the genus *Pseudomonas*. Mutated residues and their one-letter codes are indicated above. Symbols at the bottom indicate consensus (* for fully conserved,: for conservation of strongly similar chemical properties, for conservation of weakly similar chemical properties). **C)** Predicted location of mutated residues in the *Pb*INP C-terminal cap with sidechains shown and predicted H-bonds for D1208 shown as dashed lines. **D)** The ice nucleation curves for the N-and C-terminal cap mutants.

clusters were not observed in *E. coli* that overexpress INP mutants lacking R-coils and the cell envelopes stay integral after being cold acclimated over the same period as those of wild-type INP-producing *E. coli*. (**Fig. S4 A, B, C, D**).

Discussion

Previously, we showed that the *Pb*INP solenoid domain is made up of two subdomains: the larger N-terminal region of WO-coils accounting for 80-90% of the total length; and the smaller C-terminal R-coil region accounting for the remaining 10-20% (8). The length of the WO-coil region and the continuity of the water-organizing motifs were shown to directly affect ice nucleation temperature. Although the R-coil region lacks water-organizing motifs, its presence was critical for ice nucleation activity, which led us to propose a key role for this region in INP multimer formation. Here we have characterized the R-coil subdomain in terms of the attributes it needs to support INP multimerization and have shown by cryo-ET the first *in situ* view of what these multimers look like. In addition, we have advanced a working model for the INP multimer structure that is compatible with all of the known INP properties.

In the aforementioned work, we showed that removal of up to half of the *Pb*INP solenoid (reducing the number of WO-coils from 53 to 21) only dropped the ice nucleation activity by ~2 °C. Here we have confirmed this tolerance of WO-coil count variation through bioinformatic analysis of natural INPs. The majority of bacterial INPs have WO-coil counts between 30 and 70. *Pb*INP is average in this respect with 53 WO-coils. It seems counterintuitive that these bacteria have not been uniformly selected for the highest WO-coil count, which might give them an advantage in causing frost damage to plants at the highest possible temperature (33). However, it is clear that INPs are not functioning as monomers but rather as large multimers so any loss of water-organizing surface can potentially be compensated for by simply adding more monomers to the multimer.

The ability to form superstructures is a key property of INPs and centers on the R-coil subdomain. This was shown here in the same bioinformatic analysis where there is remarkably little variation to the R-coil length of 10-12 coils. The importance of a minimal R-coil region length is supported by experiments. Whereas over 30 of the WO-coils can be removed with slight loss of activity, when six of the 12 *Pb*INP R-coils were replaced by WO-coils there was a catastrophic loss of ice nucleation activity, and no activity at all with further shortening of the R-coils. We postulate that at least eight R-coils are required for efficient multimer formation and that the ice nucleation activity of a monomer is inconsequential in the natural environment.

In the absence of detailed structural information, we have probed the properties of the multimers to help develop feasible models for their structure and assembly. The location of the R-coils at the C-terminal end of the solenoid next to the highly conserved cap structure is critical, as they do not function in the middle of the WO-coil region, and only poorly at the N-terminal end. These R-coils have a strong positive charge from the Arg and Lys residues, whereas the WO-coils are negatively charged, and their interaction potentially provides an electrostatic component to the fibre assembly. As expected, changing the charge on the R-coils from positive to negative caused some loss of ice nucleation activity (~9 °C), consistent with charge repulsion between these two solenoid regions weakening the multimer structure. In wild-type INPs, the negative charges of the WO-coils are consistent throughout their length, which offers no clue as to where on the WO-coils the R-coils might interact. One possible advantage of this uniformity is that multimer assembly could still happen if the WO-coil length is appreciably shortened as it can be in nature and by experimentation (8)

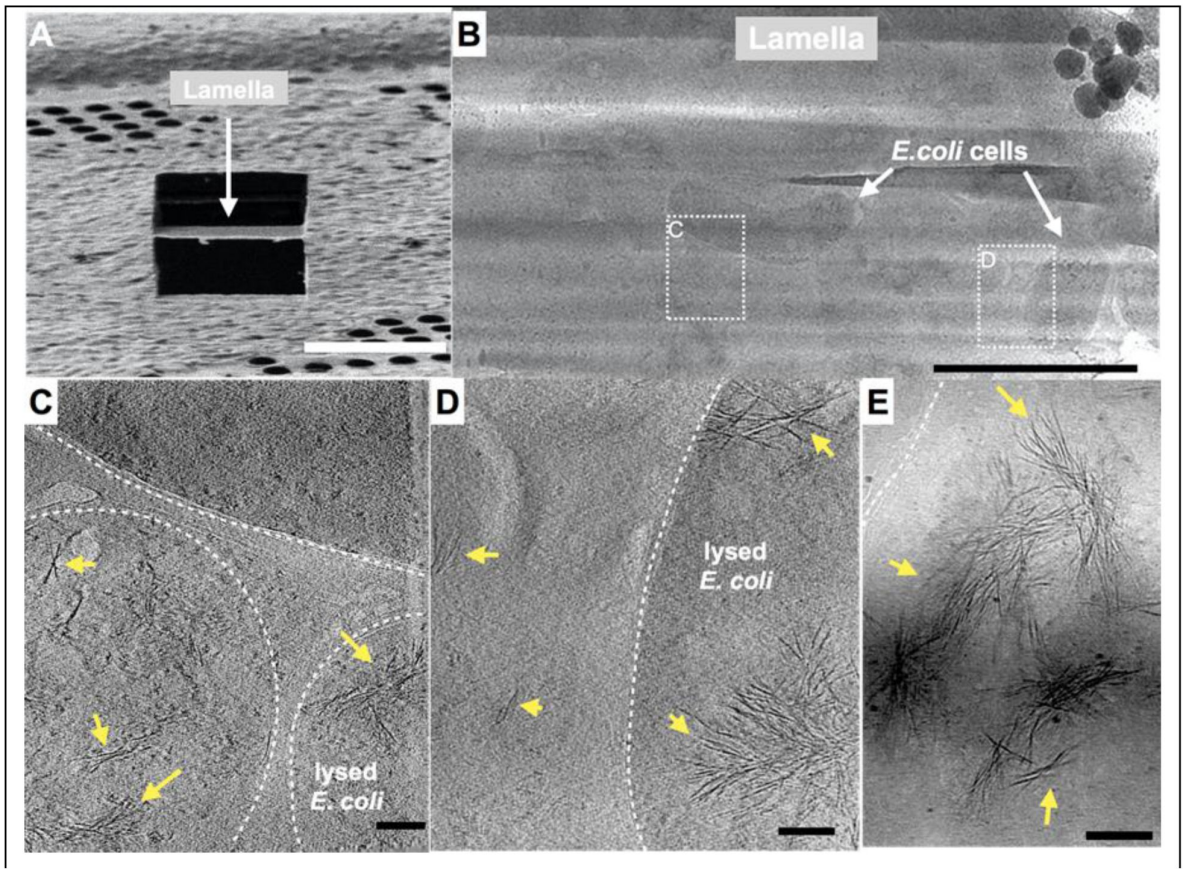

Figure 8.

Fibrous bundles observed by cryo-FIB and cryo-ET in in *E. coli* cells expressing INP.

A) Ion-beam image of a thin lamella containing *E. coli* cells expressing INP obtained from cryo-FIB milling. **B)** Zoomed-in view of a cryo-TEM image of the lamella in A). Boxes with dashed-lines indicate areas where tilt series were collected. **C) and D)** Snapshots from 3-D cryo-tomograms reconstructed from tilt series collected in the boxed regions in B) showing striking fibrous bundles (yellow arrowheads). The *E. coli* cell envelopes are indicated with thick dash-lines. **E)** Further examples of the fibrous bundles produced by INP-expressing *E. coli*. Size markers in A) is 10 μm, in B) is 2 μm and in C), D) and E) are 100 nm, respectively.

The minimal effects of pH change on native INP activity are reminiscent of the insensitivity of antifreeze activity to pH (30, 34). The ice-binding sites of AFPs are typically devoid of charged residues and there should be no effect of pH on the ability of these sites to organize ice-like waters. The same can be said for the water-organizing motifs in INPs. We noted the extraordinary heat stability of INP multimers. Even after heating to 99 °C for 10 min the bacterial extracts only lost 5 °C of ice nucleation activity, whereas the heat-stable internal GFP control was denatured at 75 °C. We cannot rule out the possibility that the INP multimers were also denatured by heat treatment but could reassemble on cooling.

Working model of the INP multimer

The fundamental unit of the INP multimer in this hypothetical model is a dimer (**Fig. 9**). The dimerization interface involves an interaction of the stacked tyrosine ladders from the two INP monomers as previously suggested (10, 24). However, in this model the INPs are aligned antiparallel to each other (**Fig. 9B**). This orientation is more likely than a parallel alignment since the R-coils and C-terminal cap structure appear to clash when modelled parallel to each other (**Movie S5**). The antiparallel dimer would not be a rigid, flat sheet but could hinge at the tyrosine ladder. Another advantage of the antiparallel arrangement is that the two dimer termini are identical allowing end-to-end linking to form a long fibre.

The end-to-end dimer associations involve electrostatic interactions between the basic side of the R-coils and the acidic side of the WO-coils. If these interactions can also form with the proteins at an approximate right-angle, it should be possible for end-linked dimers to form a compact fibre (**Fig. 9B**) with a diameter close to that seen by cryo-ET (**Fig. 8**). The antiparallel arrangement of the dimers gives a sidedness to the multimer where TxT motifs (light green) face outwards and inwards in an alternating pattern with SxT motifs (on the underside) in the opposite phase (**Fig. 9B**). Cross-sectional views of the INP fibre (**Fig. 9C**) show the interactions between the negatively (red) and positively (blue) charged regions where the dimers overlap to form a ring of four solenoids, while maintaining the interaction of the two monomers through the tyrosine ladder pairing.

Working model of the INP multimer is consistent with the properties of INPs and their multimers

We previously showed that the length of the WO-coil region can be shortened by ~60% with only a few °C decrease in ice nucleation temperature (8). The working model can accommodate these huge deletions simply by closing the gap between the dimers. For example, the deletion of 32 WO-coils leaving just 21 along with the 12 R-coils retains all the molecular interactions seen in the longer fibre but with fewer stacked tyrosine interactions. This can help explain the heat stability of the INP multimers and the minimal difference (2-3 °C) in activity loss between full-length *Pb*INP with 65 coils and the truncated version with 33 coils (**Fig. 6**). Similarly, longer WO-coil regions can be accommodated by lengthening the gap. This can explain the wide range of WO-coil lengths seen in nature (**Fig. 2**). They all fit in the same model.

Our model also shows how the interaction between the R-coils and the WO-coils of the adjacent dimers supports fibre formation. Any shortening of the R-coil subdomain jeopardizes the ability to link up the dimers. The catastrophic loss of ice nucleation activity seen below 8 R-coils is because the interacting length of R-coils and WO-coils has too few electrostatic and other interactions to bridge the dimers together. The importance of electrostatic interaction has been illustrated in this study in two ways. First, when the R-coil basic residues were replaced by acidic residues, the ice nucleation activity was severely compromised but was fully restored when the mutated residues were all converted to lysines. Second, in cell-free extracts of lysed INP-producing *E. coli* ice nucleation activity decreased by a few degrees Celsius at low pH values where the charge on acidic residue side chains was reduced or eliminated. When the carboxyl groups of aspartate and glutamate involved in electrostatic pairing lose their negative charges at low pH, they can still

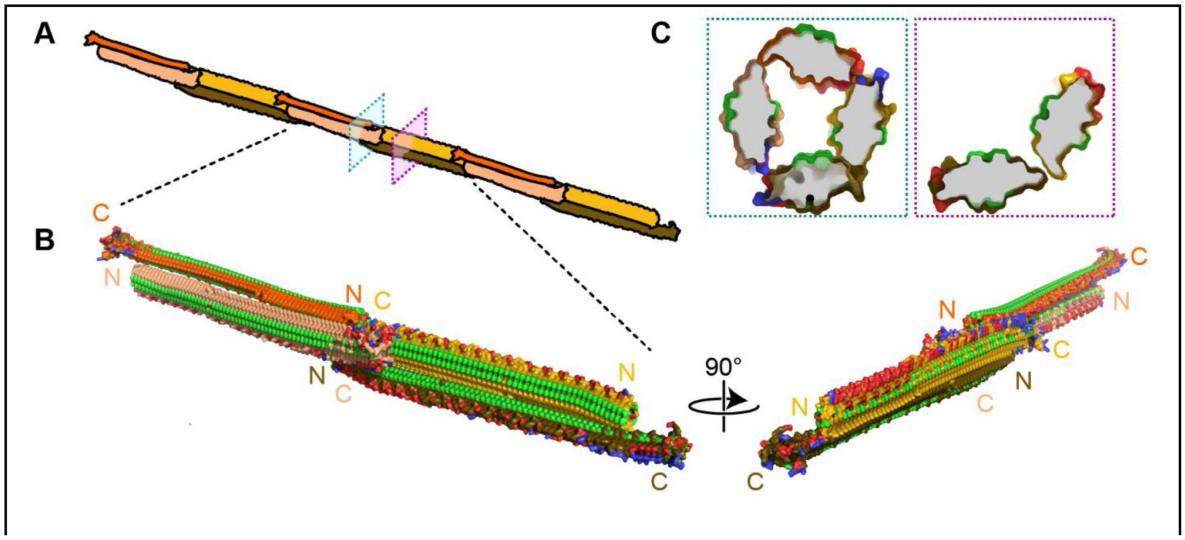

Figure 9.

Filamentous multimer model for bacterial INPs.

A) A possible assembly of INP solenoids to form long fibres composed of antiparallel INP dimers (indicated by orange and yellow pairs). **B)** Dimers are formed along the tyrosine ladder, a previoiusly proposed dimerization interface. They are joined end to end by forming electrostatic interactions between negatively (red) and positively (blue) charged surfaces. All threonines are coloured light green, displaying the arrays of TxT WO-motifs. The termini of the INP solenoids are labeled N and C and coloured to match panel A. This illustration uses a manually flattened AlphaFold model of *Pb*INP. **C**) Cross sections of the model at positions indicated in A. Monomers are rotated approximately 90° to each other and dimerized along their tyrosine ladders (purple). Toward their termini, a pair of dimers can be matched by oppositely charged electrostatic surfaces (teal).

form hydrogen bonds with basic amino acid partners, which can explain why the lowering of pH was not as disruptive as reversing the charge on these residues. Another useful test of the electrostatic component to the multimer model would be to study the effects of increasing salt concentration on ice nucleation activity of the *E. coli* extracts.

The observation that low, variable levels of ice nucleation activity remained in the construct where the R-coil basic residues were replaced by acidic residues, suggests that there are additional binding interactions between the dimers other than electrostatic ones. We suggest the involvement of the highly conserved C-terminal capping structure. When three mutations designed to disrupt the cap fold were introduced, all ice nucleation activity was lost. Also of note is the disruptive effect of extending the tyrosine ladder further into the R-coil sub-domain in the mutant where the acidic residues replaced the basic ones. The subtle details of the R-coil region will require detailed structural analysis for their elucidation.

The relocation of the R-coils to the N-terminal end of the solenoid caused a loss of just over 50% activity and it is possible to accommodate such a change in the model while retaining a charge interaction between the R-coils and WO-coils. However, the separation from the cap structure might account for some of the activity loss. Movement of the R-coils to the centre of the WO-coil region is not compatible with the model and sure enough, this construct was devoid of ice nucleation activity (**Fig. 3B**).

Other features supporting the model are that the dimer's C-and N-terminal ends are exposed and can accommodate tags and extensions without disrupting the fibre. Thus, the addition of a C-terminal GFP tag has no detrimental effect in ice nucleation activity. Nor is there any difference in activity if the N-terminal INP domain and linker region are present or not (**Fig. 7**) (8, 28). Even the incorporation of a bulky protein like mRuby into the WO-coils (8) can be accommodated because the fibre is just a dimer rather than a bundle of solenoids.

Electron microscopy of newly synthesized INPs in a cell-free system shows them as thin molecules of dimensions 4-6 nm in diameter by a few hundred nm in length (25). Negatively stained images of recombinantly produced INP multimers isolated by centrifugation and chromatography show an elongated structure ~5-7 times longer than a monomer but not much wider (24). The fibres seen *in situ* in INP-expressing *E. coli* (**Fig. 8**) are similarly long but slightly thinner, consistent with the absence of negative staining. The model in **Fig. 9** is the thinnest structure we can project for a fibrillar multimer.

Solving the structure of the INP fibres at atomic detail will be the key to understanding the remarkable ability of biological ice nucleators to start the freezing process at high sub-zero temperatures. Structures of this type offer the promise of cell-free ice nucleation for use in biotechnological and food applications where there is a need to avoid the use of bacteria.

Methods

AlphaFold prediction

The AlphaFold model for *Pb*INP was generated by Forbes *et al.* as described (6).

Bioinformatic analysis of INPs

NCBI's BLAST was accessed using the BioPython library v1.81 (35). The consensus sequence for the 16-residue coil 'AGYGSTQTAGEDSSLT' was used as the query against the non-redundant protein database. The PAM30 scoring matrix was used due to the short query. Quality control (QC) was performed using custom Python scripts, making use of BioPython's Entrez module to fetch

information on the protein, BioProject, and assembly method for each BLAST result (**Fig. S6**). Custom Python scripts were used to automatically identify the tandem repeats and classify them as WO-coils or R-coils.

Sequence logos were made using the Logomaker package v0.8 (36). Alignment of C-terminal cap sequences was performed using JalView software v2.11.2.6 (37).

Synthesis of *Pb*INP genes

Experiments for this project used a synthetic *Pb*INP gene previously developed by our group. This codon-optimized gene encodes the *P. borealis* INP gene (GenBank accession: EU573998). Additionally, the DNA sequence for enhanced green fluorescent protein (eGFP) (GenBank accession: AAB02572) was fused to the 3′-end of the *Pb*INP gene using a hexanucleotide encoding two linker residues (Asn-Ser). More details about the *Pb*INP-eGFP sequence are provided in Forbes *et al.* (8).

All mutants for this study were designed by modifying the aforementioned synthetic gene. GenScript performed all (Piscataway, NJ, USA) gene syntheses, which we subsequently cloned into the pET-24a expression vector.

Five *Pb*INP mutants were designed to test the effect of replacing the R-coils. The R-coils were incrementally replaced with the sequences of WO-coils adjacent to the R-coil region as indicated, resulting in constructs containing 10, 8, 6, 4, and 1 R-coil (**Fig. 2A**). Replacements were designed such that they maintained the periodicity of the tandem repeats. The C-terminal R-coil was left untouched to avoid disturbing possible interactions with the putative C-terminal cap structure.

The R-coils were either relocated within the protein or deleted (**Fig. 3A**), while again leaving the N-and C-terminal coils untouched to avoid interactions with adjacent domains.

Targeted mutations were introduced to the R-coil region gene to produce four additional constructs (**Fig. 4A**). For the first construct (RKH replacement), any positively charged residues (Arg, Lys, His) in positions 11, 12, or 14 in the R-coils were replaced with residues commonly found in those locations in the WO-coils (Asp, Glu, and Gly for positions 11 and 12, Ser for position 14). The second construct Y extend in **Fig. 4A** extends the stacked tyrosine ladder present at position 3 of the coils through 7 additional coils toward the C terminus of the solenoid. The third construct (RKH replacement + Y extend in **Fig. 4A** is a combination of both mutants. The fourth construct (K-coils in **Fig. 4A**) converted every Arg residue in the R-coil section to a Lys residue.

Protein expression in *E. coli*

Each *Pb*INP construct was transformed into the ArcticExpress strain of *E. coli*, since its expression of two cold-adapted chaperones, Cpn10 and Cpn60, promotes the correct folding of proteins at low temperatures (38). Transformation and induction with IPTG were performed according to the supplier's instructions (Agilent Technologies, Catalog #230192). Cells expressed at 10 °C for 24 h post-induction. The eGFP tag allowed expression to be confirmed using fluorescence microscopy (8).

Ice nucleation assays by WISDOM

Constructs were assayed on WISDOM (WeIzmann Supercooled Droplets Observation on a Microarray) (39) in a similar way as described in Forbes *et al.* (6).

Ice nucleation assays by nanoliter osmometer

Ice nucleation activity was quantified using a droplet freezing assay protocol (40) that makes use of a LabVIEW-operated nanoliter osmometer (Micro-Ice, Israel) (41). Briefly: Following induction and cold incubation, nanoliter-sized droplets of liquid cultures were pipetted into oil-filled wells resting on a cold stage. The temperature of the cold stage was lowered at a rate of 1 °C/min while a video recording was taken of the sample grid. Freezing was characterized by a distinct change in droplet appearance. After assay completion, the videos were analyzed to record the temperatures of all freezing events. The fraction of frozen droplets (f_{ice}) as a function of temperature was plotted, generating ice nucleation curves for each sample. This apparatus could not reach temperatures as low as those achieved on WISDOM, but results are in agreement between the two approaches (**Fig. 2B**).

Heat treatment and pH

To obtain cell lysates, *E. coli* cultures were centrifuged at 3,200 × g for 30 min post-induction. Cell pellets were then resuspended in a lysis buffer of 50 mM Tris-HCl, 150 mM NaCl, containing Pierce Protease Inhibitor (Thermo Scientific, Canada) before sonication at 70% amplitude for 30-s rounds. Lysate was centrifuged at 31,000 × g and the resulting supernatant was passed through a 0.2 µm filter.

For heat treatment, filtered lysate in sealed Eppendorf tubes was heated at 60 °C, 70 °C, 80 °C, 90 °C, or 99 °C for 10 min in a thermocycler and then quenched on ice prior to being assayed for activity.

For the pH experiments, aliquots of filtered lysate were diluted 50-fold in pH-adjusted buffer of 100 mM sodium citrate, 100 mM sodium phosphate, and 100 mM sodium borate following the protocol by Chao *et al.* (30). Before assaying, we verified using universal indicator strips that addition of lysate to the buffer mixtures did not meaningfully affect the pH of the final mixtures.

Preparation of the cryo-EM grids

After confirming eGFP-INP expression, the E. coli cultures were incubated at 4 °C for an additional 3 days. The E. coli cells were spun down and resuspended in PBS to an OD600 nm of ~ 3. These concentrated E. coli samples were deposited onto freshly glow-discharged QUANTIFOIL holey carbon grids (Electron Microscopy Sciences). The grids were then blotted from the back side with the filter paper for ~5 s before plunge-frozen in liquid ethane, using a manual plunger-freezing apparatus as described previously (42, 43).

Cryo-FIB milling

The plunge-frozen grids with *E. coli* cells were clipped into cryo-FIB AutoGrids and mounted into the specimen shuttle under liquid nitrogen. An Aquilos2 cryo-FIB system (Thermo Fisher Scientific) was used to mill the thick bacterial samples into lamellae of < 200 nm in thickness. The milling process was completed using a protocol as previously described (44).

Cryo-ET data acquisition and tomogram reconstruction

Grids containing the lamellae obtained from cryo-FIB milling were loaded into either a 300-kV Titan Krios electron microscope (Thermo Fisher Scientific) equipped with a Direct Electron Detector and energy filter (Gatan) or a 200-kV Glacios Electron Microscope at Yale University. The FastTOMO script was used with the SerialEM software to collect tilt series with defocus values of approximately −6 µm (45), and a cumulative dose of ~70 e^-/Å covering angles from −48° to 48° (3° tilt step). Images were acquired at 42,000 × magnification with an effective pixel size of 2.148 Å. All recorded images were first drift corrected by MotionCor2 (46), stacked by the software package IMOD (47), and then aligned by IMOD using Pt particles as fiducial markers. TOMO3D

was used to generate tomograms by simultaneous iterative reconstruction technique (SIRT) (48). In total, 10 tomograms were reconstructed with TOMO3D for the WT INP while 5 tomograms were produced for the R-coil mutant.

Data Availability

The full dataset of long-read sequences used is available in **Supplementary Data 1**.

Contributions

T.H., J.C.L., and P.L.D. planned these experiments based on the identification of the R-coil subdomain by T.H. and P.L.D. in a prior work. T.H., J.C.L., and P.L.D. wrote the original manuscript for peer review. T.H. and P.L.D. designed the model for multimerization. T.H. coded the pipeline and performed the bioinformatic analysis of INPs.

T.H. and J.C.L. designed all mutants and prepared all samples except where otherwise noted. J.C.L. prepared the buffers and samples for the pH experiments, and performed all measurements of ice nucleation activity on the nanoliter osmometer. T.H. and J.C.L. performed the heat treatment experiments.

G.O. and I.B. prepared samples whose activity was assayed by N.R. and Y.R. on WISDOM. S.G., W.G. and J.L. performed cryo-FIB and cryo-ET experiments. S.G. prepared **Fig. 8** and Movies S2 and S3. T.H. prepared all other figures.

Acknowledgements

This work was supported by CIHR Foundation Grant FRN 148422 to P.L.D., who holds the Canada Research Chair in Protein Engineering, and by an Israel Science Foundation grant to I.B. YR acknowledges support by a research grant from the Yotam project and the Weizmann Institute sustainability and energy research initiative. S.G. was supported by a CIHR Post-Doctoral Fellowship and an NIH RO1 grant (R01AI087946) of J.L. W.G. was also supported by the NIH RO1 grant (R01AI087946) of J.L. We thank Virginia K. Walker for the gift of the *Pseudomonas borealis* strain.

Supplemental information

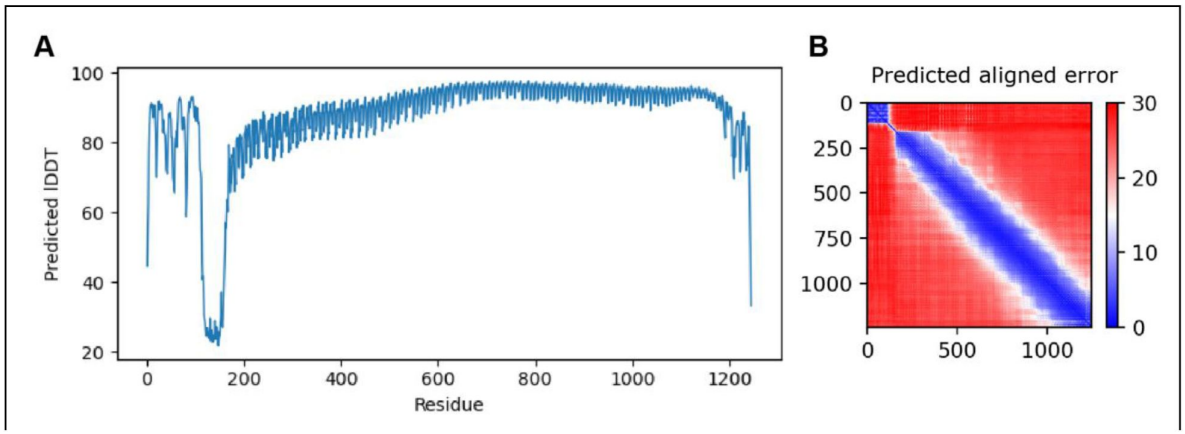

Figure S1.

AlphaFold shows high confidence in overall fold of the model.

A) Predicted location dependent difference test (pLDDT) shows the residue-by-residue confidence of the model generated by AlphaFold. Low values may indicate low-confidence or intrinsic disorder. **B)** Predicted aligned error (pAE) plots indicate the confidence in the relative orientation of the models. The x-and y-axes indicate residue position of the model, with low (blue) values indicating high confidence and high (red) indicating low. Rigid domains often appear as squares along the diagonal axis.

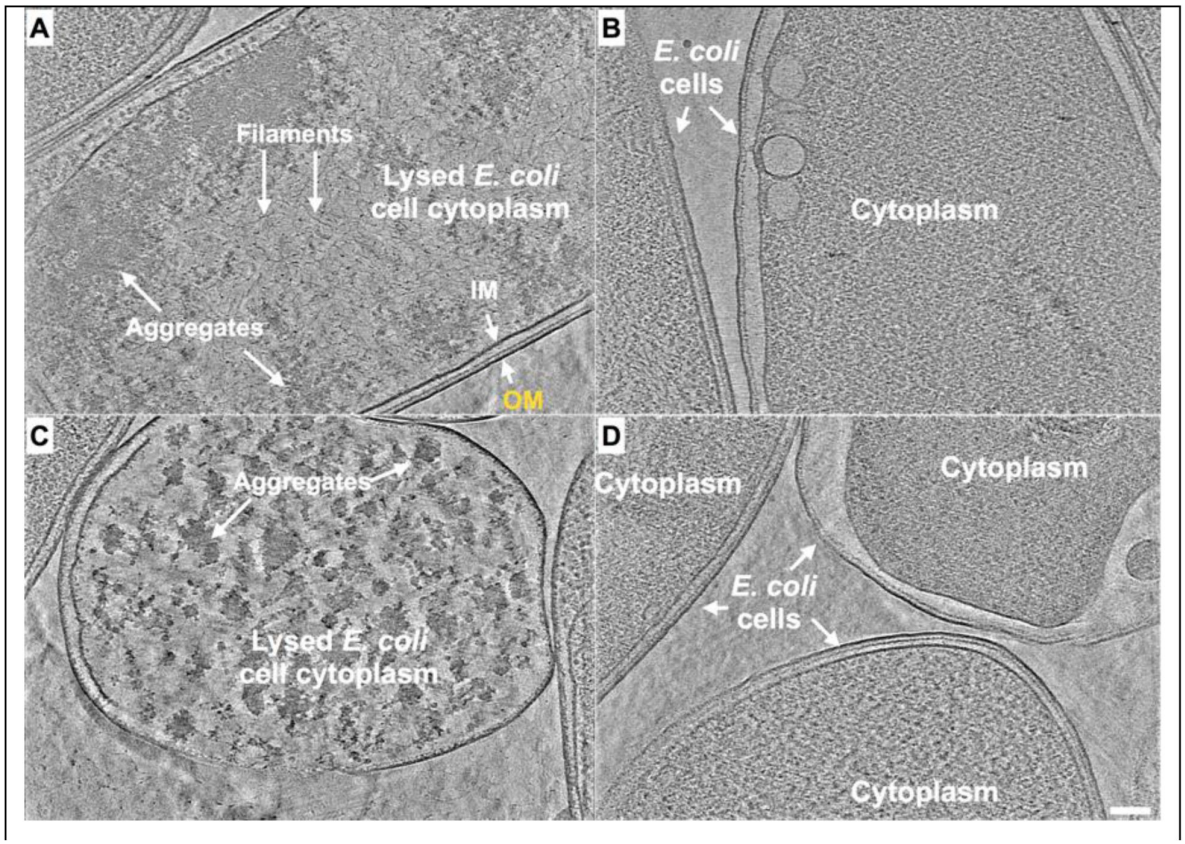

Figure S4.

***E. coli* expressing INP mutant lacking R-coils show no fibre clusters as observed in those cells overexpressing wild-type INP.**

A-D) Representative snapshots from 3-D cryo-tomograms showing cytoplasmic and extracellular features of various *E. coli* cells overexpressing an INP mutant in which all but the C-terminal R-coil have been replaced by WO-coils. All four images are in the same scale and the scale bar represents 100 nm.

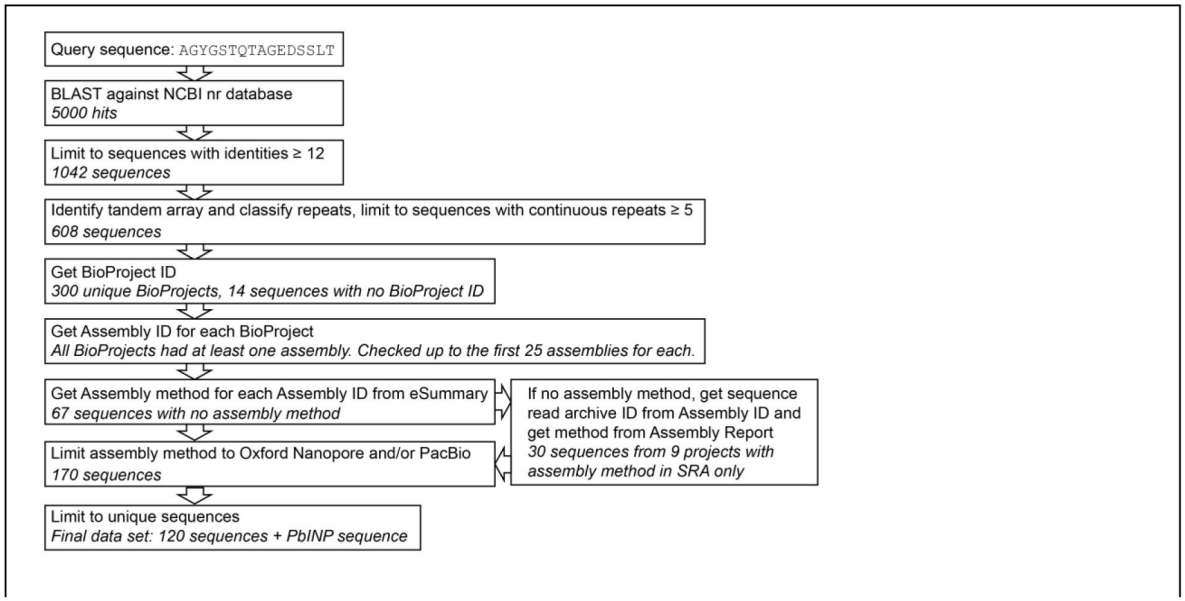

Figure S6.

Flowchart and quality control steps in sequence selection for bioinformatic analysis.

Ten known INPs from literature were used to generate a consensus sequence for WO-coils which was then used as a query in a BLAST against NCBI's non-redundant protein database to identify INPs. NCBI E-Utils were used to generate a data set using only genes from long-read DNA sequencing data.

References

1. Hoose C., Möhler O. (2012) **Heterogeneous ice nucleation on atmospheric aerosols: a review of results from laboratory experiments** *Atmospheric Chemistry and Physics* **12**:9817–9854 https://doi.org/10.5194/acp-12-9817-2012

2. Vali G., DeMott P. J., Möhler O., Whale T. F (2015) **Technical Note: A proposal for ice nucleation terminology** *Atmospheric Chemistry and Physics* **15**:10263–10270 https://doi.org/10.5194/acp-15-10263-2015

3. Lukas M., Schwidetzky R., Eufemio R. J., Bonn M., Meister K. (2022) **Toward Understanding Bacterial Ice Nucleation** *J Phys Chem B* **126**:1861–1867 https://doi.org/10.1021/acs.jpcb.1c09342

4. Lindow S. E (1983) **The Role of Bacterial ICE Nucleation in Frost Injury to Plants** *Annual Review of Phytopathology* **21**:363–384 https://doi.org/10.1146/annurev.py.21.090183.002051

5. Hill T. C., Moffett B. F., Demott P. J., Georgakopoulos D. G., Stump W. L., Franc G. D. (2014) **Measurement of ice nucleation-active bacteria on plants and in precipitation by quantitative PCR** *Appl Environ Microbiol* **80**:1256–1267 https://doi.org/10.1128/AEM.02967-13

6. Govindarajan A. G., Lindow S. E. (1988) **Size of bacterial ice-nucleation sites measured in situ by radiation inactivation analysis** *Proc Natl Acad Sci U S A* **85**:1334–1338 https://doi.org/10.1073/pnas.85.5.1334

7. Lindow S. E., Lahue E., Govindarajan A. G., Panopoulos N. J., Gies D (1989) **Localization of ice nucleation activity and the iceC gene product in Pseudomonas syringae and Escherichia coli** *Mol Plant Microbe Interact* **2**:262–272 https://doi.org/10.1094/mpmi-2-262

8. Forbes J., Bissoyi A., Eickhoff L., Reicher N., Hansen T., Bon C. G., et al. (2022) **Water-organizing motif continuity is critical for potent ice nucleation protein activity** *Nat Commun* **13** https://doi.org/10.1038/s41467-022-32469-9

9. Graether S. P., Jia Z. (2001) **Modeling Pseudomonas syringae ice-nucleation protein as a beta-helical protein** *Biophys J* **80**:1169–1173 https://doi.org/10.1016/S0006-3495(01)76093-6

10. Garnham C. P., Campbell R. L., Walker V. K., Davies P. L (2011) **Novel dimeric beta-helical model of an ice nucleation protein with bridged active sites** *BMC Struct Biol* **11** https://doi.org/10.1186/1472-6807-11-36

11. 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 https://doi.org/10.1038/35018610

12. Graham L. A., Qin W., Lougheed S. C., Davies P. L., Walker V. K (2007) **Evolution of hyperactive, repetitive antifreeze proteins in beetles** *J Mol Evol* **64**:387–398 https://doi.org/10.1007/s00239-005-0256-3

13. Kristiansen E., Ramlov H., Hojrup P., Pedersen S. A., Hagen L., Zachariassen K. E (2011) **Structural characteristics of a novel antifreeze protein from the longhorn beetle Rhagium inquisitor Insect** *Biochem Mol Biol* **41**:109–117 https://doi.org/10.1016/j.ibmb.2010.11.002

14. Garnham C. P., Campbell R. L., Davies P. L. (2011) **Anchored clathrate waters bind antifreeze proteins to ice** *Proc Natl Acad Sci U S A* **108**:7363–7367 https://doi.org/10.1073/pnas.1100429108

15. Green R. L., Warren G. J (1985) **Physical and functional repetition in a bacterial ice nucleation gene** *Nature* **317**:645–648 https://doi.org/10.1038/317645a0

16. Kajava A. V., Steven A. C (2006) **The turn of the screw: variations of the abundant beta-solenoid motif in passenger domains of Type V secretory proteins** *J Struct Biol* **155**:306–315 https://doi.org/10.1016/j.jsb.2006.01.015

17. Madzharova F., Bregnhoj M., Chatterley A. S., Lovschall K. B., Drace T., Andersen Dreyer L. S., et al. (2022) **Electrostatics Trigger Interfacial Self-Assembly of Bacterial Ice** *Nucleators Biomacromolecules* **23**:505–512 https://doi.org/10.1021/acs.biomac.1c01217

18. Lukas M., Schwidetzky R., Kunert A. T., Poschl U., Frohlich-Nowoisky J., Bonn M., et al. (2020) **Electrostatic Interactions Control the Functionality of Bacterial Ice** *Nucleators J Am Chem Soc* **142**:6842–6846 https://doi.org/10.1021/jacs.9b13069

19. Schwidetzky R., Lukas M., YazdanYar A., Kunert A. T., Poschl U., Domke K. F., et al. (2021) **Specific Ion-Protein Interactions Influence Bacterial Ice** *Nucleation Chemistry* **27**:7402–7407 https://doi.org/10.1002/chem.202004630

20. Juurakko C. L., di Cenzo G. C., Walker V. K. (2022) **Brachypodium Antifreeze Protein Gene Products Inhibit Ice Recrystallisation** *Attenuate Ice Nucleation, and Reduce Immune Response Plants (Basel* **11** https://doi.org/10.3390/plants11111475

21. Qiu Y., Hudait A., Molinero V. (2019) **How Size and Aggregation of Ice-Binding Proteins Control Their Ice Nucleation Efficiency** *J Am Chem Soc* **141**:7439–7452 https://doi.org/10.1021/jacs.9b01854

22. Roeters S. J., Golbek T. W., Bregnhoj M., Drace T., Alamdari S., Roseboom W., et al. (2021) **Ice-nucleating proteins are activated by low temperatures to control the structure of interfacial water** *Nat Commun* **12** https://doi.org/10.1038/s41467-021-21349-3

23. Schmid D., Pridmore D., Capitani G., Battistutta R., Neeser J.-R., Jann A (1997) **Molecular organisation of the ice nucleation protein InaV from Pseudomonas syringae** *FEBS Letters* **414**:590–594 https://doi.org/10.1016/s0014-5793(97)01079-x

24. Hartmann S., Ling M., Dreyer L. S. A., Zipori A., Finster K., Grawe S., et al. (2022) **Structure and Protein-Protein Interactions of Ice** *Nucleation Proteins Drive Their Activity Front Microbiol* **13** https://doi.org/10.3389/fmicb.2022.872306

25. Novikova I. V., China S., Evans J. E. (2018) **Overcoming bottlenecks for in vitro synthesis and initial structural insight of ice nucleating protein InaZ** *bioRxiv* https://doi.org/10.1101/334987

26. Li Q., Yan Q., Chen J., He Y., Wang J., Zhang H., et al. (2012) **Molecular characterization of an ice nucleation protein variant (inaQ) from Pseudomonas syringae and the analysis of its transmembrane transport activity in Escherichia coli** *Int J Biol Sci* **8**:1097–1108 https://doi.org/10.7150/ijbs.4524

27. Kassmannhuber J., Rauscher M., Schoner L., Witte A., Lubitz W (2017) **Functional display of ice nucleation protein InaZ on the surface of bacterial ghosts** *Bioengineered* **8**:488–500 https://doi.org/10.1080/21655979.2017.1284712

28. Kassmannhuber J., Mauri S., Rauscher M., Brait N., Schoner L., Witte A., et al. (2020) **Freezing from the inside: Ice nucleation in Escherichia coli and Escherichia coli ghosts by inner membrane bound ice nucleation protein** *InaZ Biointerphases* **15** https://doi.org/10.1116/1.5142174

29. Wick R. R., Judd L. M., Gorrie C. L., Holt K. E (2017) **Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads** *PLoS Comput Biol* **13** https://doi.org/10.1371/journal.pcbi.1005595

30. Chao H., Sonnichsen F. D., DeLuca C. I., Sykes B. D., Davies P. L. (1994) **Structure-function relationship in the globular type III antifreeze protein: identification of a cluster of surface residues required for binding to ice** *Protein Sci* **3**:1760–1769 https://doi.org/10.1002/pro.5560031016

31. Melnik T., Povarnitsyna T., Solonenko H., Melnik B (2011) **Studies of irreversible heat denaturation of green fluorescent protein by differential scanning microcalorimetry** *Thermochimica Acta* **512**:71–75 https://doi.org/10.1016/j.tca.2010.09.002

32. Kajava A. V., Steven A. C (2006) **Beta-rolls, beta-helices, and other beta-solenoid proteins** *Adv Protein Chem* **73**:55–96 https://doi.org/10.1016/S0065-3233(06)73003-0

33. Lindow S. E., Arny D. C., Upper C. D. (1982) **Bacterial ice nucleation: a factor in frost injury to plants** *Plant Physiol* **70**:1084–1089 https://doi.org/10.1104/pp.70.4.1084

34. Liu K., Wang C., Ma J., Shi G., Yao X., Fang H., et al. (2016) **Janus effect of antifreeze proteins on ice nucleation** *Proc Natl Acad Sci U S A* **113**:14739–14744 https://doi.org/10.1073/pnas.1614379114

35. Cock P. J., Antao T., Chang J. T., Chapman B. A., Cox C. J., Dalke A., et al. (2009) **Biopython: freely available Python tools for computational molecular biology and bioinformatics** *Bioinformatics* **25**:1422–1423 https://doi.org/10.1093/bioinformatics/btp163

36. Tareen A., Kinney J. B. (2019) **Logomaker: Beautiful sequence logos in python** *bioRxiv* https://doi.org/10.1101/635029

37. 37., Waterhouse A. M., Procter J. B., Martin D. M., Clamp M., Barton G. J. (2009) **Jalview Version 2--a multiple sequence alignment editor and analysis workbench** *Bioinformatics* **25**:1189–1191 https://doi.org/10.1093/bioinformatics/btp033

38. Belval L., Marquette A., Mestre P., Piron M. C., Demangeat G., Merdinoglu D., et al. (2015) **A fast and simple method to eliminate Cpn60 from functional recombinant proteins produced by E. coli** *Arctic Express Protein Expr Purif* **109**:29–34 https://doi.org/10.1016/j.pep.2015.01.009

39. Reicher N., Segev L., Rudich Y (2018) **The WeIzmann Supercooled Droplets Observation on a Microarray (WISDOM) and application for ambient dust Atmospheric** *Measurement Techniques* **11**:233–248 https://doi.org/10.5194/amt-11-233-2018

40. Lee J. C., Hansen T., Davies P. L (2023) **Droplet freezing assays using a nanoliter osmometer** *Cryobiology* **113** https://doi.org/10.1016/j.cryobiol.2023.104584

41. Braslavsky I., Drori R. (2013) **LabVIEW-operated novel nanoliter osmometer for ice binding protein investigations** *J Vis Exp* https://doi.org/10.3791/4189

42. Liu J., Lin T., Botkin D. J., McCrum E., Winkler H., Norris S. J. (2009) **Intact flagellar motor of Borrelia burgdorferi revealed by cryo-electron tomography: evidence for stator ring curvature and rotor/C-ring assembly flexion** *J Bacteriol* **191**:5026–5036 https://doi.org/10.1128/JB.00340-09

43. Zhao X., Zhang K., Boquoi T., Hu B., Motaleb M. A., Miller K. A., et al. (2013) **Cryoelectron tomography reveals the sequential assembly of bacterial flagella in Borrelia burgdorferi** *Proc Natl Acad Sci U S A* **110**:14390–14395 https://doi.org/10.1073/pnas.1308306110

44. Xiang Y., Surovtsev I. V., Chang Y., Govers S. K., Parry B. R., Liu J., et al. (2021) **Interconnecting solvent quality, transcription, and chromosome folding in** *Escherichia coli Cell* **184**:3626–3642 https://doi.org/10.1016/j.cell.2021.05.037

45. Xu A., Xu C. (2021) **FastTomo: A SerialEM Script for Collecting Electron Tomography Data** *bioRxiv* https://doi.org/10.1101/2021.03.16.435675

46. Zheng S. Q., Palovcak E., Armache J. P., Verba K. A., Cheng Y., Agard D. A (2017) **MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy** *Nat Methods* **14**:331–332 https://doi.org/10.1038/nmeth.4193

47. Kremer J. R., Mastronarde D. N., McIntosh J. R (1996) **Computer visualization of three-dimensional image data using** *IMOD J Struct Biol* **116**:71–76 https://doi.org/10.1006/jsbi.1996.0013

48. Agulleiro J. I., Fernandez J. J (2015) **Tomo3D 2.0--exploitation of advanced vector extensions (AVX) for 3D reconstruction** *J Struct Biol* **189**:147–152 https://doi.org/10.1016/j.jsb.2014.11.009

Article and author information

Thomas Hansen
Department of Biomedical and Molecular Sciences, Queen's University, Kingston, ON Canada K7L 3N6
ORCID iD: 0000-0001-7515-9540

Jocelyn C. Lee
Department of Biomedical and Molecular Sciences, Queen's University, Kingston, ON Canada K7L 3N6
ORCID iD: 0009-0005-4431-5439

Naama Reicher
Department of Earth and Planetary Sciences, The Weizmann Institute of Science, Rehovot 7610001, Israel
ORCID iD: 0000-0003-4555-1343

Gil Ovadia
The Robert H. Smith Faculty of Agriculture, Food and Environment, Institute of Biochemistry, Food Science, and Nutrition, The Hebrew University of Jerusalem, Rehovot 7610001, Israel

Shuaiqi Guo
Department of Microbial Pathogenesis, Yale University School of Medicine, New Haven, CT 06536
ORCID iD: 0000-0001-5763-8513

Wangbiao Guo
Department of Microbial Pathogenesis, Yale University School of Medicine, New Haven, CT 06536
ORCID iD: 0000-0001-5398-4910

Jun Liu
Department of Microbial Pathogenesis, Yale University School of Medicine, New Haven, CT 06536

Ido Braslavsky
The Robert H. Smith Faculty of Agriculture, Food and Environment, Institute of Biochemistry, Food Science, and Nutrition, The Hebrew University of Jerusalem, Rehovot 7610001, Israel
ORCID iD: 0000-0001-8985-8211

Yinon Rudich
Department of Earth and Planetary Sciences, The Weizmann Institute of Science, Rehovot 7610001, Israel
ORCID iD: 0000-0003-3149-0201

Peter L. Davies
Department of Biomedical and Molecular Sciences, Queen's University, Kingston, ON Canada K7L 3N6
For correspondence: peter.davies@queensu.ca
ORCID iD: 0000-0002-8026-7818

Copyright

Editors

Reviewing Editor
Hannes Neuweiler
University of Würzburg, Würzburg, Germany

Senior Editor
David Ron
University of Cambridge, Cambridge, United Kingdom

Reviewer #1 (Public Review):

Summary: Hansen et al. dissect the molecular mechanisms of bacterial ice nucleating proteins mutating the protein systematically. They assay the ice nucleating ability for variants changing the R-coils as well as the coil capping motifs. The ice nucleation mechanism depends on the integrity of the R-coils, without which the multimerization and formation of fibrils are disrupted.

Strengths: The effects of mutations are really dramatic, so there is no doubt about the effect. The variants tested are logical and progressively advance the story. The authors identify an underlying mechanism involving multimerization, which is plausible and compatible with EM data. The model is further shown to work in cells by tomography.

Weaknesses: The theoretical model presented for how the proteins assemble into fibrils is simple, but not supported by much data.

- https://doi.org/10.7554/eLife.91976.2.sa2

Reviewer #2 (Public Review):

Summary:
This paper further investigates the role of self-assembly of ice-binding bacterial proteins in promoting ice-nucleation. For the P. borealis Ice Nucleating Protein (PbINP) studied here, earlier work had already determined clearly distinct roles for different subdomains of the protein in determining activity. Key players are the water-organizing loops (WO-loops) of the central beta-solenoid structure and a set of non-water-organizing C-terminal loops, called the R-loops in view of characteristically located arginines. Previous mutation studies (using nucleation activity as a read-out) had already suggested the R-loops interact with the WO loops, to cause self-assembly of PbINP, which in turn was thought to lead to enhanced ice-nucleating activity. In this paper, the activities of additional mutants are studied, and a bioinformatics analysis on the statistics of the number of WO- and R-loops is presented for a wide range of bacterial ice-nucleating proteins, and additional electron-microscopy results are presented on fibrils formed by the non-mutated PbINP in E coli lysates.

Strengths:
-A very complete set of additional mutants is investigated to further strengthen the earlier hypothesis.
-A nice bioinformatics analysis that underscores that the hypothesis should apply not only to PbINP but to a wide range of (related) bacterial ice-nucleating proteins.
-Convincing data that PbINP overexpressed in E coli forms fibrils (electron microscopy on E coli lysates).

Weaknesses:
-The new data is interesting and further strengthens the hypotheses put forward in the earlier work. However, just as in the earlier work, the proof for the link between self-assembly and ice-nucleation remains indirect. Assembly into fibrils is shown for E coli lysates expressing non-mutated pbINP, hence it is indeed clear that pbINP self-associates. It is not shown however that the mutations that lead to loss of ice-nucleating activity also lead to loss of self-assembly. A more quantitative or additional self-assembly assay could shine light on this, either in the present or in future studies.

-Also the "working model" for the self-assembly of the fibers remains not more than that, just as in the earlier papers, since the mutation-activity relationship does not contain enough information to build a good structural model. Again, a better model would require different kinds of experiments, that yield more detailed structural data on the fibrils.

- https://doi.org/10.7554/eLife.91976.2.sa1

Reviewer #3 (Public Review):

Summary: in this manuscript, Hansen and co-authors investigated the role of R-coils in the multimerization and ice nucleation activity of PbINP, an ice nucleation protein identified in Pseudomonas borealis. The results of this work suggest that the length, localization, and amino acid composition of R-coils are crucial for the formation of PbINP multimers.

Strengths: The authors use a rational mutagenesis approach to identify the role of the length, localisation, and amino acid composition of R-coils in ice nucleation activity. Based on these results, the authors hypothesize a multimerization model. Overall, this is a multidisciplinary work that provides new insights into the molecular mechanisms underlying ice nucleation activity.

Weaknesses: Several parts of the work appear cryptic and unsuitable for non-expert readers. The results of this work should be better described and presented.

- https://doi.org/10.7554/eLife.91976.2.sa0

Author Response

The following is the authors' response to the original reviews.

Public Reviews:

> ***Reviewer #1 (Public Review):***
>
> *Summary: Hansen et al. dissect the molecular mechanisms of bacterial ice nucleating proteins mutating the protein systematically. They assay the ice nucleating ability for variants changing the R-coils as well as the coil capping motifs. The ice nucleation mechanism depends on the integrity of the R-coils, without which the multimerization and formation of fibrils are disrupted.*
>
> *Strengths: The effects of mutations are really dramatic, so there is no doubt about the effect. The variants tested are logical and progressively advance the story. The authors identify an underlying mechanism involving multimerization, which is plausible and compatible with EM data. The model is further shown to work in cells by tomography.*
>
> *Weaknesses: The theoretical model presented for how the proteins assemble into fibrils is simple, but not supported by much data.*

Agreed. This theoretical INP multimer model was introduced to promote discussion and elicit ideas on how to prove or disprove it. The length and width of the fibres are defined by cryo-ET results, in which the narrow width is just sufficient to accommodate a dimer of the INPs, and the long length requires that several INPs are joined end to end. Their antiparallel arrangement produces identical ends to the dimer and avoids steric clash of the C-terminal cap structures as well as the C-terminal GFP tag. This model can accommodate the wide range of INPs lengths seen in nature (due to different numbers of water-organizing coils) and introduced in mutagenesis experiments (Forbes et al. 2022). It defines a critical role for the R-

coil subdomain in joining the dimers together and explains why this region cannot be shortened by more than a few coils either in nature or by experimentation.

In response to specific criticisms of the model (Fig. 9), we have redesigned this to be less schematic and to incorporate several copies of the AlphaFold-predicted structure.

> ***Reviewer #2 (Public Review):***
>
> *Summary:*
>
> *This paper further investigates the role of self-assembly of ice-binding bacterial proteins in promoting ice-nucleation. For the P. borealis Ice Nucleating Protein (PbINP) studied here, earlier work had already determined clearly distinct roles for different subdomains of the protein in determining activity. Key players are the water-organizing loops (WO-loops) of the central beta-solenoid structure and a set of non-water-organizing C-terminal loops, called the R-loops in view of characteristically located arginines. Previous mutation studies (using nucleation activity as a read-out) had already suggested the R-loops interact with the WO loops, to cause self-assembly of PbINP, which in turn was thought to lead to enhanced ice-nucleating activity. In this paper, the activities of additional mutants are studied, and a bioinformatics analysis on the statistics of the number of WO- and R-loops is presented for a wide range of bacterial ice-nucleating proteins, and additional electron-microscopy results are presented on fibrils formed by the non-mutated PbINP in E coli lysates.*
>
> *Strengths:*
>
> *-A very complete set of additional mutants is investigated to further strengthen the earlier hypothesis.*
>
> *-A nice bioinformatics analysis that underscores that the hypothesis should apply not only to PbINP but to a wide range of (related) bacterial ice-nucleating proteins.*
>
> *-Convincing data that PbINP overexpressed in E coli forms fibrils (electron microscopy on E coli lysates).*
>
> *Weaknesses:*
>
> *-The new data is interesting and further strengthens the hypotheses put forward in the earlier work. However, just as in the earlier work, the proof for the link between self-assembly and ice-nucleation remains indirect. Assembly into fibrils is shown for E coli lysates expressing non-mutated pbINP, hence it is indeed clear that pbINP self-associates. It is not shown however that the mutations that lead to loss of ice-nucleating activity also lead to loss of self-assembly. A more quantitative or additional self-assembly assay could shine light on this, either in the present or in future studies.*

The control cryo-ET experiment where the R-coils were deleted and INP fibres were not seen is consistent with a link between the loss of ice-nucleating activity and the loss of self-assembly. However, we agree that a more direct measurement of the physical state of INP molecules is needed to prove the link.

> *-Also the "working model" for the self-assembly of the fibers remains not more than that, just as in the earlier papers, since the mutation-activity relationship does not contain enough information to build a good structural model. Again, a better model would require different kinds of experiments, that yield more detailed structural data on the fibrils.*

Reviewer #1 also raised these criticisms of the model, which we have responded to (above). Testing the model is a focus of our continuing experiments on INPs.

> ***Reviewer #3 (Public Review):***
>
> *Summary: in this manuscript, Hansen and co-authors investigated the role of R-coils in the multimerization and ice nucleation activity of PbINP, an ice nucleation protein identified in Pseudomonas borealis. The results of this work suggest that the length, localization, and amino acid composition of R-coils are crucial for the formation of PbINP multimers.*
>
> *Strengths: The authors use a rational mutagenesis approach to identify the role of the length, localisation, and amino acid composition of R-coils in ice nucleation activity. Based on these results, the authors hypothesize a multimerization model. Overall, this is a multidisciplinary work that provides new insights into the molecular mechanisms underlying ice nucleation activity.*
>
> *Weaknesses: Several parts of the work appear cryptic and unsuitable for non-expert readers. The results of this work should be better described and presented.*

In revising the manuscript for reposting we have rewritten sections to make it more accessible to the non-expert. Incorporating the detailed recommendations of the reviewers has been helpful in this effort.

Recommendations for the authors: please note that you control which revisions to undertake from the public reviews and recommendations for the authors

> ***Reviewer #1 (Recommendations For The Authors):***
>
> *Introduction: Curiously, there is no mention at all in the introduction of what the biological function of these ice-nucleating proteins is.*

We added the following text to the first paragraph of the Introduction: ”INP-producing bacteria are widespread in the environment where they are responsible for initiating frost (4) and atmospheric precipitation (5). As such, these bacteria play a significant role in the Earth’s hydrological cycle and in agricultural productivity.”

> *Line 70: TXT, SLT, and Y motifs are mentioned, but only the first is described. Also, TXT name alternates between TXT and TxT in the manuscript. (I think the latter is more correct).*

These putative water-organizing motifs are introduced in the preceding paper (new ref 8). We now use TxT consistently throughout the manuscript and have converted SLT to SxT because L is an inward-pointing residue that is not directly involved in water organization.

> *Line 236: A construct with repeats deleted is tested for thermostability, but it is not really explained what hypothesis this experiment is supposed to test.*

This is an observation that adds information about the stability of the INP multimers and will need to be explained by the structure.

> *Line 267: The authors test a mutant where the N-terminal coil is disrupted and find a big effect. Nevertheless, no conclusion is drawn. What does this result mean?*

On the contrary, INP activity is not appreciably affected by N-terminal deletion.

> *Line 269: The CryoEM begins rather abruptly with technical details. Consider introducing the paragraph with a brief statement about what you want to investigate. Also, the analysis seems a little half-hearted.*
>
> *Given that the authors describe other EM studies of fibrils of the same protein it would be nice with a clear statement about what is new in their study and how it compares to previous studies.*

We have added this statement about why we used Cryo-EM: “The idea that INPs must assemble into larger structures to be effective at ice nucleation has persisted since their discovery (6). In the interim the resolving power of cryo-EM has immensely improved. Here we elected to use cryo-electron tomography to view the INP multimers in situ and avoid any perturbation of their superstructure during isolation.”

> *Fig. 7B: Single-letter amino acid codes are always capitalized.*

We have revised this figure to use capital letters for the amino acids.

> *Fig. 9: This figure is really hard to read even though it is very simplistic. I would consider making a figure with several copies of the AlphaFold model instead. Especially panel D, I do not know what is supposed to show.*

We have followed this advice and have completely revised the figure using copies of the AlphaFold model. Panel D (now C) shows two cross-sections through the AlphaFold model.

> *Line 355 onwards: The model of the INP is the weakest part of the manuscript. This reviewer considers that the model is crude and it is unclear what information the model is supported by. The authors might want to consider running an AlphaFold multimer to get a better model of at least the dimer.*

Our objective now is to validate or disprove the model by experimentation using protein-protein cross-linking in conjunction with mass spectrometry, and higher resolution cryo-EM methods.

> ***Reviewer #2 (Recommendations For The Authors):***
>
> *I would suggest more frankly discussing the weaknesses mentioned in my public review, as well as approaches that could be used in the future to address these.*

In the cryo-ET analysis, INP mutations of the R-coils that lead to loss of ice-nucleating activity fail to show fibres in the bacteria (Fig. S4), which is consistent with the loss of self-assembly. We are working on physical methods that can assess the degree of assembly of the different INP constructs and mutations. We are working to validate and improve the working model of INP multimers.

> ***Reviewer #3 (Recommendations For The Authors):***
>
> *Abstract*
>
> *Line 18. Below 0 Celsius should be < 0 {degree sign}C.*

Done

> *Line 25. E. coli should be Escherichia coli*

Done

> *Line 29. E. coli should be in italics.*

Done

> *Introduction*
>
> *The introduction is weak and not suitable for non-expert readers. Moreover, in some parts it is cryptic and it is not clear whether the authors are describing INP in general or PbINP. The introduction should be reorganized to highlight the novelty of this paper compared to Forbes et al. 2022.*

The changes we have made to the Introduction can be seen in the 'documents compared' version where the changes are tracked.

> *Line 45. It is unclear whether this paragraph is a result reported in the literature or the result of this work. Please clarify.*

These are results reported in the literature as indicated by the references cited in the paragraph.

> *Line 54. It is not clear whether this paragraph describes PbINP or INP in general.*

This paragraph begins with INPs in general and then focuses on PbINP.

> *Results*
>
> *Line 109. This section would benefit from a paragraph in which the authors describe the rationale for this bioinformatic analysis.*

We added the following Statement: "A bioinformatic analysis of bacterial INPs was undertaken to identify their variations in size and sequence to understand what is common to all that could guide experiments to probe higher order structure and help develop a collective model of the INP multimer."

> *Some information is needed on the selected sequences such as sequence identity, what do the authors mean by nr database?*

The abbreviation nr has been replaced by 'non-redundant'. As explained in that same paragraph the sequences selected were those from long-read sequences that could be relied on to accurately count the number of solenoid coils.

> *Line 144. The standard deviation is necessary to understand whether these differences are statistically significant.*

These have been added as p values.

> *Figure 2. I noticed that the authors used GFP-tagged PbINP. Why? In addition, panel C is never mentioned in the manuscript.*

The GFP tag was used to confirm expression of the PbINP in E. coli. We have added this sentence: "As previously described these constructs were tagged with GFP as an internal control for INP production, and its addition had no measured effect on ice nucleation activity

(8).”The GFP tag was also useful as in internal control for the heat denaturation experiments featured in Fig. 6, where it lost its fluorescence between 65 and 75 °C.

Fig. 2C is now cited alongside Fig. 2B.

> *Figure 3. In my opinion, the results of the R-coil deletion should also be shown in Figure 2. Line 171. This section is cryptic. A logo sequence or an alignment of WO-coils and R-coils of PbINP could be helpful for the reader. Instead of the architecture of the whole protein, it would be useful to have the sequence of the R-coils with the residues that the authors mutagenised.*

The logo sequences are available in Fig. 1.

> *Line 202. Here, the authors describe a new experimental setup. As the Materials and Methods section follows the Discussion, the authors should state in the first paragraph of the Results section that IN activity was measured on whole cells.*

We have now modified the introductory sentence to read: “Ice nucleation assays were performed on intact E. coli expressing PbINP to assess the activity of the incremental replacement mutants.”

> *Line 202. The authors investigated the effects of pH and temperature (Line 223) on the IN activity. The authors should better introduce the rationale for these experiments and how they fit within the work.*

We have now modified the following sentence to provide the rationale: “To see how important electrostatic interactions were in the multimerization of PbINP as reflected by its ice nucleation activity, it was necessary to lyse the E. coli to change the pH surrounding the INP multimers.”

> *Line 245. This work is supported by a model provided by Alphafold. I wonder how reliable this model is; the authors should indicate the quality of the model and provide the accuracy values of the residuals.*

This information is now provided in Figure S1.

> *Line 259. Typically in mutagenesis studies, a key residue is substituted with alanine to create a loss of function variant. In this case, the authors have made the following substitutions F1204D, D1208L, and Y1230D, it is not clear to me why the authors have replaced an aromatic residue with one of aspartic acid that is negatively charged.*

We have justified these more extreme changes as follows: “For an enhanced effect of the mutations hydrophobic residues were replaced with charged ones and vice versa.”

> *Line 269. This paragraph seems completely unrelated to the section entitled: The β-solenoid of INPs is stabilized by a capping structure at the C terminus, but not at the N terminus.*

We had omitted the sub-heading “Cryo-electron tomography reveals INPs multimers form bundled fibres in recombinant cells”, which is now in place.

> *Discussion*
>
> *Overall, the discussion is too long and some parts appear cryptic, this section should be reorganized.*

The changes we have made to the Discussion can be seen in the 'documents compared' version where the changes are tracked.

> *Line 354. It is not clear what experimental evidence supports this model. In the results, this model is never mentioned and it is not clear whether it was obtained by computational analysis or not.*

The model is presented in the Discussion because it was not arrived at by experimentation but is an attempt to integrate the observations made in the Results section. The experimental evidence that supports this model is reviewed in the Discussion section: "Working model of the INP multimer is consistent with the properties of INPs and their multimers."

> *Line 354. The authors used GFP-tagged PbINP. The Authors should discuss the role of GFP in this model and IN activity. A measurement of IN activity on PbINP without GFP would be useful.*

We have previously shown in Ref 8 that the GFP tag has no detrimental effect on ice nucleation activity. Our model for the INP multimer can accommodate this C-terminal tag without any steric hindrance.

> *Line 364. The Authors hypothesize that electrostatic interactions stabilize end-to-end dimer associations. To test this hypothesis, the authors should measure the activity of IN at increasing concentrations of NaCl. It is known that high salt concentrations shield charges by preventing the formation of electrostatic intermolecular interactions.*

We have added this sentence to the Discussion: "Another useful test of the electrostatic component to the multimer model would be to study the effects of increasing salt concentration on ice nucleation activity of the E. coli extracts."

> *Line 439. Conclusions should be useful for the reader.*
>
> *Material and Methods*
>
> *In several sections, the authors refer to what has already been published in Forbes et al. However, the minimum information should also be described in this work. In addition, the Authors should indicate the number of replicates.*

The ice nucleation assays on whole cells were done on the WISDOM apparatus, which integrates 100's of individual measurements to obtain a T50 value. These T50 values were confirmed by assays on the nanoliter osmometer apparatus. The numbers of replicates used on the nanoliter osmometer apparatus are indicated by box and whisker plots in Figs. 5 & 6 with boxes and bars showing quartiles, with medians indicated by a centre line.

> *Line 500. This paragraph should be removed as the results are not described in the manuscript.*

This is a Methods section that describes how that INPs were expression in E. coli. It has details that are important for researchers who want to repeat our findings, such as the use of the Arctic Express strain for producing INP.