Competing stress-dependent oligomerization pathways regulate self-assembly of the periplasmic protease-chaperone DegP

Robert W. Harkness[a,b,c,d,1], Yuki Toyama[a,b,c,d], Zev A. Ripstein[a,b,c,d], Huaying Zhao[e], Alexander I. M. Sever[a,b,c,d], Qing Luan[f], Jacob P. Brady[a,b,c,d], Patricia L. Clark[f], Peter Schuck[e], and Lewis E. Kay[a,b,c,d,1]

[a]Department of Biochemistry, University of Toronto, Toronto, ON M5S 1A8, Canada; [b]Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada; [c]Department of Chemistry, University of Toronto, Toronto, ON M5S 3H6, Canada; [d]Program in Molecular Medicine, The Hospital for Sick Children Research Institute, Toronto, ON M5G 0A4, Canada; [e]National Institute of Biomedical Imaging and Bioengineering, NIH, Bethesda, MD 20892; and [f]Department of Chemistry & Biochemistry, University of Notre Dame, Notre Dame, IN 46556

Edited by G. Marius Clore, NIH, Bethesda, MD, and approved June 28, 2021 (received for review May 25, 2021)

DegP is an oligomeric protein with dual protease and chaperone activity that regulates protein homeostasis and virulence factor trafficking in the periplasm of gram-negative bacteria. A number of oligomeric architectures adopted by DegP are thought to facilitate its function. For example, DegP can form a "resting" hexamer when not engaged to substrates, mitigating undesired proteolysis of cellular proteins. When bound to substrate proteins or lipid membranes, DegP has been shown to populate a variety of cage- or bowl-like oligomeric states that have increased proteolytic activity. Though a number of DegP's substrate-engaged structures have been robustly characterized, detailed mechanistic information underpinning its remarkable oligomeric plasticity and the corresponding interplay between these dynamics and biological function has remained elusive. Here, we have used a combination of hydrodynamics and NMR spectroscopy methodologies in combination with cryogenic electron microscopy to shed light on the apo-DegP self-assembly mechanism. We find that, in the absence of bound substrates, DegP populates an ensemble of oligomeric states, mediated by self-assembly of trimers, that are distinct from those observed in the presence of substrate. The oligomeric distribution is sensitive to solution ionic strength and temperature and is shifted toward larger oligomeric assemblies under physiological conditions. Substrate proteins may guide DegP toward canonical cage-like structures by binding to these preorganized oligomers, leading to changes in conformation. The properties of DegP self-assembly identified here suggest that apo-DegP can rapidly shift its oligomeric distribution in order to respond to a variety of biological insults.

protein self-assembly | hydrodynamics | methyl-TROSY NMR | protein thermodynamics and kinetics

Cellular function depends on the proper folding, localization, and recycling of proteins in a highly coordinated manner. The dysregulation of these processes leads to the accumulation and aggregation of protein molecules, potentially impairing cellular function and ultimately leading to cell death (1–3). Cells maintain protein homeostasis by deploying a host of protein chaperones (for protein refolding) and proteases (for protein cleavage) (4). Interestingly, some proteins can fulfill both of these functions. For example, the widely conserved family of high temperature requirement A (HtrA) proteins exhibit dual protease and chaperone function and are critical to protein quality control (5–7). In addition to their general role in protein homeostasis, HtrA proteins are also responsible for the cleavage of regulatory proteins that control signaling pathways involved in cell motility, proliferation, and apoptosis (5).

One of the bacterial orthologs of the HtrA family is known as DegP, or the Do protease (8, 9). In gram-negative bacteria, DegP becomes localized to the periplasm through its signal sequence, which is subsequently removed to yield the mature form of the enzyme (Fig. 1 *A* and *B*, *Top*) (10). DegP then acts as a key protease and chaperone in the periplasmic protein homeostasis network (Fig. 1*A*). Highlighting its critical role in protecting bacterial function, DegP is required for growth in response to cellular stressors, such as heat (6, 8), as well as oxidative (11, 12) and osmotic shock (13), where it is thought to proteolyze misfolded or damaged client proteins (11, 14). DegP's chaperone activities, in contrast, include the sequestration and trafficking of outer-membrane proteins (OMPs) and virulence factors through the periplasm to the outer membrane (15–17). At the outer membrane, DegP maintains virulence factors in a partially unfolded state to facilitate their secretion (Fig. 1*A*) (18, 19). Modifications to the DegP gene result in the aggregation and impaired secretion of virulence factors, establishing its critical role in bacterial pathogenesis (19, 20). DegP is thus a potential drug target, as antibiotics that either disrupt its interactions with, or conversely, promote its proteolysis of virulence factors, could lead to attenuated infections (21).

Mature DegP monomers are composed of a serine protease and a pair of PDZ domains (PDZ1 and PDZ2, Fig. 1 *B*, *Top*). The basic functional unit of DegP is a homotrimer where three constituent monomers are tightly associated via their protease domains (Fig. 1 *B*, *Left*) that are oriented toward the interior of the particle, with the PDZ1 and PDZ2 domains decorating the

Significance

The pathogenesis of gram-negative bacteria relies on secretion of virulence factors to bacterial cell surfaces, where they interact with human host cells, interfering with their function. The DegP protease-chaperone regulates protein homeostasis in the periplasm of such bacteria and is a potential drug target. DegP is thought to exist in a resting hexameric state when not engaged with substrate molecules and cage-like 12 and 24 mers when processing substrates. Using a combination of biophysical methods, we show that DegP adopts an interconverting, broadly populated distribution of high-order oligomeric states in the absence of substrate molecules, which could facilitate a rapid response to protein misfolding that occurs as a result of cellular stresses, such as heat, oxidation, and osmotic shock.

Author contributions: R.W.H., Y.T., P.L.C., P.S., and L.E.K. designed research; R.W.H., Y.T., Z.A.R., H.Z., A.I.M.S., Q.L., J.P.B., and L.E.K. performed research; R.W.H. contributed new reagents/analytic tools; R.W.H., Y.T., Z.A.R., H.Z., A.I.M.S., Q.L., and P.S. analyzed data; and R.W.H. and L.E.K. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: r.harkness@utoronto.ca or kay@pound.med.utoronto.ca.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2109732118/-/DCSupplemental.

Published August 6, 2021.

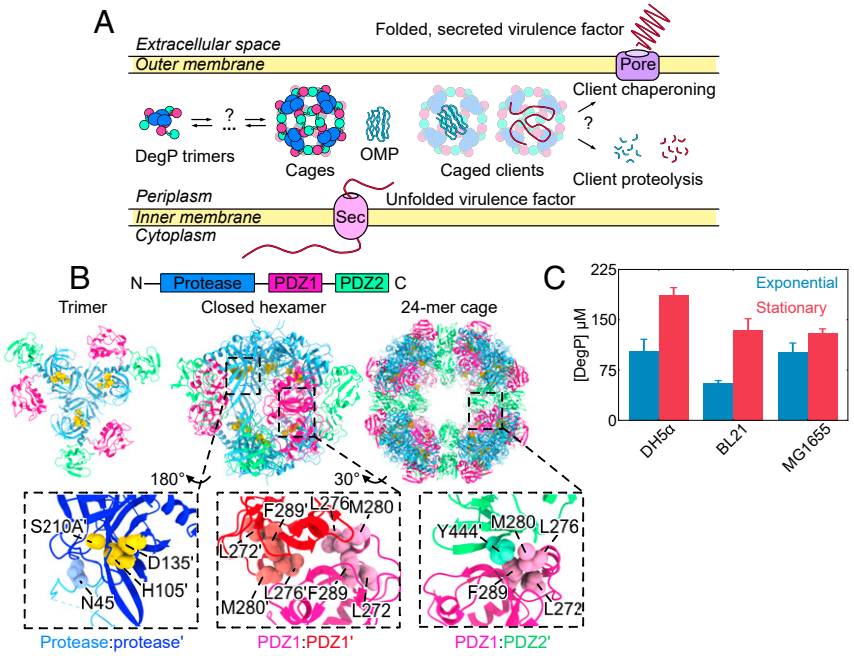

Fig. 1. Structure and function of the periplasmic protease-chaperone DegP. (*A*) DegP maintains protein homeostasis in the periplasm by forming oligomeric assemblies of trimer molecules that can proteolyze or chaperone client proteins. (*B*) DegP monomers (*Top*) contain a protease domain (blue) and tandem PDZ domains (red and green). Monomers associate into trimers (*Left*, one-half of the closed hexamer structure Protein Data Bank [PDB]: 1KY9) through the interaction of their protease domains. Trimers can associate into a variety of cage-like states depending on their substrate occupancy. The closed hexamer structure (*Middle*, PDB: 1KY9) (24) is adopted by the formation of intertrimer protease:protease′ and PDZ1:PDZ1′ contacts. Higher-order structures, such as the peptide-bound 24-mer cage (PDB: 3CS0, *Right*) (17), are facilitated by PDZ1:PDZ2′ interactions. Key interfacial residues are shown as colored balls in each case. To better visualize the trimer:trimer′ interface in the magnified views, the protease′ and PDZ1′ domains from the partner trimer in the closed hexamer structure have been colored dark blue and red, respectively. The active site catalytic triad is shown as yellow balls. (*C*) The concentration of DegP in the periplasm of three different *E. coli* strains grown at 37 °C, estimated from Western blot analysis of whole-cell DegP concentrations (28). DegP concentrations were measured in both exponential (blue) and stationary (red) phases to examine protein expression under different bacterial growth conditions.

outside of the trimer. The protease/PDZ1/PDZ2 components of DegP are connected by flexible linkers which allow PDZ1 and PDZ2 to adopt different orientations relative to the protease domain (14). This flexibility, in turn, facilitates the formation of a number of bowl or cage-like architectures (22, 23). The specific type of structure that is formed has been suggested to depend on DegP's substrate occupancy (i.e., free or bound) and the properties of the bound client (22, 24, 25). For example, the crystal structure of a DegP "closed hexamer" has been solved in the substrate-free state (apo, Fig. 1 *B*, *Middle*). The "closed hexamer" is stabilized by intertrimer protease:protease' and PDZ1:PDZ1' interactions (prime indicates that the interacting domains derive from different trimers). This structure has been proposed as DegP's resting state, with the protease active sites sequestered within the interior of the hexamer, thereby suppressing undesired cleavage of clients (24). An additional "open hexamer" structure has also been solved, which features an exposed hexamer interior that is thought to allow for initial client engagement (24). Upon substrate binding, DegP has been shown to reorganize into 12- or 24-mer cages that are mediated by intertrimer PDZ1:PDZ2' interactions (Fig. 1 *B*, *Right*) (17, 22). These forms of DegP have increased proteolytic activity toward clients (25) and can additionally encapsulate cargo proteins such as OMPs for potential chaperoning (17).

Though DegP's substrate-engaged cage structures have been well characterized by crystallography, relatively little is known about the assembly mechanism of DegP particles in the absence of substrates under biological solution conditions. For example, could DegP trimers self-assemble into high-order oligomers in the apo state? If so, how are these oligomers related to the formation of canonical cages and DegP function? To explore these questions, we have studied DegP in apo and substrate-engaged states under solution conditions that approximate those within the periplasm. We have utilized a combination of hydrodynamics methodologies such as dynamic light scattering (DLS) and analytical ultracentrifugation (AUC) along with NMR approaches, including methyl transverse relaxation–optimized spectroscopy (methyl-TROSY), and cryogenic electron microscopy (cryo-EM) to inform on DegP's self-assembly mechanism. We find that in the absence of substrates, DegP forms an interconverting and broadly populated distribution of oligomers that is mediated by assembly of trimers. Unlike canonical DegP cages, our data suggest that these oligomers do not form the full complement of available PDZ1:PDZ2' interactions and are thus distinct structurally from cage conformers. Under stress conditions where DegP becomes overexpressed, the DegP ensemble shifts toward very large assemblies. This potentially enables the rapid incorporation of client proteins, leading to concomitant conformational changes to form canonical cages inside which substrates can be processed.

Results

DegP Adopts a Distribution of Oligomeric States in the Absence of Substrate Proteins. In this study, we investigated the self-assembly mechanism of DegP from *Escherichia coli*, which is known to form a number of cage-like oligomeric structures that depend on its substrate occupancy (25). In order to understand how these structures are assembled, either in the presence or absence of substrate, an estimate of the DegP periplasmic concentration must be obtained, as the extent of molecular self-assembly at equilibrium is dictated both by the strength of interprotomer interactions (association constants) and the total concentration of monomer units, M_T (26, 27). Periplasmic concentrations on

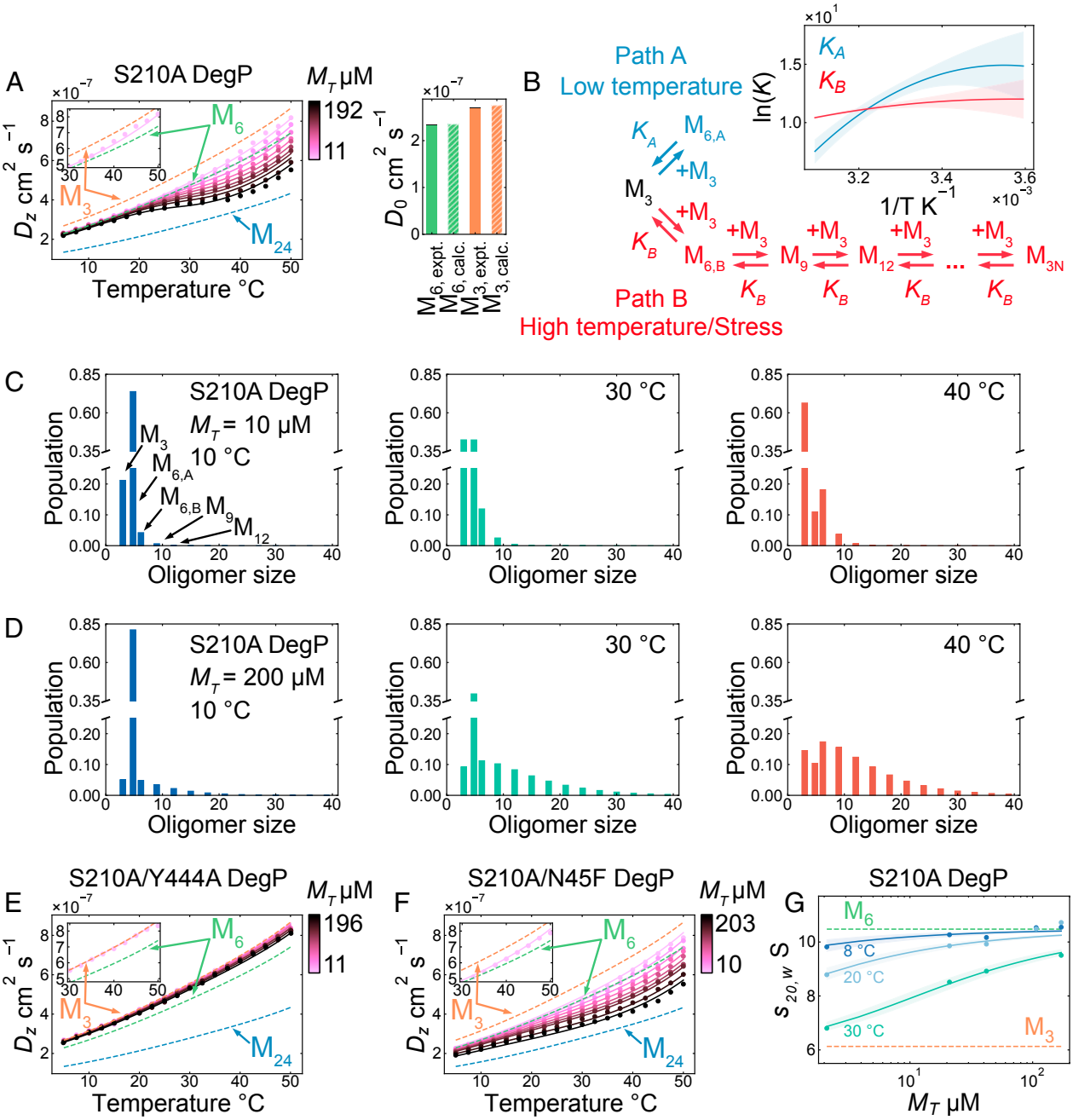

Fig. 2. DegP forms large oligomers in the absence of substrates. (*A*) 3D DLS dataset (*Left*) for refolded S210A DegP, with diffusion coefficients (D_z) measured as a function of total monomer concentration (M_T) and temperature. Experimental data from low to high M_T (11 to 192 μM) are shown as pink to black circles, respectively. Globally fitted curves using the model in *B* are shown as solid lines. Ideal diffusion coefficients (D_0) as a function of temperature for trimer (M_3, orange), hexamer (M_6, green), and 24-mer (M_{24}, blue) particles are shown as dashed lines. The line for the trimer was calculated by extrapolating the S210A/Y444A DegP D_z data in *E* to 0 M_T at 5 °C to obtain the trimer D_0 value and a hydrodynamic radius of 4.9 nm, from which trimer D_0 values were calculated over 5 to 50 °C using the Stokes–Einstein relationship (*SI Appendix*). The hexamer and 24-mer lines were calculated from the trimer D_0 value over 5 to 50 °C using a scaling law relating diffusion rate with particle size (*SI Appendix*, Eq. 15). The *Inset* shows experimental data at 11 μM M_T (pink), transitioning from the hexamer to trimer states at high temperature. Experimental D_0 values (*Right*) were estimated from the extrapolation of DLS data for S210A DegP and the S210A/Y444A trimer mutant to zero M_T at 5 °C (solid bars). Calculated D_0 values for the hexamer and trimer structures (PDB: 3MH6) (14) were obtaining by using HYDROPRO (34) and are shown for comparison (hatched bars). The 95% CIs calculated from Monte Carlo analysis (52, 53) of the experimental D_0 values are shown as black bars. (*B*) A self-assembly model for DegP in the absence of substrates. The maximum oligomer size, 3N, was set to 150 (N = 50) in the global fits shown in *A*, *E*, and *F* (*SI Appendix*). The model was globally fit to the data in *A*, and the resulting K_A and K_B values are shown in the Van 't Hoff plot (lines). The 95% CIs for each equilibrium constant were generated using Monte Carlo analyses (*SI Appendix*) and are shown as colored bands. (*C* and *D*) Mean distributions of fractional number of subunits within each S210A DegP oligomer at 10 (*C*) and 200 (*D*) μM M_T and 10, 30, and 40 °C extracted from Monte Carlo global fitting analyses of the model in *B* to the data in *A*. (*E*) 3D DLS dataset for the refolded S210A/Y444A DegP trimer mutant. (*F*) 3D DLS dataset for the refolded S210A/N45F DegP mutant. (*G*) 3D AUC dataset for S210A DegP where weighted sedimentation constants ($s_{20,w}$, corrected to 20 °C in water, colored points) were measured in an analogous manner to *A*. Global fits of the data to a trimer–hexamer model (path A only) are shown as colored lines. The 95% CIs generated using Monte Carlo analyses are shown as colored bands. Dashed lines for the trimer and hexamer *s*-values are indicated and calculated as described in *SI Appendix*.

the order of ∼40 to 100 μM were obtained during the exponential growth phase of three different strains of *E. coli*, with increases to ∼130 to 200 μM in the stationary phase (Fig. 1*C*), estimated from Western blot analyses (performed as described previously) (18, 28) of whole-cell DegP concentrations (*SI Appendix*). The increase in DegP concentration in the stationary phase is consistent with its overexpression in response to cellular stress conditions (9, 29).

Since DegP can be copurified with substrate molecules (17, 22) we next established a refolding purification protocol for isolating substrate-free protein as a starting point in elucidating DegP's self-assembly mechanism (*SI Appendix*). In addition, we worked with a mature form of DegP (residues 27 to 474, hereafter numbered from 1 to 448) where the catalytic serine was mutated to alanine (S210A) in order to abrogate autocatalysis. Notably, DegP preparations that have been refolded to remove substrates have previously been demonstrated to maintain activity (22), and we found that the oligomerization properties of refolded S210A DegP and highly purified, yet not refolded S210A DegP were identical (*SI Appendix*, Figs. S1 and S2), indicating that substrates are not bound to the S210A protein using our purification procedure, even without refolding. However, for certain DegP mutants, we do observe substrates in samples that are not refolded, as described in *Substrate Binding Remodels the DegP Ensemble toward Canonical Cage-Like Structures*. Therefore, all data presented herein, unless indicated otherwise, have been collected using refolded DegP in order to ensure that each sample is substrate free.

Having determined a biological concentration range over which to assay DegP oligomerization, we performed DLS measurements as a function of protein concentration and temperature (Fig. 2 *A*, *Left* and *SI Appendix*, Fig. S1). DLS measurements yield the z-average diffusion constant (D_z):

$$D_z = \frac{\sum_{i=1}^{N} M_i^2 c_i D_{0,i}}{\sum_{i=1}^{N} M_i^2 c_i}, \qquad [1]$$

where M_i and c_i are the molecular weight and molar particle concentration of the i^{th} diffusing species in solution, and $D_{0,i}$ is the corresponding diffusion constant for each state under ideal solution conditions (i.e., infinitely dilute concentration). The diffusion constants relate to the size of each species by the Stokes–Einstein relationship,

$$D_{0,i} = \frac{k_B T}{6\pi\eta r_i}, \qquad [2]$$

where k_B is the Boltzmann constant, T is the absolute solution temperature, η is the solution viscosity (assumed to be a function of temperature), and r_i is the hydrodynamic radius for a given species (assumed to be independent of temperature). In the case of a single diffusing species, $D_z = D_0$, and the particle hydrodynamic radius is readily estimated using Eq. **2**. For self-assembling systems such as DegP, the situation becomes more complicated, as multiple species can coexist in solution. Yet in these cases, D_z values provide a rich source of information on the oligomerization process as they are highly sensitive to trace amounts of large particles (30–32). Most importantly, the nature of an underlying oligomeric distribution can be read out in a model-free manner through variations in D_z values that are derived from changes in experimental variables such as M_T, temperature, or solution ionic strength. Given a set of thermodynamic equations that govern a self-assembling system combined with Eqs. **1** and **2**, the underlying concentrations for each molecular species that give rise to the measured D_z values can be extracted by numerical fitting of the experimental DLS dataset to an appropriate model (reference *SI Appendix*, *Materials and Methods*). This approach enabled a mapping of DegP self-assembly over a wide range of solution conditions and importantly, over temperature and concentration ranges where DegP is thought to act as a stress-protective protease and chaperone.

Three-dimensional (3D) DLS datasets (Fig. 2 *A*, *Left*) were collected on apo S210A DegP, measuring D_z as a function of M_T (∼10 to 200 μM) and temperature (5 to 50 °C) using a 384-well plate–format instrument (*SI Appendix*, Fig. S1). A 200 mM NaCl concentration was chosen for the majority of our experiments, close to the known periplasmic salt concentration range of ∼200 to 300 mOsm/L (33). Surprisingly, the DLS data so obtained were immediately suggestive of a complex temperature-dependent oligomerization process that occurs in the absence of bound substrates (Fig. 2 *A*, *Left*), with a number of key details about the nature of the self-assembly mechanism evident even without extensive fits of the experimental data. At low temperature, the D_z values are relatively insensitive to M_T, suggesting a monodisperse distribution of DegP particles. The D_0 value obtained by extrapolation to zero M_T at 5 °C is consistent with the expected D_0 value calculated for the canonical closed DegP hexamer structure using HYDROPRO (Fig. 2 *A*, *Right* and *SI Appendix*) (34). As the temperature is increased at low M_T (11 μM), D_z values (light pink line) initially vary as expected for a hexameric Stokes particle (Fig. 2 *A*, *Left*, *Inset*, green dashed line). However, at ∼30 °C, they begin to deflect upward from the hexamer line, indicating that the DegP hexamer transitions toward a smaller oligomeric species which we tentatively assigned to a DegP trimer based on the D_0 value predicted by HYDROPRO (Fig. 2 *A*, *Left*, *Inset*, and Fig. 2 *A*, *Right*). In contrast, for higher concentrations of M_T (e.g., 192 μM) the D_z values deflect downward from the hexamer line with increasing temperature (black points), as expected for the formation of larger Stokes particles (Fig. 2 *A*, *Left*, 24-mer blue dashed line). This suggests an additional oligomerization pathway where DegP spontaneously assembles into higher-order oligomers which dominate under these conditions (high temperature and M_T). Strikingly, the minimum D_z values in this dataset, where the DegP distribution is presumably shifted toward the largest oligomers, occur at biologically optimal growth temperatures of ∼30 to 40 °C. Though we have restricted our analyses to the biological concentration range of ∼10 to 200 μM M_T, we note that extending the concentration range to ∼1 mM M_T produces an extreme downward deflection in this temperature range that 1) is suggestive of DegP particles which are much larger than any known cage structures and 2) that DegP potentially oligomerizes in a continuous manner to form these states (compare Fig. 2*A* and *SI Appendix*, Fig. S1*C*). At temperatures above ∼40 °C and M_T = 192 μM, D_z values increase (Fig. 2 *A*, *Left*, black points), reflecting the dissociation of higher-order DegP oligomers. Irreversible DegP aggregation is not a factor in the DLS measurements as profiles obtained from a repeat experiment after the initial measurement over 5 to 50 °C and subsequent cooling were essentially identical, and fits of the rescanned dataset yielded similar parameters (*SI Appendix*, Fig. S3, see *SI Appendix* for fitting details). DegP oligomerization was additionally found to be equilibrated during the temperature ramp rate used in the experiments (∼0.25 °C min^{-1}, *SI Appendix*), as similar profiles and extracted parameters were obtained at a slower temperature scan rate of 0.03 °C min^{-1} (*SI Appendix*, Fig. S4).

Having established that our DLS profiles report on a system in thermodynamic equilibrium, we choose to describe the DegP self-assembly mechanism in terms of two oligomerization pathways, "path A" and "path B" (Fig. 2*B*). We assign path A as the route that forms the canonical hexameric DegP particle, which is stable at low temperatures where it sequesters the majority of trimer molecules. In path B, we assumed that DegP trimers associate into relatively weakly bound higher-order states that are

presumed to occur via intertrimer PDZ1:PDZ2' interactions (Fig. 1*B*). Although higher-order oligomers within path B are clearly observed in our DLS curves, we cannot establish their precise structural properties based on the DLS data alone. For this reason, we have chosen to describe path B in a purely phenomenological manner as a polymerization reaction (reference *SI Appendix* and Fig. 2 legend for details of the model) (26, 27). Only a pair of equilibrium constants are required for the system using this simple model.

The collection of D_z profiles was globally fit to the two-pathways model of DegP oligomerization (*SI Appendix*, Fig. 2*B*). We obtained good agreement between the experimental (points) and fitted (lines) D_z values across the entire temperature and M_T range (Fig. 2 *A*, *Left*). We note that some systematic deviation between the fitted and experimental curves was observed at high M_T, indicating that oligomerization within path B is more complicated than described here (see *Elucidating the DegP Self-Assembly Mechanism by Methyl-TROSY NMR*). Nevertheless, this relatively simple model (in terms of the number of fitted parameters) is sufficient to capture the essential features of the experimental profiles over the biological concentration range of DegP. In order to establish whether simpler models with a single equilibrium constant and fewer, lower-order oligomeric states might describe the data equally well, we additionally considered a single oligomerization pathway where hexamers are in equilibrium with 12- and 24-mer particles whose structures have been elucidated previously (17, 22, 24, 25). This model gave extremely poor fits to the data, and a slightly more complicated model including trimers also produced unsatisfactory fit results (*SI Appendix*, Fig. S5). These models could be further ruled out on the basis of subsequent analyses of DegP self-assembly by NMR (see *Elucidating the DegP Self-Assembly Mechanism by Methyl-TROSY NMR*). The thermodynamic parameters extracted from fits of the two-pathways model are reported in *SI Appendix*, Table S1, with K_i ($i \in (A, B)$) values for the oligomerization reactions $M_{3j,i} + M_3 \rightleftarrows M_{3(j+1),i}$ (j indicates the number of trimers in each oligomer; in path A, $j = 1$) shown in the form of a Van 't Hoff plot in Fig. 2*B*. Values for ΔC_p were included and optimized in the global fits, as the oligomerization reactions presumably involve the burial of large amounts of polar and nonpolar surface areas. The ΔC_p values so obtained are reasonable, differing by less than a factor of two from estimates using solvent-accessible surface areas (35, 36) calculated for the association of DegP trimers into the closed hexamer structure (ΔC_p for Path A) and the association of individual PDZ1 and PDZ2 domains into an isolated PDZ1:PDZ2 complex (ΔC_p for Path B; *SI Appendix*). The obtained association constants, K_A (path A) and K_B (path B), were used to generate distributions of the fraction of the total number of available subunits within each of the DegP oligomers as a function of M_T and temperature ($P_{3i} = \frac{3i[M_{3i}]}{\sum_{i=1}^{N} 3i[M_{3i}]}$, Fig. 2 *C* and *D*). As suggested by the model-free analysis of the raw DLS data (Fig. 2*A*) given above, $K_A > K_B$ at low temperature, leading to little polymerization of DegP (Fig. 2 *C* and *D*). As the temperature increases, K_A decreases, and eventually $K_A < K_B$, reflecting the large exothermic ΔH and negative ΔC_p associated with forming the path A hexamer. In contrast, each of the path B steps is approximately fivefold less exothermic than for path A, and the small temperature dependence of K_B ensures that it remains relatively constant over the measured temperature range. Thus, polymerization along path B occurs unabated at high temperature, while path A becomes depleted. Increasing M_T under these conditions drives DegP trimers down path B (Fig. 2*D*), leading to the observed deflections in D_z.

In order to further evaluate the robustness of our DLS data, we have studied two S210A DegP oligomerization mutants (Fig. 2 *E* and *F* and *SI Appendix*, Fig. S6 *A*–*C*), S210A/Y444A DegP and S210A/N45F DegP. The Y444A mutation has been shown to suppress oligomerization via the PDZ1:PDZ2' interface (Fig. 1 *B*, *Right*) (37), while the N45F mutation promotes oligomerization of DegP into cage structures by destabilizing protease:protease' interactions within the closed hexamer state (Fig. 1 *B*, *Middle*) (14). Notably, D_z values for S210A/Y444A DegP (Fig. 2*E* and *SI Appendix*, Fig. S6*B*) are shifted upward relative to those for S210A DegP (Fig. 2 *A*, *Left* and *SI Appendix*, Fig. S6*A*), indicating an ensemble that is skewed toward smaller particles relative to the S210A protein. The extrapolated D_0 value at 5 °C is also nearly identical to the calculated value for a DegP trimer (Fig. 2 *A*, *Right*). Importantly, the large temperature-dependent deflections in D_z recorded at high M_T for S210A DegP are not observed for the S210A/Y444A mutant, while at low M_T, the D_z values superimpose well on the curve for a trimeric Stokes particle throughout the experimental temperature range (Fig. 2*E* and *Inset*). In contrast to the Y444A mutation that suppresses path B, the DLS data of Fig. 2*F* and *SI Appendix*, Fig. S6*C* confirm that the N45F mutation increases the flux along path B at lower temperatures and decreased M_T values by effectively removing path A. Importantly, at high temperatures where the path A hexamer is not highly populated even in the S210A protein, the S210A and S210A/N45F DegP DLS datasets are essentially identical, indicating that the N45F mutation does not perturb interactions within path B. Taken together, our DLS results 1) for S210A DegP showing an initial dissociation of hexamers to trimers prior to formation of larger oligomeric species; 2) for S210A/Y444A DegP, whereby the Y444A mutation that destabilizes higher-order states prevents further oligomerization; and 3) for S210A/N45F DegP, whereby the closed hexamer is destabilized so that larger particles would be expected to form at lower temperatures, strongly support the two-pathways model of Fig. 2*B*.

In order to validate aspects of our model derived from DLS measurements, we collected a 3D AUC dataset on S210A DegP where sedimentation coefficients were measured as a function of temperature and M_T (Fig. 2*G*). Since AUC is less sensitive to the formation of large species compared to DLS, it was possible to focus on the trimer–hexamer equilibrium of path A over the M_T range of ~2 to 170 μM and for temperatures extending between 8 and 30 °C. The weighted sedimentation coefficients ($s_{20,w}$, corrected to 20 °C in water) determined as a function of M_T at 8, 20, and 30 °C clearly suggest a strongly exothermic association reaction, so that the hexameric DegP particle (Fig. 2*G*, green dashed line) transitions to a smaller species as the temperature is increased (compare blue and green curves), especially for low M_T values. The smaller species was determined to be a trimer based on an independent AUC study of the S210A/Y444A mutant (Fig. 2*G*, orange dashed line). The AUC data were globally fit to a simple trimer–hexamer association model (i.e., path A only), and the thermodynamic parameters for path A obtained from this analysis are in agreement with the values extracted from DLS fits (*SI Appendix*, Table S1), strongly supporting the formation of highly stable hexameric DegP particles at low temperature which readily dissociate as the temperature is increased.

Characterizing Interactions of DegP's Oligomerization Domains by NMR. The DLS analysis of the S210A/Y444A DegP trimer mutant described in the previous section establishes that Y444, located at the C-terminal end of the PDZ2 domain, plays an important role in the formation of substrate-free oligomeric particles (compare Fig. 2 *A* and *E*). Y444 is localized to the PDZ1:PDZ2' interface in substrate-bound crystal structures of DegP cages (Fig. 1 *B*, *Right*) (17, 22), suggesting that PDZ1:PDZ2' contacts might be critical to the formation of oligomers in the absence of substrate as well. We initially characterized the interactions between isolated PDZ domains by titrating uniformly (U)-^{15}N,^{13}C PDZ1 (100 μM) with unlabeled PDZ2 (0 to 5 mM). Amide backbone and methyl side-chain chemical shift perturbations (CSPs), defined as differences in peak positions in the absence or presence of 5 mM PDZ2, are plotted in Fig. 3*A*. Residues with large CSPs in either amide or

methyl correlation spectra such as L276, M280, and F289 have all been previously implicated in regulating DegP cage formation (37), and the PDZ1:PDZ2 interaction surface mapped by NMR is in excellent agreement with the PDZ1:PDZ2' interface in the substrate-engaged 24-mer cage crystal structure (Fig. 3*B*) (17) and other known cage architectures (22, 25). Many of the PDZ1 correlations titrated linearly with addition of PDZ2 with no changes in line widths so that the resulting CSP isotherms could be analyzed by assuming a two-state fast exchange binding model (*SI Appendix*). The isotherms were well described by this model from which fitted PDZ1:PDZ2 association constants of $(2.3 \pm 0.3) \times 10^3$ M^{-1} and $(1.4 \pm 0.2) \times 10^3$ M^{-1} were obtained at 25 and 50 °C, respectively, along with a ΔH for the PDZ1:PDZ2 domain association of approximately −15 kJ mol^{-1} (assuming $\Delta C_p = 0$, Fig. 3 *C* and *D*). Estimates of association (k_{on}) and dissociation (k_{off}) rates for the PDZ1:PDZ2 interaction at 25 and 50 °C were obtained by analysis of peak line shapes for a number of PDZ1 correlations with large shift differences between unbound and PDZ2-bound states (Fig. 3 *E* and *F*) (38). The association rate constants were diffusion limited at both temperatures ($k_{on} = 8.5 \pm 0.1 \times 10^6$ $M^{-1} \cdot s^{-1}$ and $3.9 \pm 0.2 \times 10^7$ $M^{-1} \cdot s^{-1}$ at 25 and 50 °C, respectively), while the dissociation rate constants were also rapid, in keeping with the weak PDZ1:PDZ2 interaction affinities ($k_{off} = 3.8 \pm 0.1 \times 10^3$ s^{-1} and $3.0 \pm 0.1 \times 10^4$ s^{-1} at 25 and 50 °C, respectively). Notably, $K_A = k_{on}/k_{off}$ values ($(2.2 \pm 0.4) \times 10^3$ M^{-1} and $(1.3 \pm 0.3) \times 10^3$ M^{-1} at 25 and 50 °C, respectively) were in close agreement with association constants obtained from fits of

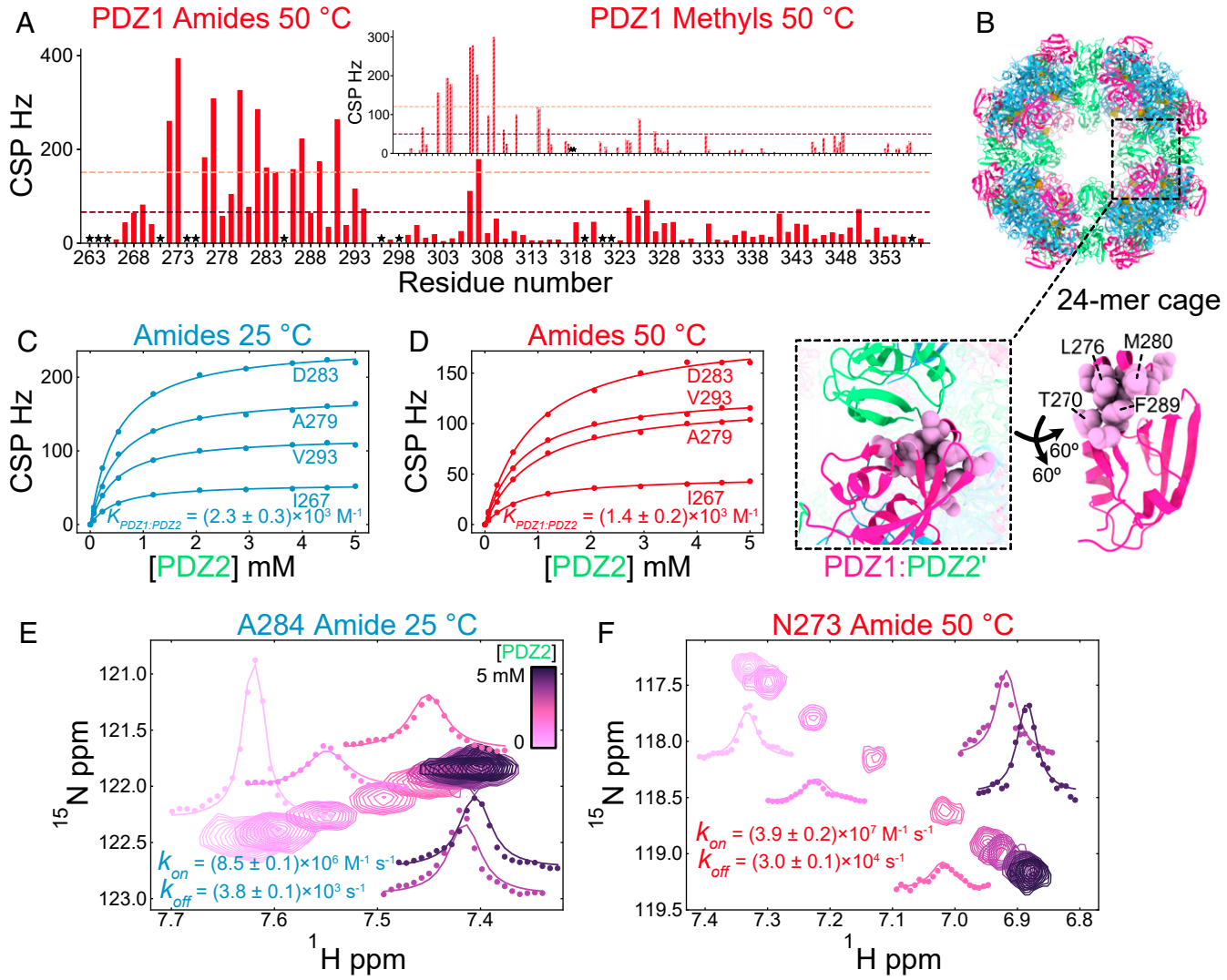

Fig. 3. Characterizing PDZ1 and PDZ2 domain interactions by solution NMR. (*A*) CSP plots from titration of U-^{15}N,^{13}C PDZ1 with unlabeled PDZ2 for PDZ1 amide and methyl groups (*Inset*) at 50 °C and 225 mM NaCl. Significant CSPs were taken to be 1 SD above (light orange dashed line) the mean value (dark red dashed line). Stars indicate residues that were unassigned, either as a result of broadening/overlap or their position adjacent to a proline residue. (*B*) The PDZ domain interactions measured by solution NMR are those predicted from the crystal structure of the peptide-bound DegP 24-mer cage (PDB: 3CS0) (17). PDZ1 residues with significant CSPs from *A* (>1 SD from mean) are at the PDZ1:PDZ2' interface in the crystal structure; residues with significant amide CSPs are shown as pink balls in the PDZ1 structure. (*C* and *D*) Binding isotherms for the PDZ1 amide groups titrated with PDZ2 at 25 and 50 °C. Each isotherm at a given temperature was individually fit to a two-state $PDZ1 + PDZ2 \rightleftarrows PDZ1:PDZ2$ model and the association constants reported as the mean ± SD. (*E* and *F*) Superposition of amide peaks for A284 and N273 of PDZ1 (100 μM) from ^{15}N-^{1}H HSQC spectra at 25 (*E*) and 50 (*F*) °C, 800 MHz, as a function of added PDZ2. Correlations in two-dimensional spectra were fit to the two-state binding model reported above using TITAN (38) with the association (k_{on}) and dissociation (k_{off}) rate constants indicated. The extracted $K_{PDZ1:PDZ2}$ association constants are $(2.2 \pm 0.4) \times 10^3$ M^{-1} and $(1.3 \pm 0.3) \times 10^3$ M^{-1} at 25 and 50 °C, respectively. Experimental one-dimensional line shapes and their fits are shown as colored points and lines, respectively.

binding isotherms for residues in the fast exchange regime (Fig. 3 *C* and *D*).

Elucidating the DegP Self-Assembly Mechanism by Methyl-TROSY NMR. Having established "signatures" for the PDZ1:PDZ2 binding interaction, we next performed a titration of S210A DegP as a function of protein concentration (25 μM to 2 mM M_T) to explore whether intertrimer contacts between PDZ1 and PDZ2' domains could be responsible for the formation of the oligomeric structures identified in our DLS analysis. Each DegP trimer has a molecular mass of ∼140 kDa, and multiple copies of the trimer associate to form larger particles. We have therefore prepared highly deuterated protein, with $^{13}CH_3$ labeling at Ile-δ, Leu-δ, Val-γ, and Met-ε positions, where only one of the pair of isopropyl methyls of Leu and Val was $^{13}CH_3$ labeled (referred to as U-^{2}H, ILVM-$^{13}CH_3$ labeling in what follows) and recorded ^{13}C-^{1}H heteronuclear multiple quantum coherence (HMQC) spectra that exploit the methyl-TROSY principle (39, 40). We chose a high temperature (50 °C) and [NaCl] = 200 mM for obtaining the DegP concentration series, as these conditions favor path B assembly (Fig. 2), and focused on the spectral positions of PDZ1 domain methyl peaks that had previously been analyzed in the context of the isolated PDZ1:PDZ2 interaction (Fig. 3). Notably, we observed PDZ1 titration profiles involving the same set of methyl groups that showed CSPs in the titration of the isolated PDZ1 and PDZ2 domains, with very similar shift trajectories in both cases (Fig. 4*A* compare corresponding *Left* and *Right* panels). Furthermore, the methyl groups with large CSPs were those that make contacts at PDZ1:PDZ2' interfaces of substrate-loaded DegP cages (Fig. 4*B*). Our data establish, therefore, that the formation of substrate-free oligomers along path B occurs by the same PDZ1:PDZ2' interactions that stabilize the substrate-bound DegP cages observed previously. Notably, the S210A DegP titration profiles are consistent with a rapid interconversion between free PDZ1 and PDZ1:PDZ2' bound states on the NMR chemical shift timescale, as observed for the individual PDZ domains (Fig. 4*A*). This presumably occurs in the context of DegP oligomers by the rapid exchange of trimers between different particles and/or transient interactions

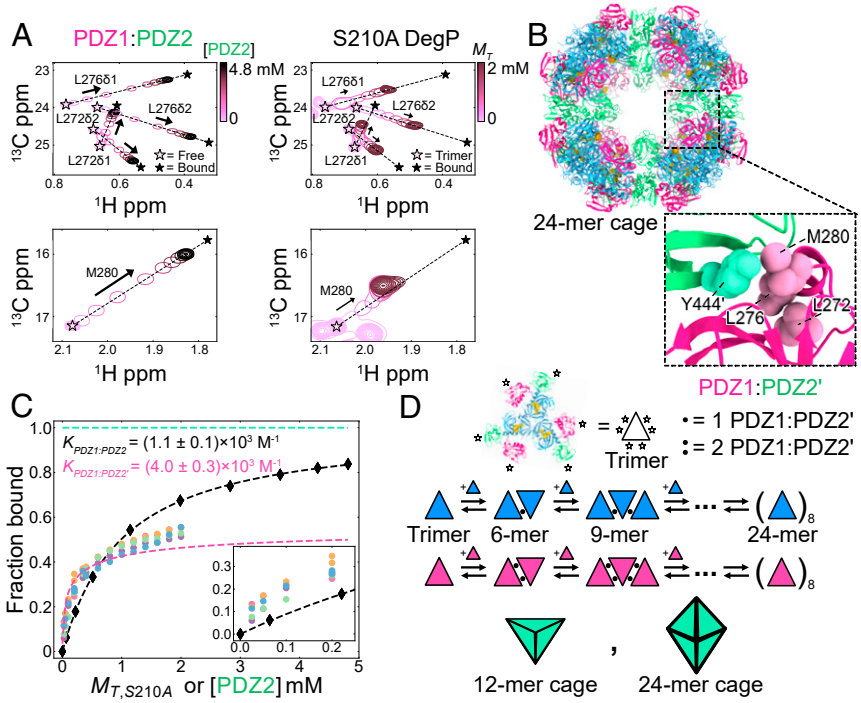

Fig. 4. Exploring the mechanism of DegP self-assembly in the absence of substrates by methyl-TROSY NMR. (*A*) Overlays of selected regions of ^{13}C-^{1}H HMQC spectra from the titration of U-^{2}H, ILVM PDZ1 (100 μM) with U-^{2}H PDZ2 from 0 to ∼4.8 mM (*Left*) and from the concentration series of U-^{2}H, ILVM S210A DegP (25 μM to 2 mM, *Right*). The residues with the largest methyl CSPs, L272, L276, and M280, are indicated. Unbound and bound endpoints are given by pink and black stars, respectively; bound endpoints were obtained from global fits of the PDZ-binding isotherms. Unbound chemical shifts for S210A DegP were taken as the shifts measured in a spectrum of the S210A/Y444A trimer mutant at 100 μM M_T. Full contours are shown for only the first and last titration points in each case. Peaks that are in the vicinity of L272, L276, and M280 have not been shown for clarity. L272 correlations have additionally been plotted at a lower contour level than those for L276 in the S210A DegP titration for better visualization. Spectra were collected in 200 mM NaCl at 50 °C, 800 MHz. (*B*) Residues with large CSPs identified in the NMR concentration series of U-^{2}H, ILVM S210A DegP match those involved in the PDZ1:PDZ2' interface in the substrate-bound 24-mer crystal structure (PDB: 3CS0) (17). (*C*) Fraction-bound profiles for the titrations in *A*. The isolated PDZ titration isotherm (black diamonds) is obtained by averaging over the profiles for individual L272, L276, and M280 methyl groups and fit to a two-state *PDZ*1 + *PDZ*2⇄*PDZ*1 : *PDZ*2 model (black dashed line) with an association constant, $K_{PDZ1:PDZ2}$, (1.1 ± 0.1) × 10^3 M^{-1} for the deuterated PDZ domains (in D_2O). Note that this value differs slightly from $K_{PDZ1:PDZ2}$ = (1.3 ± 0.3) × 10^3 M^{-1}, obtained from measurements of fully protonated PDZ domains in H_2O solvent. The S210A DegP isotherms are shown as circles, color coded to distinguish each methyl (L272δ1 in orange, L272δ2 in pink, L276δ1 in green, L276δ2 in purple, and M280 in blue). The pink dashed line is the fit of the pink self-assembly model in *D*, with the extracted $K_{PDZ1:PDZ2'}$ indicated (*SI Appendix*). Differences between $K_{PDZ1:PDZ2}$ and $K_{PDZ1:PDZ2'}$ may reflect the cooperative association of the second pair of PDZ domains in the context of the intact trimer. The green dashed line is the fraction PDZ1 bound for an ensemble comprising canonical 12- and 24-mer structures (all bound). The *Inset* highlights the deviation between the PDZ1:PDZ2 and S210A DegP titrations at low protein concentrations. (*D*) Models of DegP self-assembly with trimers depicted as triangles (each PDZ domain denoted by a white star) that assemble via their edges into linear particles with either one (blue) or two (pink) PDZ1:PDZ2' interactions per pair of trimers (denoted by black circles). Based on X-ray–derived structures, each PDZ domain from a trimer (PDZ1 or PDZ2) can, in principle, contact a domain on an adjacent trimer (either PDZ2' or PDZ1') to form PDZ1:PDZ2' or PDZ2:PDZ1' interactions (17); in the pink model, PDZ1:PDZ2' and PDZ2:PDZ1' contacts (one each) are made for each pair of trimers. Canonical DegP structures, where all contacts are formed (each trimer's three PDZ1(PDZ2) domains are bound to PDZ2'(PDZ1') domains of adjacent trimers) are shown in green.

between proximal PDZ1 and PDZ2' domains within flexible oligomeric structures. Consistent with the assumption of fast interconversion, the fraction-bound values calculated from the five DegP methyl probes used are in good agreement with each other despite chemical shift differences between bound and free states that vary twofold between methyl groups (Fig. 4*C*, circles, where profiles from each methyl group are in a different color). From the experimental CSP values, a rate of interconversion, k_{ex} ($=k_{PDZ1,bound} + k_{PDZ1,free}$) > 1,800 s^{-1} can be established, and simulations using simple exchanging systems indicate that exchange rates greater than several thousand/sec only are consistent with our data and, further, that exchange between DegP trimers and larger oligomers with higher transverse relaxation rates skew the resulting average chemical shifts very little and hence introduce little error to the calculation of the fraction of PDZ1 domains that are bound to PDZ2' (*SI Appendix*, Fig. S7). The fraction of PDZ1 domains bound to their PDZ2' partners for apo S210A DegP can therefore be calculated straightforwardly from peak positions, as was done for the titration involving isolated PDZ domains (*SI Appendix*). Fraction bound values versus M_T (S210A DegP, colored circles) or [PDZ2] (titration of individual PDZ1 and PDZ2 domains, black diamonds) are plotted in Fig. 4*C*. Note that in the physiological protein concentration range (M_T ~25 to 200 μM), the DegP titration curve increased steeply (Fig. 4 *C*, *Inset*), with a more gradual increase for the profile derived from mixing individual PDZ domains.

Although it is not possible to obtain detailed structural descriptions of the oligomers that populate path B from our titration data alone, loose limits on the number of PDZ1:PDZ2' interactions within the ensemble of conformers can be set and differences between ensembles in the absence and presence of substrate established. Fig. 4*D* highlights a number of simple path B models in which DegP trimers (denoted by blue or pink triangles) associate into linear extended conformations in which each pair of trimers are held together by either one or two PDZ1:PDZ2' interactions (black circles between triangles in the blue and pink schemes in Fig. 4*D* respectively, *SI Appendix*). These microscopic interactions that are reported by the DegP concentration series (Fig. 4*A*) are explicitly counted in these models, unlike for the phenomenological model used to fit the DLS data where only molecular size is considered (*SI Appendix*). In contrast to the linear extended models illustrated in Fig. 4*D*, all possible PDZ1:PDZ2' interactions are made in the cage structures (composed of green trimer triangles, Fig. 4*D*, with two PDZ1:PDZ2' interactions at every edge) as established by previous structural studies (17, 22). Thus, for blue and pink linear 24-mers and a green 24-mer cage, respectively, ~29, ~58, and 100% of the potential 24 PDZ1:PDZ2' interactions are established. The simple blue and pink pathways depicted in Fig. 4*D* lead to mathematically tractable models that generate simple pictures as to the extent of PDZ1:PDZ2' interactions (*SI Appendix*). Since the model in which pairs of trimers are held together by a single PDZ1:PDZ2' interaction (Fig. 4 *D*, blue) cannot describe the fraction bound profiles over the complete concentration range (as 29% is the maximum fraction bound achievable) and the green scheme consisting exclusively of cage structures (100% fraction bound) overestimates the experimental data, we therefore fit the DegP isotherms to a pathway in which two PDZ1:PDZ2' interactions are formed per pair of trimers (Fig. 4 *D*, pink), allowing the PDZ domain association equilibrium constant to float (pink dashed line). A reasonable fit was obtained, extending over the physiological range of M_T values (<200 μM) until ~1.2 mM ($K_{PDZ1:PDZ2'} = (4.0 \pm 0.3) \times 10^3$ M^{-1}). Our NMR data demonstrate that the PDZ1:PDZ2' interactions are highly dynamic in the context of apo-DegP, that the full complement of PDZ1:PDZ2' contacts where each PDZ domain is bound, observed in substrate-bound cage structures, is not formed, and that it is possible to achieve reasonable fits of the DegP isotherms in the physiological (and somewhat higher) protein concentration range assuming a simplistic linear association model with two PDZ1:PDZ2' interactions connecting each pair of trimers. We note that, in reality, it is possible for trimers to associate in many more ways than highlighted here, leading to different structural configurations such as branched conformers, bowls (23), or partly assembled cages that could exchange rapidly with each other. In addition, as the linkers connecting protease–PDZ1 and PDZ1–PDZ2 domains are flexible (22), it is likely that the PDZ1 and PDZ2' domains at a given interface can exchange between free and bound conformations within the context of an individual oligomeric species. Such models are difficult to formulate and are beyond the scope of our study here.

Substrate Binding Remodels the DegP Ensemble toward Canonical Cage-Like Structures. Having established that apo-DegP assembles into oligomeric structures in which interactions between PDZ1 and PDZ2' domains are not as extensive as in canonical cages, we next sought to understand how apo-DegP free energy landscapes would be influenced by the binding of substrates that result in cages with maximal PDZ contacts (17). Here, we have used diffusion values measured by DLS as sensitive reporters of changes to the self-assembly process (Fig. 5 *A*–*C* and *SI Appendix*, Fig. S8). During the course of this study, we found that increasing concentrations of NaCl, as for M_T (Fig. 2*A*), can also shift the apo-DegP ensemble toward larger structures, particularly at physiological temperatures of ~30 to 40 °C, as evidenced in Fig. 5*A* where D_z values have been measured in 0 to 1,000 mM NaCl at 200 μM M_T. The lowered rates of diffusion do not reflect increased solution viscosity due to added salt, as these contributions have been accounted for with plotted D_z(0 NaCl) corresponding to the "corrected" diffusion constant at 0 mM NaCl (see Fig. 5 legend). The DLS experiment of Fig. 5*A* was repeated with the addition of a small protein substrate, the DNA binding domain of the human telomere repeat binding factor (hTRF1, 54 amino acids) (41, 42) that was added in twofold excess over the concentration of DegP (Fig. 5*B*). hTRF1 is a marginally stable protein ($\Delta G(35\ °C) = -8.0$ kJ · mol^{-1}) and was shown previously to be an Hsp70 substrate (41). We reasoned, therefore, that DegP might also recognize hTRF1 given that it can interact with unfolded or partially unfolded clients (22). Notably, hTRF1 was found to shift the salt-dependent DegP distribution profiles in Fig. 5*A* so that they became nearly perfectly superimposed over all salt concentrations (0 to 1,000 mM), with the D_0 value expected for a single diffusing species corresponding to a 12-mer particle (Fig. 5 *B*, *Left*). Identical results were obtained when using ensembles prepared with different protein concentrations and at a fixed NaCl concentration of 200 mM (*SI Appendix*, Fig. S8). In order to obtain a structural model of the DegP:hTRF1 complex, we recorded a cryo-EM dataset of the sample, obtaining 62,000 final particle images. A 4.4-Å structure was generated of a tetrahedral DegP 12-mer (Fig. 5 *B*, *Right* and *SI Appendix*, Fig. S9 *A*, *B*, *E*, *F*). The four trimers in this structure are associated via 12 PDZ1:PDZ2' interactions, with their protease domain active sites pointing toward the interior of the cage, as noted in previous structural studies of DegP cages (17, 22, 25). Together, these measurements suggest that the addition of a substrate to the apo-DegP ensemble results in a significant reorganization of the energy landscape such that a single predominant conformation is obtained.

As hTRF1 is not a natural DegP substrate, we next sought to test whether binding of substrates from *E. coli* would have a similar effect on collapsing the wide distribution of apo-structures to one or two bound conformers, as for hTRF1. To this end, we purified a sample of S210A/N45F DegP that was not refolded to remove bound substrates. The N45F mutant was chosen for these experiments as it has a greater propensity to oligomerize through path B than the S210A protein (compare Fig. 2 *A* and *F*) and thus

potentially may more readily bind substrate. The DLS data acquired on this sample also suggested a single diffusing species, although with a size corresponding to an 18-mer DegP particle (Fig. 5 *C*, *Left*). The resultant 3D structure determined by cryo-EM was of a trigonal bipyramidal 18-mer cage (93% of the 30,000 final particles, Fig. 5 *C*, *Right* and *SI Appendix*, Fig. S9 *C–F*). Importantly, this structure is consistent with the 18-mer suggested on the basis of AUC measurements of DegP in the presence of substrate (22). In addition to the 18-mer, a small fraction of structures (~7%) was of the canonical 24-mer DegP cage. Notably, each of these architectures features the full complement of PDZ1:PDZ2' interactions, as observed for the 12-mer.

Discussion

DegP is a key stress-protective protease and chaperone in the periplasm of gram-negative pathogens such as *E. coli* and *Pseudomonas aeruginosa* (8, 43–45) that plays an important role in bacterial proteostasis and in the trafficking of bacterial virulence factors from the periplasm to outside the cell (16, 46). The binding of substrate proteins to DegP is known to promote assembly into cages, yet little is known about the structural ensemble formed by apo-DegP and how this might influence substrate binding. As DegP plays a key role in bacterial pathogenesis, an understanding of its structural dynamics and how these influence function could ultimately be important for the development of novel antibiotics that inhibit bacterial virulence.

Here, we have used a suite of biophysical methods to inform on the distribution of structures populated by DegP in the absence of substrate and how this distribution changes upon substrate binding, providing insight into how DegP can readily bind substrates of variable sizes. DLS data reporting on DegP self-assembly in the apo state can be well described using a model consisting of

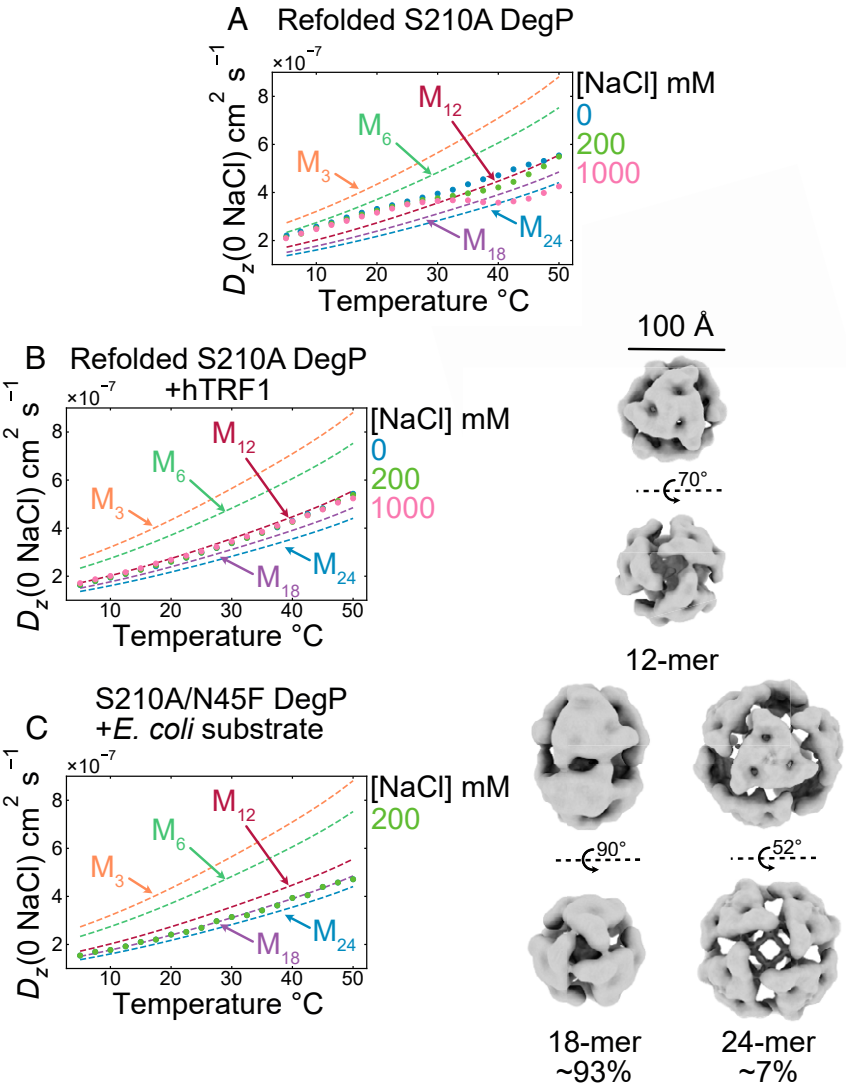

Fig. 5. Substrate binding remodels the distribution of DegP oligomers. (*A*) 3D DLS dataset for refolded S210A DegP, where D_z values have been measured as a function of temperature and [NaCl], M_T = 200 μM. D_z values at each salt condition ($D_{z,i}$) have been rescaled to [NaCl] = 0 mM ($D_{z,i}$(0 NaCl)) according to $D_{z,i}(0\ \mathrm{NaCl}) = D_{z,i}\frac{\eta_i(T)}{\eta_0(T)}$, where $\eta_i(T)$ is the viscosity at a given salt concentration, [NaCl]$_i$, and $\eta_0(T)$ is the viscosity for [NaCl] = 0 mM. Dashed colored lines are D_0 values for the trimer (orange), hexamer (green), 12-mer (red), 18-mer (purple), and 24-mer (blue) particles. (*B*) 3D DLS data for refolded S210A DegP collected under the same conditions as *A* in the presence of 400 μM hTRF1 substrate (*Left*). The corresponding 12-mer DegP cage structure solved by cryo-EM in 200 mM NaCl is shown to the *Right*. (*C*) DLS data for the S210A/N45F DegP mutant, where the sample has not been refolded to remove substrate proteins during purification. Data were collected at 100 μM M_T and 200 mM NaCl (*Left*). Structures of an 18-mer cage and the canonical 24-mer cage from cryo-EM analysis are shown to the *Right*, along with their relative particle abundancies.

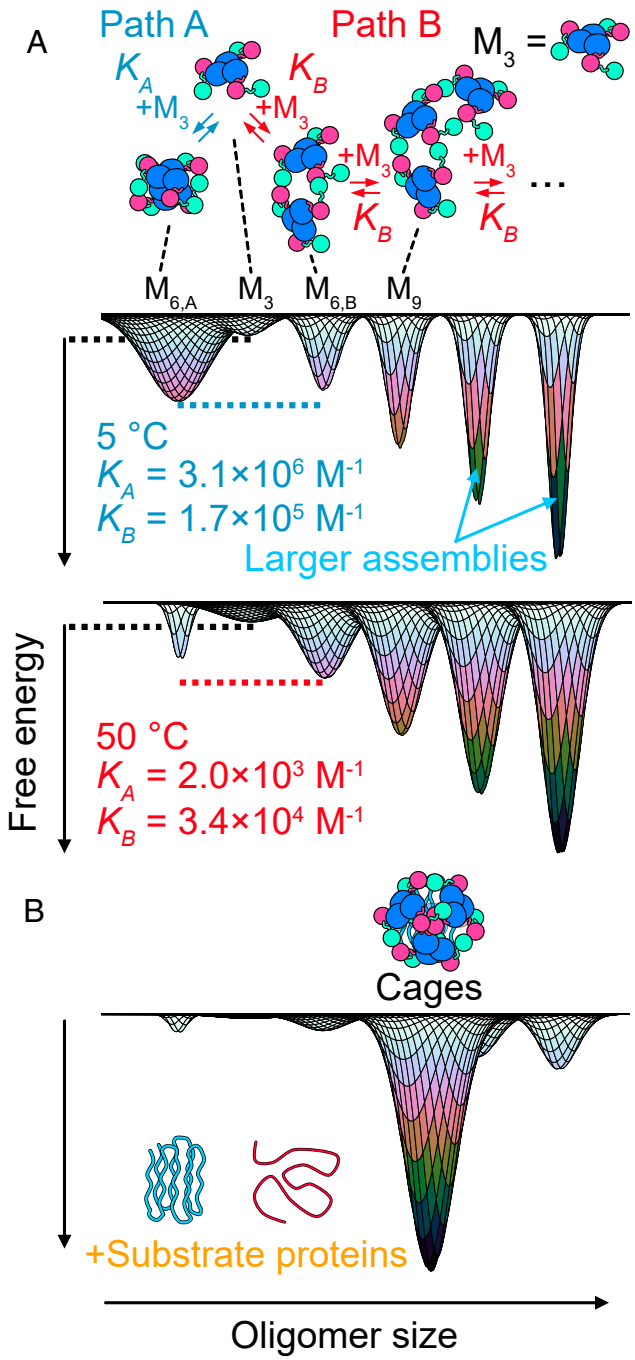

Fig. 6. A free energy landscape governing apo-DegP oligomerization and substrate binding. (*A*) The DegP energy landscape in the absence of substrates contains many wells of relatively similar energies, leading to a broadly populated distribution of assemblies at physiological protein concentrations and temperatures, whereby trimers associate via PDZ1:PDZ2' interactions. To visualize these landscapes, we represent wells in terms of two-dimensional Gaussian functions. The trimer minimum was set as the zero energy reference point, from which the heights of all other wells were calculated using the association constants from DLS global fitting as $G_{3j,i} = -(j-1)RT\ln(K_i)$, where $i \in$ (*A* and *B*), and j is the number of trimers in a given oligomer (two for the path A hexamer and a maximum of five for the truncated path B shown here). The path B wells at 5 and 50 °C appear similar in depth as a result of the weak temperature dependence of K_B. The concentration of particles within a well is determined by a combination of the well height and width (not to scale); for a given height wells with larger widths can accommodate more particles. The substantial variation in relative free energies of path A and B hexamers determines which pathway dominates. For example, at 5 °C, the difference in free energies for $M_{6,A}$ and $M_{6,B}$, $\Delta G_{6,A;6,B} = G_{6,A} - G_{6,B} = -6.7$ kJ mol^{-1} so that path A is favored, while at 50 °C, $\Delta G_{6,A;6,B} = 7.6$ kJ mol^{-1} so that path B becomes favored. (*B*) The binding of substrate proteins remodels the DegP energy landscape by stabilizing one or more canonical cage-like structures.

two competing oligomerization pathways (26, 27) in which particles grow through the binding of trimer building blocks (Fig. 2 *B*, paths A and B). Each of the two pathways in the model is parameterized in terms of a single equilibrium constant and ΔH and ΔC_p values that take into account temperature-dependent changes to the equilibria (Fig. 2*B*).

Our biophysical data paint a picture whereby the canonical hexameric DegP conformer in path A serves as the gate keeper for the formation of oligomers along path B. At low temperatures, the energy well for the path A hexamer is much deeper than for the trimer and the path B hexamer (Fig. 6 *A*, *Top*) so that the path A equilibrium is shifted far toward the formation of hexamers, with only a small fraction of trimers present. This, in turn, significantly limits the population of large oligomers along path B that are formed via trimer building blocks. At higher temperatures, the relative population of path A hexamers and trimers is shifted, with the increased concentration of trimers favoring path B oligomers (Fig. 6 *A*, *Bottom*). These findings are consistent with DegP's known role as a temperature-sensitive (6, 47, 48) and stress-protective protease-chaperone (12, 13, 21).

NMR titration studies of isolated PDZ domains in concert with concentration-dependent changes to the spectra of DegP (Fig. 4) point to an important role for PDZ1:PDZ2' interactions in driving oligomerization in the absence of substrate, as observed in structures of substrate-bound cages (17, 22, 25). A significant difference, however, is that the full complement of PDZ1:PDZ2' contacts is not formed in the apo state, suggesting that the apo conformers are more plastic and less stable than their cage counterparts, where all of the interactions are in place. In this context, it is of interest to note that discernible, well-defined structures for apo-DegP were not observed via cryo-EM, unlike for substrate-bound DegP. Notably, the fitted thermodynamic parameters for path B are consistent with the importance of PDZ1:PDZ2' interactions as well. For example, the fitted ΔH for each trimer addition along path B, -25 ± 15 kJ · mol^{-1}, is 1.7-fold more exothermic than the corresponding value obtained for binding isolated PDZ1 and PDZ2 domains (-15 ± 6 kJ · mol^{-1}); this difference likely reflects the fact that more than a single PDZ1:PDZ2' contact is made for each trimer as it is incorporated into the apo-DegP ensemble, as established by our NMR titration data. The path A association reaction, involving formation of the putative canonical hexamer, occurs through extensive protease:protease' and PDZ1:PDZ1' interactions (24), resulting in a more exothermic process ($\Delta H = -112 \pm 27$ kJ · mol^{-1}) than for path B reactions.

Molecular recognition events are most frequently modeled in terms of lock-and-key, induced fit, or conformational selection mechanisms (49–51). Flexibility is a critical component for both induced fit and conformational selection, as the receptor must be able to change structure upon ligand binding (induced fit) or populate distinct states, including the conformer that is selected by the ligand (conformational selection). Our results indicate that while conformational flexibility at the level of both apo-DegP trimers and polydisperse, higher-order assemblies is at the core of substrate recognition by DegP, the DegP–client interaction is more complex than that described by any one mechanism. The broad distribution of DegP particles at physiological temperatures and protein concentrations may lead to the facile accommodation of a variety of substrate sizes; however, these apo particles are not structurally identical to the cages that result from substrate binding. Binding of a given substrate to trimers or larger particles that comprise the apo-DegP ensemble results in structural changes to

form a canonical cage conformation (Figs. 5 and 6*B*). Our work thus provides an additional dimension to the classical DegP structure–function paradigm wherein cages are thought to form by reorganization of a resting hexameric state upon substrate engagement (25). Rather, even in the apo state, ensembles with broad distributions of particle sizes are formed, and these may well facilitate the substrate-binding process.

Materials and Methods

Details of protein expression and purification and all experiments, along with data fitting, are provided in *SI Appendix*.

Data Availability. All relevant data are included in the paper and *SI Appendix*.

ACKNOWLEDGMENTS. This research was funded by a grant from the Canadian Institutes of Health Research (CIHR), Grant No. FDN-503573 (L.E.K.) and supported by the Intramural Research Program of the National Institute of Biomedical Imaging and Bioengineering, NIH. Dr. John Rubinstein (Hospital for Sick Children, Toronto) is thanked for facilitating the use of a Tecnai F20 microscope. We thank Ashok Sekhar and Reid Alderson for providing computer programs to simulate free energy landscapes. R.W.H. is grateful to CIHR for a postdoctoral fellowship, and Y.T. was supported by a Japan Society for the Promotion of Science Overseas Research Fellowship and a Uehara Memorial Foundation post-doctoral fellowship. L.E.K. acknowledges a Canada Research Chair.

1. F. D. Schramm, K. Schroeder, K. Jonas, Protein aggregation in bacteria. *FEMS Microbiol. Rev.* **44**, 54–72 (2020).
2. H. Olzscha *et al*., Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. *Cell* **144**, 67–78 (2011).
3. E. Maisonneuve, B. Ezraty, S. Dukan, Protein aggregates: An aging factor involved in cell death. *J. Bacteriol.* **190**, 6070–6075 (2008).
4. S. Gottesman, S. Wickner, M. R. Maurizi, Protein quality control: Triage by chaperones and proteases. *Genes Dev.* **11**, 815–823 (1997).
5. T. Clausen, C. Southan, M. Ehrmann, The HtrA family of proteases: Implications for protein composition and cell fate. *Mol. Cell* **10**, 443–455 (2002).
6. C. Spiess, A. Beil, M. Ehrmann, A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. *Cell* **97**, 339–347 (1999).
7. Y. Toyama, R. W. Harkness, T. Y. T. Lee, J. T. Maynes, L. E. Kay, Oligomeric assembly regulating mitochondrial HtrA2 function as examined by methyl-TROSY NMR. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2025022118 (2021).
8. B. Lipinska, M. Zylicz, C. Georgopoulos, The HtrA (DegP) protein, essential for *Escherichia coli* survival at high temperatures, is an endopeptidase. *J. Bacteriol.* **172**, 1791–1797 (1990).
9. B. Lipinska, S. Sharma, C. Georgopoulos, Sequence analysis and regulation of the htrA gene of *Escherichia coli*: A σ^{32}-independent mechanism of heat-inducible transcription. *Nucleic Acids Res.* **16**, 10053–10067 (1988).
10. A. Karyolaimos *et al*., Enhancing recombinant protein yields in the *E. coli* periplasm by combining signal peptide and production rate screening. *Front. Microbiol.* **10**, 1511 (2019).
11. J. Skórko-Glonek *et al*., The N-terminal region of HtrA heat shock protease from *Escherichia coli* is essential for stabilization of HtrA primary structure and maintaining of its oligomeric structure. *Biochim. Biophys. Acta* **1649**, 171–182 (2003).
12. J. Skórko-Glonek *et al*., The *Escherichia coli* heat shock protease HtrA participates in defense against oxidative stress. *Mol. Gen. Genet.* **262**, 342–350 (1999).
13. M. R. Leandro *et al*., DegP protease is essential for tolerance to salt stress in the plant growth-promoting bacterium *Gluconacetobacter diazotrophicus* PAL5. *Microbiol. Res.* **243**, 126654 (2021).
14. T. Krojer, J. Sawa, R. Huber, T. Clausen, HtrA proteases have a conserved activation mechanism that can be triggered by distinct molecular cues. *Nat. Struct. Mol. Biol.* **17**, 844–852 (2010).
15. J. G. Sklar, T. Wu, D. Kahne, T. J. Silhavy, Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. *Genes Dev.* **21**, 2473–2484 (2007).
16. F. Ruiz-Perez *et al*., Roles of periplasmic chaperone proteins in the biogenesis of serine protease autotransporters of *Enterobacteriaceae*. *J. Bacteriol.* **191**, 6571–6583 (2009).
17. T. Krojer *et al*., Structural basis for the regulated protease and chaperone function of DegP. *Nature* **453**, 885–890 (2008).
18. E. Braselmann, J. L. Chaney, M. M. Champion, P. L. Clark, DegP chaperone suppresses toxic inner membrane translocation intermediates. *PLoS One* **11**, e0162922 (2016).
19. C. Baud *et al*., Role of DegP for two-partner secretion in *Bordetella*. *Mol. Microbiol.* **74**, 315–329 (2009).
20. G. E. Purdy, C. R. Fisher, S. M. Payne, IcsA surface presentation in *Shigella flexneri* requires the periplasmic chaperones DegP, Skp, and SurA. *J. Bacteriol.* **189**, 5566–5573 (2007).
21. S. Kim, R. T. Sauer, Distinct regulatory mechanisms balance DegP proteolysis to maintain cellular fitness during heat stress. *Genes Dev.* **28**, 902–911 (2014).
22. S. Kim, R. A. Grant, R. T. Sauer, Covalent linkage of distinct substrate degrons controls assembly and disassembly of DegP proteolytic cages. *Cell* **145**, 67–78 (2011).
23. Q.-T. Shen *et al*., Bowl-shaped oligomeric structures on membranes as DegP's new functional forms in protein quality control. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 4858–4863 (2009).
24. T. Krojer, M. Garrido-Franco, R. Huber, M. Ehrmann, T. Clausen, Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine. *Nature* **416**, 455–459 (2002).
25. J. Jiang *et al*., Activation of DegP chaperone-protease via formation of large cage-like oligomers upon binding to substrate proteins. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 11939–11944 (2008).
26. P. A. Korevaar *et al*., Pathway complexity in supramolecular polymerization. *Nature* **481**, 492–496 (2012).
27. D. van der Zwaag *et al*., Kinetic analysis as a tool to distinguish pathway complexity in molecular assembly: An unexpected outcome of structures in competition. *J. Am. Chem. Soc.* **137**, 12677–12688 (2015).
28. J. P. Renn, M. Junker, R. N. Besingi, E. Braselmann, P. L. Clark, ATP-independent control of autotransporter virulence protein transport via the folding properties of the secreted protein. *Chem. Biol.* **19**, 287–296 (2012).
29. P. N. Danese, W. B. Snyder, C. L. Cosma, L. J. B. Davis, T. J. Silhavy, The Cpx two-component signal transduction pathway of *Escherichia coli* regulates transcription of the gene specifying the stress-inducible periplasmic protease, DegP. *Genes Dev.* **9**, 387–398 (1995).
30. A. K. Attri, C. Fernández, A. P. Minton, Self-association of Zn-insulin at neutral pH: Investigation by concentration gradient–static and dynamic light scattering. *Biophys. Chem.* **148**, 23–27 (2010).
31. A. D. Hanlon, M. I. Larkin, R. M. Reddick, Free-solution, label-free protein-protein interactions characterized by dynamic light scattering. *Biophys. J.* **98**, 297–304 (2010).
32. A. Parupudi *et al*., Global multi-method analysis of interaction parameters for reversibly self-associating macromolecules at high concentrations. *Sci. Rep.* **11**, 5741 (2021).
33. J. B. Stock, B. Rauch, S. Roseman, Periplasmic space in *Salmonella typhimurium* and *Escherichia coli*. *J. Biol. Chem.* **252**, 7850–7861 (1977).
34. A. Ortega, D. Amorós, J. García de la Torre, Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* **101**, 892–898 (2011).
35. N. V. Prabhu, K. A. Sharp, Heat capacity in proteins. *Annu. Rev. Phys. Chem.* **56**, 521–548 (2005).
36. V. V. Loladze, D. N. Ermolenko, G. I. Makhatadze, Heat capacity changes upon burial of polar and nonpolar groups in proteins. *Protein Sci.* **10**, 1343–1352 (2001).
37. S. Kim, R. T. Sauer, Cage assembly of DegP protease is not required for substrate-dependent regulation of proteolytic activity or high-temperature cell survival. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 7263–7268 (2012).
38. C. A. Waudby, A. Ramos, L. D. Cabrita, J. Christodoulou, Two-dimensional NMR lineshape analysis. *Sci. Rep.* **6**, 24826 (2016).
39. J. E. Ollerenshaw, V. Tugarinov, L. E. Kay, Methyl TROSY: Explanation and experimental verification. *Magn. Reson. Chem.* **41**, 843–852 (2003).
40. V. Tugarinov, P. M. Hwang, J. E. Ollerenshaw, L. E. Kay, Cross-correlated relaxation enhanced $^{1}H–^{13}C$ NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J. Am. Chem. Soc.* **125**, 10420–10428 (2003).
41. A. Sekhar, R. Rosenzweig, G. Bouvignies, L. E. Kay, Mapping the conformation of a client protein through the Hsp70 functional cycle. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 10395–10400 (2015).
42. T. Nishikawa, A. Nagadoi, S. Yoshimura, S. Aimoto, Y. Nishimura, Solution structure of the DNA-binding domain of human telomeric protein, hTRF1. *Structure* **6**, 1057–1065 (1998).
43. L. F. Wood, D. E. Ohman, Independent regulation of MucD, an HtrA-like protease in *Pseudomonas aeruginosa*, and the role of its proteolytic motif in alginate gene regulation. *J. Bacteriol.* **188**, 3134–3137 (2006).
44. M. J. Pallen, B. W. Wren, The HtrA family of serine proteases. *Mol. Microbiol.* **26**, 209–221 (1997).
45. B. Lipinska, O. Fayet, L. Baird, C. Georgopoulos, Identification, characterization, and mapping of the *Escherichia coli* htrA gene, whose product is essential for bacterial growth only at elevated temperatures. *J. Bacteriol.* **171**, 1574–1584 (1989).
46. D. Frees, L. Brøndsted, H. Ingmer, "Bacterial proteases and virulence" in *Subcellular Biochemistry*, D. A. Dougan, Ed. (Springer, 2013), pp. 161–192.
47. A. Sobiecka-Szkatula *et al*., Temperature-induced conformational changes within the regulatory loops L1-L2-LA of the HtrA heat-shock protease from *Escherichia coli*. *Biochim. Biophys. Acta* **1794**, 1573–1582 (2009).
48. J. Skorko-Glonek, A. Sobiecka-Szkatula, J. Narkiewicz, B. Lipinska, The proteolytic activity of the HtrA (DegP) protein from *Escherichia coli* at low temperatures. *Microbiology (Reading)* **154**, 3649–3658 (2008).
49. D. D. Boehr, R. Nussinov, P. E. Wright, The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **5**, 789–796 (2009).
50. D. E. Koshland, Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **44**, 98–104 (1958).
51. J. Monod, J. Wyman, J. P. Changeux, On the nature of allosteric transitions: A plausible model. *J. Mol. Biol.* **12**, 88–118 (1965).
52. W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing* (Cambridge University Press, ed. 2, 1992).
53. E. Koehler, E. Brown, S. J. P. A. Haneuse, On the assessment of Monte Carlo error in simulation-based statistical analyses. *Am. Stat.* **63**, 155–162 (2009).