Published in final edited form as:
Nat Struct Mol Biol. ; 19(2): 152–157. doi:10.1038/nsmb.2210.

Newly folded substrates inside the molecular cage of the HtrA chaperone DegQ

Hélène Malet[1,5], **Flavia Canellas**[2], **Justyna Sawa**[2,5], **Jun Yan**[3], **Konstantinos Thalassinos**[3], **Michael Ehrmann**[4], **Tim Clausen**[2], and **Helen R Saibil**[1]

[1]Institute of Structural and Molecular Biology, Crystallography, Birkbeck College, Malet Street, London, WC1E 7HX, United Kingdom

[2]Research Institute of Molecular Pathology, Doktor Bohrgasse 7, A-1030 Vienna, Austria

[3]Institute of Structural and Molecular Biology, Division of Biosciences, University College London, Gower Street, London, WC1E 6BT, United Kingdom

[4]Centre for Medical Biotechnology, Faculty of Biology, University Duisburg-Essen, Universitatsstrasse, D-45117 Essen, Germany

Abstract

The HtrA protein family combines chaperone and protease activities and is essential for protein quality control in many organisms. Whereas the mechanisms underlying the proteolytic function of HtrA proteins have been analyzed in detail, their chaperone activity remains poorly characterized. Here we describe cryo-electron microscopic structures of *Escherichia coli* DegQ in its 12- and 24-mer states in complex with model substrates, providing a structural model of HtrA proteins in their chaperone mode. Up to six lysozyme substrates bind inside the DegQ 12-mer cage and are visualized in a close-to-native state. An asymmetric reconstruction reveals the binding of a well-ordered lysozyme to four DegQ protomers. DegQ PDZ domains are located adjacent to substrate density and their presence is required for chaperone activity. The substrate-interacting regions appear conserved in 12- and 24-mer cages, suggesting a common mechanism of chaperone function.

Keywords

HtrA; chaperone; protein folding; protein quality control; single particle cryo-electron microscopy

Corresponding author: Helen R. Saibil, h.saibil@mail.cryst.bbk.ac.uk.
[5]Present addresses: Unit of Virus Host-Cell Interactions, Joint International Unit 3265, 6 rue Jules Horowitz, BP 181, 38042 Grenoble Cedex 9, France (H.M.); Max F. Perutz Laboratories, University of Vienna, Dr. Bohr-Gasse 9, 1030 Vienna, Austria (J.S)

Author contributions

EM data collection and processing, fitting of atomic coordinates into EM maps and tryptophan fluorescence were carried out by H.M. under the supervision of H.R.S. Protein purification and complex formation were done by F.C., J.S. and H.M. Mass spectrometry experiments were performed by J.Y. under the supervision of K.T. The refolding assays were carried out by J.S. and F.C. under the supervision of T.C. and M.E. H.R.S. and T.C. supervised the project. H.M., H.R.S., and T.C. wrote the manuscript.

Accession codes

DegQ12–lysozyme asymmetric, DegQ12–lysozyme symmetric, DegQ12–peptide and DegQ24–casein maps and corresponding fitted atomic coordinates (C-alpha traces) have been deposited in the EMDB and PDB with accession codes EMD-1981 and PDB 4A8A, EMD-1982 and PDB 4A8B, EMD-1983 and PDB 4A8C and EMD-1984 and PDB 4A9G, respectively. A revised fit to the DegP-OMP map has been deposited as PDB 4A8D.

Competing interest statement

The authors declare that they have no competing financial interests.

Europe PMC Funders Author Manuscripts

Introduction

Cell viability depends on the proper structure and function of the proteome. For protein quality control, all cells have developed elaborate systems of molecular chaperones and proteases[1,2]. Failure of protein homeostasis leads to the accumulation of misfolded or aggregated proteins, a malfunction associated with fatal protein-folding diseases[3]. Members of the 'High temperature requirement A' (HtrA) family play a central role in protein quality control in a wide range of organisms as they combine proteolytic and remodelling activities of aberrant proteins in a highly regulated and ATP-independent mechanism[4]. Disturbances in the function of human HtrA proteins (HTRA1, HTRA2) are associated with severe disorders including Alzheimer's and Parkinson's diseases and cancers[5-7]. Prokaryotic HtrAs are essential for bacterial virulence and for survival after exposure to various environmental and cellular stresses[8]. In *Escherichia coli* three HtrA proteins contribute to maintenance of protein quality in the periplasm (DegP, DegQ and DegS)[9-13], with DegQ in particular considered as a model of the HtrA family due to its high sequence identity with many HtrA members[14].

HtrA proteins are composed of a chymotrypsin-like protease domain and one (DegS, HTRA1, HTRA2) or two PDZ domains (DegP, DegQ)[14]. Three protease domains interact tightly to constitute the trimeric building blocks of all HtrA complexes. Whereas membrane-anchored HtrA proteases such as *E. coli* DegS and human HTRA2 are active as trimers[13,15], several soluble HtrA proteins have been shown to form larger oligomers. Human HTRA1 trimers assemble into 12-mers in the presence of non-native polypeptides[16], whereas *E. coli* DegP switches between a hexameric resting state without substrate to active 12-, 15-, 18-, 24- and 30-mer states in presence of a substrate that can be refolded or degraded[9,17-20]. Similarly to *E. coli* DegP, *E. coli* DegQ changes its oligomeric state from hexamers to either 12- or 24-mers depending on the concentration of unfolded substrate. In addition, *E. coli* DegQ forms 12-mers in the absence of substrate at acidic pH[10].

Two types of higher oligomeric structure have been described for *E. coli* DegP: soluble cages (12- and 24-mer) and bowl-shaped structures bound to liposomes (12-, 15- and 18-mers)[9,17,18,20]. For *E. coli* DegQ, full length 12- and 24-mers remain uncharacterized, but the DegQΔPDZ2 12-mer has been described as a cage-like structure of 135 Å in diameter[10]. This structure provides a model for the dodecameric forms of soluble HtrA proteins containing only one PDZ domain, such as human HTRA1. In addition, the recent crystal structure of a *Legionella falloni* DegQ dodecamer shows a divergent organization and a smaller size compared to *E. coli* HtrA proteins (140 Å in diameter for $DegQ_{Lf}$ vs 165 Å for $DegP_{Ec}$)[21].

While some HtrA cage-like structures have been obtained in the presence of substrates to be folded or degraded, the protein ligands are not visible in crystallographic structures and symmetrized cryo-electron microscopic (cryo-EM) maps, likely due to conformational and positional flexibility[9,17,18,22]. Only short peptides bound to PDZ1 and protease domains have been resolved in X-ray crystal structures. An asymmetric cryo-EM reconstruction of DegP 12-mer showed a folded outer membrane protein (OMP) encapsulated within the cage, but its low resolution (28 Å) precluded analysis of the DegP–OMP interaction[9].

How the HtrA proteins bind and fold their substrates, central to understanding their chaperone activity, thus remains to be characterized. In order to probe HtrA chaperone function, we carried out a structural and biochemical analysis of the protease deficient S187A mutant of *E. coli* DegQ in complex with several model substrates. Here, we present cryo-EM structures of DegQ 12- and 24-mer cages encapsulating these substrates. Remarkably, the DegQ 12-mer can accommodate and fold up to 6 lysozymes. The

Europe PMC Funders Author Manuscripts

interaction of a bound lysozyme in close-to-native conformation with the DegQ cage is revealed by an asymmetric reconstruction of the DegQ 12-mer–lysozyme complex. PDZ domains are located close to substrate density and a refolding assay indicates that they are required for DegQ chaperone activity.

Results

Cryo-EM structure of DegQ24–β-casein complex

To gain insights into the chaperone function of DegQ, we determined the cryo-EM structures of DegQ-substrate complexes. DegQ 24-mers were formed by incubating proteolytically inactive hexameric $DegQ_{S187A}$ with β-casein (a 24 kDa natively unstructured protein used as a model substrate)[10] (Supplementary Fig. 1). Imaging the purified samples by EM reveals mainly large particles approximately 215 Å in diameter and minor complexes of 165 Å (Supplementary Fig. 2). The large complexes correspond to 24-mers and the smaller ones represent 12-mer cages. Cryo-EM images of the 24-mers exhibit 4-, 3- and 2-fold symmetry (Supplementary Fig. 2), and a three-dimensional (3D) reconstruction with octahedral symmetry was obtained by angular reconstitution at 7.5 Å resolution (Supplementary Fig. 3). Overall, the cryo-EM map displays a hollow spherical shape formed by eight trimeric building blocks, similar to the DegP 24-mer[9,17]. The quality of the map allows localization of secondary structure elements (Fig. 1a).

To obtain a pseudo-atomic model of the complex, a homology model of the DegQ PDZ2 domain was generated, based on the DegP crystal structure, and combined with the DegQ protease-PDZ1 trimer crystal structure[10,19]. Eight copies of the resulting trimer model were docked into the cryo-EM map and their positions refined by flexible fitting, allowing hinge movements between domains[23]. The pseudo-atomic model reveals a strong interaction between PDZ1 and PDZ2 from neighboring trimers, mediated mainly by hydrophobic residues. PDZ1 and PDZ2 of the same protomer are around 24 Å apart, separated by an extended 8-amino acid-long linker (residues 333-341) (Fig. 1b).

Whereas DegP and DegQ 24-mers share a similar global organization, the DegQ complexes are slightly expanded, with a diameter of 210 Å compared to 195 Å for DegP and larger pores on the 2- and 4-fold axes (Supplementary Fig. 4). These differences arise from divergences in DegP/DegQ PDZ domain orientations. When the protease domains of the two structures are superimposed, DegP PDZ1 must be rotated by 11° to fit the DegQ24 cryo-EM map (Fig. 1c). The PDZ2 domain position is substantially different, with a 70° rotation and a 5 Å translation observed between DegP and DegQ. The PDZ domain orientations found in DegQ, unambiguously identified in the cryo-EM map, create a cavity defined by protease, PDZ1 and PDZ2 domains from three different protomers. The cleft thus created is surrounded by helix 251-257 of PDZ1, loop residues 408-413 from PDZ2 and residues 31-33 and 58-62 from the protease domain (Fig. 1d). Additional density (colored in orange in Fig. 1b and d) is present in this cavity and cannot be accounted for by the DegQ model. Its volume is 7,300 A^3, corresponding to 6 kDa. We propose that this density corresponds to either (i) the LA loop comprising residues 34-57 of the protease domain, previously shown to be an important regulator of the protease activity but not modeled in DegQ 24-mer due to its flexibility or (ii) to part of the β-casein substrate. Visualization of only 6 kDa out of 24 kDa for β-casein would be consistent with the disordered, natively unstructured state of β-casein. Given that the additional density is buried in a cavity, we suggest that it is more likely to correspond to the ligand density.

Cryo-EM structure of a DegQ 12–peptide complex

To discriminate between ligand and DegQ densities, we determined 3D reconstructions of the DegQ cage with and without high molecular mass substrate. To obtain a homogeneous preparation of DegQ 12-mers devoid of substantial ligand density, we incubated DegQ hexamers with a 20 amino acid long peptide previously shown to bind to the DegQ PDZ1 domain (SPMFKGVLDMMYGGMRGYQV)[10,24]. DegQ12-peptide complexes were purified by size-exclusion chromatography (SEC) and imaged by EM, revealing a preparation of hollow, round particles with a diameter of ~165 Å (Supplementary Fig. 1 and 2). The tetrahedral symmetry of the 12-mer complexes, consistent with the observation of two- and three-fold views in cryo-EM particle averages (Supplementary Fig. 2), was unambiguously identified by angular reconstitution. A 3D cryo-EM map was reconstructed at 7.5 Å resolution (Supplementary Fig. 3) allowing precise fitting of the domains and identification of secondary structure elements (Fig. 2a and b). Previous crystallographic analyses suggest that polypeptides are cooperatively bound by PDZ1 and protease domains, based on the observation of short segments of peptide binding[10,18,24]. Our cryo-EM map is compatible with this binding mode, but the peptide was omitted from the model as the map resolution prevents peptide accurate positioning.

DegQ 12-mer formation is mediated by the interaction of four DegQ trimers via PDZ1 and PDZ2’ domains from neighboring trimers. The cage is thus formed of four structural units, each of them containing protease-PDZ1 domains of three subunits tightly bound via hydrophobic interactions to three PDZ2’ domains from neighboring protomers. The overall organization of the DegQ 12-mer is reminiscent of the arrangement previously observed in a DegP12–lysozyme cryo-EM map[17]. However there are marked differences in PDZ positions between DegP and DegQ. As in the DegQ 24-mer, PDZ2 is rotated by 70° relative to DegP (Supplementary Fig. 5). In addition, the DegQ 12-mer PDZ1 domain is rotated and shifted 7 Å towards the interior of the cage compared to its position in DegP, with appreciable flexibility of PDZ1 helix 251-257. En-bloc movement of PDZ1 relative to the protease domain is observed in both DegQ 12 and 24-mers. This conformational change was unexpected as PDZ1 was observed in a conserved position in all previously determined *E. coli* HtrA cage structures[9,10,18] (Supplementary Fig. 6).

Comparison of DegQ 12- and 24-mer cages reveals that the DegQ regions located adjacent to the suggested β-casein density in the DegQ 24-mer (34-57 from the protease domain, 251-257 from PDZ1 and 408-413 from PDZ2, see preceding section) are close to each other in the DegQ12–peptide map (Fig. 2b). They protrude towards the interior of the cage, in a position compatible with substrate binding, near the predicted peptide binding site. The DegQ12–peptide map is devoid of additional density, further suggesting that the additional density in the DegQ 24-mer corresponds to β-casein rather than to the LA loop.

Five or six folded lysozymes bind inside the DegQ 12-mer

We then investigated the positioning of substrates within the $DegQ_{S187A}$ 12-mer using reduced and chemically denatured lysozyme (14.3 kDa) as a model substrate. Upon incubation with purified hexameric $DegQ_{S187A}$ at 37°C, DegQ12–lysozyme complexes form, which were subsequently purified by SEC (Supplementary Fig. 1). Negative stain and cryo-EM images reveal the presence of cages similar in size to the ones observed for DegQ12–peptide (Supplementary Fig. 2). On the basis of this comparison, the volume of the 12-mer cage appears independent of substrate composition and molecular mass.

The DegQ12–lysozyme images show additional density filling the cages, likely representing the substrate (compare Supplementary Fig. 2b and c). A cryo-EM 3D reconstruction with tetrahedral symmetry was obtained at 13 Å resolution and shows the same DegQ structure as

in the peptide complex (Fig. 2c and d, Supplementary Fig. 3). The protease domains remain in the same positions and only a slight opening of the PDZ1-PDZ2' domain contact is observed. A difference map between DegQ12–lysozyme and DegQ12–peptide reveals the substantial lysozyme density inside the cage, interacting with the inner surface of DegQ (Fig. 2c, d, e, f). The additional density has a volume of 80,000 $Å^3$ corresponding to 66 kDa, indicating the binding of around 5 folded lysozymes (total molecular mass ~70 kDa).

To establish the exact number of lysozymes bound inside the DegQ 12-mer, we conducted mass spectrometry (MS) experiments. We initially recorded a denaturing MS spectrum of DegQ12–lysozyme to determine the mass of the DegQ monomer. The most abundant species has a molecular mass of 44,835.4 ± 11.0 Da, which is smaller than the theoretical mass, likely due to proteolysis at the unstructured termini of DegQ. To determine the precise stoichiometry of DegQ–lysozyme assemblies, we analyzed the apo- and substrate-bound DegQ 12-mer complexes by native mass spectrometry. To obtain DegQ 12-mers in the absence of substrate, we incubated DegQ at pH 5.5, yielding an equilibrium mixture of hexamers and 12-mers, as shown by SEC[10] and negative stain EM (Supplementary Fig. 2). Native mass spectrometry reveals components with the expected masses for DegQ hexamers and 12-mers, (270,135.0 ± 29.2 and 553,105.8 ± 156.5 Da respectively) (Fig. 3a). To determine the number of bound lysozymes, MS spectra of the DegQ12–lysozyme complexes were initially recorded but showed a very broad peak preventing unambiguous mass determination (data not shown). To overcome this, a tandem MS approach was applied. Tandem MS has previously been used to resolve overlapping charge states arising from polydisperse samples and from the presence of different substrate-bound complexes[25-27]. As shown in Figure 3b, the peak at 11050 m/z from the DegQ12–lysozyme precursor was selected using the quadrupole mass filter and subjected to collision-induced dissociation. The peak series in the low m/z region corresponds to a highly charged, ejected DegQ monomer subunit, and the peak series at higher m/z corresponds to the charge-stripped DegQ 11-mer with lysozyme bound. At this region there is a greater separation between charge states facilitating mass assignment[28]. Two predominant charge state series are present in the charge-stripped complex region, corresponding to a mass of 596,555 ± 249 Da and 585,788 ± 365 Da, with the former being the dominant species. These masses correspond to DegQ 11-mer bound to 6 or 5 lysozymes respectively.

Consistently, six folded lysozyme molecules can be fitted into this density without clashes (Fig. 2c, d, e, f). In order to confirm that the bound lysozyme substrates are folded as implied by the cryo-EM density, we took advantage of the absence of tryptophan in DegQ to monitor the average tryptophan fluorescence of bound lysozymes[29,30]. The maximum emission of DegQ12–lysozyme was observed at 342.5 nm, whereas the maximum emission is 341.5 nm for folded lysozyme and 352.5 nm for unfolded lysozyme (Fig. 3c). The fluorescence intensity is also increased in the unfolded state (Fig. 3c, with the same concentration of lysozyme used for all spectra). Therefore, the spectra suggest that the bound lysozymes are in a close-to-native state. The small difference in emission maximum and intensity between folded lysozyme and DegQ12–lysozyme could arise from modification of the tryptophan environment by an interaction with DegQ or from a not completely native conformation of the lysozyme. Since cage formation is triggered by denatured but not by folded lysozyme, these results indicate that lysozymes fold within the DegQ cage.

In the tetrahedrally symmetric map, the substrate density level is only about one third of that of DegQ, suggesting that the lysozyme arrangement is most likely asymmetric or disordered. The calculation of an asymmetric reconstruction of DegQ12–lysozyme is thus important for a more accurate description of lysozyme density in the DegQ 12-mer cage.

Substrate binding regions in DegQ cages

An asymmetric 3D reconstruction of DegQ12-lysozyme was obtained at 14.2 Å resolution (Fig. 4, Supplementary Fig. 3), revealing the presence of two separate substrate densities. One density is present in the middle of the complex and does not show any direct interaction with DegQ (Fig. 4a and b, density colored in purple). Its volume corresponds to 2.4 folded lysozymes and is likely to represent mobile substrates sequestered by the DegQ12-mer cage. The second density is in contact with DegQ and has a volume corresponding to 10 kDa (Fig. 4a and b, colored orange). It presents a two-lobed shape, compatible with the two domains of a folded lysozyme (Fig. 4c). A folded lysozyme can be fitted into this density without creating clashes, although it cannot be precisely positioned at the resolution of this EM map. On the other hand, the DegQ domain positions are well defined by the comparison with DegQ12–peptide. Therefore the reliable fitting of DegQ reveals the sites of its interaction with the lysozyme density. The ordered parts of the protease domain LA loop are oriented towards the lysozyme, suggesting involvement of the LA loop in lysozyme binding (Fig. 4c). In addition, lysozyme is adjacent to the protease and PDZ2 domains of two DegQ subunits, as well as the PDZ1 domains of another two DegQ subunits. The PDZ regions that appear to be involved in lysozyme binding in DegQ 12-mer, namely helix residues 251-257 of PDZ1 and loop 408-413 of PDZ2, are the ones located close to the additional density attributed to β-casein in the DegQ 24-mer (Fig. 4d).

PDZ domains are needed for DegQ chaperone activity

The proximity between PDZ regions and substrate prompted us to investigate the role of PDZ domains in DegQ chaperone function. Refolding of α-amylase (MalS), a natural periplasmic substrate of DegP, was used to monitor the chaperone activity of several DegQ constructs. In this assay, DegQ was incubated with chemically denatured MalS and the MalS substrate *p*-nitrophenylhexaoside (PNP6). When folded, MalS cleaves PNP6 yielding the chromogenic p-nitrophenol that absorbs at 405 nm. The rate of MalS folding in the presence of HtrA chaperones can thus be readily monitored and compared to its spontaneous folding. With this refolding assay, we show that, under the *in vitro* conditions used, the protease-deficient $DegQ_{S187A}$ mutant has a slightly higher chaperone activity than its ortholog $DegP_{S210A}$ (Fig. 5a and b). Deletion mutants of $DegQ_{S187A}$ (ΔPDZ2 and ΔPDZ1+2) have a much lower chaperone activity than $DegQ_{S187A}$. A role for PDZ domains in chaperone activity is in accordance with the observation of substrate densities adjacent to these domains in the cryo-EM structures of *E. coli* DegQ cages. The low chaperone activity of $DegQ_{S187A}$ΔPDZ1+2 may also arise from its inability to form cages[10].

We then attempted to determine which PDZ domains residues are implicated in substrate binding and folding. PDZ2 loop 408-413 contains poorly conserved hydrophobic residues. Therefore, PDZ2 might be dispensable for substrate binding in other organisms, as reported for *L. fallonii* DegQ[21]. We thus focused our analysis on the PDZ1 helix 251-257, which contains two hydrophobic residues, namely I253 and F257. In addition, we noticed that residue F266 was correctly oriented to potentially interact with substrates. Unfortunately, the triple mutant I253A F257A F266A, the double I253A F257A mutant and the three corresponding single mutants interfere with 12-mer formation in the presence of lysozyme substrate, preventing the analysis of DegQ12 mutant chaperone activity (Supplementary Fig. 1).

Discussion

The DegQ–substrate complexes presented here reveal new information about the chaperone function of HtrA proteins. The combination of single particle cryo-EM, native mass spectrometry and fluorescence provides strong evidence for the folding of 5 or 6 lysozymes

inside the DegQ 12-mer. It suggests that, in the context of the cell, DegQ can capture and enclose multiple small, unfolded substrates that are subsequently refolded within its cavity. Remarkably, six folded lysozymes can fit inside the DegQ 12-mer complex. It is instructive to compare DegQ12–lysozyme with the structure of a closed chaperonin cage containing a newly folded substrate, GroEL–gp31 bound to the T4 bacteriophage capsid protein gp23 (gp31 is the T4 bacteriophage homolog of GroES)[32]. The mechanisms of cage assembly differ: chaperonin cage formation is regulated by ATP binding, whereas DegQ cage assembly is ATP-independent and triggered by substrate binding. The cage architecture is also different, with two compartments alternately used for folding in chaperonins, while DegQ forms a single, larger cage. Substrate packing inside the molecular chaperone cages is very dense for both chaperonins and DegQ. The fraction of substrate that is visible, presumably because of ordered packing in a restricted volume, is also comparable, around 70% for GroEL-gp23 and 78 % for DegQ12–lysozyme. A similar packing density has been observed for tubulin inside the CCT chaperonin[31]. However, chaperonins encapsulate only one substrate at a time per compartment, in contrast to the present finding of up to 6 lysozymes inside one DegQ cage. Folding of multiple substrates within the same compartment, along with the combination of proteolytic and chaperone activities, might have evolved in response to the direct exposure of the periplasm to environmental stresses.

The asymmetric map of DegQ12–lysozyme identifies regions adjacent to lysozyme (Fig. 4c). They originate from protease, PDZ1 and PDZ2 domains of four different protomers. In addition, we show that PDZ-deletion mutants of DegQ have low chaperone activity in MalS refolding assays (Fig. 5a and b). These data suggest that cage formation and/or interaction of PDZ domains with the substrates are required for chaperone activity. It is striking that the regions of DegQ close to lysozyme (helix 251-257 of PDZ1, loop 408-413 of PDZ2 and loop LA of the protease domain) are also adjacent to the proposed β-casein density in DegQ 24-mer (Fig. 4d). Thus our maps suggest that not only the global structural organization, but also the binding mode of chaperone substrates are conserved in the two cages. The cryo-EM maps of DegQ assemblies reveal a conserved organization of cage-like complexes in *E. coli* DegP and DegQ. PDZ1 and PDZ2' from different protomers are involved in 12- and 24-mer cage assembly. In a previous study, we proposed a different domain arrangement, based on the fitting of OMP-bound DegP into a low-resolution, asymmetric cryo-EM map. In light of the results presented here and the DegP 12-mer structures[9,17], we have revised our fitting of the OMP-bound DegP 12-mer, leading to a consensus for *E. coli* DegP and DegQ cage architecture (Supplementary Fig. 7) (4A8D for PDB accession code of new fit).

Interestingly, while the overall organization of *E. coli* DegQ and DegP 12-mer structure is conserved, the recently published DegQ 12-mer structure of *Legionella fallonni* shows a different assembly[21]. Whereas a PDZ2 domain of $DegQ_{Ec}$ only interacts with a PDZ1' domain of another protomer, a PDZ2 domain in the 12-mer of $DegQ_{Lf}$ additionally interacts with two PDZ2 and one protease domain of neighboring protomers. Consequently, the $DegQ_{Lf}$ 12-mer forms a smaller cage (140 Å in diameter vs 165 Å for $DegQ_{Ec}$). It will be interesting to determine the physiological implications of these architectural differences between $DegQ_{Lf}$ and $DegQ_{Ec}$ in future studies.

Although the global structure of *E. coli* DegP and DegQ cages is preserved, the positions of their PDZ domains differ. PDZ1 domains of DegQ 12- and 24-mer deviate from their positions in DegP cages, with marked rotation relative to the protease domain. The PDZ1 densities are less well defined in the DegQ 12-mer asymmetric map, implying that they are mobile. PDZ2 orientation differs by 70° between DegP and DegQ. As a consequence, DegP and DegQ cages differ slightly in shape, size and electrostatic potential. These structural divergences might be related to the differences in DegP and DegQ function, for example regarding OMP biogenesis in *E. coli*. Indeed, we observe that isolated outer membranes

from *degQ*- null strain show no detectable alterations of OMP composition compared to wild type *E. coli* cells, unlike the *degP*-null mutant, in which the levels of some OMPs including OmpA, OmpC and OmpF are markedly decreased (Fig. 5c). Consistent with this observation, DegQ is also reported to be dispensable for OMP folding in *Neisseria meningitidis*[33]. Our data thus support the model of divergence of substrate specificity between DegP and DegQ and suggest directions for further investigation.

In conclusion, this study pinpoints substrate-binding regions within the cavity of *E. coli* DegQ cages. As many HtrA members are DegQ homologs, the results presented here provide insights on how the HtrA protein family encapsulate and fold substrates.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Daniel Clare, Natalya Lukoyanova and Elena Orlova for advice on EM data collection and processing, Luchun Wang, David Houldershaw and Richard Westlake for computing and EM support, and Tina Daviter for help with fluorescence spectroscopy. This work was supported by Wellcome Trust and European Science Foundation grants to H.R.S., by ERA-Net NEURON, FWF I 235-B09 to F.C. and T.C., Wellcome Trust and Institute of Structural and Molecular Biology grants to K.T. The Research Institute of Molecular Pathology is funded by Boehringer Ingelheim.

Appendix

Methods

Specimen preparation and EM data collection

DegQ24–casein, DegQ12–lysozyme and DegQ12–peptide were prepared as described previously[1]. For cryo-EM data collection, DegQ24–casein, DegQ12–lysozyme and DegQ12–peptide were diluted to 0.2 mg.ml^{-1} in a buffer containing 20 mM HEPES/NaOH pH 7.5, 150 mM NaCl. Buffer was supplemented with 100 μM of the PDZ1-binding peptide for the DegQ12–peptide sample. 4 μl samples were applied to glow discharged C-flat grids (CF-2/2-4C-100; Protochips). After 30s, excess solution was blotted and the grid frozen in liquid ethane. Cryo-EM was collected on a Tecnai F20 microscope (FEI) operated at 200 kV under low-dose conditions. Micrographs were recorded on Kodak SO-163 film at 50,000 magnification with defocus ranging from 1 to 3 μm.

Image processing

Micrographs were digitized on a SCAI microdensitometer (Zeiss) at 1.4 Å per pixel. A total of 12312 (DegQ24–casein) and 16790 (DegQ12–lysozyme) particles were manually picked in Ximdisp[2]. 36790 particles of DegQ12–peptide were semi-automatically picked with BOXER[3]. The defocus and astigmatism of the images were determined using CTFFIND3 and corrected for the effect of the contrast transfer function by phase flipping[4]. Full CTF correction was applied at the final stage for DegQ12–peptide and DegQ24–casein reconstructions. Images were filtered between 230-4 Å (DegQ24–casein), 165-4 Å (DegQ12–lysozyme, DegQ12–peptide) and normalized, using SPIDER[5]. Image processing and 3D reconstructions were done in SPIDER and in IMAGIC-5[6]. Multivariate statistical analysis and eigenimage analysis revealed the presence of 4-, 3- and 2-fold symmetry in DegQ24–casein particles and 3- and 2-fold symmetry in DegQ 12-mer complexes. Angular reconstitution 3D models clearly indicated that Deg 12-mer is tetrahedral and DegQ 24-mer is octahedral. The 3D maps were refined using angular reconstitution and projection matching. The asymmetric 3D map of DegQ12–lysozyme was calculated by projection

matching using the final DegQ12–lysozyme symmetrised map as a starting model. Resolution of the reconstructions was assessed by Fourier shell correlation at 0.5 correlation. A more comprehensive description of data processing procedures can be found in the Supplementary Methods.

Fitting

E. coli DegQ protease-PDZ1 crystal structure and a homology model of DegQ PDZ2 were used during the fitting procedure (homology model generated with MODELLER[7] based sequence-structure alignment with *E. coli* DegP pdb code 3CS0[8]). Rigid body and flexible fitting were performed using UCSF Chimera[9] 'fit-in-map module' and Flex-EM[10] as described in the Supplementary Methods. Improvement of cross-correlation between atomic models and cryo-EM maps is shown on Supplementary Table 1. Z-scores revealing uniqueness of fits are indicated on Supplementary Table 2.

Mass spectrometry

For native MS experiments, DegQ complexes with and without bound lysozyme were buffer exchanged into 100 mM ammonium acetate and concentrated to 15 μM using Amicon Ultra 0.5 ml centrifugal filters (Millipore UK Ltd, Watford UK). For denaturing MS experiments, DegQ complexes were buffer exchanged into 49:49:2 (v:v:v) water: ethanol: acetic acid.

MS experiments were carried out on a Synapt HDMS (Waters UK Ltd, Manchester, UK) Quadrupole-TOF mass spectrometer[11]. Samples (2-3 μl aliquots) were introduced to the mass spectrometer by means of nano electrospray ionisation using gold-coated capillaries, prepared in house. Typical instrumental parameters were as follows: source pressure 6 mbar, capillary voltage 1.0-1.3 kV, cone voltage 150-200 V, trap energy 20 V, transfer energy 10 V, bias 2.0 V, and trap pressure 3.6×10^{-2} mbar. For tandem MS experiments, the bias voltage was increased to 80 V.

Mass spectra were smoothed and peak-centered in MassLynx. Mass assignment was achieved by the method described in Tito et al.[12], whereby the charge is iterated over the measured mass value and each time the standard deviation for a given charge state series is calculated. The solution is the series that gives rise to the lowest standard deviation.

Fluorescence

Intrinsic tryptophan fluorescence was excited at 295 nm (to exclude tyrosine fluorescence) and monitored between 300-400 nm, with a slit width of 0.7 mm, using a FluoroMax®-3 spectrofluorometer (Horiba). The same concentration of lysozyme was used for all the experiments (0.01 mg.ml^{-1}). Unfolded lysozyme was prepared by incubation with 10 mM DTT and 8 M urea at 37° for 30 min. Folded lysozyme and SEC purified DegQ12–lysozyme were buffered with 10 mM HEPES/NaOH pH 7.5 in 150 mM NaCl. The low background signals from the buffers were subtracted for analysis of the spectra, and each measurement was repeated six times.

MalS refolding assays

MalS, DegP, DegQ purifications and MalS refolding assays were carried out as described previously[1,8,13]. The point mutations were introduced using a QuikChange site-directed mutagenesis kit (Stratagene) and the constructs verified by DNA sequence analysis. To determine the effect of DegQ and DegP on MalS refolding we preincubated 2 μM of DegQ wild-type, DegQ$_{S187A,}$ DegP$_{S210A,}$ DegQ $_{S187A}$ ΔPDZ1+2 or DegQ $_{S187A}$ ΔPDZ2 with 2 mM PNP6 in 250 mM NaH_2PO4 pH 7.5. After 5 min, unfolded MalS was added to a final concentration of 0.13 μM. The activity of the refolded amylase was determined using *p*-nitrophenylhexaoside (PNP6, 2 mM final concentration) as a substrate. The release of *p*-

Europe PMC Funders Author Manuscripts

nitrophenol from PNP6 by MalS was monitored at 405 nm with a microplate reader. Assays were carried out in a total volume of 100 μl at 22°C. Lysozyme was used as a negative control as it supports a lower rate of MalS refolding in a concentration-independent manner, thus reflecting non-specific interactions.

Outer membrane isolation

Outer membranes of *E. coli* wild-type, *degQ*-null (MG1655 *degQ::Tn5 KanR)* and *degP*-null (CLC198, *degP*::Tn*10*) mutant strain were prepared as described previously[14], with minor modifications detailed in the Supplementary Methods.

Method only references

1. Sawa J, et al. Molecular adaptation of the DegQ protease to exert protein quality control in the bacterial cell envelope. J Biol Chem. 2011; 286:30680–90. [PubMed: 21685389]
2. Crowther RA, Henderson R, Smith JM. MRC image processing programs. J Struct Biol. 1996; 116:9–16. [PubMed: 8742717]
3. Ludtke SJ, Baldwin PR, Chiu W. EMAN: semiautomated software for high-resolution single-particle reconstructions. J Struct Biol. 1999; 128:82–97. [PubMed: 10600563]
4. Mindell JA, Grigorieff N. Accurate determination of local defocus and specimen tilt in electron microscopy. J Struct Biol. 2003; 142:334–47. [PubMed: 12781660]
5. Frank J, et al. SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. J Struct Biol. 1996; 116:190–9. [PubMed: 8742743]
6. van Heel M, Harauz G, Orlova EV, Schmidt R, Schatz M. A new generation of the IMAGIC image processing system. J Struct Biol. 1996; 116:17–24. [PubMed: 8742718]
7. Sali A. Comparative protein modeling by satisfaction of spatial restraints. Mol Med Today. 1995; 1:270–7. [PubMed: 9415161]
8. Krojer T, et al. Structural basis for the regulated protease and chaperone function of DegP. Nature. 2008; 453:885–90. [PubMed: 18496527]
9. Goddard TD, Huang CC, Ferrin TE. Visualizing density maps with UCSF Chimera. J Struct Biol. 2007; 157:281–7. [PubMed: 16963278]
10. Topf M, et al. Protein structure fitting and refinement guided by cryo-EM density. Structure. 2008; 16:295–307. [PubMed: 18275820]
11. Pringle SD, et al. An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. International Journal of Mass Spectrometry. 2007; 261:1–12.
12. Tito MA, Tars K, K V, Hajdu J, Robinson CV. Electrospray time of flight mass spectrometry of the intact MS2 virus capsid. J Am Chem Soc. 2000; 122:3550–3551.
13. Spiess C, et al. Biochemical characterization and mass spectrometric disulfide bond mapping of periplasmic alpha-amylase MalS of Escherichia coli. J Biol Chem. 1997; 272:22125–33. [PubMed: 9268356]
14. Rouviere PE, Gross CA. SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins. Genes Dev. 1996; 10:3170–82. [PubMed: 8985185]

Main references

1. Gottesman S, Wickner S, Maurizi MR. Protein quality control: triage by chaperones and proteases. Genes Dev. 1997; 11:815–23. [PubMed: 9106654]
2. Wickner S, Maurizi MR, Gottesman S. Posttranslational quality control: folding, refolding, and degrading proteins. Science. 1999; 286:1888–93. [PubMed: 10583944]
3. Selkoe DJ. Folding proteins in fatal ways. Nature. 2003; 426:900–4. [PubMed: 14685251]
4. Clausen T, Kaiser M, Huber R, Ehrmann M. HTRA proteases: regulated proteolysis in protein quality control. Nat Rev Mol Cell Biol. 2011; 12:152–62. [PubMed: 21326199]

5. Chien J, Campioni M, Shridhar V, Baldi A. HtrA serine proteases as potential therapeutic targets in cancer. Curr Cancer Drug Targets. 2009; 9:451–68. [PubMed: 19519315]
6. Grau S, et al. Implications of the serine protease HtrA1 in amyloid precursor protein processing. Proc Natl Acad Sci U S A. 2005; 102:6021–6. [PubMed: 15855271]
7. Plun-Favreau H, et al. The mitochondrial protease HtrA2 is regulated by Parkinson's disease-associated kinase PINK1. Nat Cell Biol. 2007; 9:1243–52. [PubMed: 17906618]
8. Ingmer H, Brondsted L. Proteases in bacterial pathogenesis. Res Microbiol. 2009; 160:704–10. [PubMed: 19778606]
9. Krojer T, et al. Structural basis for the regulated protease and chaperone function of DegP. Nature. 2008; 453:885–90. [PubMed: 18496527]
10. Sawa J, et al. Molecular adaptation of the DegQ protease to exert protein quality control in the bacterial cell envelope. J Biol Chem. 2011; 286:30680–90. [PubMed: 21685389]
11. Spiess C, Beil A, Ehrmann M. A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. Cell. 1999; 97:339–47. [PubMed: 10319814]
12. Walsh NP, Alba BM, Bose B, Gross CA, Sauer RT. OMP peptide signals initiate the envelope-stress response by activating DegS protease via relief of inhibition mediated by its PDZ domain. Cell. 2003; 113:61–71. [PubMed: 12679035]
13. Wilken C, Kitzing K, Kurzbauer R, Ehrmann M, Clausen T. Crystal structure of the DegS stress sensor: How a PDZ domain recognizes misfolded protein and activates a protease. Cell. 2004; 117:483–94. [PubMed: 15137941]
14. Kim DY, Kim KK. Structure and function of HtrA family proteins, the key players in protein quality control. J Biochem Mol Biol. 2005; 38:266–74. [PubMed: 15943900]
15. Li W, et al. Structural insights into the pro-apoptotic function of mitochondrial serine protease HtrA2/Omi. Nat Struct Mol Biol. 2002; 9:436–41. [PubMed: 11967569]
16. Truebestein L, et al. Substrate-induced remodeling of the active site regulates human HTRA1 activity. Nat Struct Mol Biol. 2011; 18:386–8. [PubMed: 21297635]
17. Jiang J, 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. 2008; 105:11939–44. [PubMed: 18697939]
18. Kim S, Grant RA, Sauer RT. Covalent Linkage of Distinct Substrate Degrons Controls Assembly and Disassembly of DegP Proteolytic Cages. Cell. 2011; 145:67–78. [PubMed: 21458668]
19. Krojer T, Garrido-Franco M, Huber R, Ehrmann M, Clausen T. Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine. Nature. 2002; 416:455–59. [PubMed: 11919638]
20. Shen QT, 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. 2009; 106:4858–63. [PubMed: 19255437]
21. Wrase R, Scott H, Hilgenfeld R, Hansen G. The Legionella HtrA homologue DegQ is a self-compartmentizing protease that forms large 12-meric assemblies. Proc Natl Acad Sci U S A. 2011; 108:10490–5. [PubMed: 21670246]
22. Merdanovic M, et al. Determinants of structural and functional plasticity of a widely conserved protease chaperone complex. Nat Struct Mol Biol. 2011; 17:837–43. [PubMed: 20581826]
23. Topf M, et al. Protein structure fitting and refinement guided by cryo-EM density. Structure. 2008; 16:295–307. [PubMed: 18275820]
24. Krojer T, Sawa J, Huber R, Clausen T. HtrA proteases have a conserved activation mechanism that can be triggered by distinct molecular cues. Nat Struct Mol Biol. 2010; 17:844–52. [PubMed: 20581825]
25. Aquilina JA, Benesch JL, Bateman OA, Slingsby C, Robinson CV. Polydispersity of a mammalian chaperone: mass spectrometry reveals the population of oligomers in alphaB-crystallin. Proc Natl Acad Sci U S A. 2003; 100:10611–6. [PubMed: 12947045]
26. McCammon MG, Hernandez H, Sobott F, Robinson CV. Tandem mass spectrometry defines the stoichiometry and quaternary structural arrangement of tryptophan molecules in the multiprotein complex TRAP. J Am Chem Soc. 2004; 126:5950–1. [PubMed: 15137744]
27. Sharon M, et al. 20S proteasomes have the potential to keep substrates in store for continual degradation. J Biol Chem. 2006; 281:9569–75. [PubMed: 16446364]

28. Sobott F, Hernandez H, McCammon MG, Tito MA, Robinson CV. A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. Anal Chem. 2002; 74:1402–7. [PubMed: 11922310]

29. Itzhaki LS, Evans PA, Dobson CM, Radford SE. Tertiary interactions in the folding pathway of hen lysozyme: kinetic studies using fluorescent probes. Biochemistry. 1994; 33:5212–20. [PubMed: 8172895]

30. Lin JL, Ruaan RC, Hsieh HJ. Refolding of partially and fully denatured lysozymes. Biotechnol Lett. 2007; 29:723–9. [PubMed: 17310324]

31. Munoz IG, et al. Crystal structure of the open conformation of the mammalian chaperonin CCT in complex with tubulin. Nat Struct Mol Biol. 2011; 18:14–9. [PubMed: 21151115]

32. Clare DK, Bakkes PJ, van Heerikhuizen H, van der Vies SM, Saibil HR. Chaperonin complex with a newly folded protein encapsulated in the folding chamber. Nature. 2009; 457:107–10. [PubMed: 19122642]

33. Volokhina EB, et al. Role of the periplasmic Chaperones Skp, SurA, and DegQ in outer membrane protein biogenesis in Neisseria meningitidis. J Bacteriol. 2011; 193:1612–21. [PubMed: 21296967]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

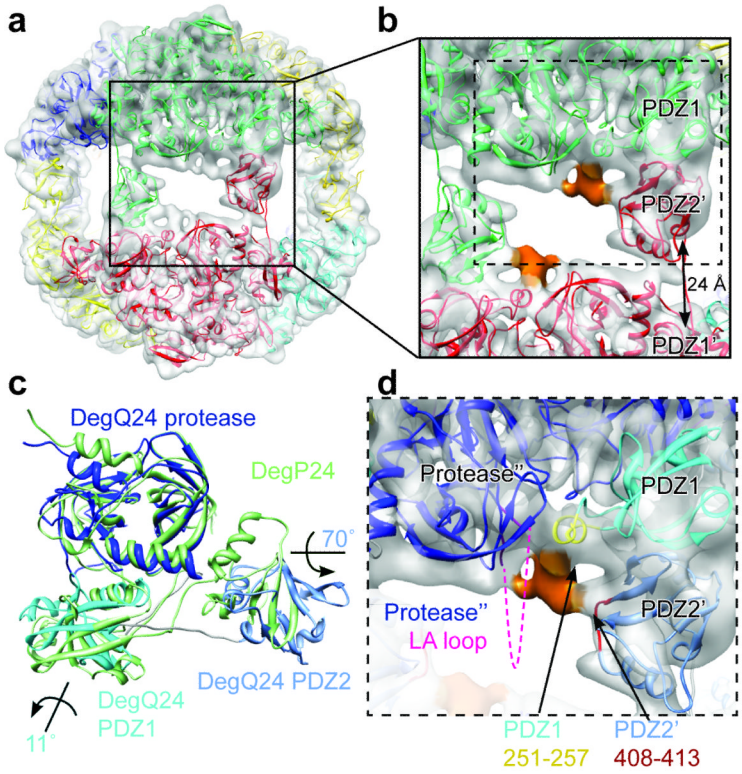

Figure 1. The DegQ 24-mer cryo-EM map reveals a potential β-casein binding site

(a) Two-fold view of the DegQ24 atomic model fitted into the 7.5 Å DegQ24–casein cryo-EM reconstruction with octahedral symmetry. Each trimer is displayed in a different color. **(b)** Enlargement of contact regions. Cage formation is mediated by a tight interaction between PDZ1 and PDZ2' domains from adjacent trimers. PDZ1 and PDZ2 domains from the same protomer are connected by an elongated linker spanning 24 Å. Density which is not accounted for by the fitted DegQ atomic model is shown in orange. **(c)** Overlay of DegP and Q protomers from the 24-mer cages, aligned via their protease domains. DegQ is colored by domain with PDZ1 in cyan, PDZ2 in blue and protease in dark blue and DegP is shown in light green. The 11° and 70° orientation differences observed between DegP/DegQ PDZ1 and PDZ2 domains are indicated. **(d)** Enlargement of the boxed area in **b**. DegQ is colored by domain as in **c**. Residues in the vicinity of the additional orange density belong to domains of three different protomers (labeled domain, domain' and domain"), namely helix 251-257 of PDZ1 (colored in yellow), loop 408-413 of PDZ2' (colored in brown) and potentially the LA loop of the protease" domain (residues 31-33 and 58-62 colored in magenta and the LA loop shown as a magenta dashed line).

Europe PMC Funders Author Manuscripts

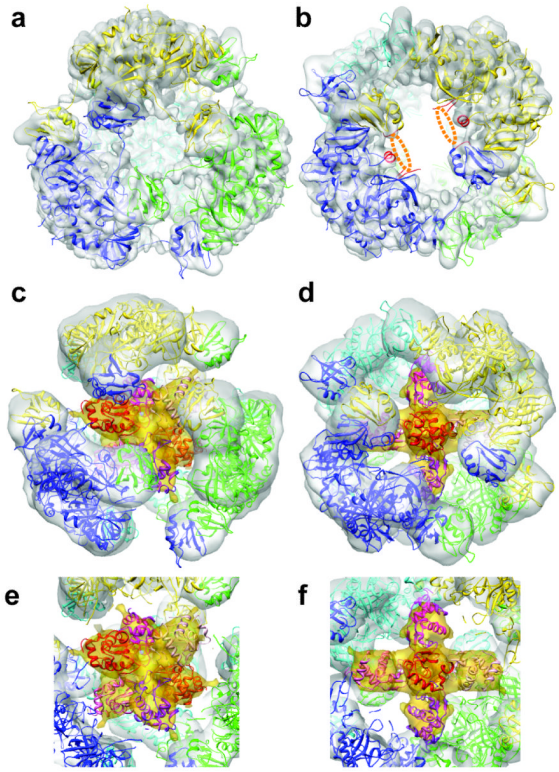

Figure 2. DegQ 12-mer complexes with a peptide and lysozyme

(a,b) Three-fold **(a)** and two-fold **(b)** views of the DegQ12 atomic model, colored by trimer, fitted into the 7.5 Å resolution DegQ12–peptide cryo-EM map with tetrahedral symmetry. Regions in the vicinity of the suggested β-casein density of DegQ 24-mer are shown in red. No substrate is visible at this position in the DegQ12–peptide complex, indicated by dashed orange ellipses. The position of the peptide C-terminus from the crystal structure[24] is indicated with a star (*). **(c,d)** Three-fold **(c)** and two-fold **(d)** views of the tetrahedral map of DegQ12–lysozyme at 13 Å resolution. The extra density relative to DegQ12–peptide is colored in orange. Six folded lysozymes are fitted into this additional density. They are colored in different shades of purple, magenta and orange. **(e,f)** Cut-away of three-fold **(e)** and two-fold view **(f)** highlighting the lysozyme fitting.

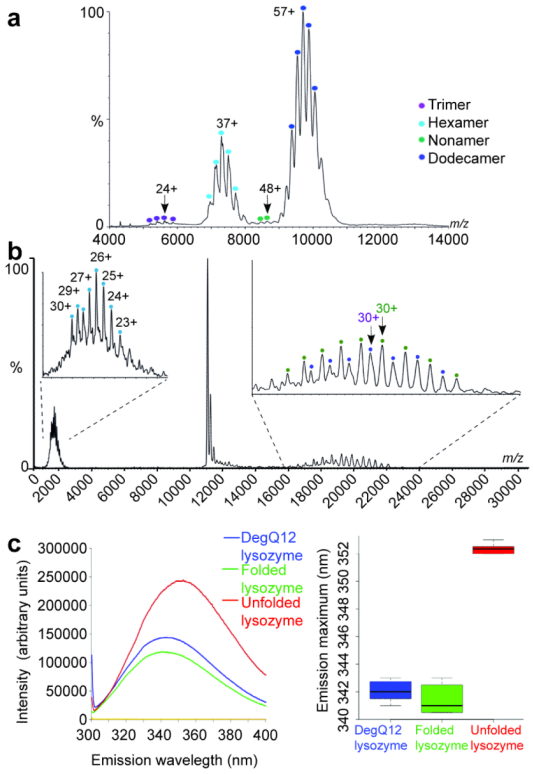

Figure 3. Five or six folded lysozymes are bound to the DegQ 12-mer

(a) Mass spectrum of the DegQ protein without substrate bound. Peaks corresponding to different multimeric states of apo-DegQ are labeled in different colors. The most abundant charge state for each of the peak series is indicated. **(b)** Tandem MS spectrum of the DegQ protein bound to lysozyme. The peak at m/z 11050 corresponds to the precursor ion that was selected for dissociation. The peak series in the lower m/z region corresponds to the ejected DegQ monomer. The two peak series at higher m/z represent charge-stripped DegQ 11-mer bound to 6 (major), and 5 (minor) lysozyme molecules, and are labeled in green and purple respectively. **(c)** Tryptophan fluorescence spectra (left panel) and box plot of the maximum emission wavelength (right panel) for folded lysozyme (green), unfolded lysozyme (red) and DegQ12–lysozyme (blue). As expected, there is no detectable fluorescence from DegQ alone (yellow), indicating that DegQ12–lysozyme signal corresponds to lysozyme fluorescence.

Europe PMC Funders Author Manuscripts

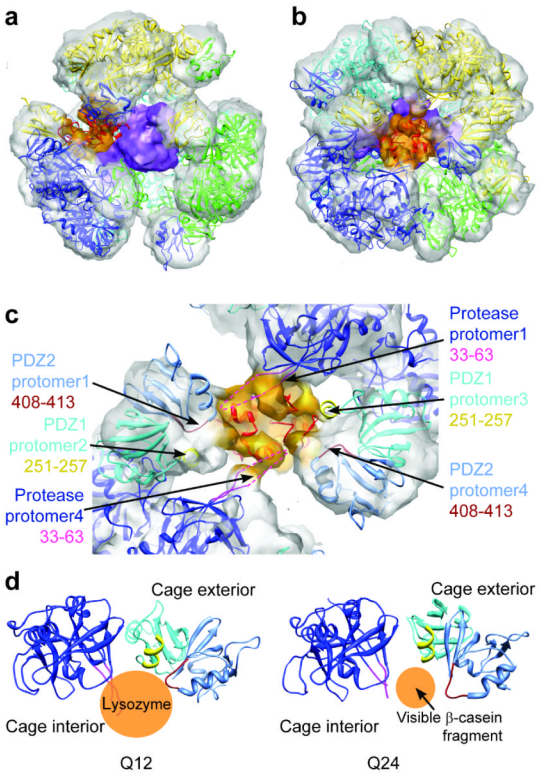

Figure 4. Lysozyme–DegQ interaction

(a,b) Three-fold **(a)** and two-fold **(b)** views of the asymmetric map of DegQ–lysozyme at 14.2 Å resolution. Two internal densities corresponding to lysozymes are visible inside DegQ and are colored in orange and purple. A folded lysozyme colored in red is fitted into the orange density. **(c)** Zoomed-in view of the interaction between DegQ and lysozyme in the same orientation as in **b**, with DegQ colored by domain as in Figure 1**d**. **(d)** Regions close to substrates are identical in DegQ 12-mers (left) and 24-mers (right). They involve protease, PDZ1 and PDZ2 domains from different subunits. Substrate positions are indicated by orange circles.

Europe PMC Funders Author Manuscripts

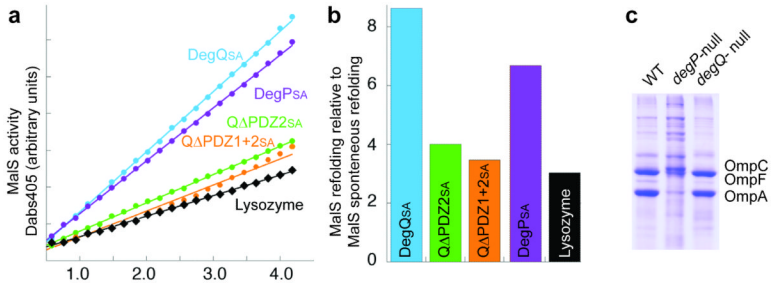

Figure 5. Chaperone activity of DegQ

(a) MalS refolding assay. The effects of protease-deficient $DegP_{S210A}$/$DegQ_{S187A}$ constructs on MalS refolding were examined: full length $DegQ_{SA}$ in blue, $DegQ_{SA}\Delta PDZ2$ in green, $DegQ_{SA}\Delta PDZ1+PDZ2$ in orange and full length $DegP_{SA}$ in purple. Lysozyme (in black) was used as a non-specific solute for the negative control. The linear time course of 405 nm absorbance indicates that the rate of MalS refolding is constant over time. **(b)** MalS refolding activity of different DegQ and DegP constructs, and lysozyme control, relative to spontaneous MalS refolding in buffer. **(c)** Steady-state levels of OmpA, OmpC and OmpF in wild type (WT), *degP*-null and *degQ*-null mutant strains. Outer membranes were prepared from equivalent numbers of wild-type, *degP*-null and *degQ*-null cells.

Europe PMC Funders Author Manuscripts