Structural analysis of the *Legionella pneumophila* Dot/Icm Type IV Secretion System Core Complex

Clarissa L. Durie[1]†, Michael J. Sheedlo[2]†, Jeong Min Chung[1], Brenda G. Byrne[3], Min Su[1], Thomas Knight[3], Michele S. Swanson[3*], D. Borden Lacy[2,4*], and Melanie D. Ohi[1,5*]

†These authors contributed equally to this work

*To whom correspondence should be addressed

Author Affiliations

Life Sciences Institute, University of Michigan, Ann Arbor MI, USA[1]

Department of Pathology, Microbiology, and Immunology, Vanderbilt University Medical Center, Nashville TN, USA[2]

Department of Microbiology and Immunology, University of Michigan, Ann Arbor MI, USA[3]

The Veterans Affairs Tennessee Valley Healthcare System, Nashville TN, USA[4]

Department of Cell and Developmental Biology, University of Michigan, Ann Arbor MI, USA[5]

Abstract

Legionella pneumophila is an opportunistic pathogen that causes the potentially fatal pneumonia Legionnaires' Disease. This infection and subsequent pathology require the Dot/Icm Type IV Secretion System (T4SS) to deliver effector proteins into host cells. Compared to prototypical T4SSs, the Dot/Icm assembly is much larger, containing ~27 different components including a core complex reported to be composed of five proteins: DotC, DotD, DotF, DotG, and DotH. Using single particle cryo-electron microscopy (cryo-EM), we report reconstructions of the core complex of the Dot/Icm T4SS that includes a symmetry mismatch between distinct structural features of the outer membrane cap (OMC) and periplasmic ring (PR). We present models of known core complex proteins, DotC, DotD, and DotH, and two structurally similar

proteins within the core complex, DotK and Lpg0657. This analysis reveals the stoichiometry and contact interfaces between the key proteins of the Dot/Icm T4SS core complex and provides a framework for understanding a complex molecular machine.

Introduction

The Type IV Secretion System (T4SS) is a potent weapon used by some bacteria to infect their host and can deliver effector proteins into eukaryotic cells as well as DNA and/or toxins into bacterial neighbors[1-3]. T4SSs are deployed by a variety of human pathogens, such as *Legionella pneumophila, Helicobacter pylori, and Bordetella pertussis*[1-3]. In the Gram negative pathogen *L. pneumophila*, the Dot/Icm Type IV Secretion System (T4SS)[4] delivers ~300 effector proteins to the cytoplasm of host cells, in some cases causing Legionnaires' Disease[5,6]. The T4SSs of Gram-negative bacteria are organized into an inner membrane complex, a core complex that spans the periplasmic space, and in some species an extracellular pilus[1,3,7-9]. These T4SSs vary in complexity with some species requiring only 12 components to assemble the complete apparatus. Some systems, such as the Dot/Icm T4SS of *L. pneumophila* and the Cag T4SS of *H. pylori*, are much larger and are constructed from over 20 different proteins, many of which are species-specific[10,11].

The current structural understanding of the large T4SSs has been limited to comparisons to minimized, prototype systems from *Xanthomonas citri* and the pKM101 conjugation system (referred to herein as pKM101). These studies have revealed homologous core structures that are comprised of three components known as VirB7, VirB9 and VirB10[10,12-14]. Previous cryo-electron tomography (cryo-ET) studies on the Dot/Icm T4SS have suggested a similar arrangement of some components of this system though no high-resolution data of the intact complex have been obtained to date[15-18].

Results & Discussion

Towards obtaining a high resolution understanding of the Dot/Icm T4SS we purified from *L. pneumophila* intact core complex particles as evident from negative stain electron microscopy, as previously described[19] (Figure 1 – figure supplement 1A). Central to assembly of the apparatus are five proteins that define the core complex: DotC, DotD, DotF, DotG, and DotH[15,16,20,21]. Mass spectrometry analysis of the purification verified the presence of these predicted core components [16,19,20,22], as well as additional proteins identified in *dot* (defect in organelle trafficking) or *icm* (intra-cellular multiplication) genetic screens[23-25] (Table 1). We vitrified this sample and, although particles adopt a preferred orientation in vitrified ice, both *en face* and side views are observed, allowing for 3D reconstruction (Figure 1A, Figure 1 – figure supplement 1B,C, and Figure 1 – figure supplement 2). The Dot/Icm T4SS is ~400 Å wide and ~165 Å long, consistent in shape and size with T4SS complexes visualized in intact *L. pneumophila* using cryo-ET[15-18] (Figure 1B). The global resolution of the map without imposed symmetry is 4.6 Å, with the highest resolution regions near its center (Figure 1 – figure supplement 1D,E). The map can be divided into two major regions: an outer membrane cap (OMC) and a hollow periplasmic ring (PR). The OMC can be further subdivided into two features, a central dome and a flat disk containing 13 arms that extend radially outward (Figure 1A,C,D). While cryo-ET analysis of the Dot/Icm T4SS in intact cells included a stalk bridging the PR and the inner membrane[16], this portion of the complex is not observed in the reconstruction of the purified T4SS, likely due to dissociation during purification. An axial section through the map in Figure 1B,D reveals a large cavity running through the T4SS, starting from the bottom of the PR and extending to the OMC region that spans the outer membrane, although there appears to be a density in the central cavity closest to the outer membrane. This is perhaps the "plug" seen in the cryo-ET analysis of *in situ* T4SS[16].

The dome of the T4SS OMC is positioned within the center of the map and is about ~50 Å high and ~100 Å wide. Attempts to refine the dome by imposing different symmetries did not improve the resolution; therefore no clear symmetry was defined. In other T4SSs that have

been structurally characterized, the dome is a contiguous part of the OMC, shares the same symmetry, and is clearly composed of organized α-helices[10,12,14]. While we do not see individual helices in our map, in the C1 reconstruction the narrow opening of the *L. pneumophila* dome is ~40 Å in diameter, a dimension within the range of pore sizes observed in the OMC of other species (Figure 1A)[10,12,14].

Using symmetry and focused refinement, we determined a 3.5 Å resolution map of the OMC disk and a 3.7 Å map of the PR (Figure 1C, Figure 1 – figure supplement 2, and Figure 1 – figure supplement 3). Notably, while the disk exhibits the expected 13-fold symmetry observed previously[15-18,26], the PR contains 18-fold symmetry (Figure 1C and Figure 1 – figure supplement 3A,D). While the possibility of this symmetry mismatch was postulated from low-resolution *in situ* structures of the Dot/Icm T4SS[18], the symmetry of the different regions was not determined. Interestingly, a similar symmetry mismatch occurs between the *H. pylori* T4SS OMC and PR: its OMC contains 14-fold symmetry while the PR has 17-fold symmetry[10]. The resolution of the Dot/Icm T4SS OMC disk and PR maps made it possible to construct models of the proteins in these regions (Figure 1C,D).

The OMC disk makes up the pinwheel-shaped portion of the T4SS and is organized into a thick central region with 13 arms extending radially outward (Figure 2A). The disk is ~75 Å along the axial dimension with an interior chamber ~150 Å wide. The disk is also thin compared to other structurally characterized T4SS OMCs and contains no distinct inner or outer layers[10,12,14]. Within the disk, we unambiguously traced and identified DotC, DotD, DotK and DotH along with the protein Lpg0657[27] (Figure 2B, Figure 2 – figure supplement 1 and 2). Of these, only structures of the C-terminal domain of DotD (PDB 3ADY) and the structure of Lpg0657 (PDB 3LDT) had been previously reported[28]. Notably, rather than an equimolar ratio, the components of the OMC exist at a ratio of 2:1:1:1:1 (DotD:DotC:DotH:DotK:Lpg0657) (Figure 2B).

At the center of the OMC is an elongated fold that is comprised of α-helices and β-strands that we have identified as DotC (residues 58-161 and 173-268). DotC is folded such that two large α-helices protrude toward the outer membrane and are flanked on either side by two β-strands (Figure 3A). The two β-strands adjacent to the central α-helices fold into a nearly uninterrupted β-sheet that is formed between asymmetric units and consists of a generally hydrophilic surface that runs about the central cavity (Figure 3B). When docked into the asymmetric reconstruction, this β-sheet lines the poorly resolved central section of the map. From these data, the only clear contact that is made between DotC and the central pore is a small interface at the top of the two long α-helices, which may explain why this portion is not well resolved in the maps (Figure 3C). A search of the protein data bank yielded a wide array of potential structural homologs, including a number of channels and transporters; most of these share little homology to DotC overall[29]. On the peripheral side, DotC makes contact with two copies of DotD, which we have called $DotD_1$ and $DotD_2$ (Figure 4A-C). The core folds of these proteins are similar to that of other components of large bacterial complexes such as VirB7 homologs from other T4SSs as well as components of type 4 pilus systems (Figure 4D). The interface between the two copies of DotD is mediated by electrostatic interactions and hydrogen bonds (Figure 4E). The N-terminus, which was not fully visualized in the previously reported DotD crystal structure[28], is α-helical and extends from the middle of the disk toward the pore, forming a dimer that interacts with the central α-helices of DotC (Figure 4F).

Adjacent to the N-terminus of DotD we have modelled a small globular fold that we have identified as DotK. DotK is positioned adjacent to the outer membrane, consistent with previous studies (Figure 5A,B)[16]. By chain tracing, we identified a second, similar fold near DotK. After docking the structure of DotK into the density, we noted that although the model fits well globally, some portions of the model were not supported by the density locally. Therefore, we initiated a DALI search for structurally similar molecules that identified Lpg0657 (PDB 3LDT) as a structural homolog and candidate for this density[29] (Figure 5A-D). Though Lpg0657 has not

been shown to directly interact with the Dot/Icm T4SS in prior studies, it is vital for *L. pneumophila* replication *in vitro*[27]. In fact, Goodwin and colleagues hypothesized that Lpg0657 might interact with the Dot/Icm T4SS[27]. Upon refining this model into the density, we noted that all features of Lpg0657 fit well into the density both globally and locally (Figure 5E). The presence of Lpg0657 is corroborated by mass spectrometry data of this sample (Table 1). The structures of both DotK and Lpg0657 resemble peptidoglycan binding domains[30]; however, density corresponding to peptidoglycan was not observed within the binding cleft in either DotK or Lpg0657, and several key residues known to mediate peptidoglycan interactions are not present. Thus, we suspect that neither protein mediates direct interaction with peptidoglycan. Notably, although the two proteins share a similar fold, they make contact with different members of the T4SS. DotK contacts the C-terminal domain of $DotD_2$, whereas Lpg0657 interacts with the N-terminal α-helices of both $DotD_1$ and $DotD_2$. This architecture is likely due to differences in the primary sequences of DotK and Lpg0657 that dictate the arrangement of these two similar proteins within the apparatus.

On the periplasmic side of the OMC is another small globular fold that we identified as the C-terminal domain of DotH. DotH consists of two β-sheets that are arranged in a β-sandwich fold (Figure 2B, 6A). Our subsequent structural search identified the VirB9 homolog TraO as the protein with the highest degree of structural similarity to DotH (PDB 3JQO, Figure 6B). Indeed, TraO harbors a conserved fold that is also observed in similar proteins such as VirB9 and CagX (Figure 6C). DotH could not have been predicted as a VirB9 homolog based on its primary structure, as very little conservation is observed between the two proteins (Figure 6D). Similar to its counterparts in other species[10,12-14], the C-terminal domain of DotH begins with an α-helix positioned near the center of the map which extends outward from the periplasm (Figure 6E).

Within the OMC disk we traced three poly-alanine chains that could not be unambiguously identified as any component of the Dot/Icm T4SS (Figure 2B, 7A). The first (chain 1) is a two-lobed polypeptide where ~70 residues form a series of loops (Figure 7B). This

unknown protein is positioned atop the OMC making contact with DotC, DotD, DotH, and presumably the outer membrane (Figure 7A). Located on the periphery of the map and extending outward radially, we have modelled a 22-strand β-helix (chain 2) (Figure 7C). Bound to this β-helix is a small fold consisting of six strands (chain 3). It is currently unclear if this domain represents an insertion in the β-helix or a distinct protein (Figure 7D). Based on previous reports it is tempting to speculate that the β-helix structure may correspond to DotG[16]. However, a register could not be identified in this part of the density due to low resolution, and, additionally, tomography studies have suggested the penta-peptide repeats reside in the stalk of the Dot/Icm T4SS[16].

To test whether DotG may be localized to the dome region, the radial arms, or both, we isolated and structurally characterized the T4SS in a *Lp* mutant lacking DotG (Δ*dotG*) (Figure 8A, Figure 8 – figure supplement 1, 2, and 3). Although Δ*dotG* mutant bacteria assemble a Dot/Icm T4SS, they are defective for secretion and replication in host cells[25]. Mass spectrometry from this mutant purification confirms DotG is absent (Table 2). The ΔDotG T4SS also lacks DotF (Table 2). Since Western blotting analysis of purified complexes from a Δ*dotG* mutant bacteria showed wild-type levels of DotF[20], the *dotG* deletion-insertion allele analyzed here may be polar on expression of the downstream *dotF* gene. Our structural analysis of ΔDotG T4SS shows that while it contains the OMC disk and the 13 extended arms, the complex lacks both the dome and the PR (Figure 8A and Figure 8 – figure supplement 1D). This finding is in agreement with the proposed model for the overall organization of both ΔDotG T4SS and ΔDotFΔDotG T4SS complexes predicted from immunoblot analysis and images of negatively stained T4SS complexes lacking either DotF or DotG[20]. All components modeled in the OMC from the wild-type T4SS were also present within the ΔDotG T4SS complexes, supporting the identifications described above (Table 2, Figure 8B). In other T4SSs, the dome region of the complex is comprised of homologous proteins known as VirB10 (*X. citri*), CagY (*H. pylori*), or TraF (pKM101). By sequence homology, the C-terminus of DotG is predicted to be structurally

similar to these components (Figure 8 – figure supplement 4A,C) [10,12-15,21,26]. Thus, we propose that the C-terminus of DotG makes up the dome of the *Lp* T4SS (Figure 8 – figure supplement 4B). In agreement with this, we note that a model of the C-terminus of DotG based on the structure of CagY (generated in Swiss Model) fits into the dome density (Figure 8 – figure supplement 4B), though its identity as DotG needs to be confirmed[31]. The rest of DotG and DotF may contribute to the structural interface between the OMC and the PR and/or form a portion of the structure of the PR.

Our predicted placement of DotF and DotG is consistent with two previous reports [16,22]. Biochemical studies showed that DotG associates closely with DotH and DotC, shown here to form part of the OMC (Figure 2B), and that both DotF and DotG are integral inner membrane proteins that also associate with the outer membrane[22]. A cryo-ET analysis of the T4SS in a Δ*dotG* strain reported missing density from the stalk, plug, and dome compared to subtomogram averages from T4SS complexes in a wild-type strain[16]. Moreover, these cryoET studies showed that the T4SS subtomogram averages in a Δ*dotF* strain lack density in the periplasmic region compared to complex in a wild-type strain[16].

The PR has been observed in the recently characterized single particle cryo-EM reconstruction of the *H. pylori* Cag T4SS and tomography studies of both *H. pylori* and *L. pneumophila* T4SSs[10,15-18,26,32]. The resolution in this region of our Dot/Icm T4SS map was sufficient to model two distinct polyalanine chains within the PR (Figure 9A,B). The backbone trace of one of these chains revealed a structure homologous to the C-terminus of *X. citri* VirB9 and similar to a polyalanine model of the PR constructed from the *H. pylori* T4SS (Figure 9C). The other, as of yet unidentified, density within the PR is comprised of a single α-helix followed by an extended loop that spans the entire length of the PR. Although the identity of either protein is currently not clear, prime candidates are either DotG or DotF, two proteins present in our preparations but not confidently localized in our maps (Table 1, Figure 2B). In addition, the entire PR is missing from the ΔDotG T4SS (Table 2, Figure 7A). As was reported previously for

208 the *H. pylori* Cag T4SS[10], while the *Lp* OMC and PR make physical contact in the lower
209 resolution map with no applied symmetry (Figure 9D), the connections are lost in the refined
210 structures due to the symmetry mismatch.

211 The high-resolution structure of the *L. pneumophila* Dot/Icm T4SS allows us to compare
212 the structural organization shared between the *H. pylori* Cag T4SS and the *L. pneumophila*
213 Dot/Icm T4SS (Figure 10). Both structures display different symmetry within the OMC and PR[10];
214 however, the symmetry mismatch itself is conserved, suggesting that this feature is important to
215 the function of secretion systems. However, many questions remain about which T4SS
216 components are most important for translocation efficiency, how substrates are recognized, and
217 how effectors are engaged. To address these remaining questions in the context of the Dot/Icm
218 T4SS, future studies need to characterize the presently unidentified chains and determine the
219 roles and locations of DotF and DotG in this complex. As our high-resolution structure of the
220 OMC revealed unexpected core components (DotK and Lpg0657), additional insights into the
221 stalk and coupling proteins of the Dot/Icm T4SS are also needed. This first high-resolution
222 structure of the Dot/Icm T4SS shows the importance of complex intermolecular interactions
223 between core components to build a large OMC and highlights the conservation of symmetry
224 mismatch in complex T4SSs, suggesting that both structural features are important for the
225 function of these very large transport systems.

226

227 **Materials and Methods**

Key Resources Table				
Reagent type (species) or resource	**Designation**	**Source or reference**	**Identifiers**	**Additional information**
strain, strain background (*Legionella pneumophila*)	Lp02; WT	PMID: 23717549		

strain, strain background (*Legionella pneumophila*)	Δ*dotG*	This paper		Δ*dotG* deletion-insertion mutant
Recombinant DNA reagent	pKD3 (plasmid)	Sigma Aldrich	RRID:AddGene_45604	template plasmids for frt-flanked cat cassette
sequence-based reagent	dotG-F	This paper	PCR primers	aaagcactccacctaagcctacag
sequence-based reagent	dotG-R	This paper	PCR primers	aaaaattagccaagcccgacctg
sequence-based reagent	dotG-P0	This paper	PCR primers	aaatcatgcaactcaaggtagaagggttataagcaaatgtgtgtaggctggagctgcttc
sequence-based reagent	dotG-P2	This paper	PCR primers	tatccgccatcaaattaaattgttgtaacatcctggcatatgaatatcctccttagttcc
commercial assay or kit	ProteoSilver ™ Plus Silver Stain Kit	Sigma Aldrich	PROTSIL2-1KT	
software, algorithm	Leginon	PMID: 15890530		
software, algorithm	MotionCor2	PMID: 28250466		
software, algorithm	CTFFind4	PMID: 26278980		
software, algorithm	cryoSPARC	PMID: 28165473		

software, algorithm	RELION	PMID: 27685097 PMID: 30412051		
software, algorithm	Coot	PMID: 20383002		
software, algorithm	UCSF Chimera	PMID: 15264254 PMID: 29340616		
software, algorithm	PHENIX	PMID: 29872004		
software, algorithm	DALI server	PMID: 31263867		

Preparation of Strains

L. pneumophila was cultured in ACES (Sigma)-buffered yeast extract broth at pH 6.9 supplemented with 0.1 mg/ml thymidine, 0.4 mg/ml l-cysteine, and 0.135 mg/ml ferric nitrate or on solid medium of this broth supplemented with 15 g/liter agar and 2 g/liter charcoal. The *L. pneumophila* laboratory strain Lp02, a thymidine auxotroph derived from the clinical isolate Philadelphia-1[33], was utilized as the wild-type strain. The *dotG* locus of Lp02 was replaced with a *cat* cassette encoding chloramphenicol resistance by homologous recombination as previously described[34]. The wild-type and Δ*dotG* alleles were amplified using primers dotG-F (5'-aaagcactccacctaagcctacag-3') and dotG-R (5'-aaaaattagccaagcccgacctg-3'). The *cat* cassette was amplified from plasmid pKD3[35] using primers dotG-P0 (5' aaatcatgcaactcaaggtagaagggttataagcaaatgtgtgtaggctggagctgcttc-3') and dotG-P2 (5'-tatccgccatcaaattaaattgttgtaacatcctggcatatgaatatcctccttagttcc-3'). The Δ*dotG* deletion-insertion mutant was selected and purified on medium supplemented with 5 μg/ml chloramphenicol.

Complex Isolation

Complexes were isolated from wild-type *L. pneumophila* strain Lp02 and the Δ*dotG* mutant strain as described[19,20]. Cells were suspended in 140 mL of buffer containing 150 mM Trizma base pH 8.0, 500 mM NaCl, and EDTA-free Complete protease inhibitor (Roche) at 4 ºC. The suspension was incubated on the benchtop, with stirring, until it reached ambient temperature. PMSF (final concentration 1 mM), EDTA (final concentration 1 mM), and lysozyme (final concentration 0.1 mg/mL) were added and the suspension was incubated at ambient temperature for an additional 30 minutes. Bacterial membranes were lysed using detergent and alkaline lysis. Triton X-100 (20% w/v) with AG501-X8 resin (BioRad) was added dropwise, followed by $MgSO_4$ (final concentration 3 mM), DNaseI (final concentration 5 μg/mL), and EDTA (final concentration 10 mM), and then the pH was adjusted to 10.0 using NaOH. The remaining steps were conducted at 4 ºC. The cell lysate was subjected to centrifugation at 12,000 x g for 20 minutes to remove unlysed material. The supernatant was then subjected to ultracentrifugation at 100,000 x g for 30 minutes to pellet membrane complexes. The membrane complex pellets were resuspended and soaked overnight in a small volume of TET buffer (10 mM Trizma base pH 8.0, 1 mM EDTA, 0.1% Triton X-100). The resuspended sample was then subjected to centrifugation at 14,000 x g for 30 minutes to pellet debris. The supernatant was subjected to ultra-centrifugation at 100,000 x g for 30 minutes. The resulting pellet was resuspended in TET and complexes were further separated by Superose 6 10/300 column chromatography in TET buffer with 150 mM NaCl using an AKTA Pure system (GE Life Sciences). The sample collected from the column was used for microscopy and visualized by SDS-PAGE with silver staining (ProteoSilver™ Plus Silver Stain Kit). Mass spectrometry analysis was performed as described [36].

Cryo-EM Data Collection and Map Reconstruction – Wild type T4SS

For cryo-EM, 4 μL of the isolated Dot/Icm T4SS sample was applied to a glow discharged ultrathin continuous carbon film on Quantifoil 2/2 200 mesh copper grids (Electron Microscopy Services). The sample was applied to the grid five consecutive times and incubated for ~ 60 s after each application. The grid was then rinsed in water to remove detergent before vitrification by plunge-freezing in a slurry of liquid ethane using a FEI vitrobot at 4°C and 100% humidity.

The images of the T4SS complexes from wild type cells were collected on the Thermo Fisher 300 kV Titan Krios with Gatan K2 Summit Direct Electron Detector (DED) camera having a nominal pixel size of 1.64 Å. Micrographs were acquired using Leginon software[37]. The total exposure time was 16 s, and frames were recorded every 0.2 s, resulting in a total accumulated dose of 58.1 e^{-} $Å^{-2}$ using a defocus range of -1 to -3 μm.

The video frames were first dose-weighted and aligned using Motioncor2[38]. The contrast transfer function (CTF) values were determined using CTFFind4[39]. Image processing was carried out using cryoSPARC and RELION 3.0[40-42]. Using the template picker in cryoSPARC, 771,806 particles were picked from 3,594 micrographs and extracted using a 510 pixel box size (1.64 Å/pixel). The extracted particles were used to generate representative 2D classes in cryoSPARC and approximately 20,800 particles were kept in good classes. These particles were then used for an *ab initio* model in cryoSPARC, which was then used as the reference for 3D auto-refinement with and without C13 symmetry (lowpass filtered to 40 Å). Finally, a solvent mask and B-factor were applied to improve the overall features and resolution of the 3D maps with and without C13 symmetry, resulting in reconstruction of 3D maps with a global resolution of 4.55 Å and 3.60 Å, respectively.

The C13 refined volume and corresponding particles were then exported to RELION for focused refinements. Estimation of beam-tilt values (CTF-refinement) was applied to the selected particles using RELION. With the CTF-refined particle stack, C13 symmetry-imposed

refinement with a soft mask around the core complex was done, resulting in a 5.0 Å resolution 3D map.

For focused refinement of the OMC disk, signal subtraction for each particle containing the OMC disk was used with a soft mask. The subtracted particles were subjected to alignment-free focused 3D classification (three classes). The best 3D class of the OMC (~12,200 particles) was then subjected to a masked 3D refinement with local angular searches using C13 symmetry resulting in a 4.60 Å resolution. Estimation of per-particle defocus values (CTF-refinement) was applied to the selected particles using RELION. With the CTF-refined particle stack, C13 symmetry-imposed refinement with a soft mask around the OMC disk region of the Dot/Icm T4SS core complex was done, resulting in a 4.33 Å resolution 3D map that contained improved features. Post-processing resulted in the final OMC disk map with 3.5 Å resolution.

The same steps were followed for focused refinement of the PR, starting with signal subtraction for each particle containing the PR with a soft mask. The subtracted particles were subjected to alignment-free focused 3D classification (three classes). The best 3D class of the PR (~6,850 particles) was selected based on class distribution (particle distribution), estimated resolution, and comparison of the 3D density maps. This class was then subjected to a masked 3D refinement with local angular searches using C18 symmetry resulting in a 7.54 Å resolution. Estimation of per-particle defocus values (CTF-refinement) was applied to the selected particles using RELION. With the CTF-refined particle stack, C18 symmetry-imposed refinement with a soft mask around the PR region of the Dot/Icm T4SS core complex was done, resulting in a 7.40 Å resolution 3D map that contained improved features. Post-processing resulted in the final PR map with 3.7 Å resolution.

Cryo-EM Data Collection and Map Reconstruction – ΔDotG T4SS

For cryo-EM, 4 μL of the isolated ΔDotG T4SS sample was applied to a glow discharged Quantifoil 2/2 200 mesh copper grid with ultrathin (2 nm) continuous carbon film (Electron Microscopy Services). The sample was applied to the grid five consecutive times and incubated for ~ 60 s after each application. The grid was rinsed in water to remove detergent before vitrification by plunge-freezing in a slurry of liquid ethane using a FEI vitrobot at 22 °C and 100% humidity.

The images of the T4SS complexes purified from the Δ*dotG* cells were collected by personnel at the National Center for CryoEM and Training (NCCAT) on the Thermo Fisher 300 kV Titan Krios with Gatan K2 Summit Direct Electron Detector (DED) camera having a nominal pixel size of 1.07 Å. Micrographs were acquired using Leginon software[37]. The total exposure time was 8 s and frames were recorded every 0.2 s, resulting in a total accumulated dose of ~65 e^{-} $Å^{-2}$ using a defocus range of -1.5 to -2.5 μm.

The video frames were first dose-weighted and aligned using Motioncor2[38]. The contrast transfer function (CTF) values were determined using CTFFind4[39]. Image processing was carried out using cryoSPARC and RELION 3.0[40-42]. 120,367 particles were picked manually using Relion from 6990 micrographs and extracted using a 640 pixel box size (1.07 Å/pixel). The particles were imported into cryoSPARC, and the remaining processing steps were performed in cryoSPARC. The extracted particles were used to generate representative 2D classes and 9,619 particles were kept in good classes. These particles were used to generate two 3D *ab initio* models with C13 symmetry, the better of which contained 6,342 particles. Homogeneous refinement of this model, also with C13 symmetry, resulted in a map with 4.2 Å resolution.

Model Building and Refinement

A model was constructed from the OMC disk by first tracing all chains within the asymmetric unit using Coot[43]. We identified two folds that were similar to the crystal structure of DotD and thus docked the corresponding crystal structure (PDB 3ADY) into the map using UCSF Chimera[28]. All other chains were then iteratively built *de novo* in Coot and refined in PHENIX[44]. During subsequent rounds of model building and refinement it was noted that a second fold which was similar to DotK was present in the EM map which could not be identified as any of the known core components. The structure of DotK was then subjected to a protein fold analysis using the DALI server which returned Lpg0657 as a potential candidate[29]. This crystal structure of Lpg0657 (PDB 3LDT) was then docked into the map using UCSF Chimera and the entire asymmetric unit was refined. The asymmetric unit was then duplicated in UCSF Chimera and each asymmetric unit docked into the map to generate a model of the entire OMC[45,46]. This structure was then refined in PHENIX with secondary structure and Ramachandran restraints applied[44]. During iterative rounds of refinement the nonbonded weighting parameter within PHENIX was optimized. A polyalanine model of the PR was constructed *de novo* in Coot and was refined using a similar protocol to that which was outlined above[43].

To generate a model of the entire core complex, maps from the focused refinement of the OMC disk and PR were aligned in UCSF Chimera using the asymmetric reconstruction as a guide. The maps were then combined in PHENIX using phenix.combine_focus_maps with the models of the OMC disk and PR provided as additional templates[44]. The entire complex was then subjected to a round of refinement in PHENIX with secondary structure and Ramachandran restraints applied. Figure 2 – figure supplement 1 shows the Fourier shell correlations (FSCs) of the half maps against the refined model agree with each other, suggesting that the models are not over-refined.

To model the OMC reconstructed from the Δ*dotG* strain, the OMC disk from the wild type strain was docked into the map using UCSF Chimera[45]. The model was then refined in PHENIX

following a similar protocol to that which was outlined above[44]. Figures were made in part using ChimeraX[46].

Acknowledgements

The work presented here was supported by NIH R01AI118932 (MDO), F32 AI150027-01 (CLD), NIH 2T32DK007673 (MJS), S10OD020011 and the University of Michigan Department of Microbiology & Immunology (Swanson). Some of this work was performed at the National Center for CryoEM Access and Training (NCCAT) and the Simons Electron Microscopy Center located at the New York Structural Biology Center, supported by the NIH Common Fund Transformative High Resolution Cryo-Electron Microscopy program (U24 GM129539,) and by grants from the Simons Foundation (SF349247) and NY State. A portion of the molecular graphics and analyses was performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at UC-San Francisco, with support from NIH P41-GM103311. Mass spectrometry experiments were performed by the University of Michigan Proteomics Resource Facility. The authors acknowledge Young-In (Eva) Kwon for assistance with figures. We thank the Cianfrocco, Cover, Lacy, and Ohi labs for helpful discussions. We acknowledge the use of the U-M LSI cryo-EM facility, managed by M Su, A Bondy, and L. Koepping, and U-M LSI IT support. We thank U-M BSI and LSI for significant support of the cryo-EM facility.

Data Availability

All cryo-EM data included in this manuscript are available through the Electron Microscopy Data Bank (EMD-22068, EMD-22069, EMD-22070 and EMD-22071). All models that were constructed from these data are available via the Protein Data Bank (PDB 6x62, 6x64, 6x65, and 6x66).

392
393

References

1 Grohmann, E., Christie, P. J., Waksman, G. & Backert, S. Type IV secretion in Gram-negative and Gram-positive bacteria. *Mol Microbiol* **107**, 455-471, doi:10.1111/mmi.13896 (2018).

2 Cascales, E. & Christie, P. J. The versatile bacterial type IV secretion systems. *Nat Rev Microbiol* **1**, 137-149, doi:10.1038/nrmicro753 (2003).

3 Fronzes, R., Christie, P. J. & Waksman, G. The structural biology of type IV secretion systems. *Nat Rev Microbiol* **7**, 703-714, doi:10.1038/nrmicro2218 (2009).

4 Schroeder, G. N. The toolbox for uncovering the functions of *Legionella* Dot/Icm type IVB secretion system effectors: Current state and future directions. *Front Cell Infect Microbiol* **7**, 528, doi:10.3389/fcimb.2017.00528 (2017).

5 Swanson, M. S. & Hammer, B. K. *Legionella pneumophila* pathogenesis: A fateful journey from amoebae to macrophages. *Annu Rev Microbiol* **54**, 567-613, doi:10.1146/annurev.micro.54.1.567 (2000).

6 Molofsky, A. B. & Swanson, M. S. Differentiate to thrive: Lessons from the *Legionella pneumophila* life cycle. *Mol Microbiol* **53**, 29-40, doi:10.1111/j.1365-2958.2004.04129.x (2004).

7 Waksman, G. From conjugation to T4S systems in Gram-negative bacteria: A mechanistic biology perspective. *EMBO Rep* **20**, doi:10.15252/embr.201847012 (2019).

8 Low, H. H., Gubellini, F., Rivera-Calzada, A., Braun, N., Connery, S., Dujeancourt, A., Lu, F., Redzej, A., Fronzes, R., Orlova, E. V. & Waksman, G. Structure of a type IV secretion system. *Nature* **508**, 550-553, doi:10.1038/nature13081 (2014).

9 Christie, P. J., Whitaker, N. & Gonzalez-Rivera, C. Mechanism and structure of the bacterial type IV secretion systems. *Biochim Biophys Acta* **1843**, 1578-1591, doi:10.1016/j.bbamcr.2013.12.019 (2014).

10 Chung, J. M., Sheedlo, M. J., Campbell, A. M., Sawhney, N., Frick-Cheng, A. E., Lacy, D. B., Cover, T. L. & Ohi, M. D. Structure of the *Helicobacter pylori* Cag type IV secretion system. *Elife* **8**, doi:10.7554/eLife.47644 (2019).

11 Purcell, M. & Shuman, H. A. The *Legionella pneumophila icmGCDJBF* genes are required for killing of human macrophages. *Infect Immun* **66**, 2245-2255 (1998).

12 Sgro, G. G., Costa, T. R. D., Cenens, W., Souza, D. P., Cassago, A., Coutinho de Oliveira, L., Salinas, R. K., Portugal, R. V., Farah, C. S. & Waksman, G. Cryo-EM structure of the bacteria-killing type IV secretion system core complex from *Xanthomonas citri*. *Nat Microbiol* **3**, 1429-1440, doi:10.1038/s41564-018-0262-z (2018).

13 Rivera-Calzada, A., Fronzes, R., Savva, C. G., Chandran, V., Lian, P. W., Laeremans, T., Pardon, E., Steyaert, J., Remaut, H., Waksman, G. & Orlova, E. V. Structure of a bacterial type IV secretion core complex at subnanometre resolution. *EMBO J* **32**, 1195-1204, doi:10.1038/emboj.2013.58 (2013).

14 Chandran, V., Fronzes, R., Duquerroy, S., Cronin, N., Navaza, J. & Waksman, G. Structure of the outer membrane complex of a type IV secretion system. *Nature* **462**, 1011-1015, doi:10.1038/nature08588 (2009).

15 Ghosal, D., Chang, Y. W., Jeong, K. C., Vogel, J. P. & Jensen, G. J. *In situ* structure of the *Legionella* Dot/Icm type IV secretion system by electron cryotomography. *EMBO Rep* **18**, 726-732, doi:10.15252/embr.201643598 (2017).

16 Ghosal, D., Jeong, K. C., Chang, Y. W., Gyore, J., Teng, L., Gardner, A., Vogel, J. P. & Jensen, G. J. Molecular architecture, polar targeting and biogenesis of the *Legionella* Dot/Icm T4SS. *Nat Microbiol* **4**, 1173-1182, doi:10.1038/s41564-019-0427-4 (2019).
17 Chetrit, D., Hu, B., Christie, P. J., Roy, C. R. & Liu, J. A unique cytoplasmic atpase complex defines the *Legionella pneumophila* type IV secretion channel. *Nat Microbiol* **3**, 678-686, doi:10.1038/s41564-018-0165-z (2018).
18 Park, D., Chetrit, D., Hu, B., Roy, C. R. & Liu, J. Analysis of Dot/Icm type IVB secretion system subassemblies by cryoelectron tomography reveals conformational changes induced by DotB binding. *mBio* **11**, doi:10.1128/mBio.03328-19 (2020).
19 Kubori, T. & Nagai, H. Isolation of the Dot/Icm type IV secretion system core complex from *Legionella pneumophila*. *Methods Mol Biol* **1921**, 241-247, doi:10.1007/978-1-4939-9048-1_15 (2019).
20 Kubori, T., Koike, M., Bui, X. T., Higaki, S., Aizawa, S. & Nagai, H. Native structure of a type IV secretion system core complex essential for *Legionella* pathogenesis. *Proc Natl Acad Sci U S A* **111**, 11804-11809, doi:10.1073/pnas.1404506111 (2014).
21 Nagai, H. & Kubori, T. Type IVB secretion systems of *Legionella* and other Gram-negative bacteria. *Front Microbiol* **2**, 136, doi:10.3389/fmicb.2011.00136 (2011).
22 Vincent, C. D., Friedman, J. R., Jeong, K. C., Buford, E. C., Miller, J. L. & Vogel, J. P. Identification of the core transmembrane complex of the Legionella Dot/Icm type IV secretion system. *Mol Microbiol* **62**, 1278-1291, doi:10.1111/j.1365-2958.2006.05446.x (2006).
23 Segal, G., Purcell, M. & Shuman, H. A. Host cell killing and bacterial conjugation require overlapping sets of genes within a 22-kb region of the *Legionella pneumophila* genome. *Proc Natl Acad Sci U S A* **95**, 1669-1674 (1998).
24 Segal, G. & Shuman, H. A. *Legionella pneumophila* utilizes the same genes to multiply within *Acanthamoeba castellanii* and human macrophages. *Infect Immun* **67**, 2117-2124 (1999).
25 Vogel, J. P., Andrews, H. L., Wong, S. K. & Isberg, R. R. Conjugative transfer by the virulence system of *Legionella pneumophila*. *Science* **279**, 873-876 (1998).
26 Hu, B., Khara, P., Song, L., Lin, A. S., Frick-Cheng, A. E., Harvey, M. L., Cover, T. L. & Christie, P. J. *In situ* molecular architecture of the *Helicobacter pylori* Cag type IV secretion system. *mBio* **10**, doi:10.1128/mBio.00849-19 (2019).
27 Goodwin, I. P., Kumova, O. K. & Ninio, S. A conserved OmpA-like protein in *Legionella pneumophila* required for efficient intracellular replication. *FEMS Microbiol Lett* **363**, doi:10.1093/femsle/fnw173 (2016).
28 Nakano, N., Kubori, T., Kinoshita, M., Imada, K. & Nagai, H. Crystal structure of *Legionella* DotD: Insights into the relationship between type IVB and type II/III secretion systems. *PLoS Pathog* **6**, e1001129, doi:10.1371/journal.ppat.1001129 (2010).
29 Holm, L. Benchmarking fold detection by DaliLite v.5. *Bioinformatics* **35**, 5326-5327, doi:10.1093/bioinformatics/btz536 (2019).
30 Lin, X., Ye, F., Lin, S., Yang, F., Chen, Z., Cao, Y., Chen, Z., Gu, J. & Lu, G. Crystal structure of PA0833 periplasmic domain from *Pseudomonas aeruginosa* reveals an unexpected enlarged peptidoglycan binding pocket. *Biochem Biophys Res Commun* **511**, 875-881, doi:10.1016/j.bbrc.2019.02.104 (2019).
31 Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R. & Schwede, T. Swiss-

Model: Homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296-W303, doi:10.1093/nar/gky427 (2018).
32 Chang, Y. W., Shaffer, C. L., Rettberg, L. A., Ghosal, D. & Jensen, G. J. *In vivo* structures of the *Helicobacter pylori* Cag type IV secretion system. *Cell Rep* **23**, 673-681, doi:10.1016/j.celrep.2018.03.085 (2018).
33 Rao, C., Benhabib, H. & Ensminger, A. W. Phylogenetic reconstruction of the *Legionella pneumophila* Philadelphia-1 laboratory strains through comparative genomics. *PLoS One* **8**, e64129, doi:10.1371/journal.pone.0064129 (2013).
34 Bryan, A., Abbott, Z. D. & Swanson, M. S. Constructing unmarked gene deletions in *Legionella pneumophila*. *Methods Mol Biol* **954**, 197-212, doi:10.1007/978-1-62703-161-5_10 (2013).
35 Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640-6645, doi:10.1073/pnas.120163297 (2000).
36 Anwar, T., Arellano-Garcia, C., Ropa, J., Chen, Y. C., Kim, H. S., Yoon, E., Grigsby, S., Basrur, V., Nesvizhskii, A. I., Muntean, A., Gonzalez, M. E., Kidwell, K. M., Nikolovska-Coleska, Z. & Kleer, C. G. p38-mediated phosphorylation at T367 induces EZH2 cytoplasmic localization to promote breast cancer metastasis. *Nat Commun* **9**, 2801, doi:10.1038/s41467-018-05078-8 (2018).
37 Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C. S. & Carragher, B. Automated molecular microscopy: The new Leginon system. *J Struct Biol* **151**, 41-60, doi:10.1016/j.jsb.2005.03.010 (2005).
38 Zheng, S. Q., Palovcak, E., Armache, J. P., Verba, K. A., Cheng, Y. & Agard, D. A. MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332, doi:10.1038/nmeth.4193 (2017).
39 Rohou, A. & Grigorieff, N. CtfFind4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192**, 216-221, doi:10.1016/j.jsb.2015.08.008 (2015).
40 Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. CryoSPARC: Algorithms for rapid unsupervised cryo-em structure determination. *Nat Methods* **14**, 290-296, doi:10.1038/nmeth.4169 (2017).
41 Bharat, T. A. & Scheres, S. H. Resolving macromolecular structures from electron cryo-tomography data using subtomogram averaging in RELION. *Nat Protoc* **11**, 2054-2065, doi:10.1038/nprot.2016.124 (2016).
42 Zivanov, J., Nakane, T., Forsberg, B. O., Kimanius, D., Hagen, W. J., Lindahl, E. & Scheres, S. H. New tools for automated high-resolution cryo-em structure determination in RELION-3. *Elife* **7**, doi:10.7554/eLife.42166 (2018).
43 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).
44 Afonine, P. V., Poon, B. K., Read, R. J., Sobolev, O. V., Terwilliger, T. C., Urzhumtsev, A. & Adams, P. D. Real-space refinement in PHENIX for cryo-em and crystallography. *Acta Crystallogr D Struct Biol* **74**, 531-544, doi:10.1107/S2059798318006551 (2018).
45 Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612, doi:10.1002/jcc.20084 (2004).

46 Rodriguez-Guerra Pedregal, J. & Marechal, J. D. PyChimera: Use UCSF Chimera modules in any Python 2.7 project. *Bioinformatics* **34**, 1784-1785, doi:10.1093/bioinformatics/bty021 (2018).

Table 1. Dot proteins present in isolated complex sample isolated from wild type strain (Lp02)

Identified Proteins	Gene Number[a]	Spectral Counts[b] Prep 1	Prep 2	Prep 3
DotG[c]	Q5ZYC1	112	114	195
DotF	Q5ZYC0	94	69	101
DotA	Q5ZS33	38	65	60
DotO	Q5ZYB6	37	38	47
DotH	Q5ZYC2	28	19	28
IcmF	Q5ZYB4	15	18	36
IcmX	Q5ZS30	19	13	28
DotL	Q5ZYC6	10	26	20
DotC	Q5ZS44	9	11	16
DotD	Q5ZS45	11	9	14
DotB	Q5ZS43	16	11	7
Lpg0657	Q5ZXS4	6	4	16
DotM	Q5ZYC7	2	14	3
DotK	Q5ZYC5	2	6	10
IcmW	Q5ZS31	5	7	5
DotN	Q5ZYB7	2	6	4
DotI	Q5ZYC3	1	3	5
IcmS	Q5ZYD0		4	1
IcmT	Q5ZYD1	1	1	3
IcmV	Q5ZS32	1	1	2

[a]UniProtKB Accession Number

[b]Proteins were identified by searching the MS/MS data against *L. pneumophila* (UniProt; 2930 entries) using Proteome Discoverer (v2.1, Thermo Scientific). Search parameters included MS1 mass tolerance of 10 ppm and fragment tolerance of 0.1 Da. False discovery rate (FDR) was determined using Percolator and proteins/peptides with a FDR of ≤1% were retained for further analysis. Complete results are in Supplementary File 1 in supplementary material.

[c]Predicted core T4SS components and additional components identified in this structure are in bold.

Table 2. Dot proteins present in isolated complex sample isolated from Δ*dotG* deletion strain

Identified Proteins	Gene Number[a]	Spectral Counts[b] Prep 1	Prep 2	Prep 3
IcmF	Q5ZYB4	28	85	72
DotA	Q5ZS33	20	94	63
IcmX	Q5ZS30	15	88	25
DotH[c]	Q5ZYC2	11	81	25
DotC	Q5ZS44	12	58	17
DotO	Q5ZYB6	16	27	38
Lpg0657	Q5ZXS4	7	43	13
DotL	Q5ZYC6	8	15	32
DotB	Q5ZS43	10	23	13
DotK	Q5ZYC5	6	26	6
DotD	Q5ZS45	3	27	6
IcmW	Q5ZS31	5	9	9
DotI	Q5ZYC3	1	10	4
IcmV	Q5ZS32	2	1	11
DotU	Q5ZYB3	1	5	2
DotM	Q5ZYC7	0	2	6
IcmS	Q5ZYD0	2	2	4

[a]UniProtKB Accession Number

[b]Proteins were identified by searching the MS/MS data against *L. pneumophila* (UniProt; 2930 entries) using Proteome Discoverer (v2.1, Thermo Scientific). Search parameters included MS1 mass tolerance of 10 ppm and fragment tolerance of 0.1 Da. False discovery rate (FDR) was determined using Percolator and proteins/peptides with a FDR of ≤1% were retained for further analysis. The complete results are shown in Supplementary File 2 in the supplemental material.

[c]Predicted core T4SS components and additional components identified in this structure are in bold.

Figures

Figure 1. Cryo-EM structure of the *L. pneumophila* Dot/Icm T4SS. (a) Reconstruction of the *L. pneumophila* Dot/Icm T4SS particles at 4.6 Å with no symmetry applied reveals two parts, the outer membrane cap (OMC), composed of a central dome and flat disk, and a periplasmic ring (PR). (b) Comparison of the central sections through the longitudinal plane of the T4SS 3D density determined by cryo-ET of intact *L. pneumophila* (left panel, EMD 0566)[16] with arrow indicating plug density or by cryo-EM of purified particles (this study) (right panel). OM, Outer Membrane, IM, Inner Membrane, Scale bar 10 nm. (c) Combined high resolution structures of the *L. pneumophila* Dot/Icm T4SS that include the 3.5 Å OMC disk (blue) with 13-fold symmetry and the 3.7 Å PR (green) with 18-fold symmetry. (d) Central axial slice view showing how atomic models of the OMC disk (blue) and PR (green) fit into the C1 3D map of the Dot/Icm T4SS (light gray).

Figure 1 – figure supplement 1 – Asymmetric reconstruction of the Dot/Icm T4SS. (a) A representative image of the *Lp* Dot/Icm T4SS particles in negative stain EM (scale bar = 50 nm) and subsequent 2D class averages (scale bar = 20 nm). (b) A representative image of the *Lp* Dot/Icm T4SS particles in cryo-EM (scale bar = 50 nm) and subsequent 2D class averages (scale bar = 10 nm). (c-d) The Dot/Icm T4SS map was reconstructed with no symmetry applied to a resolution of 4.6 Å and reveals the outer membrane cap (OMC), composed of a dome and a flat disk, and the periplasmic ring (PR). (d) The local resolution of the C1 map extends to ~4.0 Å in the best resolved portions of the map.

Figure 1 – figure supplement 2 – Flow chart of cryo-EM processing steps. Using template picking in cryoSPARC, ~770,000 Dot/Icm T4SS particles were selected. The processing steps done in cryoSPARC are on a gray background, and the processing steps done using RELION are on a tan background. After 2D and 3D classification, the best class of ~20,000 particles was chosen for further refinement with and without C13 symmetry, resulting in reconstruction of 3D maps at 4.6 Å and 3.6 Å, respectively. The particle stack used in the refinement job in cryoSPARC was exported into Relion for further processing, such as CTF-refinement and focused refinement. The 3.6 Å 3D model with C13 symmetry was used as an initial model for 3D structure determination in Relion using auto-3D refinement with C13 symmetry, resulting in reconstruction of 3D map of 5.5 Å resolution. The CTF-refinement beam tilt estimation was applied, followed by another round of 3D refinement, improving the resolution to 5.0 Å. Focused 3D classification (without alignment) was used to determine higher resolution maps of the OMC disk (with 13-fold symmetry) and the PR (with 18-fold symmetry). The maps of the OMC disk and PR were further refined using focused 3D refinement (with local refinement), resulting in 3D maps of the OMC disk (13-fold symmetry) and the PR (18-fold symmetry) at 3.5 Å and 3.7 Å, respectively.

Figure 1 – figure supplement 3 – Symmetric reconstructions of the OMC and PR of the Dot/Icm T4SS. (a-b) The OMC disk has been reconstructed to 3.5 Å resolution using C13 symmetry. The cryo-EM density map of the Dot/Icm T4SS OMC disk showing the side, bottom, and top faces of the OMC. (c) The local resolution of the OMC extends to 3.0 Å in the best resolved portions of the map. (d-e) The PR has been reconstructed to 3.7 Å using C18 symmetry. (f) The local resolution of the PR extends to ~4.0 Å in the best resolved regions of the map. (g) A channel runs through the center of the Dot/Icm T4SS that is ~150 Å in diameter in the OMC and ~100 Å in diameter in the PR.

Figure 2. The OMC disk of the Dot/Icm T4SS. (a) The OMC disk of the Dot/Icm T4SS was reconstructed from samples that were purified from the WT strain (b) We have defined an

asymmetric unit that is comprised of DotC (brown), $DotD_1$ (red), $DotD_2$ (salmon), DotH (orange), DotK (cyan), Lpg0657 (blue), and three unknown chains (gray).

Figure 2 – figure supplement 1 – Correlation between the OMC cryo-EM map and models. (a) Statistics for each model that was constructed indicate their relative quality. (b) A model-map correlation curve for the entire OMC is shown with the unmasked data shown in green and the masked data shown in black. (c-i) The model-map correlation was also determined for each individual protein structure that was solved within the OMC.

Figure 2 - figure supplement 2 – Model-map correlation for each protein within the OMC. (a-f) The correlation coefficient for each protein within the OMC is depicted as a heat map on the corresponding structure and is plotted against the residue number.

Figure 3. DotC forms a nearly uninterrupted β-sheet about the center of the complex. (a) DotC is composed of two large α-helices and two β-sheets. (b)These sheets are formed between asymmetric units and comprise a predominantly electrostatic structure that lines the central chamber. (c) DotC is adjacent to the poorly resolved portion of the C1 map (circled with a red dotted line, left). The positioning of DotC is such that a gap exists between it and the asymmetric portion of the map (brown dotted line, center). This results in a small predicted interface between DotC and the central dome (right).

Figure 4 – Components in the center of the Dot/Icm T4SS. (a-b) On the peripheral side, DotC is observed with two copies of DotD, shown in red and salmon. (c) The two copies of DotD vary little in their overall organization, deviating by an RMSD of only 0.6 Å within the C-terminal domain as shown in the inset. (d) Both copies of DotD observed within this map adopt a fold that is similar to the previously reported crystal structure as well as related VirB7 homologs. (e) The interface that is formed between $DotD_1$ and $DotD_2$ is formed by a number of electrostatic and polar interactions as shown in the inset. (f) The N-terminal α-helices of $DotD_1$ and $DotD_2$ form the bulk of the interface observed with DotC, made up of a hydrophobic interface as shown in the inset.

Figure 5 – Components of the Dot/Icm T4SS near the outer membrane. (a) Two nearly identical folds were discovered on the outer membrane side of the OMC that correspond to DotK (cyan) and Lpg0657 (blue). (b) DotK and Lpg0657 arrange such that each contacts both copies of DotD within the asymmetric unit, although they use different surfaces to mediate this interaction. (c) The two folds have a nearly identical topology. Proteins colored as in panel (a). (d) Although these folds are very similar, at least three regions could be used to distinguish between the two proteins in this map. (e) Characteristic density is observed for each protein and supports the assignment of each protein within the map.

Figure 6 – DotH is a VirB9 homolog and positioned in the center of the map. (a) The C-terminal domain of DotH was discovered within the center of the OMC map and consists of a β-sandwich fold. (b) The structure is similar to other T4SS proteins such as the C-termini of TraO, VirB9, and CagX. (c) Though the structures are similar, little sequence similarity (d) is observed throughout the family. (e) DotH is positioned such that the N-terminus (consisting of 272 residues) is positioned between the OMC and PR.

Figure 7 – Three unknown components were modelled as polyalanine chains. (a) Three chains were constructed as polyalanine models, denoted as chain 1-3 and shown in gray. (b) Chain 1 is observed near the OMC and consists loops forming two lobes. (c) Chain 2 consists of

a long β-helix that interacts with chain 3. (d) Chain 3 is a small globular fold that is comprised of only β-strands.

Figure 8. The OMC disk of the ΔDotG T4SS. (a)The OMC disk of the ΔDotG T4SS was reconstructed from samples that were purified from a strain lacking DotG (Δ*dotG*). (b) All proteins within the asymmetric unit adopt nearly identical orientations in complexes identified from the WT and mutant strain. (Colored as in panel a).

Figure 8 – figure supplement 1 – 3D reconstruction of the T4SS purified from a Δ*dotG* mutant. (a) A representative image of the *Lp* ΔDotG T4SS particles in cryo-EM (scale bar = 50 nm) and (b) subsequent 2D reconstructions (Scale bar = 10 nm). (c) Silver-stained gel from the T4SS samples used in this study. Lanes (L to R) are molecular weight markers, wild type, and ΔDotG complexes. (d-e) The map was reconstructed with C13 symmetry applied to a resolution of 4.2 Å and reveals only the outer membrane cap (OMC) disk region (f) The local resolution of the map extends to ~4.1 Å in the best resolved portions of the map.

Figure 8 – figure supplement 2 – Correlation between the *Lp* ΔDotG OMC cryo-EM map and models. (a) Statistics for each model that was constructed indicate their relative quality. (b) A model-map correlation curve for the entire *Lp* ΔDotG OMC is shown with the unmasked data shown in green and the masked data shown in black. (c-i) The model-map correlation was also determined for each individual protein structure that was solved within the *Lp* ΔDotG OMC.

Figure 8 - figure supplement 3 – Model-map correlation for each protein within the *Lp* ΔDotG OMC. (a-f) The correlation coefficient for each protein within the *Lp* ΔDotG OMC is depicted as a heat map on the corresponding structure and is plotted against the residue number.

Figure 8 – figure supplement 4 – Homology model of proposed dome component DotG. (a) The C-terminus of DotG is one of the few components of the Dot/Icm T4SS that is predicted to have homology to other systems, in this case to *H. pylori* CagY and *X. citri* VirB10. (b) Atop the OMC is a low-resolution dome that contains no clear symmetry (gray). A model of the C-terminus of DotG with C13 symmetry applied fits well into this dome suggesting a pore size of ~40 Å (yellow). (c) Sequence homology suggests the C-terminus of DotG is structurally similar to *H. pylori* CagY and *X. citri* VirB10.

Figure 9. The PR of the Dot/Icm T4SS. (a) The Dot/Icm T4SS is comprised of two features known as the outer membrane cap (OMC) and the periplasmic ring (PR). Within the asymmetric unit of the PR we have modelled two chains (shown in yellow and green). (b) Chain 1 of the PR, shown in yellow, consists of two β-sheets and two α-helices. Chain 2, shown in green, consists of one β-strand and one α-helix. (c) Chain 1 of the Dot/Icm T4SS contains a globular fold that is similar to the C-terminus of VirB9 from *X. citri* and the core fold of the polyalanine model that was modelled in the PR of the Cag T4SS from *H. pylori.* (d) The physical connection between the OMC disk (blue) and PR (green) is shown with atomic models fit into the cryoEM reconstruction with no symmetry applied (gray).

Figure 10. Comparison of *H. pylori* Cag T4SS and *L. pneumophila* Dot/Icm T4SS. (a) Surface view of atomic models for *Hp* Cag T4SS (left) and *Lp* Dot/Icm T4SS (right). OMC, blue; PR, green. For *Lp* Dot/Icm T4SS, C1 EM density is shown in gray for density not included in atomic models. (b) Secondary structure of *Hp* Cag T4SS (left) and *Lp* Dot/Icm T4SS (right). For *Lp* Dot/Icm T4SS, dome region outline is shown as a dotted line.

Additional Data

Table 1 – Map Reconstruction and Model Refinement

Cryo-EM data collection and processing

EMDB Accession Codes	Dot/Icm T4SS OMC WT EMD-22068	Dot/Icm T4SS PR WT EMD-22069	Dot/Icm T4SS EMD-22070	Dot/Icm T4SS OMC ΔDotG EMD-22071
Data Collection and Processing				
Magnification	18,000	18,000		29,000
Voltage (kV)	300	300		300
Total Electron Dose (e-/Å2)	58.1	58.1		~65.0
Defocus Range (μM)	-1 ~ -3	-1 ~ -3		-1.5 ~ -2.5
Pixel Size (Å)	1.64	1.64		1.07
Processing Software	cryoSPARC / Relion	cryoSPARC / Relion		cryoSPARC / Relion
Symmetry	C13	C18		C13
Initial Particles (number)	771,806	771,806		120,367
Final Particles (number)	12,200	6,850		6,342
Map Sharpening B Factor	-40.63	-52.51		-55.30
Map Resolution (Å)	3.5	3.7		4.2
FSC Threshold	0.143	0.143		0.143

Model Refinement and Validation

Refinement				
Initial Model Used	3ADY, 3LDT	N/A	N/A	N/A
Model Resolution				
FSC (0.5)	3.4	3.9	3.6	7.3
FSC (0.143)	3.2	3.6	3.2	4.2
	15,548	3,168	18,716	15,548
Model Composition (Residues)				
DotC	58-161, 173-268	-	58-161, 173-268	58-161, 173-268
$DotD_1$	25-159	-	25-159	25-159
$DotD_2$	2-161	-	2-161	2-161
DotH	273-361	-	273-361	273-361
DotK	41-180	-	41-180	41-180
Lpg0657	99-234	-	99-234	99-234
Unk OMC Chain 1	1-70	-	1-70	1-70
Unk OMC Chain 2	1-127	-	1-127	1-127
Unk OMC Chain 3	1-52	-	1-52	1-52
Unknown PR	-	1-32, 1-144	1-32, 1-144	-
Bond RMSD				
Bond Length (Å)	0.014	0.007	0.01	0.016
Bond Angle (°)	1.868	1.06	1.27	1.723
Validation				
Molprobity Score	2.41	1.17	2.70	2.44
Clashscore	6.90	0.47	6.94	21.63
Ramachandran (%)				
Favored	91.5	91.5	92.0	87.7
Allowed	8.5	8.5	8.0	12.3
Outliers	0.0	0.0	0.0	0.0
B Factor	94.6	60.7	152.1	341.9
PDB Accession Codes	6X62	6X64	6X65	6X66

Supplementary File 1. Proteins that copurify with Dot T4SS

[a]UniProtKB Accession Number

[b]Proteins were identified by searching the MS/MS data against *L. pneumophila* (UniProt; 2930 entries) using Proteome Discoverer (v2.1, Thermo Scientific). Search parameters included MS1 mass tolerance of 10 ppm and fragment tolerance of 0.1 Da. False discovery rate (FDR) was determined using Percolator and proteins/peptides with a FDR of ≤1% were retained for further analysis. The complete results are shown in Supplementary File 2 in the supplemental material.

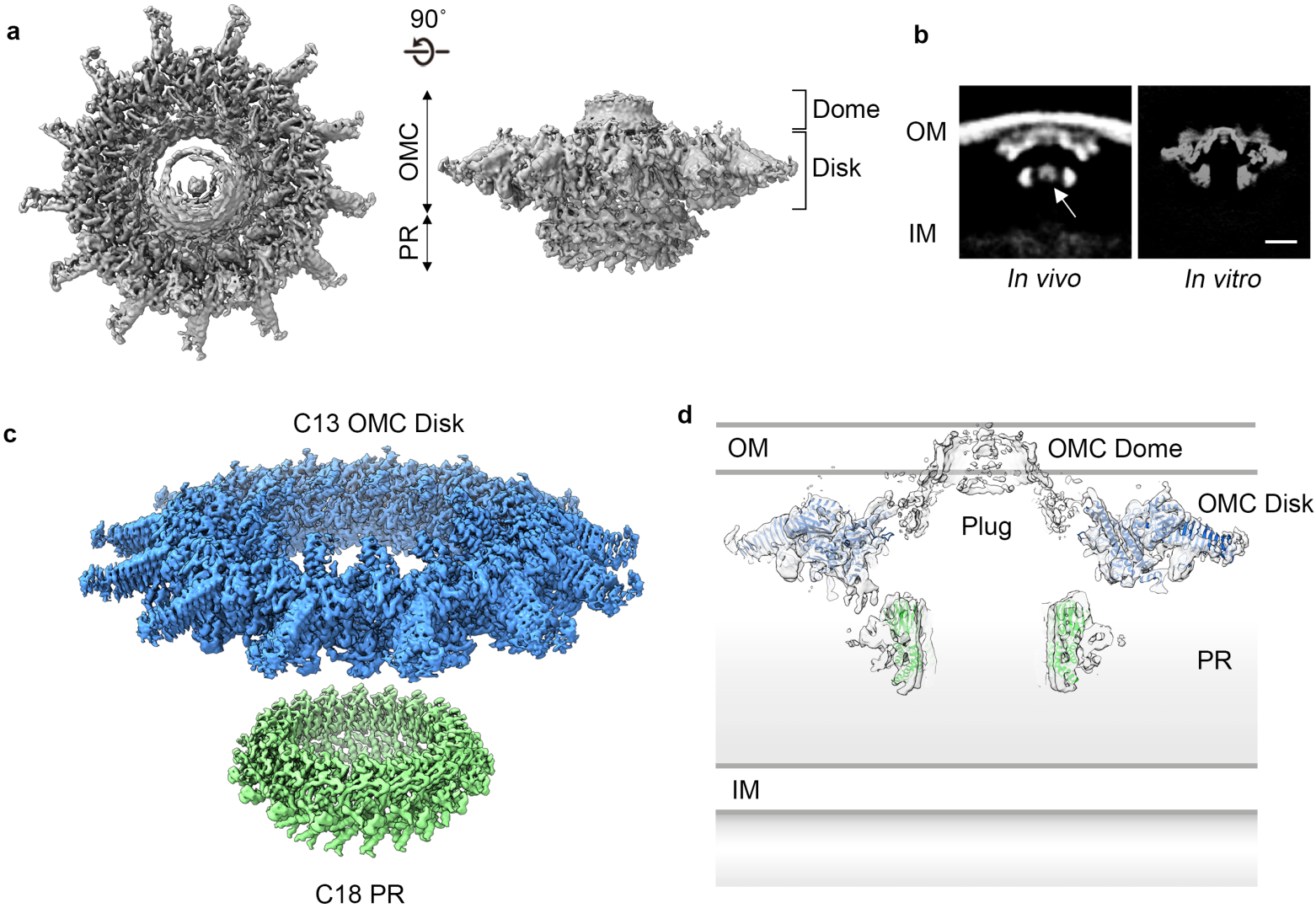
[c]Dot/Icm T4SS components are in bold.

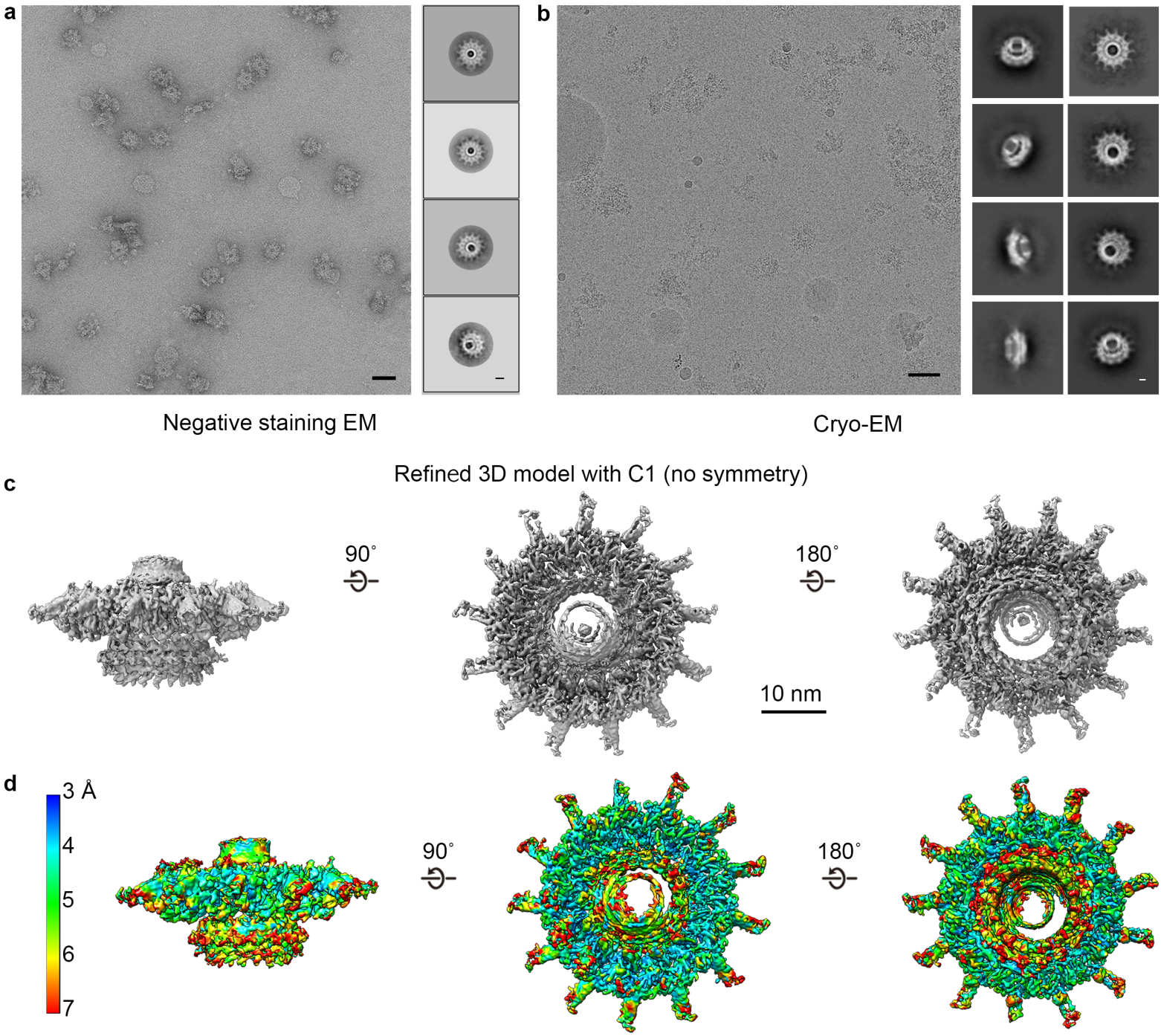
Supplementary File 2. Proteins that copurify with ΔDotG T4SS

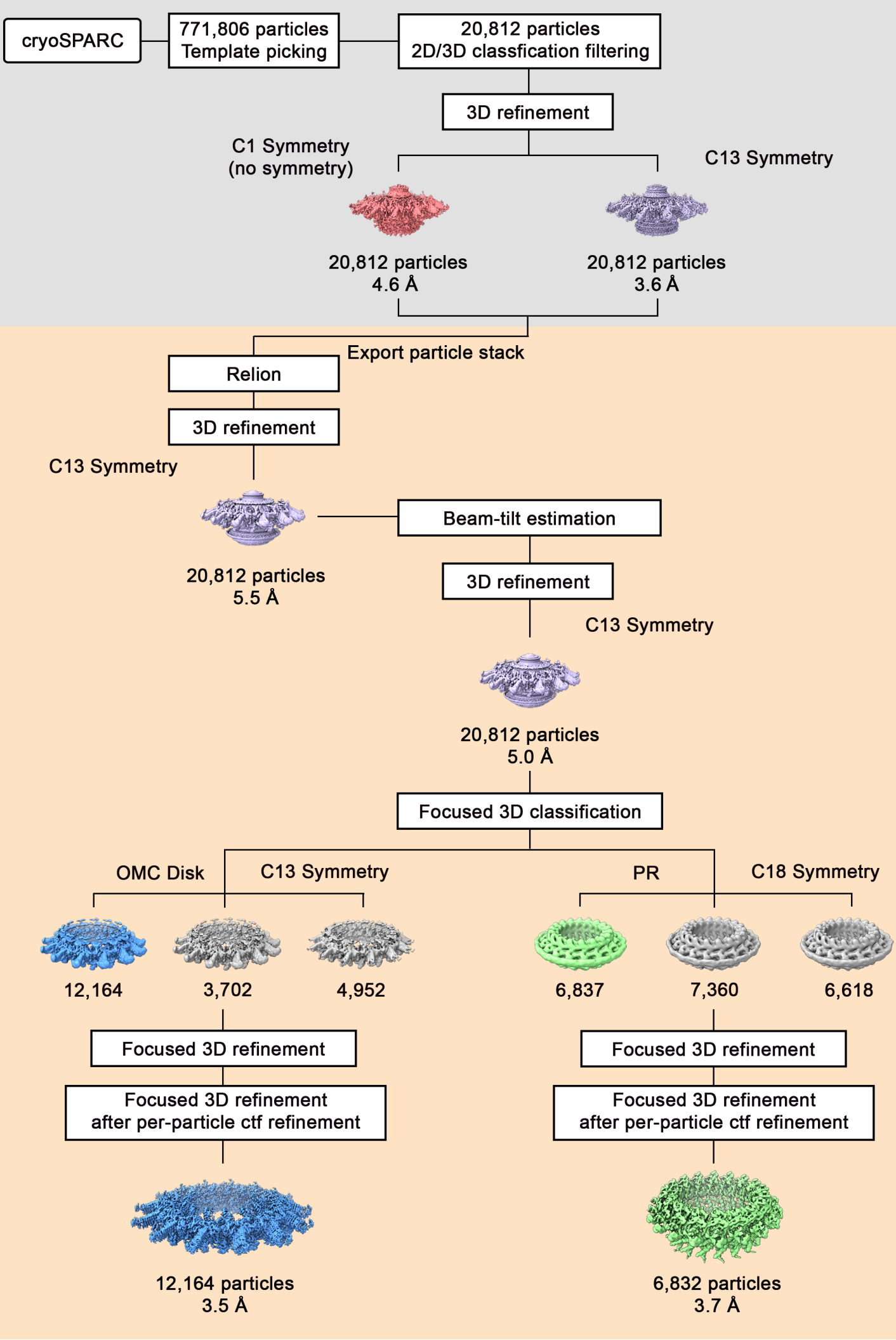
[a]UniProtKB Accession Number

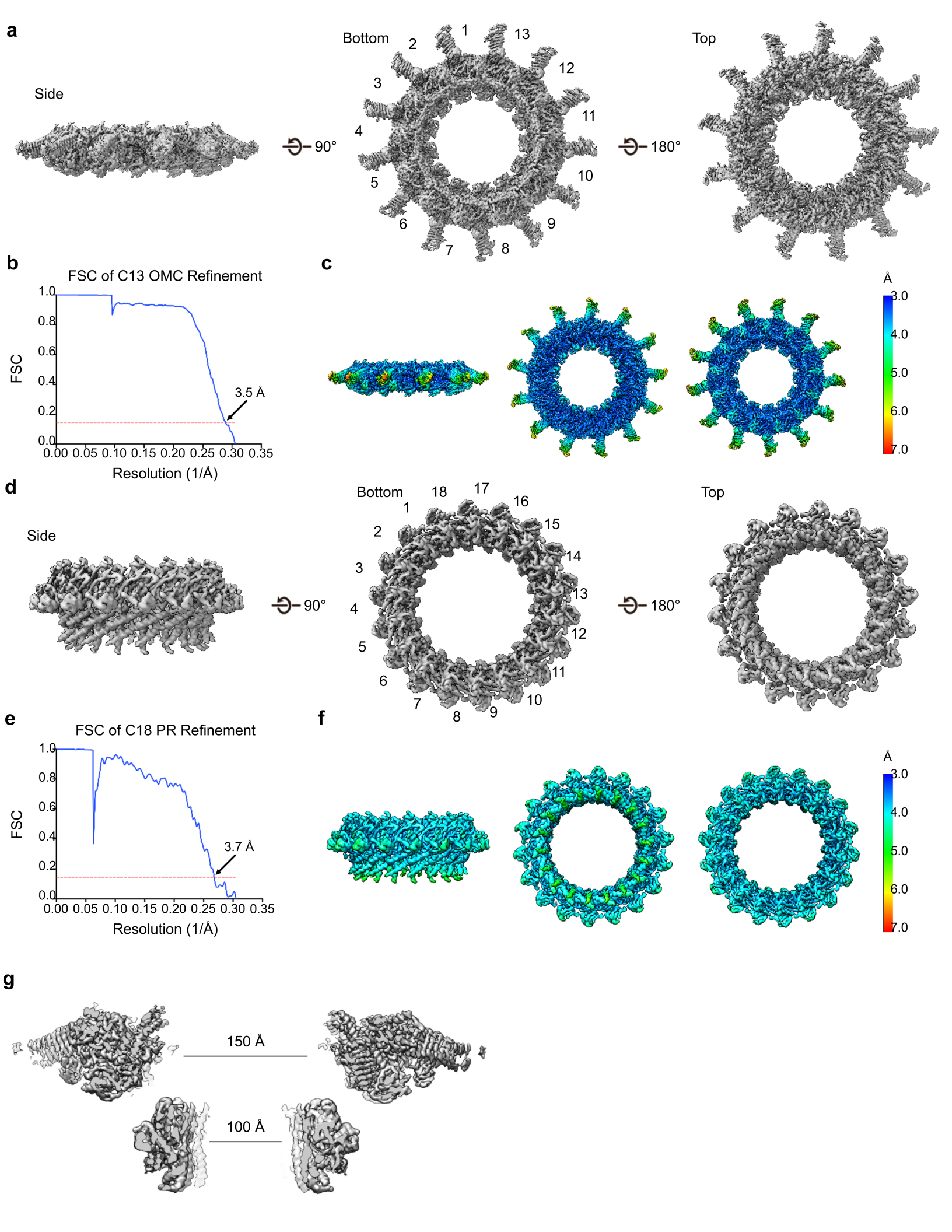
[b]Proteins were identified by searching the MS/MS data against *L. pneumophila* (UniProt; 2930 entries) using Proteome Discoverer (v2.1, Thermo Scientific). Search parameters included MS1 mass tolerance of 10 ppm and fragment tolerance of 0.1 Da. False discovery rate (FDR) was determined using Percolator and proteins/peptides with a FDR of ≤1% were retained for further analysis. The complete results are shown in Supplementary File 2 in the supplemental material.

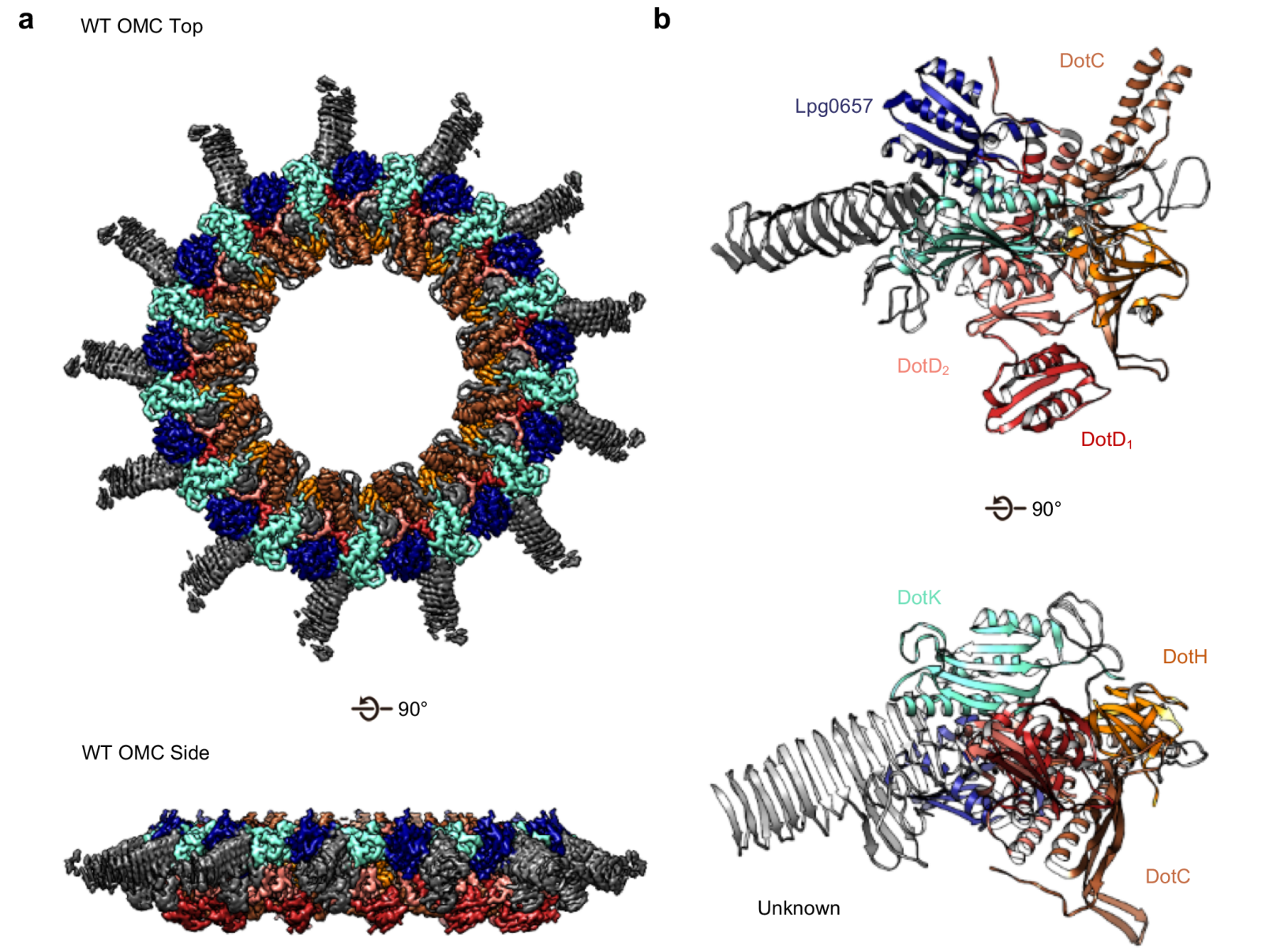
[c]Dot/Icm T4SS components are in bold.


a
90˚
OMC
PR
Dome
Disk
b
OM
IM
In vivo
In vitro
c
C13 OMC Disk
C18 PR
d
OM
OMC Dome
OMC Disk
Plug
PR
IM


a
b
Negative staining EM
Cryo-EM
c
Refined 3D model with C1 (no symmetry)
90˚
180˚
10 nm
d
3 Å
4
5
6
7
90˚
180˚
e
FSC of C1 refinement
1.0
0.8
0.6
0.4
0.2
0.0
FSC
4.6 Å
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Resolution (1/Å)

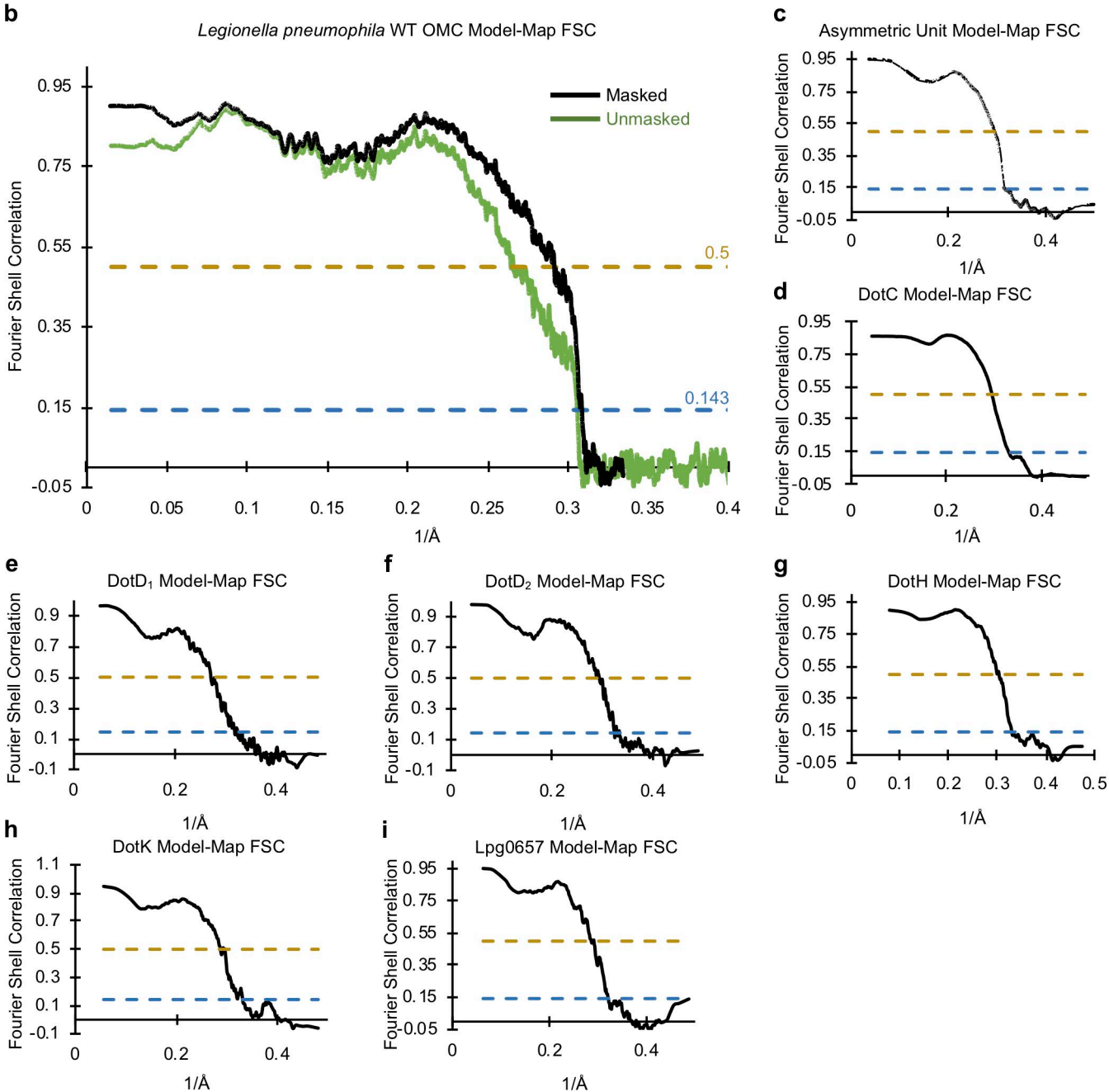

cryoSPARC
771,806 particles
Template picking
20,812 particles
2D/3D classfication filtering
3D refinement
C1 Symmetry
(no symmetry)
C13 Symmetry
20,812 particles
4.6 Å
20,812 particles
3.6 Å
Export particle stack
Relion
3D refinement
C13 Symmetry
20,812 particles
5.5 Å
Beam-tilt estimation
3D refinement
C13 Symmetry
20,812 particles
5.0 Å
Focused 3D classification
OMC Disk
C13 Symmetry
PR
C18 Symmetry
12,164
3,702
4,952
6,837
7,360
6,618
Focused 3D refinement
Focused 3D refinement
after per-particle ctf refinement
Focused 3D refinement
Focused 3D refinement
after per-particle ctf refinement
12,164 particles
3.5 Å
6,832 particles
3.7 Å

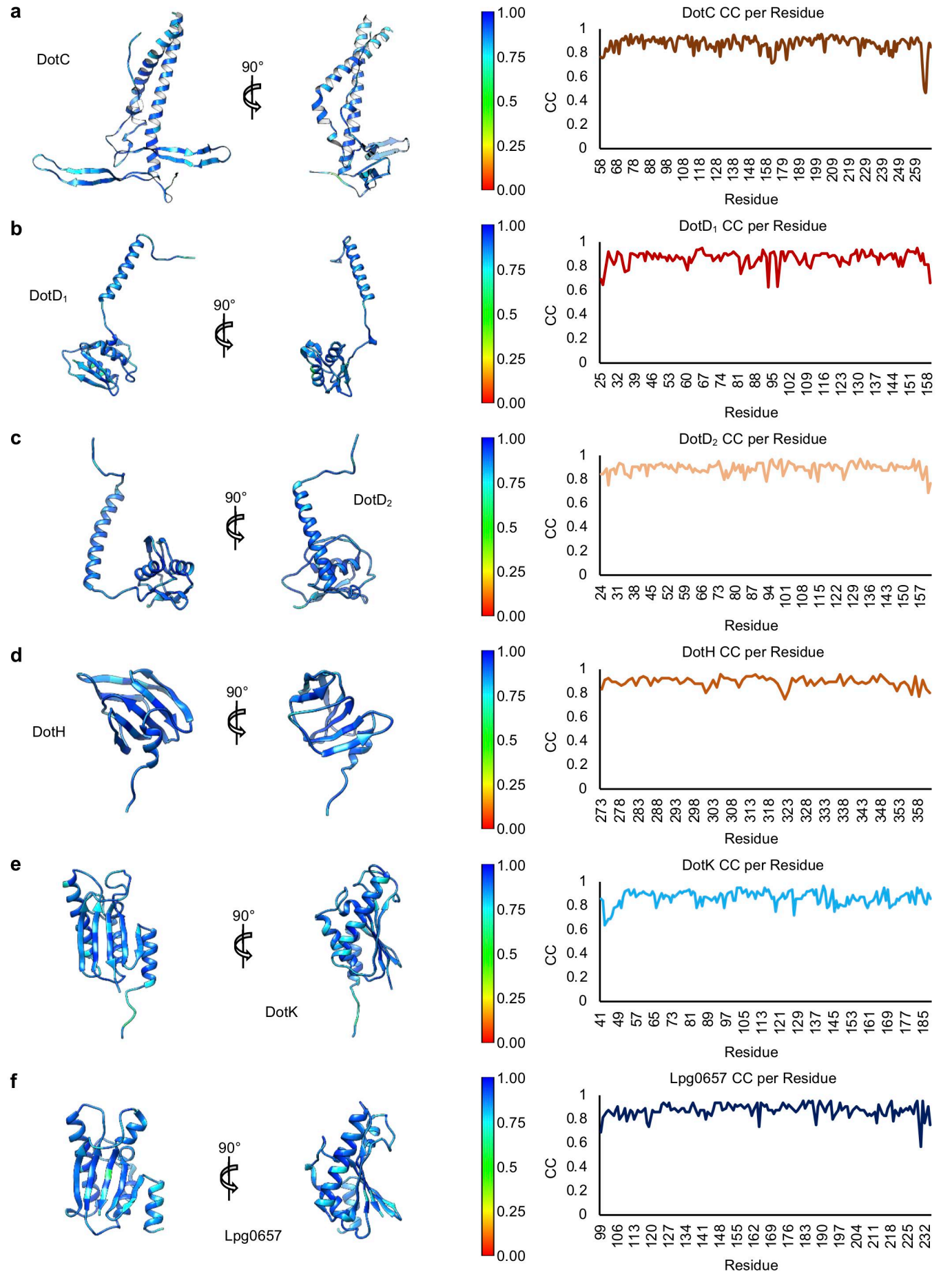

a
Side
90°
Bottom
1
2
3
4
5
6
7
8
9
10
11
12
13
180°
Top
b
FSC of C13 OMC Refinement
FSC
1.0
0.8
0.6
0.4
0.2
0.0
3.5 Å
0.00 0.05 0.10 0.15 0.20 0.25 0.30 0.35
Resolution (1/Å)
c
Å
3.0
4.0
5.0
6.0
7.0
d
Side
90°
Bottom
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
180°
Top
e
FSC of C18 PR Refinement
FSC
1.0
0.8
0.6
0.4
0.2
0.0
3.7 Å
0.00 0.05 0.10 0.15 0.20 0.25 0.30 0.35
Resolution (1/Å)
f
Å
3.0
4.0
5.0
6.0
7.0
g
150 Å
100 Å

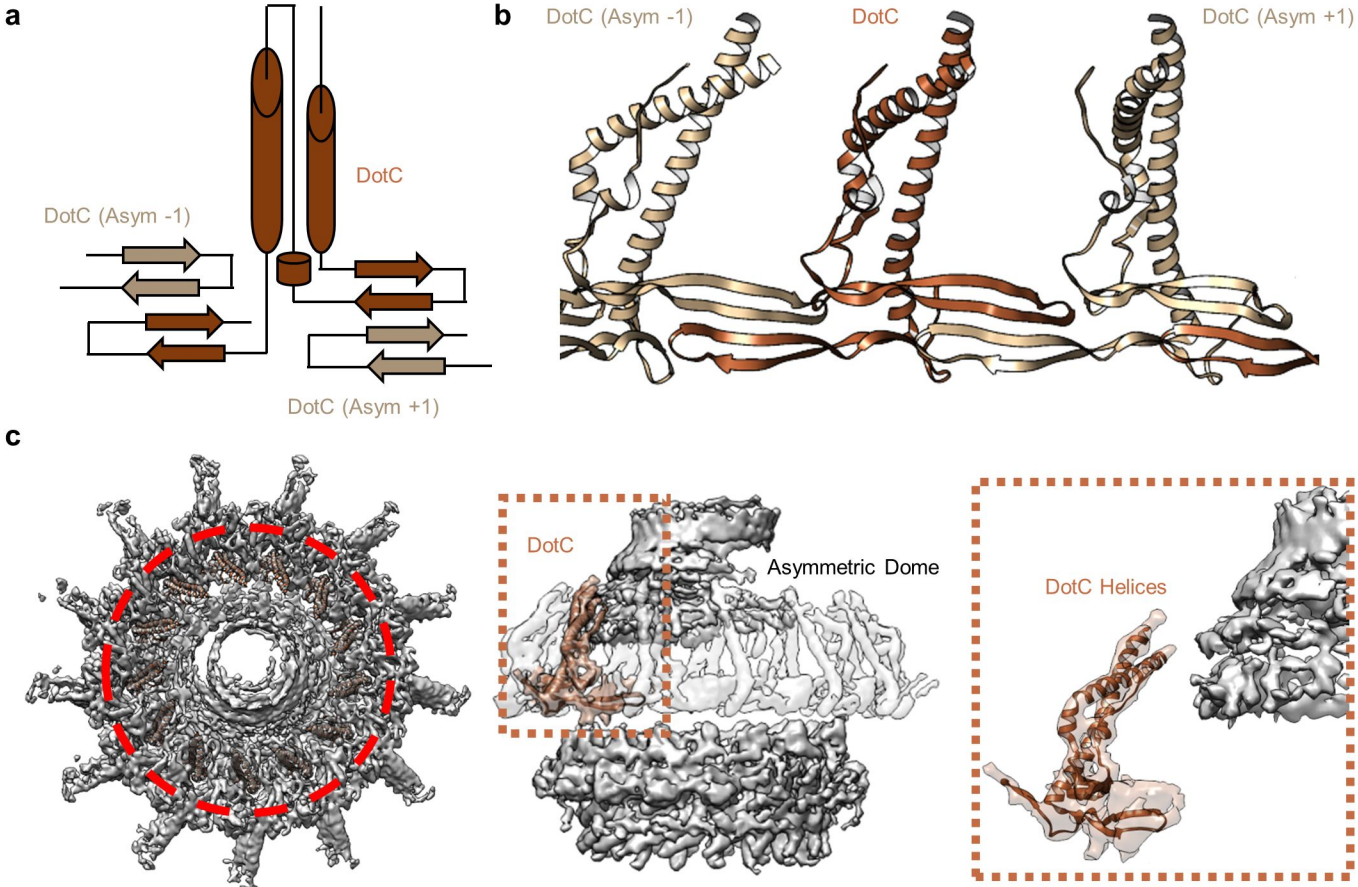

a
WT OMC Top
90°
WT OMC Side
b
DotC
Lpg0657
DotD2
DotD1
90°
DotK
DotH
DotC
Unknown

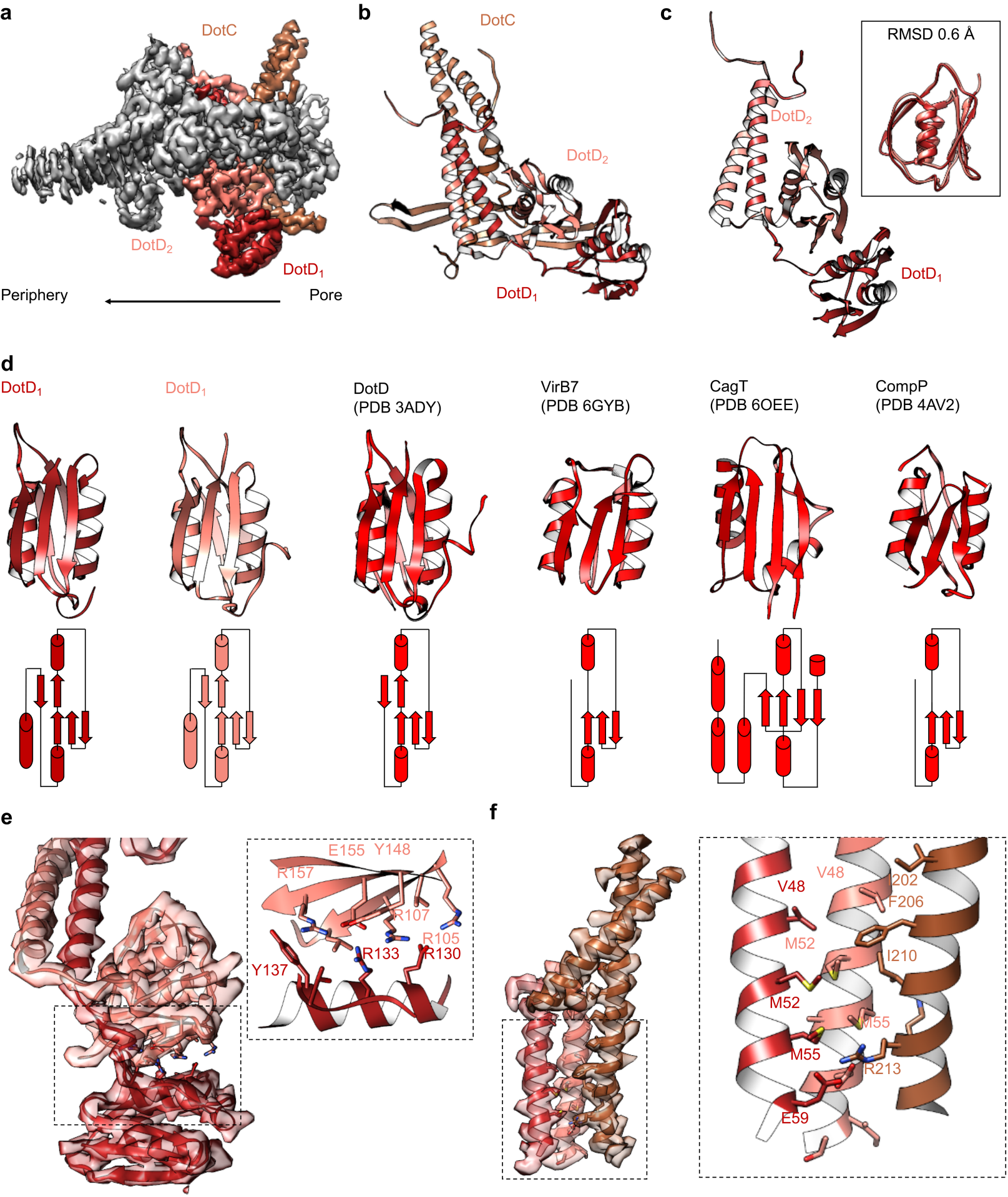
a

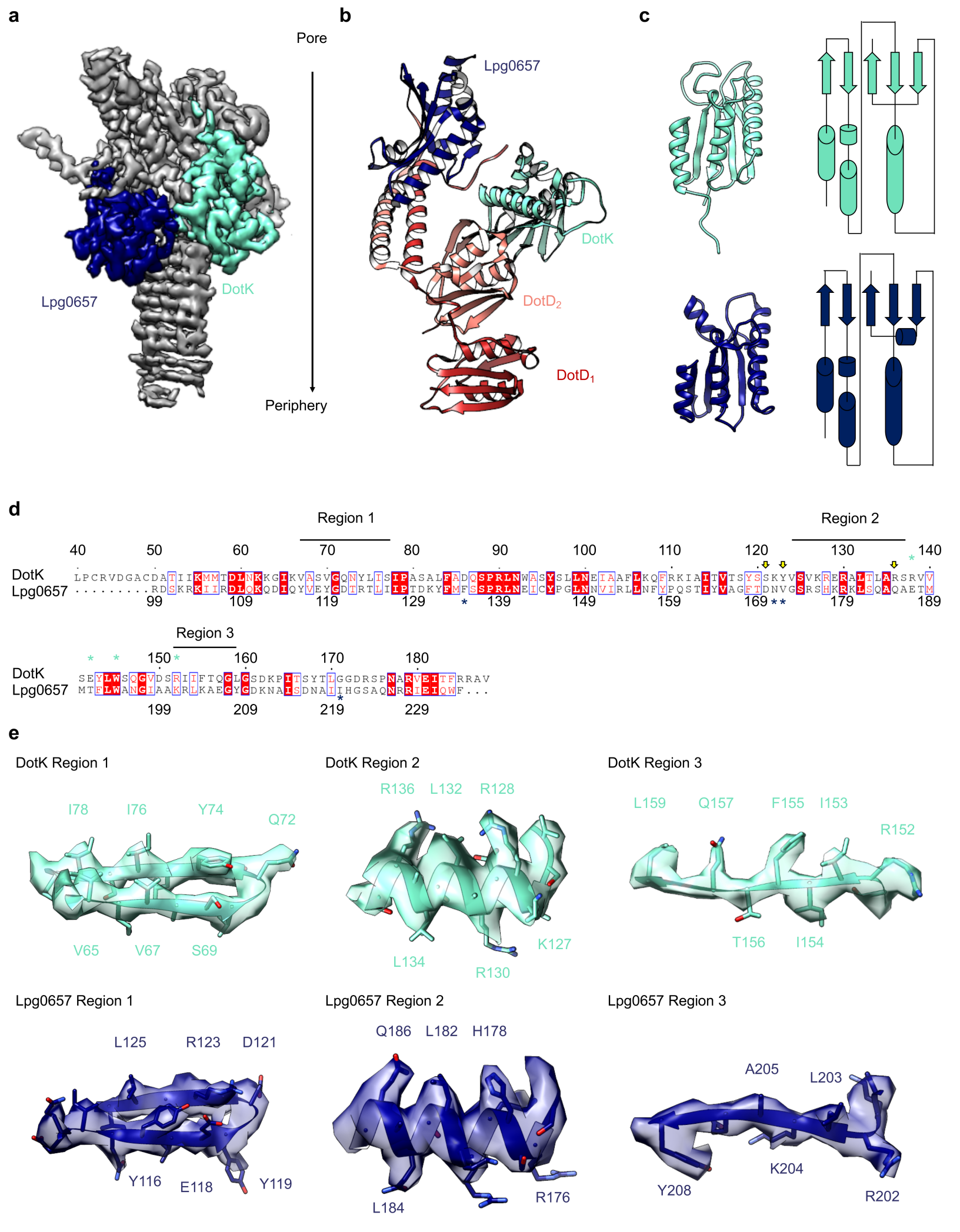
	DotC	DotD$_1$	DotD$_2$	DotH	DotK	Lpg0657	Unk1	Unk2	Unk3
RMSD									
Bond Length	0.015	0.024	0.019	0.015	0.014	0.013	0.005	0.006	0.005
Bond Angle	2.047	2.726	2.379	2.139	2.100	1.954	1.382	1.280	1.460
Ramachandran									
Favored	94.9	90.2	94.1	90.8	93.8	91.8	88.9	80.0	88.2
Allowed	5.1	9.8	5.9	9.2	6.2	8.2	11.1	20.0	11.8
Outliers	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Model-Map Correlation									
FSC (0.5)	3.4	3.7	3.4	3.3	3.4	3.4	3.5	4.2	3.5
FSC (0.143)	3.0	3.7	3.0	3.0	3.0	3.1	3.2	3.8	3.0
Molprobity Score	2.24	1.62	2.40	2.61	2.40	2.45	1.94	1.69	1.47
Clashscore	4.07	2.85	6.83	7.32	6.83	6.77	6.58	1.92	1.42
CC									
Main Chain	0.85	0.84	0.84	0.87	0.84	0.86	0.80	0.80	0.79
Side Chain	0.83	0.79	0.80	0.86	0.80	0.80	-	-	-

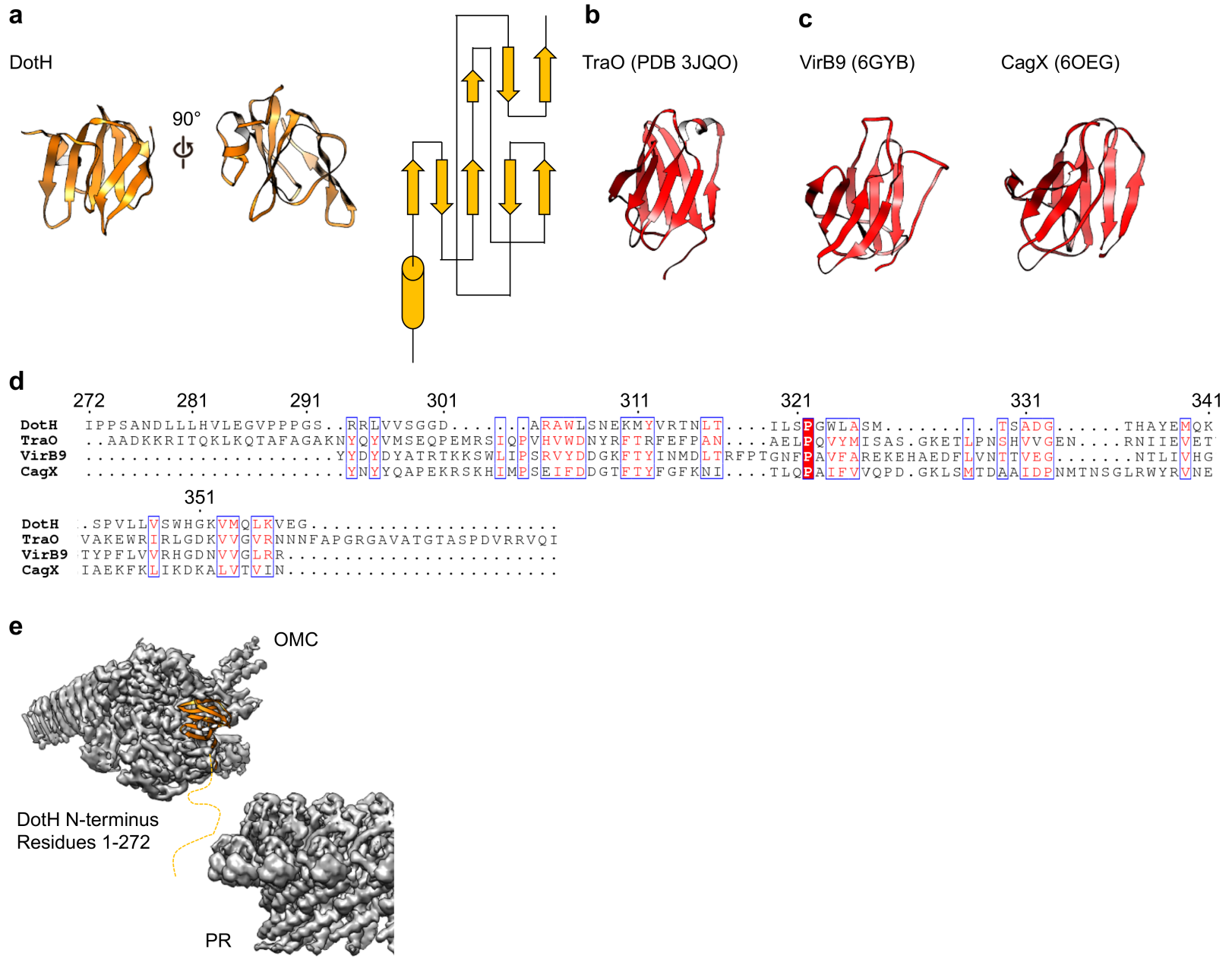


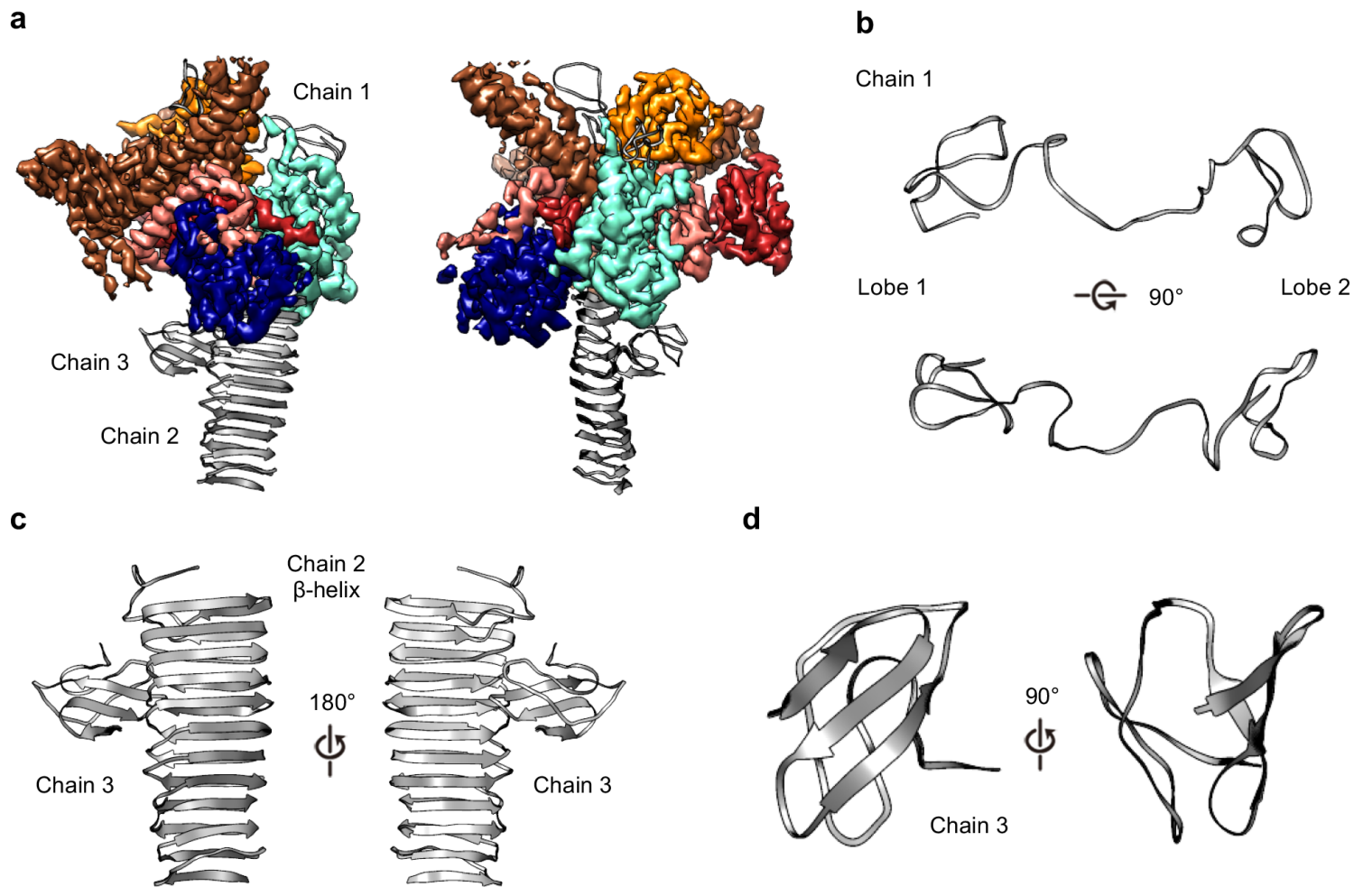

a
DotC
90°
1.00
0.75
0.5
0.25
0.00
DotC CC per Residue
CC
Residue
b
DotD1
90°
DotD1 CC per Residue
c
90°
DotD2
DotD2 CC per Residue
d
DotH
90°
DotH CC per Residue
e
90°
DotK
DotK CC per Residue
f
90°
Lpg0657
Lpg0657 CC per Residue

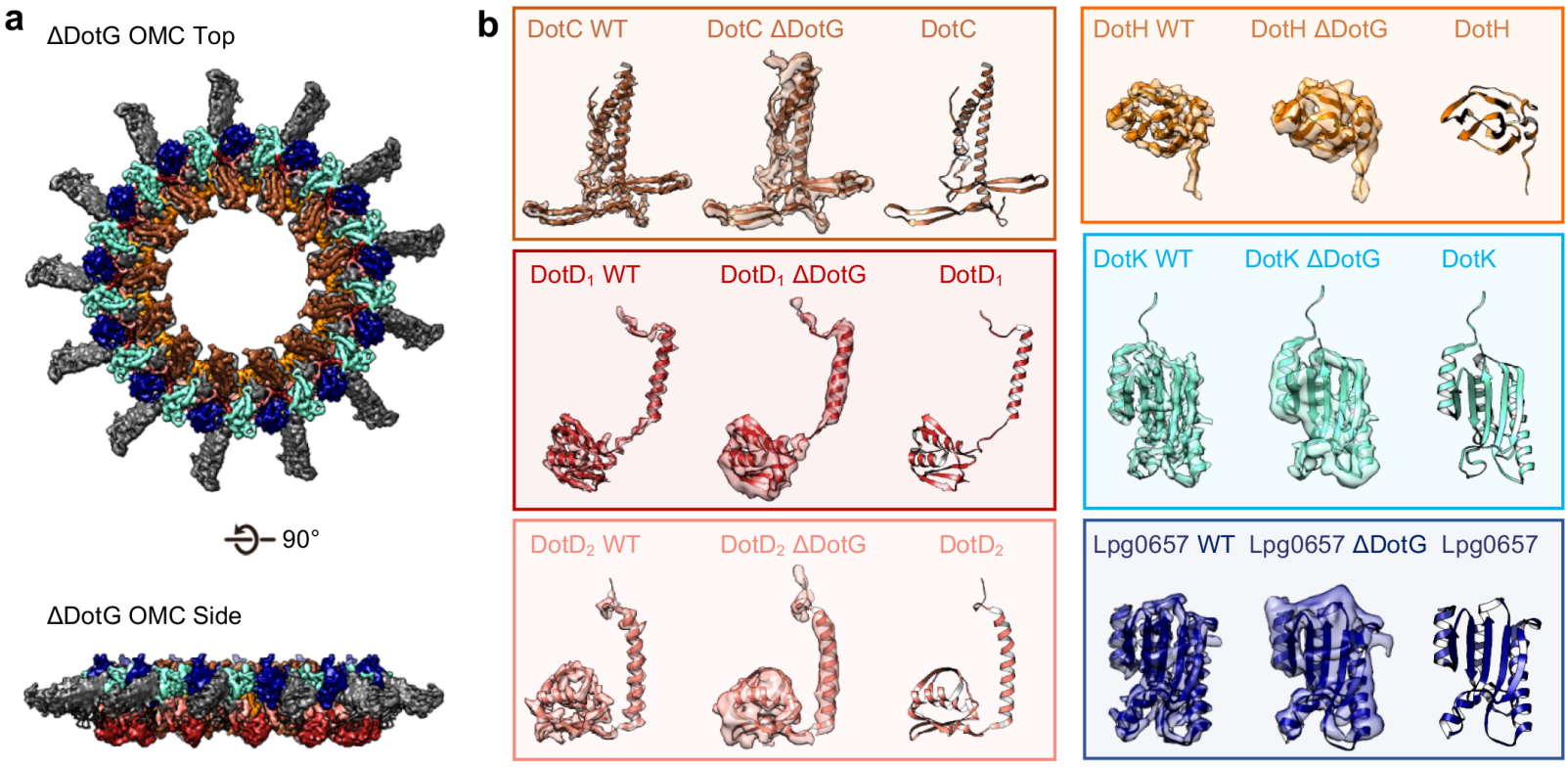

a
DotC
DotC (Asym -1)
DotC (Asym +1)
b
DotC (Asym -1)
DotC
DotC (Asym +1)
c
DotC
Asymmetric Dome
DotC Helices

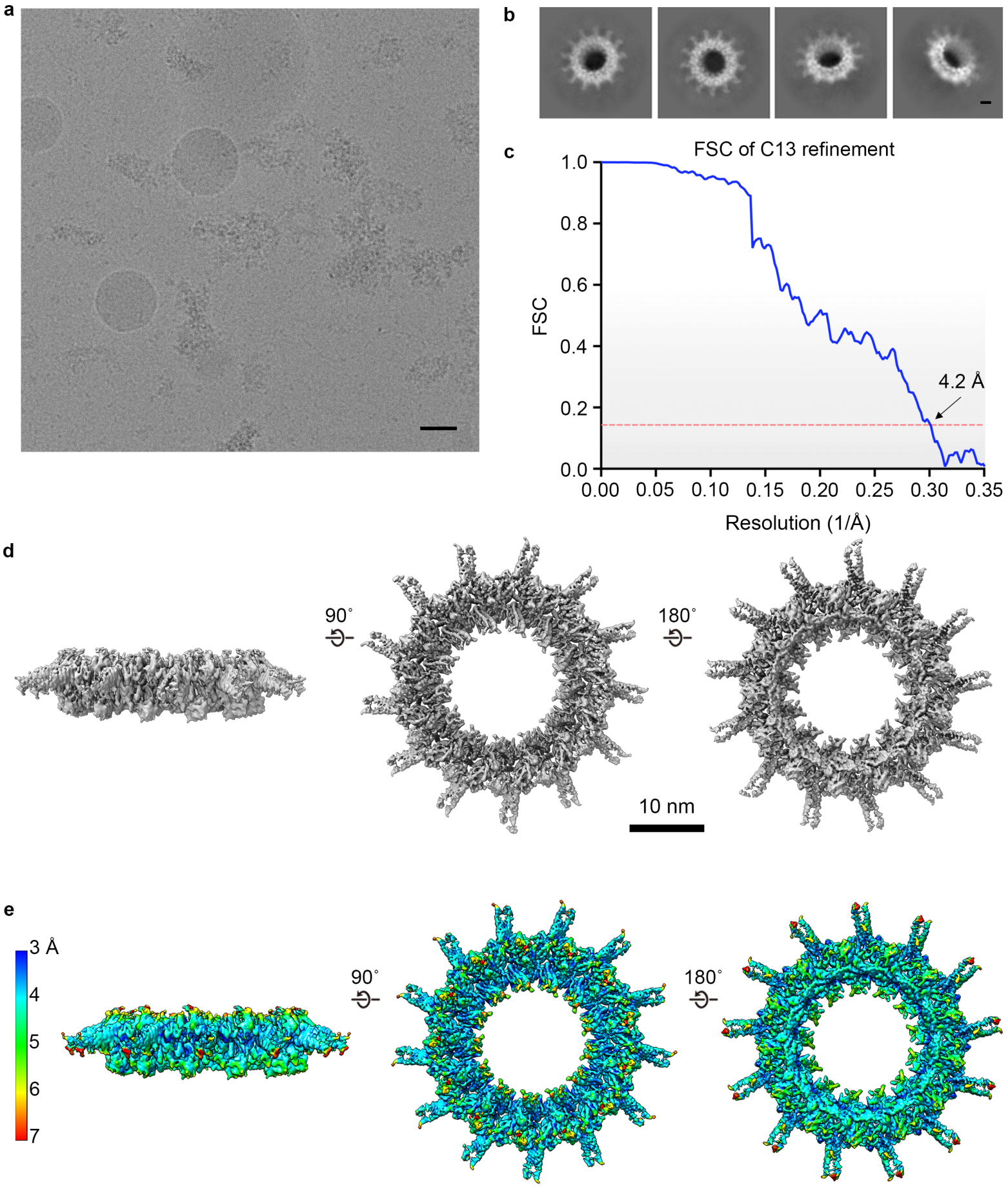

a
DotC
DotD2
DotD1
Periphery
Pore
b
DotC
DotD2
DotD1
c
RMSD 0.6 Å
DotD2
DotD1
d
DotD1
DotD1
DotD
(PDB 3ADY)
VirB7
(PDB 6GYB)
CagT
(PDB 6OEE)
CompP
(PDB 4AV2)
e
E155
Y148
R157
R107
R105
R133
R130
Y137
f
V48
V48
I202
F206
M52
I210
M52
M55
M55
R213
E59


a
Pore
Lpg0657
DotK
Periphery
b
Lpg0657
DotK
DotD2
DotD1
c
d
Region 1
Region 2
Region 3
DotK
Lpg0657
e
DotK Region 1
I78
I76
Y74
Q72
V65
V67
S69
DotK Region 2
R136
L132
R128
K127
L134
R130
DotK Region 3
L159
Q157
F155
I153
R152
T156
I154
Lpg0657 Region 1
L125
R123
D121
Y116
E118
Y119
Lpg0657 Region 2
Q186
L182
H178
L184
R176
Lpg0657 Region 3
A205
L203
K204
Y208
R202


a
DotH
90°
b
TraO (PDB 3JQO)
c
VirB9 (6GYB)
CagX (6OEG)
d
272
281
291
301
311
321
331
341
DotH IPPSANDLLLHVLEGVPPPGS..RRLVVSGGD....A.A.ARAWLSNEKMYVRTNLT....ILSPGWLASM.....A...TSADG.....THAYEMQK
TraO ..AADKKRITQKLKQTAFAGAKNYQYVMSEQPEMRSIQPVHVWDNYRFTRFEFPAN....AELPQVYMISAS.GKETLPNSHVVGEN....RNIIEVET
VirB9 .....YYDYDYATRTKKSWLIPSRVYDDGKFTYINMDLTRFPTGNFPAVFAREKEHAEDFLVNTTVEG.....NTLIVHG
CagX .....YNYYQAPEKRSKHIMPSEIFDDGTFTYFGFKNI....TLQPAIFVVQPD.GKLSMTDAAIDPNMTNSGLRWYRVNE
351
DotH .SPVLLVSWHGKVMQLKVEG.....
TraO VAKEWRIRLGDKVVGVRNNNFAPGRGAVATGTASPDVRRVQI
VirB9 TYPFLVVRHGDNVVGLRR.....
CagX IAEKFKLIKDKALVTVIN.....
e
OMC
DotH N-terminus
Residues 1-272
PR

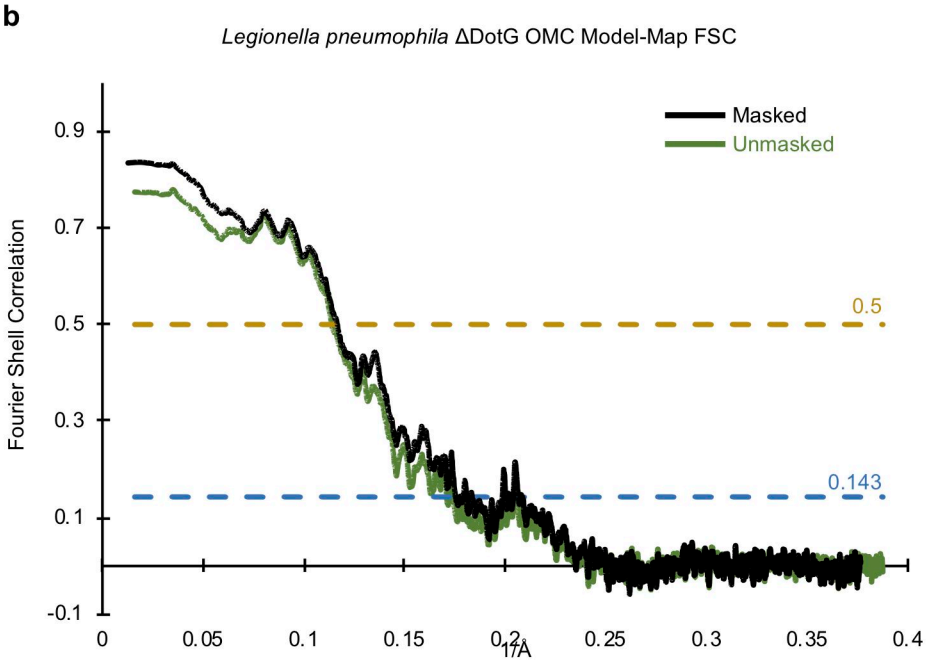

a
Chain 1
Chain 3
Chain 2
b
Chain 1
Lobe 1
90°
Lobe 2
c
Chain 2
β-helix
180°
Chain 3
Chain 3
d
90°
Chain 3

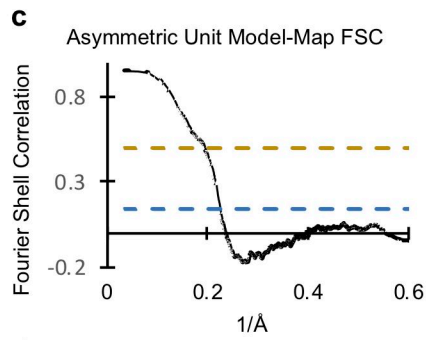

a
ΔDotG OMC Top
90°
ΔDotG OMC Side
b
DotC WT
DotC ΔDotG
DotC
DotH WT
DotH ΔDotG
DotH
DotD1 WT
DotD1 ΔDotG
DotD1
DotK WT
DotK ΔDotG
DotK
DotD2 WT
DotD2 ΔDotG
DotD2
Lpg0657 WT
Lpg0657 ΔDotG
Lpg0657

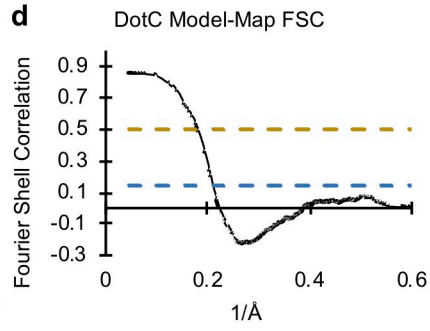

a
b
c
FSC of C13 refinement
1.0
0.8
0.6
0.4
0.2
0.0
FSC
4.2 Å
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
Resolution (1/Å)
d
90˚
180˚
10 nm
e
3 Å
4
5
6
7
90˚
180˚

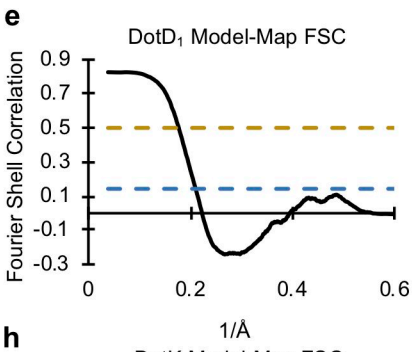
a

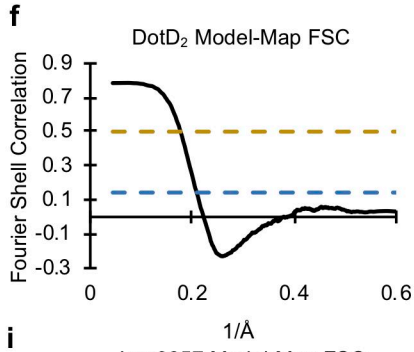
	DotC	$DotD_1$	$DotD_2$	DotH	DotK	Lpg0657	Unk1	Unk2	Unk3
RMSD									
Bond Length	0.006	0.006	0.007	0.006	0.006	0.005	0.005	0.005	0.006
Bond Angle	1.173	1.259	1.472	1.400	1.121	1.130	1.109	1.190	1.464
Ramachandran									
Favored	96.4	98.5	95.6	90.8	97.3	97.0	91.2	92.0	78.0
Allowed	3.6	1.5	4.4	9.2	2.7	30	8.8	8.0	22.0
Outliers	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Model-Map Correlation									
FSC (0.5)	5.6	5.6	5.7	5.3	5.4	5.4	6.0	7.2	6.6
FSC (0.143)	4.8	4.8	4.8	4.5	4.7	4.6	4.8	5.3	5.3
Molprobity Score	1.38	1.25	1.54	1.94	1.53	1.43	1.39	1.80	2.27
Clashscore	1.88	4.75	4.61	5.85	4.69	4.97	1.42	5.70	9.58
CC									
Main Chain	0.82	0.87	0.85	0.82	0.84	0.85	-	-	-
Side Chain	0.85	0.88	0.87	0.85	0.85	0.86	-	-	-

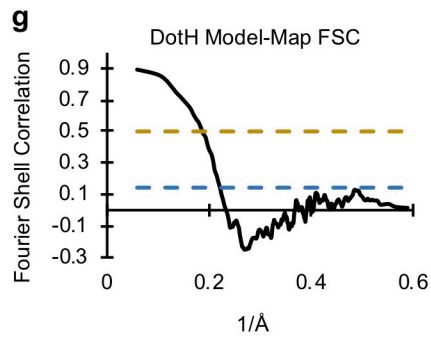


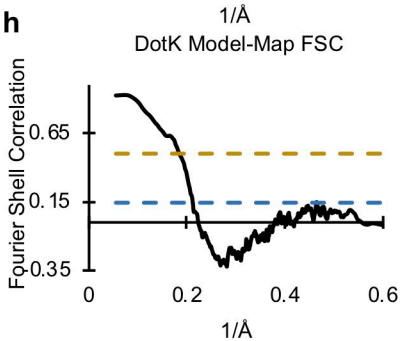


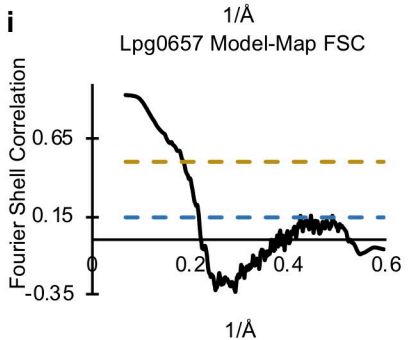


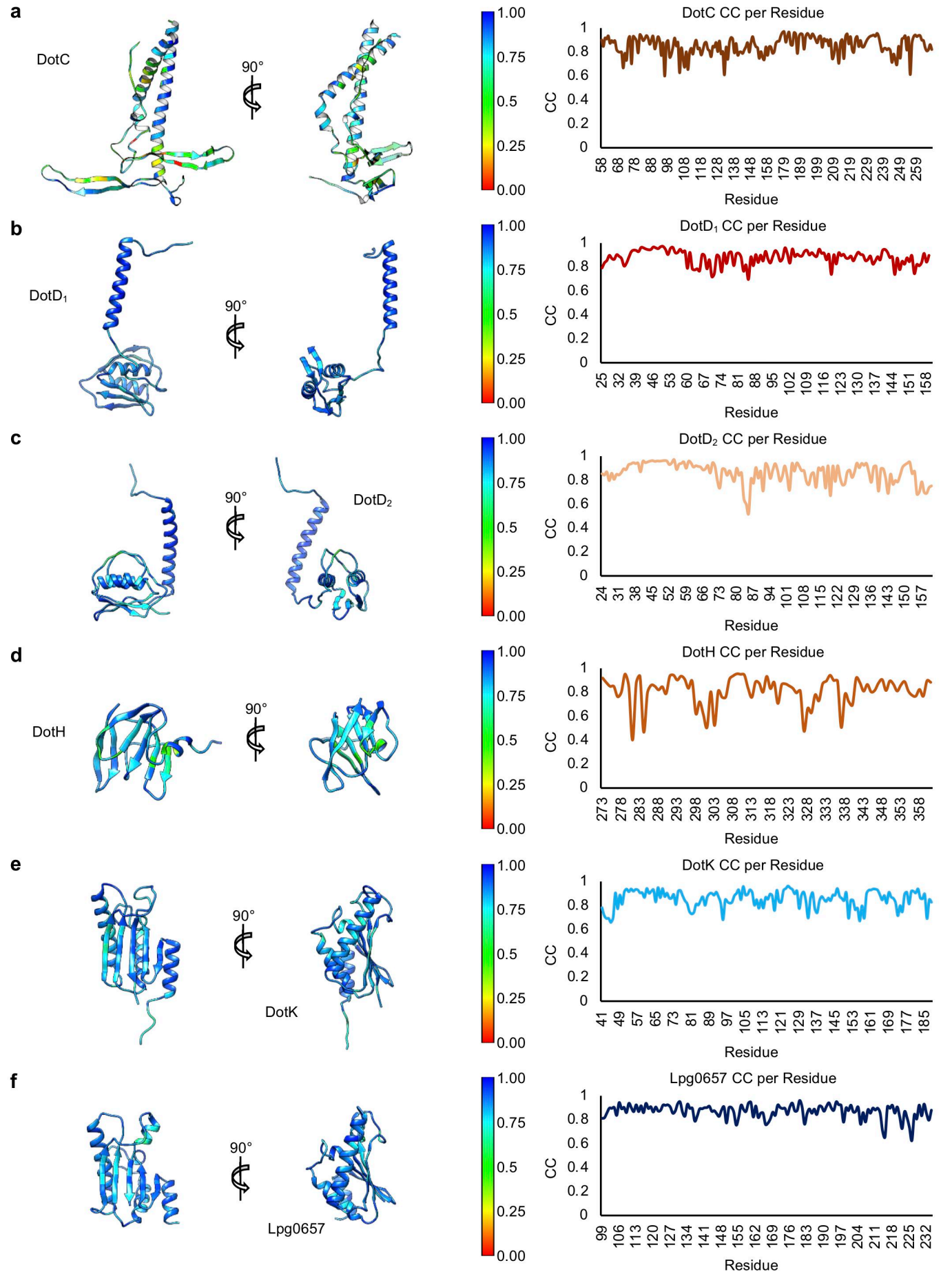


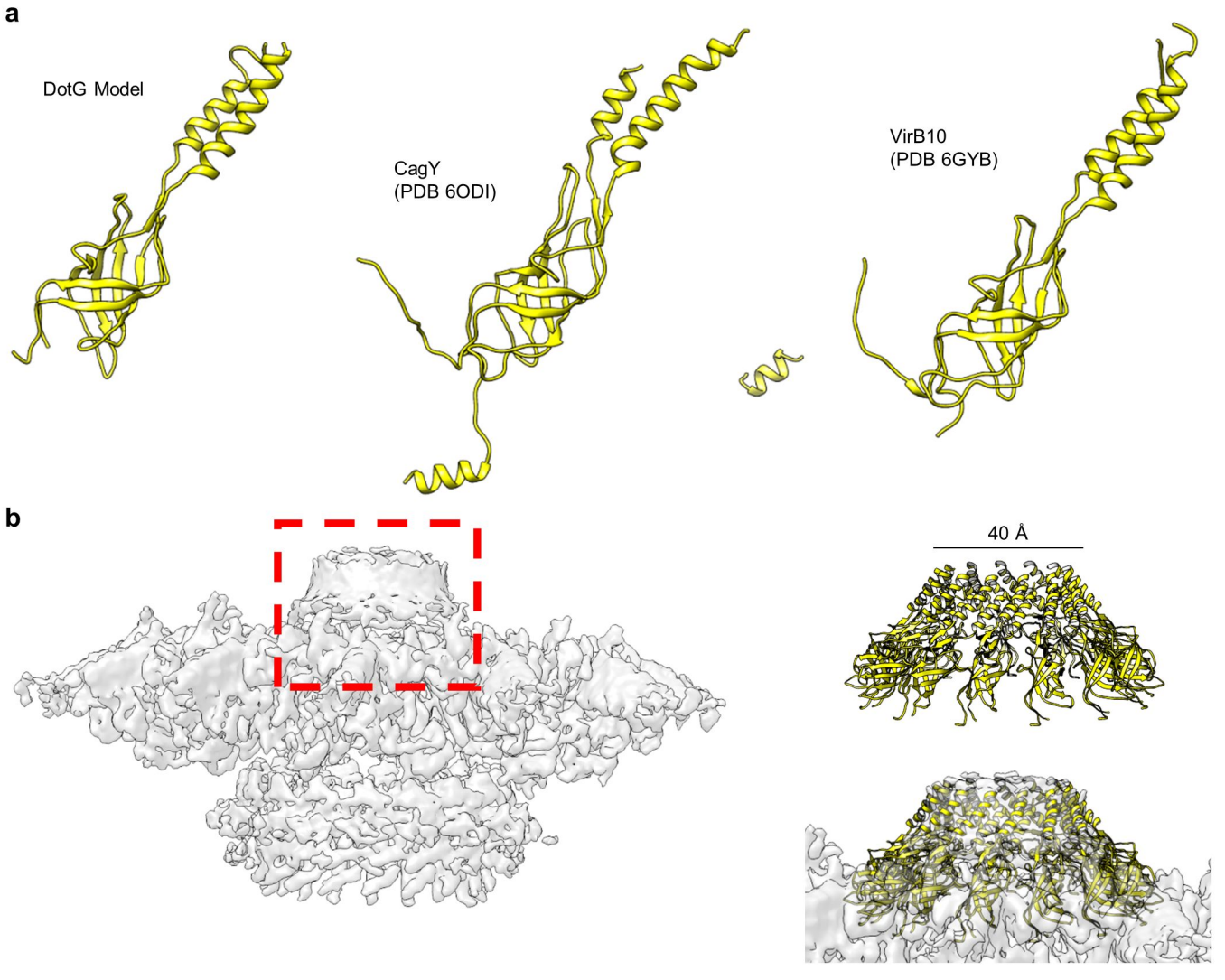


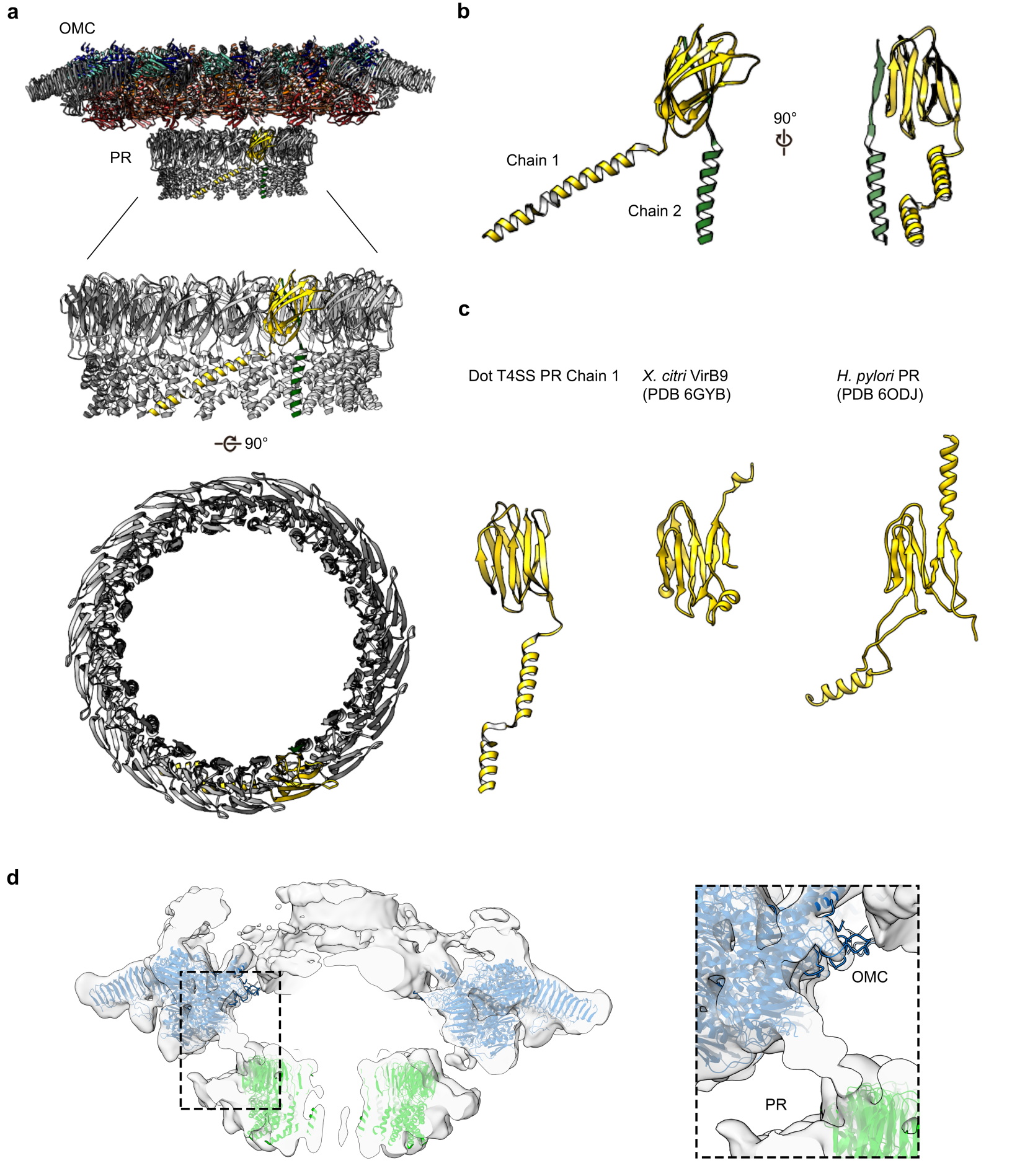


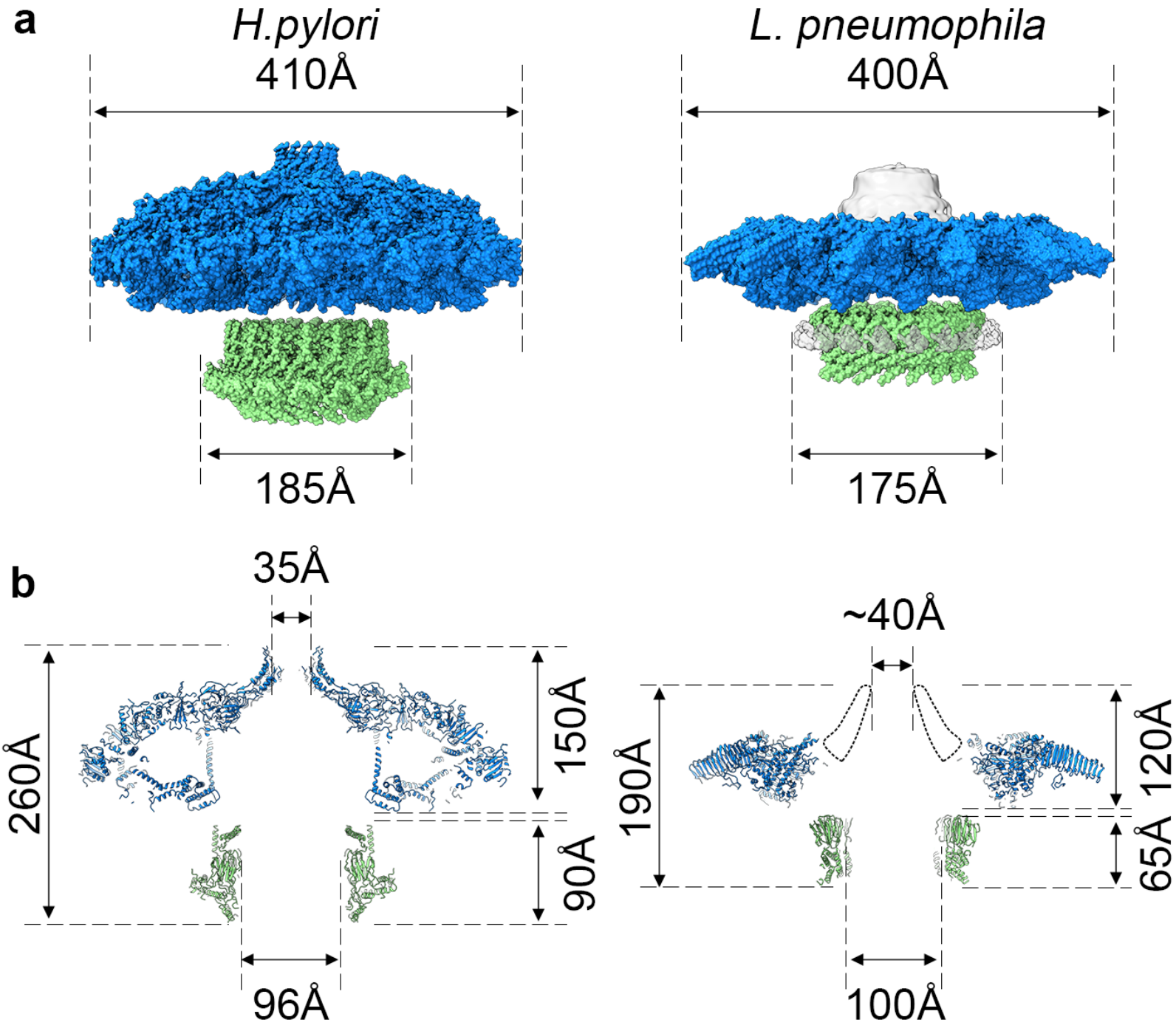





a
DotC
90°
1.00
0.75
0.5
0.25
0.00
DotC CC per Residue
CC
1
0.8
0.6
0.4
0.2
0
58 68 78 88 98 108 118 128 138 148 158 179 189 199 209 219 229 239 249 259
Residue
b
DotD1
90°
DotD1 CC per Residue
25 32 39 46 53 60 67 74 81 88 95 102 109 116 123 130 137 144 151 158
c
90°
DotD2
DotD2 CC per Residue
24 31 38 45 52 59 66 73 80 87 94 101 108 115 122 129 136 143 150 157
d
DotH
90°
DotH CC per Residue
273 278 283 288 293 298 303 308 313 318 323 328 333 338 343 348 353 358
e
90°
DotK
DotK CC per Residue
41 49 57 65 73 81 89 97 105 113 121 129 137 145 153 161 169 177 185
f
90°
Lpg0657
Lpg0657 CC per Residue
99 106 113 120 127 134 141 148 155 162 169 176 183 190 197 204 211 218 225 232


a
DotG Model
CagY
(PDB 6ODI)
VirB10
(PDB 6GYB)
b
40 Å
c
859 869 879 889 899 909 919 929
DotG .NQAIIKTGDIMFAVLDTSVNSDEPGPILATIVTGKL..........KGSKLIGSFN....LPSNADKMVITFFNTMSIPGAEKTI
CagY .PTFVLAQYTPIEITLTSKVDATLTG.IVSGVVAKDVWNMNGTMILLDKGTKVYGNYQSVKGGTPIMTRLMIVFTKAITP.DGVII
VirB10 NPDGLLVRGTYIRCILETRISDFGG.YTSCIVTEPVYSINGHNLLLPKGSKMLGQYS...AGEPTSHRLQVVWDRVTTP.TGLDV
939 949 959 969 979 989 999 1009
DotG SISAYAIDPNTARTALASRTNHHYLMRYGSLFASSFLQGFGNAFQSANTTITIGGTGGGNNITVANGVGRSTLENAVIGLATVGKA
CagY PLANAQAAGMLGEAGVDGYVNNHFMKRIGFAVIASVVNSFLQTAPIIALDKLIGLGKGRSERTPEFNYALGQAINGSM..QSSAQM
VirB10 TLMGPGID.TLGSSGHPGNYNAHWGNKIASALFISLLSDAFKYAAAEYGPETTTIGVGSGIVTQQP......FESNTA..RSMQQL
1019 1029 1039
DotG WSQQAQQLFNTPTTVEVYSGTGLGILFTQDVTT......
CagY SNQILGQLMNIPPSFYKNEGDSIKILTMDDIDFSGVY..
VirB10 AEQAVEKSGRRPATLTINQGTVLNVYVAKDVDFSAVLPK


a
OMC
PR
90°
b
Chain 1
Chain 2
90°
c
Dot T4SS PR Chain 1
X. citri VirB9
(PDB 6GYB)
H. pylori PR
(PDB 6ODJ)
d
OMC
PR


a
H.pylori
410Å
185Å
L. pneumophila
400Å
175Å
b
35Å
260Å
150Å
90Å
96Å
~40Å
190Å
120Å
65Å
100Å