Structural characterization of an essential lipoprotein of *Mycoplasma pneumoniae*

Irem Keles[1#], Sina Manger[1#], Mbuso S. Mantanya[1], Achilleas S. Frangakis[1*]

[1]Buchmann Institute for Molecular Life Sciences and Institute for Biophysics, Goethe University, Frankfurt am Main, Germany

[#]These authors contributed equally: Irem Keles, Sina Manger

*Correspondence: A.S.F. (achilleas.frangakis@biophysik.org)

Abstract

Mycoplasma pneumoniae is a human pathogen causing atypical community-acquired pneumonia. It is a model for a minimal cell, known for its non-canonical use of surface proteins for host-cell adhesion through ectodomain shedding and antigenic variation of surface proteins to evade the host cell immune response. Mpn444 is an essential mycoplasma surface protein implicated in both processes. It is one of 46 lipoproteins of *M. pneumoniae*, none of which have been structurally or functionally characterized. Here, we report the structure of Mpn444 at 3.04 Å as well as the molecular architecture of the trimeric Mpn444 complex. Our experimental structure displays striking similarity to structure predictions of several other essential lipoproteins in *M. pneumoniae* and other related *Mycoplasma* species, suggesting it to have a specialized and conserved function. The essentiality and involvement of Mpn444 in host immune evasion makes our structure a target for the development of new treatment strategies against mycoplasma infections.

Introduction

Mycoplasma pneumoniae is a human pathogen responsible for up to 40% of global cases of atypical community-acquired pneumonia. While mycoplasma infections can be treated by macrolide antibiotics, macrolide resistance is on the rise and no approved vaccines exist to date (1). *Mycoplasma* are minimal cell model organisms characterized by the absence of a cell wall and the smallest genome size known for any self-replicating organism (2). *M. pneumoniae* encodes 688 ORFs and lacks genes needed for essential metabolic pathways such as biosynthesis of a cell wall, amino acids and nucleotides (3) (4). Furthermore, many mycoplasma proteins are multifunctional such that some cytosolic proteins moonlight as cell adhesion proteins during host invasion (5) (6). Surface protein antigenic variation is an immune evasion mechanism used by *M. pneumoniae*. This mechanism has been reported for the major adhesion complex as well as several lipoproteins in other *Mycoplasma* species (7). Cell surface proteins undergo ectodomain shedding, a process that has been described for Mpn142, the major adhesion complex of *M. pneumoniae*, and several uncharacterized lipoproteins (4). Lipoproteins in different *Mycoplasma* species have been shown to regulate host cell immune response (8) (9). Some lipoproteins have also been reported to be genomically associated with ABC transporter genes or share sequence similarity with ABC transporter genes, suggesting involvement in the transport of nutrients into the cell (7).

Mpn444 is an essential *M. pneumoniae* lipoprotein of unknown function with an approximate molecular weight of 145 kDa (10). It carries a lipoprotein signal sequence at the N-terminus, directly followed by a conserved cysteine that can be lipidated (as a N-palmitoyl cysteine or a S-diacylglycerol cysteine) and serves to anchor the protein to the mycoplasma membrane (11). The gene, *mpn444* is one of 46 lipoprotein genes

in six multigene lipoprotein families of *M. pneumoniae*, all of unknown structure and function (7). Among this group, only five proteins, including Mpn444, are essential (10).

Mpn444 undergoes ectodomain shedding (4), though the process has not been mapped and requires further studies. It also contains short peptide variable-number tandem repeats (VNTRs) (12), and thereby plays a role in host immune evasion. Its homologue in *M. genitalium* MG309, likewise an uncharacterized lipoprotein, has been shown to activate host innate immune response (13). The canonical function of Mpn444 on the mycoplasma surface, however, remains unclear. Mpn444 shares an operon with deaD, an RNA helicase associated with ribosome assembly (10). Functional studies have shown Mpn444 to crosslink with the uncharacterized transmembrane protein secD and the uncharacterized lipoproteins Mpn489 and Mpn436 (14). Both Mpn489 and Mpn436 belong to the small group of essential lipoproteins and are paralogs of Mpn444 (10) (15). Mpn436 is also the closest related homolog of Mpn444 in *Mycoplasma pneumoniae* with 24% sequence identity (7). AlphaFold3 structure predictions of both proteins display significant similarities to that of Mpn444. Here, we present the novel cryo-EM structure of Mpn444 at 3.04 Å resolution, the first experimental structure of a *Mycoplasma pneumoniae* lipoprotein crucial for offering new insights into *M. pneumoniae* pathogenesis and for advancing drug discovery efforts.

Results

Single-particle cryo-EM of Mpn444

We truncated the lipoprotein signal peptide of Mpn444 (aa 1-27) and expressed the truncated construct in *E. coli*. The protein was purified using affinity and size

exclusion chromatography, and its identity and purity were confirmed by mass spectrometry.

Two-dimensional (2D) classes of Mpn444 displayed very strong orientation bias as the majority of classes showed uniform views (Figure 1a). The replacement of the detergent LMNG with CHAPSO led to distinctly different orientations at the cost of lower resolution in the 2D classes (Figure 1b). Particles from both datasets were pooled, but sufficient sampling could not be achieved, as particle side views were severely underrepresented and bottom or top views were absent (Figure 1 a & b). Thus, we acquired datasets with a fixed tilted stage at 20°, 30° and 40° tilt angles, yielding unique particle views at every tilt angle (Figure 1c - e). However, the quality of the particles declined with increasing stage tilt. This was evident by the resolution of the 2D classes (Figure 1c - e) and the declining CTF fit of the particle sets. The CTF was fitted to a range of 2.0 Å - 4.8 Å at 0° tilt, with the maximum fit increasing to 6.25 Å at 20° tilt, 7.9 Å at 30° tilt and 8.4 Å at 40° tilt. Hence, we limited the tilt angle to 40°. This approach allowed us to keep the average data quality high while collecting particles in all orientations except for direct top views (Figure 1f). While our final dataset still oversampled particle front and back views, we achieved sufficient signal contribution from all viewing orientations. (Figure 1g, Supplemental Figure 1). We solved the structure of the Mpn444 monomer at 3.04 Å resolution (according to the gold-standard criterion of FSC = 0.143; EMD-51459; Figure 1h, Supplemental Table 1).

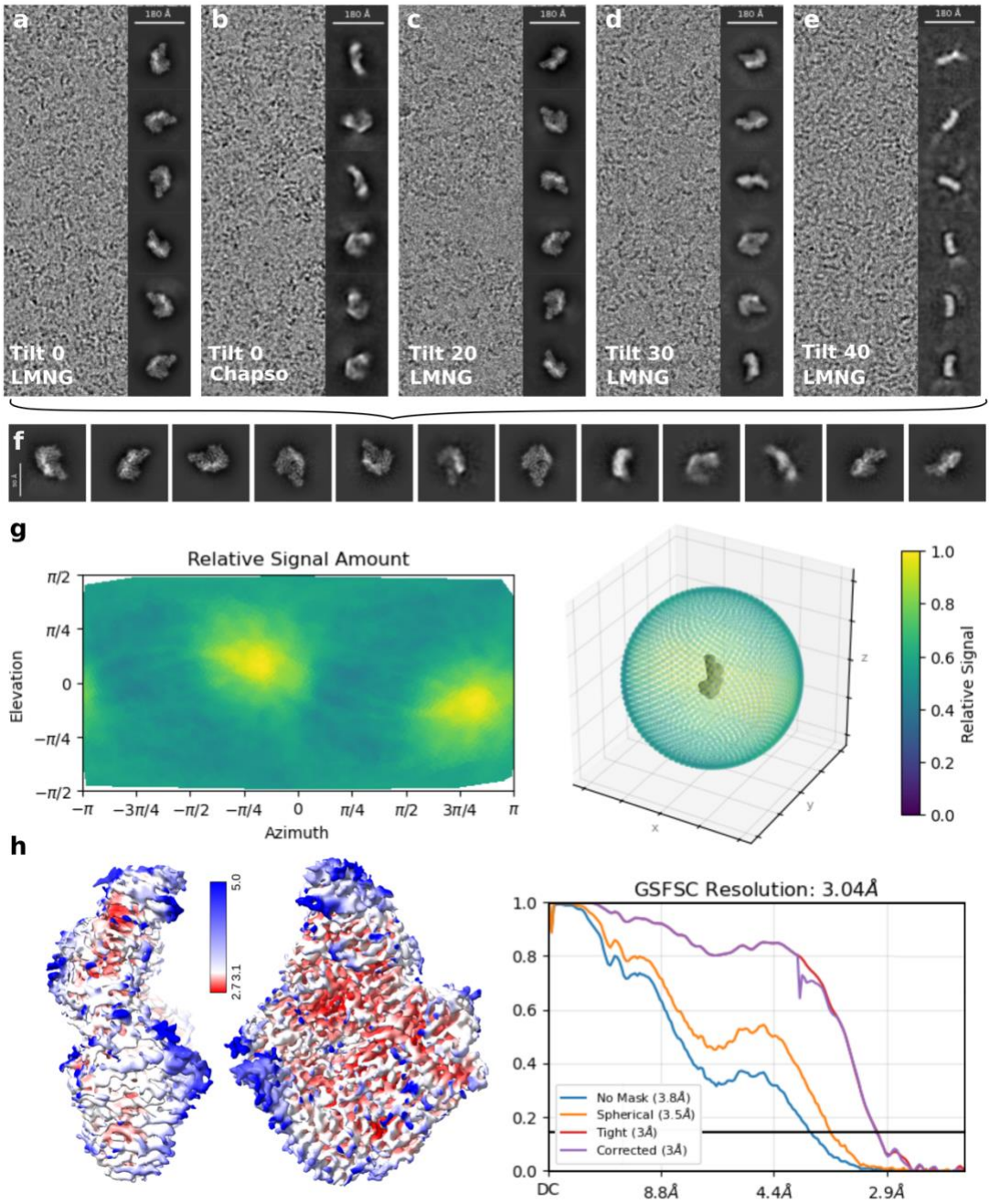

Figure 1: Processing pipeline of Mpn444

a - e Representative micrographs at tilts 0°, 20°, 30° and 40° with CHAPSO or LMNG (0.92 Å/pix) with corresponding 2D classes (scale bar 180 Å).

f 2D classes (scale bar 90 Å) of Mpn444 from the pooled particle set and processing pipeline.

g Orientation diagnostics of the pooled particle set from cryoSPARC (16). The diagram displays the relative signal amount with relation to the viewing direction.

h Density map of the Mpn444 monomer at 3.04 Å, colored by local resolution (left) and the corresponding FSC curve from cryoSPARC (16) (right).

Molecular architecture of Mpn444

Using the resolved cryo-electron density map, the atomic model of Mpn444 was built using the AlphaFold3 prediction as the starting model (Supplemental Table 1). The fold loosely resembles a ribeye steak with one convex and one concave interface that contains a central opening. The front view spans in the long and short axes approximately 120 Å and 90 Å, respectively. In the concave view, the bottom of the fold under the opening is formed by alpha-helices on the left and a small beta-sheet on the right side (core beta-sheet). The fold above the opening is formed by 7 smaller alpha-helices. A patch of basic residues spans the opening on both the concave and convex surfaces. The tip consists of three short alpha-helices with an additional beta-sheet below the tip on the right side. N- and C-termini are both located at the left side edge of the fold (Figure 2a).

A structure similarity search of either the entire fold or smaller fragments yielded no relevant results in the PDB, highlighting the novelty of the structure. However, multiple structure predictions from AlphaFold (17) were similar (Supplemental Figure 2) (18) and belong to uncharacterized (lipo)proteins from different mycoplasma species, including the essential *M. pneumoniae* lipoproteins Mpn436 and Mpn489 (Figure 2b), as well as their essential *M. genitalium* homologs, MG309, MG307 and MG338 (19) (15).

While the secondary structure elements are very stable throughout all 3D classes of the cryo-EM density maps, 20% of the amino acid sequence is located on large, highly flexible loops, for which no defined density could be discerned (Supplemental Figure 3a). A sequence alignment with homologues of other closely related *Mycoplasma* species revealed that residues within the stable fold are significantly conserved (11% identity, 92% conservation) (Figure 2c) than residues

situated on flexible loops (0% identity, 14% conservation) (Supplemental Table 2, Supplemental Figure 4). Moreover, loop regions of Mpn444 are predominantly absent in the homologue sequences, indicating that those regions are not crucial for the main function of the protein. Instead, they could play a role in the formation of complexes with other mycoplasma proteins or in the recognition of host cell factors.

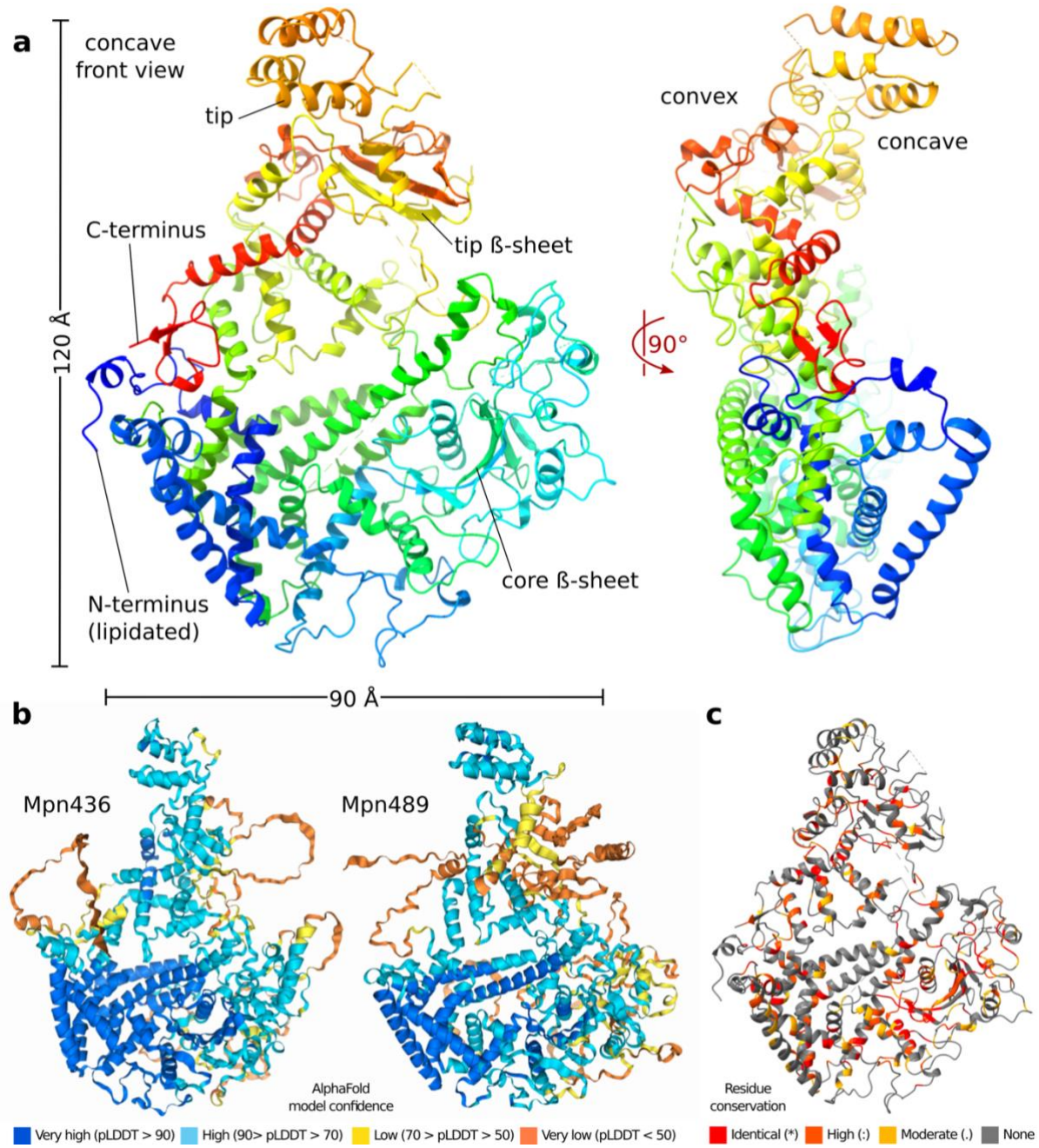

Figure 2: Molecular architecture of Mpn444
a Structure of Mpn444 as a ribbon model, colored in a rainbow gradient from N- to C-terminus. The model is 80% complete, as loops with no discernable density were not modelled (Supplemental Figure 2).
b AlphaFold3 predictions of the structures of Mpn436 and Mpn489 from *M. pneumoniae*. Model confidence is indicated by color (dark blue for very high confidence, light blue for high confidence, yellow for low confidence, orange for very low confidence). The predicted fold of both proteins appears similar to that of Mpn444.
c Sequence conservation of Mpn444. Ribbon model colored by conservation (red is identical, dark orange is highly conserved, light orange is moderately conserved, grey is not conserved).

Mpn444 forms trimers with C3 symmetry

In addition to the 2D class averages of Mpn444 monomers (approximately 60% of particles), we routinely observed trimeric complexes (approximately 20% of particles, Supplemental Figure 5a) as well as dimeric complexes (approximately 20% of particles, Supplemental Figure 5b). The dimers appeared to be incomplete trimeric assemblies as they displayed the same subunit architecture as the trimers. While the trimer exhibited C3 symmetry, we confirmed this by comparing a non-symmetrized (C1) reconstruction to a C3-symmetrized reconstruction of the same particle set. C1-symmetric trimer showed the same overall domain architecture, but with weaker density in one monomer, likely originating from the persistent presence of monomers and dimers in the particle set even after several rounds of 2D classification (Supplemental Figure 5c). All following refinement steps were done with C3 symmetry.

We determined the structure of the homotrimer at a global resolution of 5.4 Å (6.31 Å according to the EMDB validation) with a resolution range of 4 to 9 Å (Supplemental Figure 5d & e, Supplemental Table 1, PDB-9GNH, EMD-51470). Three Mpn444 monomers assemble into a C3 symmetric complex (Figure 3a & b, Supplemental Figure 5c) with the central interface forming a channel with a diameter of 23 Å in the top view and 11.5 Å in the bottom view. The monomeric convex surfaces

face toward the outside of the complex, with the tip of one monomer linking the lower concave surface of the juxtaposed monomers, forming the lateral dimer interface (Figure 3b & c). The N-termini of the monomers which contain the lipidated cysteine residues serving as lipid anchors, all face the same side of the complex; thus, corresponding to the membrane facing side (Figure 3c). This architecture is considerably different from the AlphaFold3 prediction of the trimer (Supplemental Figure 6), in which the monomers are stacked together like rose petals with the bottom of one monomer nested into the convex interface of its neighbor. We evaluated our homotrimer assembly with reference to the available crosslinking-data and found six out of seven amino acids involved in external crosslinks to be located at the periphery of the complex, thus available for protein-protein interactions (Supplemental Figure 7).

The central interface is composed of three short loops (Figure 3d), that are conserved in the closely related homologue MG309 from *M. genitalium*. The lateral dimer interfaces and the central trimer interface (Figure 3b) have additional densities unaccounted for by rigid body docking of the monomers (Supplemental Figure 3b). Specifically, the central trimer interface contains a loop that could not be resolved in the monomeric structure (GLY 649 – ASN 665: GTSTVKTTSSNTGQTKN, Supplemental Figure 3c), making it likely to be involved in trimer formation or stabilization. The additional density in the central channel of the central interface may be an artifact corresponding to incomplete orientational sampling as there are no unmodelled residues in this region.

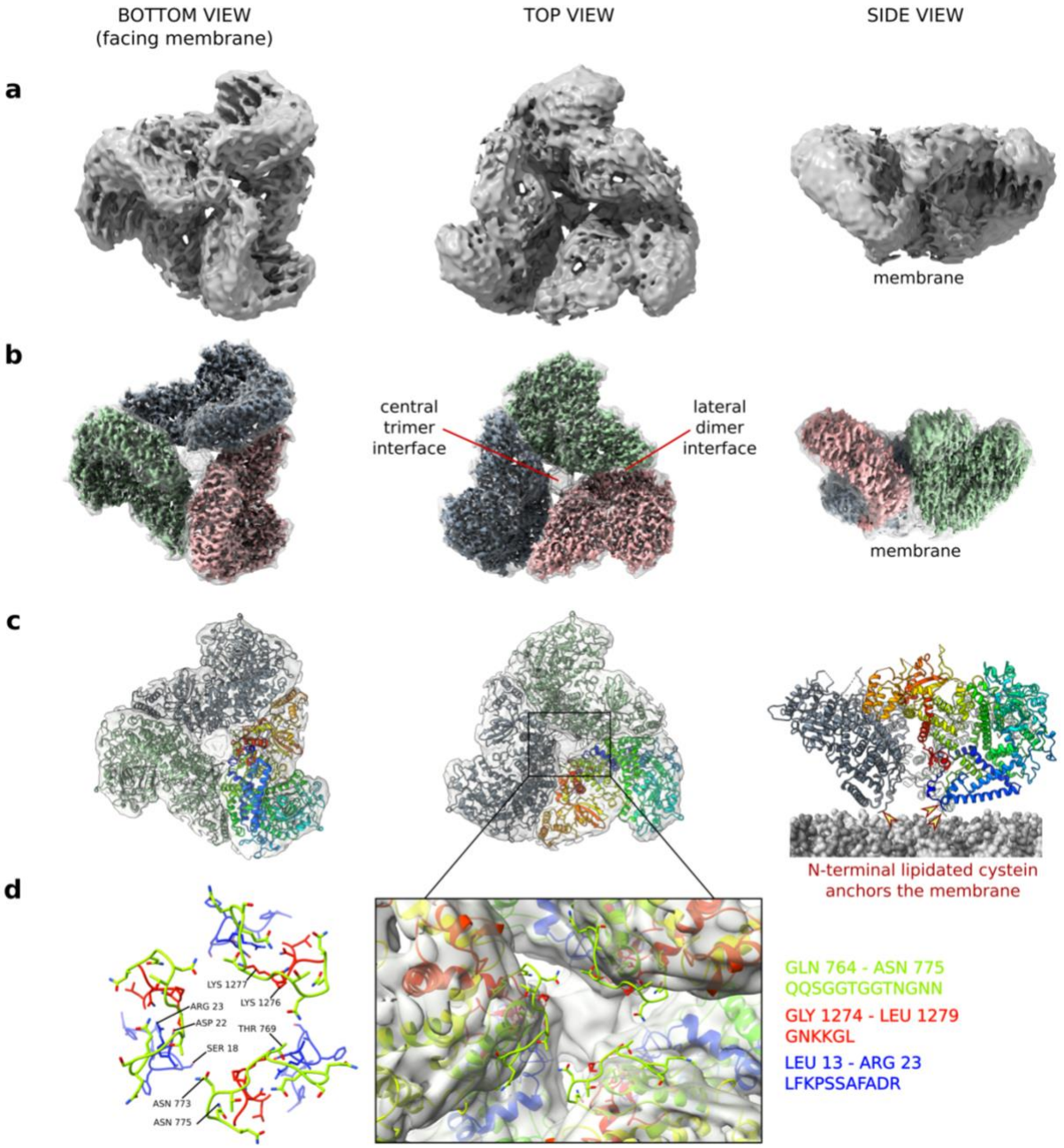

Figure 3: Multimeric organization of Mpn444

a Density map of the Mpn444 trimer at 5.4 Å (grey surface). (left) bottom view, (middle) top view, (right) side view with the position of the membrane indicated

b Density map of the Mpn444 trimer (transparent grey surface) fitted with the density maps of the Mpn444 monomer (dark slate grey, rose, sea green).

c Density map of the Mpn444 trimer (transparent grey surface) fitted with the molecular models of the Mpn444 monomer (ribbon models in dark slate grey, sea green, and a rainbow gradient from N- to C-terminus). The C-terminal lipidated cysteines mediating membrane anchoring are indicated with red and yellow arrows.

d The central trimer interface is formed by three loops. The amino acid composition can be found on the right side and individual amino acids are labelled on the left side. The interface surface is hydrophilic and moderately basic.

Discussion

Here, we present the structure of a previously uncharacterized mycoplasma membrane protein, as well as the molecular architecture of its trimeric complex. It is the first experimental structure of any *M. pneumoniae* lipoprotein. Comparison with AlphaFold predictions suggests that the fold is conserved through multiple *Mycoplasma* species (Supplemental Figure 2). *M. pneumoniae* and its closest relative *M. gentitalium* are known for their minimal set of genes, yet they each use three genes for essential proteins sharing the same structure (Mpn444 with Mpn436 and Mpn489 in *M. pneumoniae* and MG309, MG307 and MG338 in *M. genitalium*) (Figure 2b, Supplemental Figure 2), strongly suggesting a specialized function of these proteins. Considering the structural similarity of Mpn444 with Mpn436 and Mpn489 as well as their matching expression levels (48, 57 and 48 copies per cell on average), (10) we infer that they might form heterotrimeric complexes *in vivo*. While functional studies have shown the existence of multiple crosslinks between each of the three proteins (14), whose position indicates that a heterotrimeric complex might have a different architecture than the homotrimeric Mpn444 complex. However, the total number of crosslinks (2 between Mpn444 and Mpn436, 2 between Mpn444 and Mpn489, 1 between Mpn436 and Mpn489) is low when compared to the heterotetrameric major adhesion complex (14 crosslinks between Mpn141/mgpA and Mpn141), challenging the formation of *in situ* heterotrimers.

Our structure of the Mpn444 trimer, whose architecture implicates its orientation towards the mycoplasma membrane, is valuable as a template to identify Mpn444 on tomograms of mycoplasma cells in future research. Collecting information about the cellular context, including co-localizing cytoplasmic and membrane proteins as well as other interaction partners, may facilitate the elucidation of the protein function.

Classification of sub-tomograms of Mpn444 could open a route to investigate and potentially map ectodomain shedding of the protein.

While *Mycoplasma pneumoniae* infections are treatable with antibiotics, macrolide resistance is on the rise and no approved vaccines exist to date (20). Considering that Mpn444 is an essential protein of *Mycoplasma pneumoniae* that is also involved in host immune evasion through antigenic variation, our structure may open opportunities to elucidate the pathogenesis of *M. pneumoniae* community-acquired pneumonia and for the development of new treatment strategies against this pathogen.

Material & Methods

Expression and purification of Mpn444

For expression, a pET-21(+) vector containing a codon optimized *mpn444* construct without the N-terminal signal sequence (leaving aa 28 – 1325) but with a C-terminal His6-tag was used (Twist Bioscience, San Francisco, USA). *E. coli* BL21 (DE3) cells harboring the vector were grown overnight at 37°C in LB medium. The overnight culture was inoculated in in 2x YT medium autoinduction medium supplemented with 0.2% lactose, 0.05% glucose and 0.6% glycerol grown overnight at 37°C with mild shaking. Cells were harvested, lysed by sonication in lysis buffer (50 mM Tris-HCl at pH 7.4, 200 mM NaCl, 20 mM imidazole, 1 mM PMSF, 0 – 0.5% LMNG) and centrifuged at 15,000 x g (Centrifuge 5804 R, Eppendorf, Hamburg, Germany) for 15 minutes at 4°C and again at 20,400 x g in a tabletop centrifuge for 30 min at 4°C. The supernatant was loaded onto a HisTrap 5 mL column (GE Healthcare, Chicago, USA) that was pre-equilibrated with binding buffer (20 mM Tris-HCl, pH 7.4, 200 mM NaCl, 20 mM imidazole, 0 – 0.5% LMNG), thoroughly washed with binding buffer and eluted with elution buffer (20 mM Tris-HCl pH 7.4, 200 mM NaCl, 400 mM imidazole, 0 – 0.5% LMNG). The eluate was loaded onto a Superose 6 Increase 10/300 GL column (Cytivia, Marlborough, USA) in protein buffer (20 mM TRIS-HCl, pH 7.4, 200 mM NaCl, 0 – 0.005% LMNG). The expression yield was approximately 2.5 mg/L. Fractions were pooled and stored at –80°C. The identity and purity of Mpn444 was confirmed by SDS-PAGE and mass spectrometry (carried out by the functional genomics center at ETH Zurich, Switzerland).

Single-particle cryo-EM

A 3.5 µL drop of purified Mpn444 (200 µg/mL in buffer containing 20 mM Tris pH 7.4, 200 mM NaCl and 0 – 0.01% LMNG or 1 mM CHAPSO) was applied to a 45 s glow-discharged R1.2/1.3 C-flat grid (Electron Microscopy Science, Hatfield, USA) or R1.2/1.3 300 copper mesh quantifoil grid (Quantifoil Micro Tools, Großlöbichau, Germany). Grids were plunge-frozen in liquid ethane (Vitrobot Mark IV; Thermo Fischer Scientific) at 4°C, 100% relative humidity, with a nominal blot force of –3, a wait time of 10 s, and a blotting time of 14 s. Before freezing, Whatman 595 filter papers were incubated for 1 h in the Vitrobot chamber at 100% relative humidity and 4°C.

Dose-fractionated movies of Mpn444 were collected with SerialEM v4.1.0beta (21) at a nominal magnification of 105,000 (0.92Å per pixel) in nanoprobe EFTEM mode at 300 kV with a Titan Krios (Thermo Fischer Scientific) electron microscope equipped with a GIF Quantum S.E. post-column energy filter in zero-loss peak mode and a K3 Summit detector (Gatan, Pleasanton, USA). The camera was operated in dose-fractionation counting mode with a dose rate of ~16 electrons per $Å^2$ s^{-1}, resulting in a total dose of 50 electrons per $Å^2$. The frame time was adjusted to 1 frame per 1 electron per $Å^2$. A total of 7,811 micrographs were recorded at defocus values ranging from –0.8 to –3.5 µm. Micrographs were acquired at 0°, 20°, 30° and 40° stage tilt to overcome the evident preferred orientation of the particles.

CryoSPARC v4.5.3 (16) was used to process the cryo-EM data, unless stated otherwise. Beam-induced motion correction and contrast transfer function (CTF) estimation were performed separately for each stage tilt dataset using cryoSPARC's own implementation. For the monomers, particles were initially selected with the Blob picker using a particle diameter of 100–200 Å. Particles were then subjected to unsupervised 2D classification. The selected 2D classes were used as templates to

retrieve more particles from the micrographs. A total of 5,832,694 particles were extracted with a box size of 480 pixels and subjected to unsupervised 2D classification. Selected 2D classes were uniformly rebalanced (using 8 superclasses and a rebalancing factor of 1). The remaining 497,320 particles were used to generate an *ab initio* reconstruction. Subsequent non-uniform refinement was followed by local and global CTF refinement and a non-uniform refinement (22). A final 3D classification resulted in a global monomeric resolution of 3.04 Å (3.80 Å according to EMDB validation) from 156,400 particle images with a resolution range of 2.7 – 5.5 Å (Figure 1, Supplemental Table 1, EMD-51459, PDB-9GN0).

For the homotrimers, particles were initially selected with the blob picker using a particle diameter of 150–300 Å. Particles were then subjected to unsupervised 2D classification. Particles were also selected with the template picker, using the generated 2D averages as templates. After removal of duplicate particles, a total of 3,376,184 particles were extracted with a box size of 480 pixels. False-positive picks were removed by two rounds of unsupervised 2D classification. The remaining 626,110 particles were used to generate *ab initio* reconstructions with C1 and C3 symmetry. The non-symmetrized reconstructions displayed the same overall domain architecture as the C3-symmetrized reconstruction but with weaker density in one monomer that we attribute to the persistent presence of monomers and dimers in the particle set after 2D classification. All following refinement steps were done with C3 symmetry.

The CTF was refined locally and globally, followed by a non-uniform refinement (22).The obtained global resolution of the trimer was 5.4 Å (6.31 Å according to EMDB validation) with a resolution range of 4 to 9 Å (Supplemental Figure 5, Supplemental Table 1, PDB-9GNH, EMD-51470).

Model building

The quality of our maps was initially assessed using ModelAngelo (23) with and without FASTA sequence. A reduction of map anisotropy was attempted using preferred orientation-correction with spIsoNet (24) and anisotropic sharpening using the Phenix v1.21.1-5286 suite (25), but neither tool improved map quality for model building. An input AlphaFold3 predicted model (26) was used in the Predict and Built tool in Phenix (27) to generate an initial model. Missing loops were modelled with the Fit Loops tool in Phenix. Loops for which no defined density could be discerned were not modelled, yielding a sequence coverage of 80%. The initial model was iteratively refined between the Real-Space Refinement tool in Phenix (global), the Real-Space Refinement Zone tool in Coot v0.9.8 (28) (per residue) and manual rotamer modification in Coot. The whole Mpn444 model was then refined using the Real-Space Refinement tool in Phenix, and the final model was deposited in the PDB under the accession code 9GN0 (Supplemental Table 1).

The resolution of the Mpn444 trimer cryo-EM density map was too low to allow for ab-initio model building. Thus, the model of the trimeric assembly was built by first rigid-body fitting the cryo-EM density map of the monomer into the cryo-EM density map of the trimer and then rigid-body fitting the monomer model into the cryo-EM density map of the monomer and is deposited in the PDB under the accession code 9GNH (Supplemental Table 1).

Data Availability

The cryo-EM structures solved in this study are available at http://www.emdataresource.org/ under EMD-51459, PDB-9GN0, EMD-51470, and PDB-9GNH (Supplemental Table 1). The vector is available upon request.

Acknowledgements and Funding

We would like to thank Margot P. Scheffer for the critical reading of the manuscript. We thank the Frankfurt Center for Electron Microscopy for measurement time. We thank the Pos lab (Goethe University Frankfurt) for their support with the ÅKTA system. A. S. F. acknowledges support from the Deutsche Forschungsgemeinschaft and the Research Training Group iMOL (GRK 2566/1 for S.M., I.K. and M.S.M.).

Author contributions

I.K.: Expressed and purified Mpn444, prepared samples for single-particle analysis.

S.M.: Processed single-particle datasets of Mpn444 monomers and trimers, wrote the manuscript, with contributions from all authors.

I.K. and S.M.: Designed the expression plasmid for Mpn444, recorded single-particle datasets of Mpn444, built and refined the molecular model of Mpn444.

M.S.M.: Assisted in processing single-particle datasets of Mpn444 monomers as well as building and refining of the molecular model.

A.S.F.: Designed and supervised research.

Competing interests

The authors declare no competing interests.

References

1. J. L. Waller *et al.*, Detection and Characterization of *Mycoplasma pneumoniae* during an Outbreak of Respiratory Illness at a University. *Journal of Clinical Microbiology* **52**, 849-853 (2014).
2. M. F. Balish, *Mycoplasma pneumoniae*, an Underutilized Model for Bacterial Cell Biology. *Journal of Bacteriology* **196**, 3675-3682 (2014).
3. M. A. Meseguer *et al.*, *Mycoplasma pneumoniae*: a reduced-genome intracellular bacterial pathogen. *Infect Genet Evol* **3**, 47-55 (2003).
4. M. Widjaja, I. J. Berry, E. J. Pont, M. P. Padula, S. P. Djordjevic, P40 and P90 from Mpn142 are Targets of Multiple Processing Events on the Surface of *Mycoplasma pneumoniae*. *Proteomes* **3**, 512-537 (2015).
5. Y. Li *et al.*, DnaK Functions as a Moonlighting Protein on the Surface of *Mycoplasma hyorhinis* Cells. *Frontiers in Microbiology* **13** (2022).
6. M. Widjaja *et al.*, Elongation factor Tu is a multifunctional and processed moonlighting protein. *Scientific Reports* **7** (2017).
7. K. M. Hallamaa, G. F. Browning, S. L. Tang, Lipoprotein Multigene Families in *Mycoplasma pneumoniae*. *Journal of Bacteriology* **188**, 5393-5399 (2006).
8. A. Christodoulides *et al.*, The Role of Lipoproteins in *Mycoplasma*-Mediated Immunomodulation. *Frontiers in Microbiology* **9** (2018).
9. K. M. Hallamaa, S.-L. Tang, N. Ficorilli, G. F. Browning, Differential expression of lipoprotein genes in *Mycoplasma pneumoniae* after contact with human lung epithelial cells, and under oxidative and acidic stress. *BMC Microbiology* **8** (2008).
10. C. Elfmann *et al.*, MycoWiki: Functional annotation of the minimal model organism *Mycoplasma pneumoniae*. *Front Microbiol* **13**, 935066 (2022).
11. A. Bateman *et al.*, UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research* **51**, D523-D531 (2023).
12. J. Zhang *et al.*, Inter- and Intra-Strain Variability of Tandem Repeats in *Mycoplasma Pneumoniae* Based on Next-Generation Sequencing Data. *Future Microbiology* **12**, 119-129 (2016).
13. C. L. McGowin, L. Ma, D. H. Martin, R. B. Pyles, *Mycoplasma genitalium* -Encoded MG309 Activates NF-κB via Toll-Like Receptors 2 and 6 To Elicit Proinflammatory Cytokine Secretion from Human Genital Epithelial Cells. *Infection and Immunity* **77**, 1175-1181 (2009).
14. F. J. O'Reilly *et al.*, In-cell architecture of an actively transcribing-translating expressome. *Science* **369**, 554-557 (2020).
15. J. C. C. Somody (2023) Leveraging deep learning for identification and structural determination of novel protein complexes from in situ electron cryotomography of *Mycoplasma pneumoniae*.
16. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods* **14**, 290-296 (2017).
17. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
18. M. van Kempen *et al.*, Fast and accurate protein structure search with Foldseek. *Nature Biotechnology* **42**, 243-246 (2023).
19. J. I. Glass *et al.*, Essential genes of a minimal bacterium. *Proceedings of the National Academy of Sciences* **103**, 425-430 (2006).

20. Y.-C. Chen, W.-Y. Hsu, T.-H. Chang, Macrolide-Resistant *Mycoplasma pneumoniae* Infections in Pediatric Community-Acquired Pneumonia. *Emerging Infectious Diseases* **26**, 1382-1391 (2020).
21. D. N. Mastronarde, SerialEM: A Program for Automated Tilt Series Acquisition on Tecnai Microscopes Using Prediction of Specimen Position. *Microscopy and Microanalysis* **9**, 1182-1183 (2003).
22. A. Punjani, H. Zhang, D. J. Fleet, Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat Methods* **17**, 1214-1221 (2020).
23. K. Jamali *et al.*, Automated model building and protein identification in cryo-EM maps. *Nature* **628**, 450-457 (2024).
24. Y. T. Liu, H. Fan, J. J. Hu, Z. H. Zhou, Overcoming the preferred orientation problem in cryoEM with self-supervised deep-learning. *bioRxiv* 10.1101/2024.04.11.588921 (2024).
25. D. Liebschner *et al.*, Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861-877 (2019).
26. J. Abramson *et al.*, Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493-500 (2024).
27. T. C. Terwilliger, P. D. Adams, P. V. Afonine, O. V. Sobolev, A fully automatic method yielding initial models from high-resolution cryo-electron microscopy maps. *Nat Methods* **15**, 905-908 (2018).
28. A. Casanal, B. Lohkamp, P. Emsley, Current developments in Coot for macromolecular model building of Electron Cryo-microscopy and Crystallographic Data. *Protein Sci* **29**, 1069-1078 (2020).