Structural Analysis of the *Helicobacter pylori* Pore-Forming Toxin, VacA

by

Amanda L. Erwin

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Cell and Developmental Biology)
in the University of Michigan
2020

Doctoral Committee:

     Professor Lois S. Weisman, Chair
     Assistant Professor Michael Cianfrocco
     Associate Professor Melanie D. Ohi
     Assistant Professor Randy Buzzell Stockbridge

Amanda L. Erwin

erwinal@umich.edu

ORCID iD: 0000-0003-1308-5405

"It takes a village."

This dissertation is dedicated to my village of family and friends.

Acknowledgements

First, I would like to thank Dr. Melanie Ohi for her mentorship and infinite optimism, patience, and trust. Thank you for helping me grow as a scientist by presenting me with challenging and exciting opportunities and advocating for my career and ideas.

I would like to thank my committee members Dr. Lois Weisman, Dr. Michael Cianfrocco, Dr. Randy Stockbridge for their feedback and encouragement during my studies at University of Michigan. Scientific and personal discussions with Dr. Lois Weisman in the LSI 6th floor kitchen were always motivating and enriched my PhD experience and training. I would like to thank Dr. Mike Cianfrocco for productive discussions regarding my research and career and increasing the efficiency of my cryo-EM imaging sessions. I would like to thank Dr. Randy Stockbridge for sharing her excitement for unique membrane proteins and her insightful scientific and career advice.

Thank you to past and present members of the Ohi lab, especially Dr. Louise Chang, Dr. Eddy Chung, Sarah Connolly, Dr. Clarissa Durie, Dr. Nora Foegeding, Jessica Hanks, Dr. Kelly Manthei, Ashley Schork, and Rachel Torrez - I am grateful to have had the privilege to work with such fun, kind, knowledgeable, and passionate scientists. Thank you to Dr. Louise Chang for being my lab mom and providing encouragement and endless wisdom during my PhD studies. Thank you to my labmates

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AdiC   Arginine/agmatine antiporter

BabA   Blood group antigen-binding adhesion

CagA   Cytotoxin-associated antigen A

CCD   Charge-coupled device

CDC   Cholesterol-dependent cytolysin

cisTEM   Computational imaging system for transmission electron microscopy

cNW50   Circularized 50 nm nanodisc

CMC   Critical micelle concentration

Cryo-EM   Cryo-electron microscopy

CTF   Contrast transfer function

CYMAL-6   6-Cyclohexyl-1-Hexyl-β-D-Maltoside

CYMAL-7   7-Cyclohexyl-1-Heptyl-β-D-Maltoside

DOPS   1,2-dioleoyl-sn-glycero-3-phospho-L-serine sodium

DDM   N-Dodecyl-B-D-Maltoside

DMSO   Dimethyl sulfoxide

DTT   Dithiothreitol

EDTA   Ethylenediaminetetraacetic acid

eggPC   Egg L-a-phosphatidylcholine

EGFR   Epidermal growth factor receptor

EM	Electron microscopy
Fab	Fragment antigen-binding
FSC	Fourier Shell Correlation
FFT	Fast fourier transform
Fluc	Fluoride channel
Fluc-Bpe	*Bordetella pertussis* fluoride channel)
Fluc-Ec2	*Escherichia coli* fluoride channel
goCTF	Geometrically optimized contrast transfer function
GlpF	Glycerol uptake facilitator protein
H. pylori	*Helicobacter pylori*
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
LE	Late endosome
LMNG	Lauryl Maltose Neopentyl Glycol
LRP1	Low-density lipoprotein receptor-related protein 1
LUV	Large unilamellar vesiscle
OM	Outer membrane
POPC	1-Palmitoyl-2-Oleoyl-Glycero-3-Phosphocholine
POPG	1-Palmitoyl-2-Oleoyl-sn-Glycero-3-Phosphoglycerol
MPER	Membrane-proximal external region
NH_4Cl	Ammonium chloride
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PFT	Pore forming toxin

RELION   REgularized LIkelihood OptimizatioN

ResMap   Resolution Map

RPTP   Receptor protein tyrosine phosphatase

RT   Room temperature

scFv   Single-chain variable-domain antibody fragment

SDS   Sodium dodecyl sulfate

SEC   Size exclusion chromatography

SUV   Small unilamellar vesicle

TEM   transmission electron microscopy

TMH   Transmembrane helix

Tris   Tris(hydroxymethyl)aminomethane

VacA   Vacuolating cytotoxin A

Abstract

Helicobacter pylori is a Gram-negative bacterium that persistently colonizes the stomachs of >50% of the human population, with a prevalence as high as 90% in developing nations. *H. pylori* infection causes gastritis and can lead to the development of peptic ulcer disease and gastric adenocarcinoma in a subset of infected individuals. Gastric cancer is a leading cause of cancer-related deaths worldwide and the World Health Organization has classified H. pylori as a type 1 carcinogen.

H. pylori has evolved an arsenal of virulence factors that enhance infection, including the secreted pore forming toxin vacuolating cytotoxin A (VacA). VacA has important roles in *H. pylori* colonization of the human stomach and the pathogenesis of *H. pylori*-related gastroduodenal diseases. VacA causes multiple cellular effects including vacuolation, membrane permeabilization, mitochondrial dysfunction, cell death, and autophagy. These cellular effects are attributed to VacA oligomerizing and forming a transmembrane anion channel in various cell membranes.

The mature 88-kDa VacA toxin contains two regions (p33 and p55); the N-terminal p33 region contains hydrophobic amino acids required for channel formation and sections within both p33 and p55 mediate VacA oligomerization and binding to host cells. While oligomerization and membrane insertion are essential for VacA pore forming activity, the mechanism of how VacA oligomerizes and forms a pore in cell membranes has not been fully characterized.

To understand how VacA oligomerizes and associates with membrane, we determined structures of a soluble and membrane-bound VacA monomers and oligomers using single-particle cryo-electron microscopy. We mapped the regions of p33 and p55 involved in soluble VacA hexamer assembly, modeled how flexible interactions between protomers could support heptamer formation, and analyzed p33 and p55 residues that contribute to oligomerization using disulfide mutants. Additionally, we utilized negative stain and cryo-electron microscopy to structurally analyze VacA in a liposome bound and detergent-solubilized state resulting in a pre-pore structure of VacA associated with membrane. Together, this work provides structural insights into the process of VacA oligomerization and pore formation and identifies regions of VacA that undergo conformational changes upon membrane association. Since the molecular mechanisms by which VacA elicits its variety of cellular responses are not fully elucidated, these structural studies can be used to understand VacA function within the context of cells.

Structural analysis of membrane proteins like VacA can be challenging, especially when imaged in presence of detergent. In this work, we identified and characterized the helical MPER-epitope tag as a strategy to complex membrane proteins with existing, easy to produce MPER-targeting antibody fragments for improved structural analysis with X-ray crystallography and single-particle electron microscopy. Altogether, this dissertation provides 1) important structural insights into how VacA oligomerizes and begins to form a pore in membrane and 2) a tool for future structural analysis of membrane proteins in general.

Chapter I

Introduction

Pore-forming toxins

One defining feature of a cell is the plasma membrane, which serves as a selective barrier between the cell and its environment. Disrupting the integrity of the plasma membrane can compromise cellular ion homeostasis and may lead to cell death. From bacteria to metazoans, organisms from all kingdoms have evolved proteins that are able to target this plasma membrane barrier. The ancient pore-forming protein family includes some of the proteins that target the plasma membrane and can alter membrane permeability[1].

In general, pore-forming proteins are produced by an organism as a soluble monomer that can oligomerize, insert, and form a pore on the plasma membrane and elicit a variety of downstream cellular effects[2]. The pore-forming protein family includes pore-forming toxins, components of the type III and IV secretion systems, and the B subunits of AB toxins[3-5]. Eukaryotes such as parasites, fungi, earthworms, and plants produce pore-forming proteins as a defense mechanism or for nutrient access[2, 6, 7]. Vertebrates also produce pore-forming proteins to target bacteria (i.e. C9), infected/malignant cells (i.e. perforin), or mitochondria (i.e. Bax, Bak)[8-10].

The largest and most well characterized class of pore forming proteins are the pore-forming toxins which are expressed by many pathogenic bacteria to kill other bacteria (i.e. colicins) or as virulence factors that aid in the colonization and infection of

hosts[1, 11]. This family of proteins forms pores with distinct properties in terms of molecular weight, target membranes, target cell receptors, stoichiometry, assembly mechanism, and induced cellular responses [1]. The two major classes of bacterial pore forming toxins, α and β, are distinguished by the secondary structure of the transmembrane pore[12-14]. Over the last few decades, structural work on monomers, oligomers in solution and in the context of membrane (e.g. liposomes, detergents, nanodiscs) have shed light on the dynamic rearrangements of this protein family from monomers to pre-pores to oligomeric transmembrane pores on membrane [7, 15-26].

Cryo-electron microscopy of membrane proteins: pore forming toxins

Cryo-electron microscopy has been particularly critical for structural analysis of pore forming toxins, as it is well-suited for the analysis of heterogeneous samples. Pore forming toxins can often form multiple oligomeric assemblies in solution, ruling out X-ray crystallography as a useful technique for understanding the oligomerization and dynamic pore formation process of some of these proteins. Advances in cryo-EM direct electron detectors, micrograph alignment programs, and 3D classification algorithms have enabled detailed analyses of heterogeneous prepore and pore assemblies, allowing insights into pore formation[27-33].

While cryo-EM has made studying the structures of pore-forming toxins more feasible, there remain many challenges to high resolution structure determination of these proteins at several steps of the cryoEM workflow, especially at the sample preparation and 3D refinement steps. Pore forming toxins are especially prone to aggregation when removed from a membrane environment and can adopt preferred

views at the air-water interface. These sample preparation issues can be overcome with extensive detergent and grid type optimization[7, 23]. Recent membrane protein structure papers show the potential for structural analysis of membrane proteins incorporated into liposomes through use of thin carbon-coated or graphene oxide coated grids, extensive classification, and signal subtraction[9, 20, 34]. These are exciting advances that will enable the structure determination of pore forming toxins in a near native membrane environment with the potential to manipulate the pH or ionic content on both sides of the bilayer, which can be important for pore forming toxin function.

H. pylori infection and virulence factors

Helicobacter pylori is a Gram-negative bacterium that persistently colonizes the stomachs of >50% of the human population, with a prevalence as high as 90% in developing nations[35]. *H. pylori* infection causes gastritis and can lead to the development of peptic ulcer disease and gastric adenocarcinoma in a subset of infected individuals. Gastric cancer is a leading cause of cancer-related deaths worldwide and the World Health Organization has classified *H. pylori* as a type 1 carcinogen[36-38]. *H. pylori* is proposed to be transmitted from person-to-person contact[39]. Early detection of *H. pylori* infection and a combined treatment of a proton-pump inhibitor and multiple antibiotics can prevent *H. pylori*-associated gastroduodenal diseases[40-42]. Growing resistance to antibiotics used to treat *H. pylori* necessitates identifying specific *H. pylori* proteins as future drug targets[43].

H. pylori is able colonize the harsh, acidic environment of the human stomach (pH < 4) due to their motility, chemotaxis behavior, and virulence factors. *H.pylori* uses its flagella and spiral shape to move toward the more neutral mucosal layer and reach the gastric epithelium (Figure 1.1)[44]. *H. pylori* expresses the virulence factor urease, which buffers the local pH environment around it by hydrolyzing urea to ammonia and carbamate, resulting in a net increase in pH[45]. Other virulence factors that enhance infection include BabA (blood group antigen-binding adhesin), CagA (cytotoxin associated gene A), and VacA (vacuolating cytotoxin A)[46]. BabA is an *H. pylori* outer membrane protein proposed to be important for *H. pylori* adhesion to the gastric epithelium[47]. CagA is translocated into host cells through the Cag Type IV Secretion System where it is phosphorylated and leads to multiple host cell morphology changes and alterations in cell signaling pathways[48-51]. VacA is secreted as an 88kDa pore forming toxin that causes multiple cellular effects, including vacuolation and membrane permeabilization, that are attributed to its ability to form an anion channel in various cell membranes (Figure 1.2)[52].

Structure and function of *H. pylori* pore forming toxin VacA

VacA has important roles in *H. pylori* colonization of the human stomach and the pathogenesis of *H. pylori*-related gastroduodenal diseases. VacA has been reported to cause a wide range of cellular alterations including vacuolation, membrane permeabilization, mitochondrial dysfunction, cell death, and autophagy (Figure 1.2)[52]. These cellular effects are attributed to VacA oligomerizing and forming a

4

transmembrane anion channel in the host cell plasma membrane, endosomal membranes, and mitochondrial membranes (Figure 1.2)[52].

VacA does not share sequence homology with any other known bacterial pore forming toxins. VacA is translated as a 140 kDa protoxin that is comprised of an N-terminal 3 kDa signal peptide, 88 kDa mature toxin, 12 kDa peptide of unknown function, and C-terminal 35 kDa domain that has been predicted to be a beta-barrel autotransporter domain (Fig 1.3)[52-55]. Protoxin translocation through the *H. pylori* inner membrane is predicted to be initiated by the signal peptide binding the Sec machinery[56]. Subsequently, the 33 residue signal peptide is cleaved off in the periplasm[56]. VacA (1-1287) is proposed to be secreted from *H. pylori* via a Type Va secretion mechanism, using the predicted C-terminal autotransporter domain for protein translocation through the *H. pylori* outer membrane (Fig 1.3A)[57]. The mature 88 kDa toxin (residues 1-821) and a 12kDa peptide of unknown function are then cleaved by an unidentified protease[57].

The mature toxin VacA p88 forms an elongated, parallel right-handed beta-helix, a structure shared with Gram-negative bacteria passenger domains secreted through the Type Va autotransporter mechanism[24, 26, 58, 59]. p88 is susceptible to limited proteolysis by trypsin which yields a 33 kDa fragment (p33, residues 1-311) and a 55 kDa fragment (p55, residues 312-821). These two distinct regions, p33 and p55, are bridged by a hydrophilic loop (~300-334) that has been predicted to be flexible in solution (Fig 1.3C)[54, 60]. There are three main regions of sequence variation that have been identified in the *vacA* gene: an N-terminal p33 region (s1 or s2), an intermediate region (i1 or i2) located near the C-terminus of p33, and a p55 middle-

5

region (m1 or m2)[61-64]. The s2-type VacA has a 12 residue hydrophilic extension (NTPNDPIHSESR) that prevents VacA pore formation and cell vacuolation[65, 66]. Variation in the p33 i-region and p55 m-region (~250-300 residues) results in differences in VacA binding to cells[62, 67, 68]. All *H. pylori* strains have the *vacA* gene with multiple combinations of the s-, i-, and m- type sequences[69]. Humans infected with strains containing *vacA* s1/i1/m1 have a higher risk of developing *H. pylori*-associated gastroduodenal diseases compared to strains containing *vacA* s2/i2/m2[38, 61, 63, 70, 71].

The use of various VacA mutants has shown that oligomerization and the formation of active transmembrane channels is critical for VacA activity. It has been shown that when expressed inside the cell, the minimum region required for VacA-induced cell vacuolation comprises residues 1-433 (all of p33 and 122 residues of p55)[77]. Residues from both the p33 an p55 regions are required for the formation of oligomers in solution and hexameric/heptameric oligomers on membrane[24, 26, 58, 59, 74-76]. The non-oligomerizing mutant Δ346-347 (with two p55 residues deleted) is able to bind membrane, but not able to oligomerize in solution or on membrane[76, 78]. The N-terminal section of p33 contains three tandem hydrophobic GXXXG motifs (residues 14-26) proposed to form a transmembrane helical bundle that functions as an anion channel (Figure 1.3C)[72, 73]. Deletion of the GXXXG motif region (Δ6-27) results in VacA that oligomerizes and binds membrane, but is unable to vacuolate cells in culture or conduct chloride across planar lipid bilayers[72, 76, 79]. Mutations in the GXXXG region (proline 9, glycine 14, and glycine 18) decreases VacA channel formation on planer lipid bilayers and cell vacuolation[72].

In order to form an active anion selective channel in host cell membranes to elicit a wide range of downstream cellular effects, VacA must oligomerize and insert into the membrane bilayer. While several structural studies of VacA oligomers have determined the type of oligomers present in solution (hexamers, heptamers, dodecamers, tetradecamers) and bound to membrane (predominantly hexamers), as well as the beta-helical nature of this passenger domain, the structural basis and mechanism of VacA oligomerization and membrane insertion, critical steps for VacA activity in gastroepithelial cells, are not fully understood.

Unanswered questions in VacA biology

There are several remaining unanswered questions in VacA biology. It has not been clearly shown how VacA is secreted from H*. pylori.* After secretion, VacA adopts hexamers and heptamers on membrane (like all pore forming toxins, many of which can form multiple co-existing assemblies). The functional significance of hexameric and heptameric oligomers on membrane is not clear and must be further defined. How and where VacA oligomerization and pore formation occur are not well understood – further exploration of requirements for this process including, but not limited to pH, substrates, receptors, intermediate states are needed. Finally, given that the *vacA* gene is associated with enhancing *H. pylori* colonization of the human stomach, why are there variants like s2/i2/m2 that are not as active? Why do these *vacA* variants exist in nature if not advantageous for bacterium?

Summary

Structural biology techniques have been used to further understand the diverse mechanisms of pore forming toxins. VacA is a unique pore-forming toxin that is involved in H. pylori colonization of the human stomach and is associated with the development of *H. pylori*-related gastroduodenal diseases. While VacA requires oligomerization and pore formation for its wide range of cellular alterations in gastroepithelial cells, the structural basis for these steps is not well understood. In my dissertation studies, I aimed to determine the structure of the *H. pylori* VacA pore forming toxin in a variety of states using cryo-electron microscopy. I sought to understand how VacA oligomerizes in solution (Chapter II), and forms a pore in a membrane-like environment (Chapter III). I also aimed to address a key challenge in the membrane protein structural biology field through a collaborative work aimed at improving the structural analysis of small membrane protein complexes using a helical epitope tag (Chapter IV).
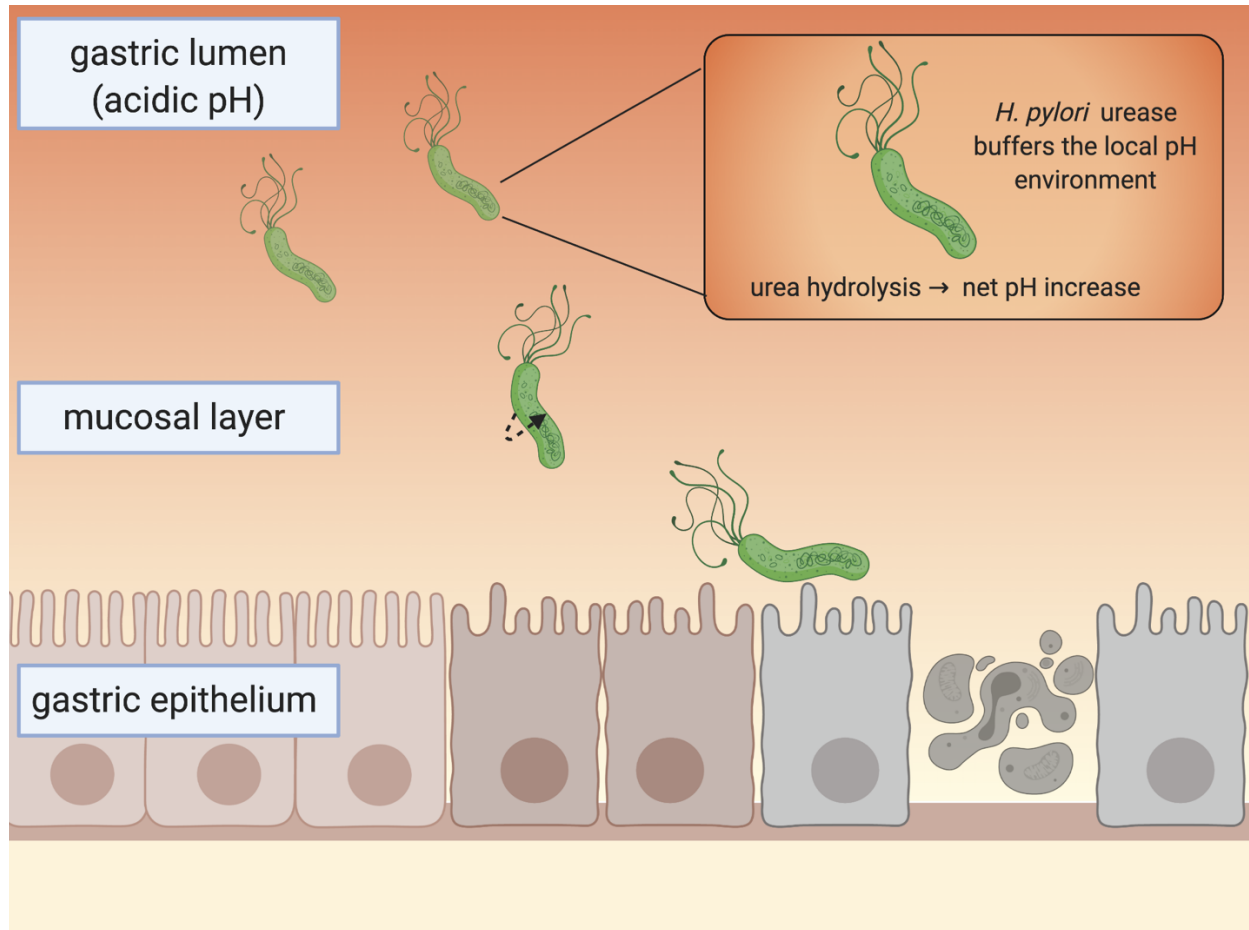
Figure 1.1. *H. pylori* infection of the gastric epithelium. *H. pylori* expresses the virulence factor urease to buffer its local pH environment. Urease hydrolyzes urea to ammonia and carbamate, which decomposes to another molecule of ammonia and carbonic acid, resulting in a net increase in pH. *H. pylori* uses it flagella and spiral shape to move through the mucosal layer to reach the gastric epithelium. Made with BioRender.
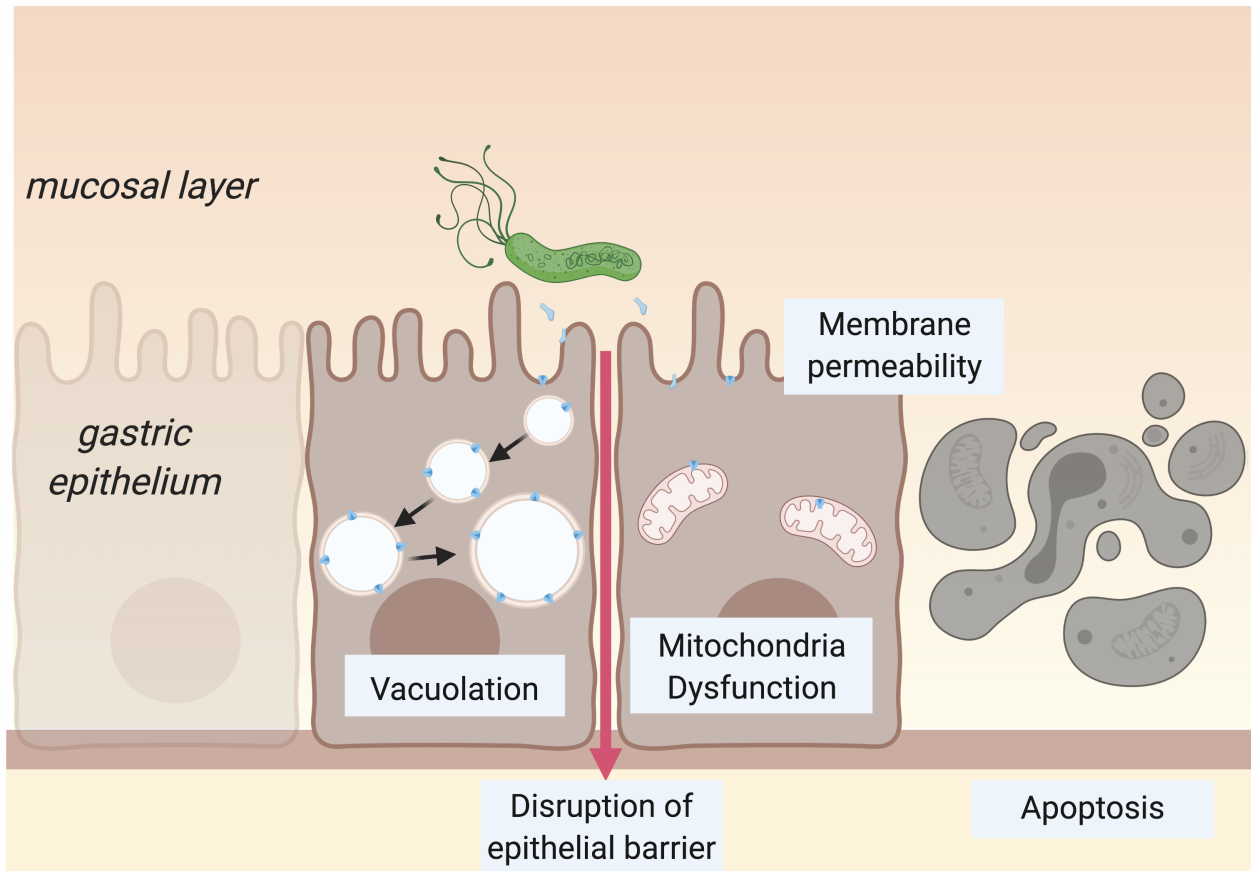
Figure 1.2. *H. pylori* **VacA causes multiple cellular effects.** VacA is secreted from *H. pylori* as p88 (88 kDa) monomers that bind to the plasma membrane of gastric epithelial cells and oligomerize on various cell membranes (VacA monomers and oligomers depicted in light blue). After oligomerization, VacA has been shown to elicit a number of cellular effects including vacuolation, mitochondrial dysfunction, disruption of cell-cell junctions, and apoptosis. Many of its effects on host cells are attributed to VacA oligomers forming channels in intracellular sites such as endosomes and the mitochondria. Made with BioRender.

A **Type Va Secretion of VacA**

OM

periplasm

B **140 kDa protoxin**

12 kDa cleaved peptide

-33 1   1287

signal peptide for Sec-pathway secretion
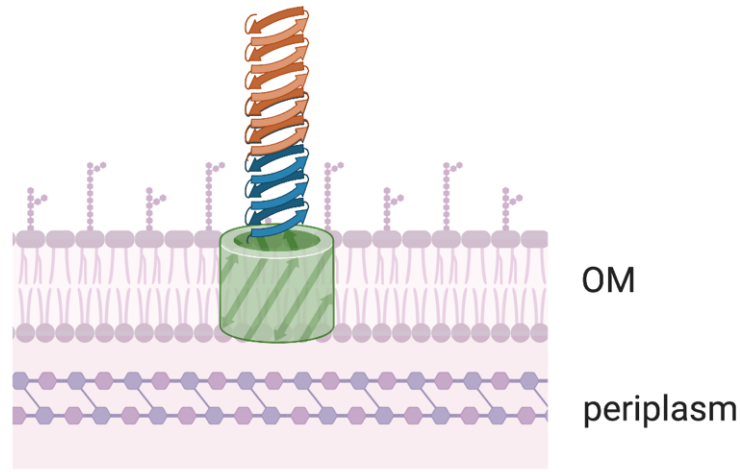
autotransporter domain for OM secretion
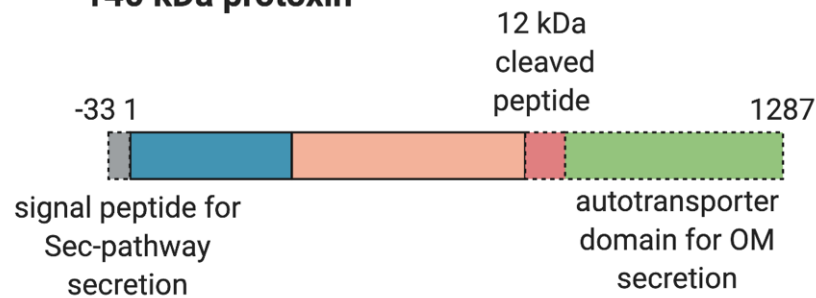
C **88 kDa mature toxin (p88)**

1   311 312   821

p33   p55
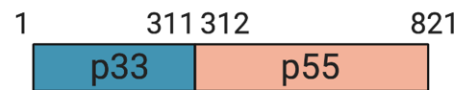
GXXXG motif region*

Hydrophilic loop^

* [1]AFFTTVIIPAIVGGIATGTAVGTVSGL[27]

^ [300]DKPNNTPSQSGAKNDKQESSQNNSNTQVINPPNST[334]

Figure 1.3. VacA secretion and domain organization. A) VacA is secreted (partial secretion depicted by orange and teal beta strands) through the outer membrane (OM) of *H. pylori* via a Type Va secretion mechanism, utilizing the C-terminal autotransporter domain (green) for protein translocation. The mature 88 kDa toxin and a 12kDa peptide

of unknown function are then cleaved by an unidentified protease. B) VacA is translated as a 140kDa protoxin comprised of a 3 kDa signal peptide for Sec-pathway secretion (grey, dashed), 88kDa mature toxin (teal and orange), 12 kDa cleaved peptide of unknown function (red, dashed), and 35 kDa predicted autotransporter beta-barrel domain for secretion through the *H. pylori* outer membrane (green, dashed). C) The 88 kDa mature VacA toxin (p88) is comprised of two distinct regions, p33 (teal) and p55 (orange), that are bridged by a hydrophilic loop. The N-terminal section of p33 contains hydrophobic GXXXG motifs proposed to form a transmembrane pore. Sequences of GXXXG motif region (*) and the hydrophilic loop (^) region from VacA s1/i1/m1 are listed. The three tandem GXXXG motifs are underlined. Made with BioRender.

References

[1] Dal Peraro M, van der Goot FG. Pore-forming toxins: ancient, but never really out of fashion. Nat Rev Microbiol. 2016;14:77-92.

[2] Bischofberger M, Iacovache I, van der Goot FG. Pathogenic pore-forming proteins: function and host response. Cell Host Microbe. 2012;12:266-75.

[3] Green ER, Mecsas J. Bacterial Secretion Systems: An Overview. Microbiol Spectr. 2016;4.

[4] Los FC, Randis TM, Aroian RV, Ratner AJ. Role of pore-forming toxins in bacterial infectious diseases. Microbiol Mol Biol Rev. 2013;77:173-207.

[5] Odumosu O, Nicholas D, Yano H, Langridge W. AB toxins: a paradigm switch from deadly to desirable. Toxins (Basel). 2010;2:1612-45.

[6] Rojko N, Kristan KC, Viero G, Zerovnik E, Macek P, Dalla Serra M, et al. Membrane damage by an alpha-helical pore-forming protein, Equinatoxin II, proceeds through a succession of ordered steps. J Biol Chem. 2013;288:23704-15.

[7] Bokori-Brown M, Martin TG, Naylor CE, Basak AK, Titball RW, Savva CG. Cryo-EM structure of lysenin pore elucidates membrane insertion by an aerolysin family protein. Nat Commun. 2016;7:11293.

[8] Dudkina NV, Spicer BA, Reboul CF, Conroy PJ, Lukoyanova N, Elmlund H, et al. Structure of the poly-C9 component of the complement membrane attack complex. Nat Commun. 2016;7:10588.

[9] Pang SS, Bayly-Jones C, Radjainia M, Spicer BA, Law RHP, Hodel AW, et al. The cryo-EM structure of the acid activatable pore-forming immune effector Macrophage-expressed gene 1. Nat Commun. 2019;10:4288.

[10] Westphal D, Dewson G, Czabotar PE, Kluck RM. Molecular biology of Bax and Bak activation and action. Biochim Biophys Acta. 2011;1813:521-31.

[11] Cascales E, Buchanan SK, Duche D, Kleanthous C, Lloubes R, Postle K, et al. Colicin biology. Microbiol Mol Biol Rev. 2007;71:158-229.

[12] Gouaux E. Channel-forming toxins: tales of transformation. Curr Opin Struct Biol. 1997;7:566-73.

[13] Lesieur C, Vecsey-Semjen B, Abrami L, Fivaz M, Gisou van der Goot F. Membrane insertion: The strategies of toxins (review). Mol Membr Biol. 1997;14:45-64.

[14] Parker MW, Feil SC. Pore-forming protein toxins: from structure to function. Prog Biophys Mol Biol. 2005;88:91-142.

[15] Brauning B, Bertosin E, Praetorius F, Ihling C, Schatt A, Adler A, et al. Structure and mechanism of the two-component alpha-helical pore-forming toxin YaxAB. Nat Commun. 2018;9:1806.

[16] Gatsogiannis C, Lang AE, Meusch D, Pfaumann V, Hofnagel O, Benz R, et al. A syringe-like injection mechanism in Photorhabdus luminescens toxins. Nature. 2013;495:520-3.

[17] Gatsogiannis C, Merino F, Prumbaum D, Roderer D, Leidreiter F, Meusch D, et al. Membrane insertion of a Tc toxin in near-atomic detail. Nat Struct Mol Biol. 2016;23:884-90.

[18] Gatsogiannis C, Merino F, Roderer D, Balchin D, Schubert E, Kuhlee A, et al. Tc toxin activation requires unfolding and refolding of a beta-propeller. Nature. 2018;563:209-13.

[19] Peng W, de Souza Santos M, Li Y, Tomchick DR, Orth K. High-resolution cryo-EM structures of the E. coli hemolysin ClyA oligomers. PLoS One. 2019;14:e0213423.

[20] Piper SJ, Brillault L, Rothnagel R, Croll TI, Box JK, Chassagnon I, et al. Cryo-EM structures of the pore-forming A subunit from the Yersinia entomophaga ABC toxin. Nat Commun. 2019;10:1952.

[21] Roderer D, Hofnagel O, Benz R, Raunser S. Structure of a Tc holotoxin pore provides insights into the translocation mechanism. Proc Natl Acad Sci U S A. 2019;116:23083-90.

[22] Savva CG, Clark AR, Naylor CE, Popoff MR, Moss DS, Basak AK, et al. The pore structure of Clostridium perfringens epsilon toxin. Nat Commun. 2019;10:2641.

[23] Schubert E, Vetter IR, Prumbaum D, Penczek PA, Raunser S. Membrane insertion of alpha-xenorhabdolysin in near-atomic detail. Elife. 2018;7.

[24] Su M, Erwin AL, Campbell AM, Pyburn TM, Salay LE, Hanks JL, et al. Cryo-EM Analysis Reveals Structural Basis of Helicobacter pylori VacA Toxin Oligomerization. J Mol Biol. 2019;431:1956-65.

[25] van Pee K, Neuhaus A, D'Imprima E, Mills DJ, Kuhlbrandt W, Yildiz O. CryoEM structures of membrane pore and prepore complex reveal cytolytic mechanism of Pneumolysin. Elife. 2017;6.

[26] Zhang K, Zhang H, Li S, Pintilie GD, Mou TC, Gao Y, et al. Cryo-EM structures of Helicobacter pylori vacuolating cytotoxin A oligomeric assemblies at near-atomic resolution. Proc Natl Acad Sci U S A. 2019;116:6800-5.

[27] Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, Agard DA. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods. 2017;14:331-2.

[28] Scheres SH. RELION: implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol. 2012;180:519-30.

[29] Scheres SH. Processing of Structurally Heterogeneous Cryo-EM Data in RELION. Methods Enzymol. 2016;579:125-57.

[30] Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife. 2018;7.

[31] Nakane T, Kimanius D, Lindahl E, Scheres SH. Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. Elife. 2018;7.

[32] Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods. 2017;14:290-6.

[33] Takizawa Y, Binshtein E, Erwin AL, Pyburn TM, Mittendorf KF, Ohi MD. While the revolution will not be crystallized, biochemistry reigns supreme. Protein Sci. 2017;26:69-81.

[34] Yao X, Fan X, Yan N. Cryo-EM analysis of a membrane protein embedded in the liposome. Proc Natl Acad Sci U S A. 2020;117:18497-503.

[35] Suerbaum S, Michetti P. Helicobacter pylori infection. N Engl J Med. 2002;347:1175-86.

[36] Moller H, Heseltine E, Vainio H. Working group report on schistosomes, liver flukes and Helicobacter pylori. Int J Cancer. 1995;60:587-9.

[37] Fuchs CS, Mayer RJ. Gastric carcinoma. N Engl J Med. 1995;333:32-41.

[38] Cover TL. Helicobacter pylori Diversity and Gastric Cancer Risk. mBio. 2016;7:e01869-15.

[39] Cover TL, Blaser MJ. Helicobacter pylori in health and disease. Gastroenterology. 2009;136:1863-73.

[40] Gisbert JP, Gonzalez L, Calvet X, Garcia N, Lopez T, Roque M, et al. Proton pump inhibitor, clarithromycin and either amoxycillin or nitroimidazole: a meta-analysis of eradication of Helicobacter pylori. Aliment Pharmacol Ther. 2000;14:1319-28.

[41] Calvet X, Garcia N, Lopez T, Gisbert JP, Gene E, Roque M. A meta-analysis of short versus long therapy with a proton pump inhibitor, clarithromycin and either metronidazole or amoxycillin for treating Helicobacter pylori infection. Aliment Pharmacol Ther. 2000;14:603-9.

[42] Scott DR, Sachs G, Marcus EA. The role of acid inhibition in Helicobacter pylori eradication. F1000Res. 2016;5.

[43] Kim SY, Choi DJ, Chung JW. Antibiotic treatment for Helicobacter pylori: Is the end coming? World J Gastrointest Pharmacol Ther. 2015;6:183-98.

[44] Josenhans C, Suerbaum S. The role of motility as a virulence factor in bacteria. Int J Med Microbiol. 2002;291:605-14.

[45] Huang JY, Sweeney EG, Sigal M, Zhang HC, Remington SJ, Cantrell MA, et al. Chemodetection and Destruction of Host Urea Allows Helicobacter pylori to Locate the Epithelium. Cell Host Microbe. 2015;18:147-56.

[46] Chang WL, Yeh YC, Sheu BS. The impacts of H. pylori virulence factors on the development of gastroduodenal diseases. J Biomed Sci. 2018;25:68.

[47] Styer CM, Hansen LM, Cooke CL, Gundersen AM, Choi SS, Berg DE, et al. Expression of the BabA adhesin during experimental infection with Helicobacter pylori. Infect Immun. 2010;78:1593-600.

[48] Frick-Cheng AE, Pyburn TM, Voss BJ, McDonald WH, Ohi MD, Cover TL. Molecular and Structural Analysis of the Helicobacter pylori cag Type IV Secretion System Core Complex. MBio. 2016;7:e02001-15.

[49] Hatakeyama M. Helicobacter pylori CagA and gastric cancer: a paradigm for hit-and-run carcinogenesis. Cell Host Microbe. 2014;15:306-16.

[50] Chung JM, Sheedlo MJ, Campbell AM, Sawhney N, Frick-Cheng AE, Lacy DB, et al. Structure of the Helicobacter pylori Cag type IV secretion system. Elife. 2019;8.

[51] Sheedlo MJ, Chung JM, Sawhney N, Durie CL, Cover TL, Ohi MD, et al. Cryo-EM reveals species-specific components within the Helicobacter pylori Cag type IV secretion system core complex. Elife. 2020;9.

[52] Foegeding NJ, Caston RR, McClain MS, Ohi MD, Cover TL. An Overview of Helicobacter pylori VacA Toxin Biology. Toxins (Basel). 2016;8.

[53] Cover TL, Blaser MJ. Purification and characterization of the vacuolating toxin from Helicobacter pylori. J Biol Chem. 1992;267:10570-5.

[54] Telford JL, Ghiara P, Dell'Orco M, Comanducci M, Burroni D, Bugnoli M, et al. Gene structure of the Helicobacter pylori cytotoxin and evidence of its key role in gastric disease. J Exp Med. 1994;179:1653-58.

[55] Bumann D, Aksu S, Wendland M, Janek K, Zimny-Arndt U, Sabarth N, et al. Proteome analysis of secreted proteins of the gastric pathogen Helicobacter pylori. Infect Immun. 2002;70:3396-403.

[56] Schmitt W, Haas R. Genetic analysis of the Helicobacter pylori vacuolating cytotoxin: structural similarities with the IgA protease type of exported protein. Mol Microbiol. 1994;12:307-19.

[57] Nguyen VQ, Caprioli RM, Cover TL. Carboxy-terminal proteolytic processing of Helicobacter pylori vacuolating toxin. Infect Immun. 2001;69:543-6.

[58] Gonzalez-Rivera C, Campbell AM, Rutherford SA, Pyburn TM, Foegeding NJ, Barke TL, et al. A Nonoligomerizing Mutant Form of Helicobacter pylori VacA Allows Structural Analysis of the p33 Domain. Infect Immun. 2016;84:2662-70.

[59] Gangwer KA, Mushrush DJ, Stauff DL, Spiller B, McClain MS, Cover TL, et al. Crystal structure of the Helicobacter pylori vacuolating toxin p55 domain. Proc Natl Acad Sci U S A. 2007;104:16293-8.

[60] Burroni D, Lupetti P, Pagliaccia C, Reyrat JM, Dallai R, Rappuoli R, et al. Deletion of the major proteolytic site of the Helicobacter pylori cytotoxin does not influence toxin activity but favors assembly of the toxin into hexameric structures. Infect Immun. 1998;66:5547-50.

[61] Atherton JC, Cao P, Peek RM, Jr., Tummuru MK, Blaser MJ, Cover TL. Mosaicism in vacuolating cytotoxin alleles of Helicobacter pylori. Association of specific vacA types with cytotoxin production and peptic ulceration. J Biol Chem. 1995;270:17771-7.

[62] Pagliaccia C, de Bernard M, Lupetti P, Ji X, Burroni D, Cover TL, et al. The m2 form of the Helicobacter pylori cytotoxin has cell type-specific vacuolating activity. Proc Natl Acad Sci U S A. 1998;95:10212-7.

[63] Rhead JL, Letley DP, Mohammadi M, Hussein N, Mohagheghi MA, Eshagh Hosseini M, et al. A new Helicobacter pylori vacuolating cytotoxin determinant, the intermediate region, is associated with gastric cancer. Gastroenterology. 2007;133:926-36.

[64] Gangwer KA, Shaffer CL, Suerbaum S, Lacy DB, Cover TL, Bordenstein SR. Molecular evolution of the Helicobacter pylori vacuolating toxin gene vacA. J Bacteriol. 2010;192:6126-35.

[65] Letley DP, Atherton JC. Natural diversity in the N terminus of the mature vacuolating cytotoxin of Helicobacter pylori determines cytotoxin activity. J Bacteriol. 2000;182:3278-80.

[66] McClain MS, Cao P, Iwamoto H, Vinion-Dubiel AD, Szabo G, Shao Z, et al. A 12-amino-acid segment, present in type s2 but not type s1 Helicobacter pylori VacA proteins, abolishes cytotoxin activity and alters membrane channel formation. J Bacteriol. 2001;183:6499-508.

[67] Wang WC, Wang HJ, Kuo CH. Two distinctive cell binding patterns by vacuolating toxin fused with glutathione S-transferase: one high-affinity m1-specific binding and the other lower-affinity binding for variant m forms. Biochemistry. 2001;40:11887-96.

[68] Gonzalez-Rivera C, Algood HM, Radin JN, McClain MS, Cover TL. The intermediate region of Helicobacter pylori VacA is a determinant of toxin potency in a Jurkat T cell assay. Infect Immun. 2012;80:2578-88.

[69] Cover TL, Tummuru MK, Cao P, Thompson SA, Blaser MJ. Divergence of genetic sequences for the vacuolating cytotoxin among Helicobacter pylori strains. J Biol Chem. 1994;269:10566-73.

[70] Figueiredo C, Machado JC, Pharoah P, Seruca R, Sousa S, Carvalho R, et al. Helicobacter pylori and interleukin 1 genotyping: an opportunity to identify high-risk individuals for gastric carcinoma. J Natl Cancer Inst. 2002;94:1680-7.

[71] Memon AA, Hussein NR, Miendje Deyi VY, Burette A, Atherton JC. Vacuolating cytotoxin genotypes are strong markers of gastric cancer and duodenal ulcer-associated Helicobacter pylori strains: a matched case-control study. J Clin Microbiol. 2014;52:2984-9.

[72] McClain MS, Iwamoto H, Cao P, Vinion-Dubiel AD, Li Y, Szabo G, et al. Essential role of a GXXXG motif for membrane channel formation by Helicobacter pylori vacuolating toxin. J Biol Chem. 2003;278:12101-8.

[73] Kim S, Chamberlain AK, Bowie JU. Membrane channel structure of Helicobacter pylori vacuolating toxin: role of multiple GXXXG motifs in cylindrical channels. Proc Natl Acad Sci U S A. 2004;101:5988-91.

[74] Leunk RD, Johnson PT, David BC, Kraft WG, Morgan DR. Cytotoxic activity in broth-culture filtrates of Campylobacter pylori. J Med Microbiol. 1988;26:93-9.

[75] Genisset C, Galeotti CL, Lupetti P, Mercati D, Skibinski DA, Barone S, et al. A Helicobacter pylori vacuolating toxin mutant that fails to oligomerize has a dominant negative phenotype. Infect Immun. 2006;74:1786-94.

[76] Pyburn TM, Foegeding NJ, Gonzalez-Rivera C, McDonald NA, Gould KL, Cover TL, et al. Structural organization of membrane-inserted hexamers formed by Helicobacter pylori VacA toxin. Mol Microbiol. 2016;102:22-36.

[77] Ye D, Willhite DC, Blanke SR. Identification of the minimal intracellular vacuolating domain of the Helicobacter pylori vacuolating toxin. J Biol Chem. 1999;274:9277-82.

[78] Ivie SE, McClain MS, Torres VJ, Algood HM, Lacy DB, Yang R, et al. Helicobacter pylori VacA subdomain required for intracellular toxin activity and assembly of functional oligomeric complexes. Infect Immun. 2008;76:2843-51.

[79] Vinion-Dubiel AD, McClain MS, Czajkowsky DM, Iwamoto H, Ye D, Cao P, et al. A dominant negative mutant of Helicobacter pylori vacuolating toxin (VacA) inhibits VacA-induced cell vacuolation. J Biol Chem. 1999;274:37736-42.

Chapter II

Cryo-EM Analysis Reveals Structural Basis of *Helicobacter Pylori* VacA Toxin

Oligomerization

This chapter is a modified version of a published article with authors listed as follows:

M.S., A.L.E., T.M.P collected the EM data and determined the structures. A. M. C. purified VacA. J.H., and A.L.E. carried out biochemical work. A.L.E. built the model and L.E.S., D.L.A., and D.L.B. provided feedback on the model. T.L.C. and M.D.O. conceived the project and provided feedback on all experiments and data processing. All authors interpreted the data and wrote the manuscript.

Abstract

Helicobacter pylori colonizes the human stomach and contributes to the development of gastric cancer and peptic ulcer disease. *H. pylori* secretes a pore-forming toxin called vacuolating cytotoxin A (VacA), which contains two domains (p33 and p55) and assembles into oligomeric structures. Using single particle cryo-electron microscopy, we have determined low-resolution structures of a VacA dodecamer and heptamer, as well as a 3.8 Å structure of the VacA hexamer. These analyses show that VacA p88 consists predominantly of a right-handed beta-helix that extends from the p55 domain into the p33 domain. We map the regions of p33 and p55 involved in hexamer oligomerization, model how interactions between protomers support heptamer formation, and identify surfaces of VacA that likely contact membrane. This work provides structural insights into the process of VacA oligomerization and identifies regions of VacA protomers that will contact the host cell surface during channel formation.

Introduction

Helicobacter pylori is a Gram-negative bacterium that persistently colonizes the stomachs of >50% of the human population, resulting in a gastric mucosal inflammatory response. While most *H. pylori*-positive individuals remain asymptomatic, *H. pylori* is the strongest known risk factor for development of gastric adenocarcinoma and contributes to the pathogenesis of peptic ulcer disease and gastric lymphoma [1, 2]. Gastric cancer is the third leading cause of cancer-related death worldwide and the World Health Organization has classified *H. pylori* as a type 1 carcinogen [3-5]

There is a high level of genetic diversity among *H. pylori* strains from unrelated individuals, and strain-specific variation in the production of secreted proteins is an important factor that determines whether or not symptomatic disease develops [6-10]. One secreted *H. pylori* protein is VacA (vacuolating cytotoxin A), named for its ability to induce vacuolation in cultured eukaryotic cells [11]. VacA can cause a wide range of cellular effects in addition to cellular vacuolation, including cell death, depolarization of membrane potential, mitochondrial dysfunction, autophagy, activation of mitogen-activated protein kinases, inhibition of T cell activities, and other immunomodulatory effects (reviewed in, [12]).

VacA is secreted from *H. pylori* as an 88 kDa monomer (p88), which shares very little sequence similarity to any characterized proteins from other bacterial species. The toxin can undergo limited proteolysis to yield a 33 kDa N-terminal fragment (p33) and a 55 kDa C-terminal fragment (p55) [13]. Regions within both domains contribute to VacA binding to cells [10, 14-18]. Experiments with intracellularly expressed toxin show that the minimum portion of VacA required for cell-vacuolating activity contains p33 and 110 amino acids within p55 (residues 1-422) [19]. The sequence of VacA is variable in three polymorphic regions, an N-terminal signal region (s1 or s2) [6], an intermediate region (i1 or i2) located near the C-terminus of p33 [10], and a p55 mid-region (m1 or m2) [6, 14]. Sequence variation in all three regions influences the capacity of VacA to cause cellular alterations [14-16, 20, 21]. Humans infected with *H. pylori* strains that secrete s1/i1/m1 forms of VacA have an increased risk of developing peptic ulcer disease or gastric cancer compared to those infected with *H. pylori* strains secreting s2/i2/m2 forms of VacA [6-10].

Planar lipid bilayer experiments and patch clamping experiments have shown that VacA can form anion-selective membrane channels [22-26]. Mutant forms of VacA defective in channel formation are defective in cell-vacuolating activity [27, 28]. Most cellular alterations caused by VacA are attributed to membrane channel formation, either in the plasma membrane or in the membranes of endosomes or lysosomes [22-24, 26, 29, 30].

The 88 kDa VacA monomers secreted by VacA can assemble into an assortment of water-soluble oligomeric structures, including hexamers, heptamers, dodecamers, and tetradecamers [31-35]. Mutant forms of VacA that are defective in oligomerization are also defective in cell-vacuolating activity, and dominant negative mutant forms of VacA can inhibit the activity of wild-type VacA through a process that involves formation of mixed oligomers [27, 29, 30]. These data suggest that oligomerization is required for VacA toxin activity. Water-soluble oligomeric forms of VacA are presumed to be structurally related to the membrane channels formed by VacA [22, 26, 33, 36]. Water-soluble VacA predominantly organizes into double-layered oligomeric structures, whereas membrane-bound VacA predominantly organizes into hexameric oligomers, with some heptamers also present [37]. Comparison of the two-dimensional (2D) averages of membrane-bound and soluble VacA hexamers generated using negative stain single particle electron microscopy reveals a structural difference in the central region of the oligomers (corresponding to the p33 region), suggesting that membrane association triggers a structural change in the N-terminus of VacA [37].

A 2.4 Å crystal structure of the majority of p55 (residues 355-821) showed that this region is composed predominantly of a right-handed beta-helix, a structural

arrangement adopted by passenger domains of many proteins secreted by type V secretion systems in Gram-negative bacteria [38]. A lower resolution (4.2 Å) crystal structure of a non-oligomerizing VacA mutant protein (VacA Δ346-47) showed that this structural fold extends at least partway into p33 [39]. Low-resolution structures of different oligomer types using negative stain EM have confirmed that p55 is in the peripheral "arm" region of the oligomers and p33 is located toward the center of the oligomeric structure [34, 35]. An N-terminal hydrophobic region containing multiple GXXXG motifs, required for toxin activity and membrane channel formation, localizes within a central region of the oligomers and likely corresponds to the channel of VacA membrane pores [27, 28]. While it is known that deletion of p33 residues 49-57 and p55 residues 346-347 ablate VacA oligomerization [29, 30], there is currently no structural model for how VacA oligomerizes and how different oligomeric states are accommodated. In addition, although there are negative stain 2D averages of VacA bound to membranes [37], there is no 3D model for how VacA interacts with membrane. In the current study, we use single particle cryo-electron microscopy (cryo-EM) to determine a 3.8 Å structure of a VacA s1/i1/m1 hexamer and lower resolution structures of a VacA heptamer (8.5 Å) and dodecamer (12 Å). These structures provide molecular insights into how VacA oligomerizes and interacts with membranes, two requirements for its cellular activity.

Results

Cryo-electron microscopy reconstruction of VacA oligomers.

To investigate the structural basis of VacA oligomerization, we purified wild-type

s1/i1/m1 VacA secreted by *H. pylori* and analyzed the protein by single particle cryo-EM

(Fig. 2.1A). Similar to what was observed in negative stain using random conical tilt and

cryo-negative stain methods [34, 35], 2D averages of VacA in vitrified ice showed a

mixture of oligomers including hexamers, heptamers, dodecamers, and tetradecamers

(Fig. 2.2A). In vitrified ice, hexamers and dodecamers are the most prevalent oligomer

in the sample (~77%), with heptamers and tetradecamers making up the rest of the

particles (23%) (Fig. 2.3). Just as was seen in negative stain and cryo-negative stain

[35], the dodecamers clearly adopt multiple conformations (Fig. 2.2A). While we

determined a 3D structure of a VacA dodecamer (Fig. 2.1B), the structural

heterogeneity of the dodecamers limited the resolution to ~12 Å (Fig. 2.1C).

In the presence of membrane, VacA organizes into single-layers (hexamers and

heptamers) [37]. Therefore, we were interested in determining structures of soluble

single-layered VacA oligomers. While there were *en face* views of hexamers and

heptamers in the dataset (Fig. 2.2A), the side views were not easily visible. Additionally,

although there are clearly 2D averages showing side views of dodecamers and

tetradecamers (Fig. 2.2A, classes labeled with "*"), it is not possible by eye alone to

differentiate dodecamers and tetradecamers. Previous studies either proposed [33] or

showed [34] that double-layer VacA oligomers (dodecamers or tetradecamers) were

formed by the apposition of two face-to-face hexamers or heptamers with a free plane

of rotation. For this reason, to determine the structure of a hexamer and heptamer we

focused on the 2D averages corresponding to side views of the dodecamers and/or

tetradecamers and used particle signal subtraction to remove the signal from one half of

the double-layer of each particle. The signal subtracted images were then combined

with the *en face* views of hexamers and heptamers and then classified and aligned in

three dimensions (3D). This led to low resolution structures of a hexamer and heptamer

that were then further refined using RELION and cisTEM [40, 41] (Fig. 2.3). While the

arms of the oligomers appear symmetrical even with no symmetry imposed, the

densities in the central regions are less-defined and do not appear to be well-structured

(Fig. 2.4A). With no symmetry enforced, the average resolution of the VacA hexamer

was 8.3 Å (Fig. 2.4B). After imposing six-fold symmetry during the final 3D refinement

steps, the average resolution of the 3D reconstruction was 3.8 Å (Fig. 2.2B, 2.5A). We

also determined a 3D structure of a heptamer using a similar approach (Fig. 2.3, 2.5A).

The 8.5 Å resolution density of the heptamer is shown in Figure 2.2C. The orientations

of the particles found in each structure and the local resolution map of the hexamer with

applied 6-fold symmetry are shown in Fig. 2.5B-D. We have colored one p88 in the

VacA hexamer to highlight how individual protomers are organized in the oligomer (Fig

2.2B). Interestingly, a density extends from p55 of each protomer to contact the

adjacent protomer (Fig. 2.2B).

Structure of VacA p88

The p88 protomers in the 3.8 Å cryo-EM density map of the VacA hexamer are

composed of a straight stretch of rolling β-strands extending from the middle of the

oligomer and ending in a "hook-shaped" tip (Fig. 2.2B). From the 2.4 Å crystal structure

of p55 (amino acid (a. a.) 355-811, PDB 2QV3) and negative stain 3D maps of a VacA

Δ6-27 mutant protein [34, 35, 38], we know that the VacA C-terminus is located in the

"hook-shaped" periphery of the hexamer arms, while the N-terminus is located in the

center of the oligomer. The 3.8 Å map of the VacA hexamer allows us to visualize

individual β-strands in each protomer (Fig. 2.2B); however, the resolution is not high

enough to confidently define the register of amino acids directly from the map. For this

reason, we were not able to build a near-atomic model of p88 *de novo*. There are two

available high resolution structures of portions of VacA: 1) a 2.4 Å crystal structure of a

majority of p55 that mapped residues 355-811 [38], and 2) a 4.2 Å crystal structure of

VacA Δ346-347, a non-oligomerizing mutant, where residues 348-811 were mapped

into the density and an additional 165 residues mainly from the p33 domain were built

into the electron density [39]. Combining the mapped residues in these structures

allowed for the generation of a VacA model containing residues 348-811 and an alanine

backbone for an additional 165 residues. We docked this model into a protomer of the

3.8 Å cryo-EM map of the hexamer using the program Phenix [42]. We were then able

to build an additional 65 alanine residues into the cryo-EM density map, including a

prominent alpha-helix that sits perpendicular to the β-strands located in the p33 region

and a string of residues that extends from p55 to contact an adjacent protomer (Fig.

2.6A,C-D, 2.9).

The model from our cryo-EM analysis shows that the structure of VacA is

composed predominantly of rolling β-strands connected by flexible loops (Fig. 2.6A,B).

Combining the crystal structures with the cryo-EM density map, the p55 domain can be

traced from Threonine 811 (T811) to Glutamine 340 (Q340) (Fig. 2.6A). We could not

determine the sequence of the three additional beta strands that make up an additional

turn in the beta-helix visible in the p33 domain, so these are modeled as an alanine

backbone. In total, ~115 amino acids are missing in the VacA model built from the 3.8 Å cryo-EM map of a VacA hexamer. We predict that these residues are located in flexible loop regions, as well as in the unstructured density seen in the center of the C1 3D reconstruction (Fig. 2.4A). We overlaid the density of one p88 promoter extracted from the hexamer density and the crystal structure of VacA Δ346-347 to compare these structures (Fig. 2.6C,D). The biggest differences between the VacA hexamer cryo-EM map and the crystal structure of VacA Δ346-347 are the presence of density extending from p55 (a.a. 340-348) that contacts the next protomer, better resolved β-strands extending into the p33 region, and a better defined alpha-helix positioned where the VacA protomers interact (Fig. 2.6A,D).

VacA oligomerization is mediated by p33-p55 interactions of neighboring protomers.

The 3.8 Å resolution cryo-EM structures of the soluble VacA hexamer allow us to structurally define the points of interaction between the "arms" of VacA protomers. In the VacA hexamer, p88-p88 interactions are mediated by residues in both the p33 and p55 regions (Fig. 2.7A). Each protomer-protomer interaction involves contacts between one surface in protomer 1 and two surfaces in protomer 2 (Fig. 2.7A,B). In protomer 1, the major region of interaction is in the part of the density map that did not have a clearly traceable secondary structure; therefore, we cannot predict which residues are involved other than that these residues are within the p33 domain (Fig. 2.7A,B). In protomer 2, there are two major regions involved in protomer-protomer interactions that include: 1) the short α-helix that sits perpendicular to p33 beta-strands, and 2) residues 340-348, which extend from a beta-strand in the p55 density (Fig. 2.7A,B).

26

In order to understand the difference between p88 protomer interactions in hexamers versus heptamers, we docked the VacA p88 map built using the 3.8 Å VacA hexamer density into the protomers of the 8.5 Å VacA heptamer (Fig. 2.7C). The p88 map built from the hexamer was able to fit directly into the "arms" of the heptamer, indicating that p88 does not undergo any major conformational changes that depend on the type of oligomer. Similar to what was observed in the hexamer, heptamer protomer-protomer interactions are mediated by one major region in p33 from protomer 1; however, in protomer 2 the angle of the two major contact regions has slightly changed (Fig. 2.7C,D). This small change in angle at the binding surfaces leads to a shift in protomer angle that propagates into an ~18 Å displacement of each protomer at the p88 C-terminal tip of the heptamer when compared to the hexamer (Fig. 2.7D).

Model of VacA dodecamer predicts p88 regions that bind membrane

It has been proposed that soluble VacA dodecamers and tetradecamers form as a result of interactions between surfaces of hexamers or heptamers that would normally interact with the lipid membrane and/or cellular receptors if they were present [33, 37]. This model is supported by negative stain EM analysis that shows VacA organizes on membranes as mostly hexamers with some heptamers [37], as well as single channel and computational modeling studies that suggest chloride channels are formed by hexameric VacA [24, 43]. To more clearly define the VacA regions involved in double-layer oligomer formation, and thus also the regions important for interacting with the lipid bilayer and/or cellular receptors, we fit two identical models of the VacA hexamer into the low resolution VacA dodecamer map (Fig. 2.8A). The VacA hexamers fit directly

into the VacA dodecamer, without any detectable structural differences at the 3.8 Å

resolution of our map. However, it is possible that the middle region of the hexamer as

compared to the middle region of the dodecamers (and similarly, the middle region of

the heptamer compared to the middle region of a tetradecamer) are not similar at this

resolution. One possible explanation is that in solution the central region is not well

structured and thus are not the same when comparing single layer and double layer

hexamer-dodecamers and heptamers-tetradecamers respectively. This may be one of

the reasons we do not see structured density in the middle of the oligomers. The model

shows that dodecamers form through interactions between two hexamers and are

mediated by 18 loop regions extending from the β-helix strands in both the p55 and p33

regions (Fig. 2.8B,C).

VacA contains three polymorphic regions (known as s, i and m regions). Amino

acids within the intermediate region (i1 or i2) located near the C-terminus of p33 and

within a middle region (m1 or m2) in p55 are predicted to be involved in receptor binding

and/or cellular tropism [6, 10, 14]. The amino acids that encompass the m-region span

have not been completely defined, but span up to several hundred amino acids from

approximately amino acids 460-736 [6, 15]. The position of the m region is shown in

Figure 4b with brackets and an arrow marks the position where a cluster of about 23

amino acids would be inserted in some m2 forms of VacA [15, 44]. We predict that the

loops and strands involved in interactions between the two layers of the dodecamer are

involved in interactions with cellular membranes, and that the corresponding loops and

strain variations found in the i- and m- regions are particularly important in making

contacts with lipids and/or receptors, thereby influencing toxin activity and cellular tropism (Fig. 2.8D).

Discussion

We have presented a 3.8 Å structure of a VacA hexamer that provides molecular insight into how soluble VacA oligomerizes and predicts regions of VacA that interact with cellular membranes. The structure shows that VacA is composed of mostly of rolling beta strands connected by flexible loops. The protomers interact through residues in both the p33 and p55 regions. These include an alpha-helix sandwiched between protomers and an extended p55 loop from one protomer that contacts the p33 region from an adjacent protomer. Our structural analysis also shows that a major difference between hexamers and heptamers is the angle of protomer-protomer interactions, with the binding interface between protomers in hexamers being larger than the binding interface supporting heptamers. This difference in binding areas helps explain previous observations that dodecamers and hexamers are more abundant in solution than heptamers and tetradecamers [35].

The model built from our cryo-EM density map shows a number of structural differences compared to the crystal structure of the non-oligomerizing mutant [39]. These include the presence of density extending from p55 (a.a. 340-348) that contacts the adjacent protomer, a better defined alpha helix that sits between protomers, and better definition of three additional beta-strands in p33. These differences may simply reflect the use of two disparate methodologies for structural analysis, or alternatively, some of the differences reflect structural changes associated with VacA oligomerization,

especially the difference involving a strand of residues that extends from p55 and contacts the neighboring protomer.

VacA must oligomerize and insert into the lipid-bilayer to form ion channels [22, 26, 33, 36]. It has been proposed that double layer oligomers found in solution result from interactions between the surfaces of p88 molecules that would contact host cell membranes [33].Therefore, we predict that the loop and beta strand regions that make contact in the dodecamer represent the same regions of VacA that interact with cellular membranes (Fig. 2.8D). Sequence variations among VacA proteins in both the p33 (i1 and i2) and p55 (m1 and m2) domains influence the cellular tropism and toxicity of VacA [6, 10, 14], likely due to altering the ability of the toxin to bind cellular receptors and/or lipids. We localized the m-region in the cryo-EM map and found that the sequence variations are not confined to individual loops or beta strands (Fig. 2.8D); however, it is likely that sequence variations, especially deletions and insertions, alter the loops that interact with the surface of the cell. The repeating beta-strand-loop-beta-strand pattern of most of VacA also helps explain why it has been difficult to inhibit VacA binding to lipid bilayers [37]. Our model predicts that approximately 18 loops contact the cell membrane, and thus the combined contribution of these protein-lipid interactions along the length of p88 increases the avidity of p88 binding to membrane. In summary, these structures provide new insights into the basis of VacA oligomerization and define regions that are likely important for VacA interactions with cell membranes, two functions that are each required for VacA activity.
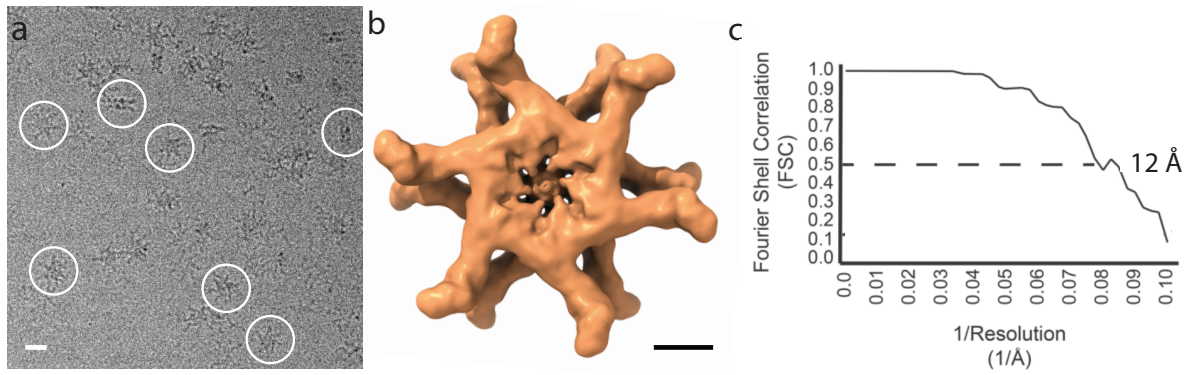
Figure 2.1. Single particle cryo-EM analysis of *H. pylori* VacA. (a) Representative area of a motion corrected image of vitrified *H. pylori* VacA. Examples of particles are circled in white. Scale bar, 20 nm (b) Cryo-EM density map of a VacA dodecamer at 12 Å resolution. Scale bar, 50 Å. (c) Fourier shell correlation (FSC) curve of a VacA dodecamer with C6 applied symmetry.
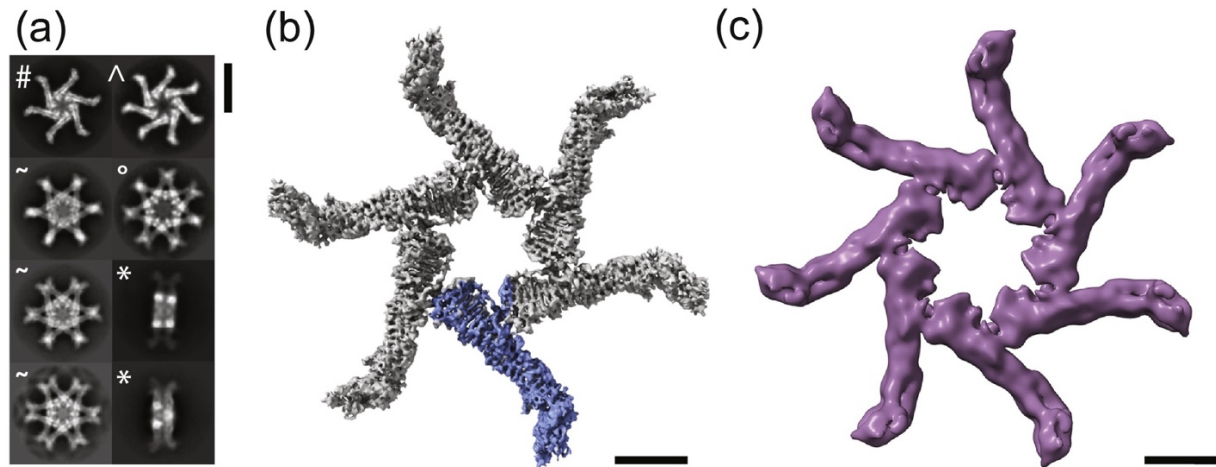
Figure 2.2. Cryo-EM analysis of VacA s1m1 oligomers. (a) Representative 2D class averages of VacA oligomer particles. Scale bar, 20 nm. #, *en face* hexamer class; ^, *en face* heptamer 2D class; ~, *en face* dodecamer 2D class; °, *en face* tetradecamer 2D class; *, side view of double layer oligomer 2D class. (b) 3.8 Å cryo-EM density map of hexameric VacA. One p88 protomer is colored blue. Scale bar, 50 Å (c) 8.5 Å resolution cryo-EM density map of heptameric VacA. Scale bar, 50 Å.
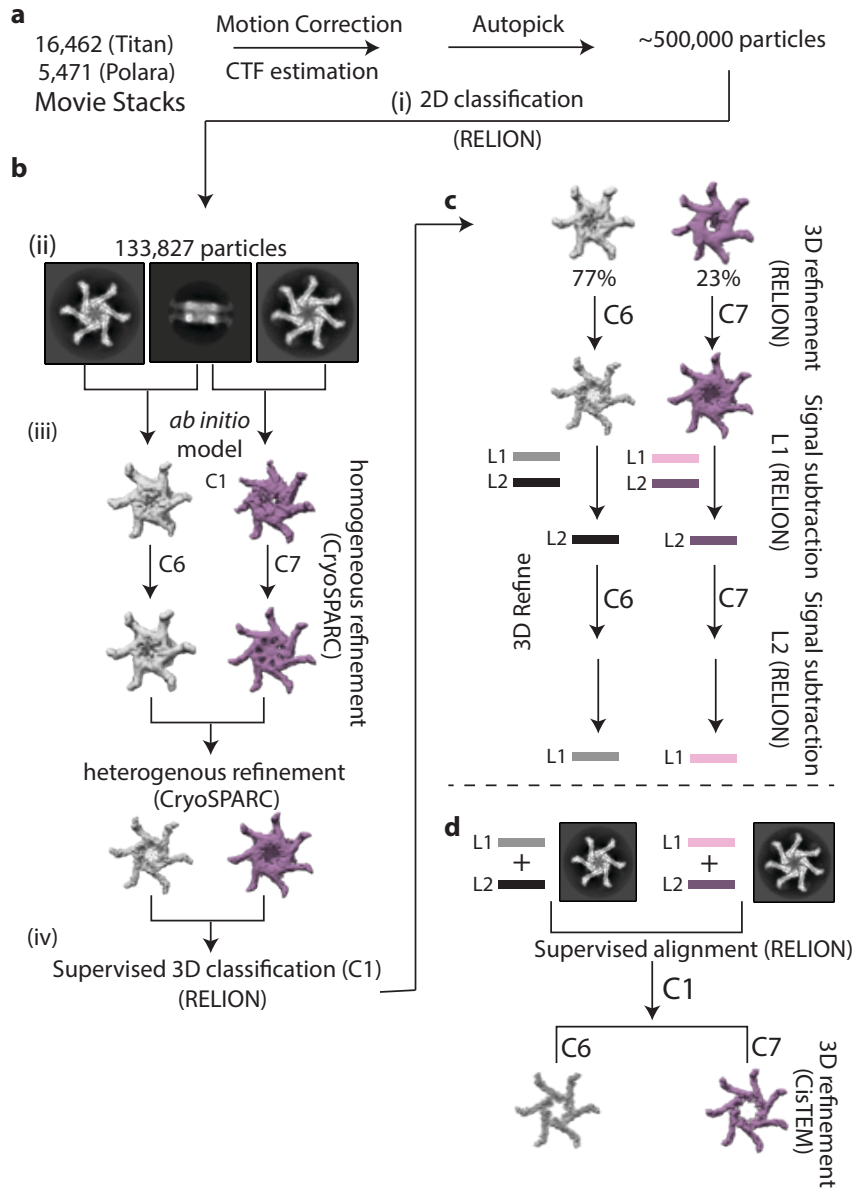
Figure 2.3. Flow chart of cryo-EM processing steps for the hexamer and heptamer. (a) Reconstruction of initial models for hexamer and heptamer models using the following steps: i) 2D classification in RELION. (b) After 2D averaging using RELION ii) All en face views of hexamers and heptamers, and side-views of double layer oligomers were selected and combined. Figure shows representative examples of selected class averages; iii) *Ab initio* model calculation for heptamer and hexamer in CryoSPARC (C1) followed by C6 or C7 homogenous refinement; iv) Particle stack and models moved to RELION for C1 supervised 3D classification (2 designated classes) in RELION. (c) Signal subtraction of each side of the side views of the double layer to expand particle stack size. L1, layer 1 in double layer oligomer; L2, layer 2 in double layer oligomer. (d) 3D refinement (C1) using the *en face* views of the hexamer and heptamer combined with the signal subtracted side views of the VacA double layers

using RELION. Final 3D refinements were done using cisTEM applying either C6 or C7 symmetry.
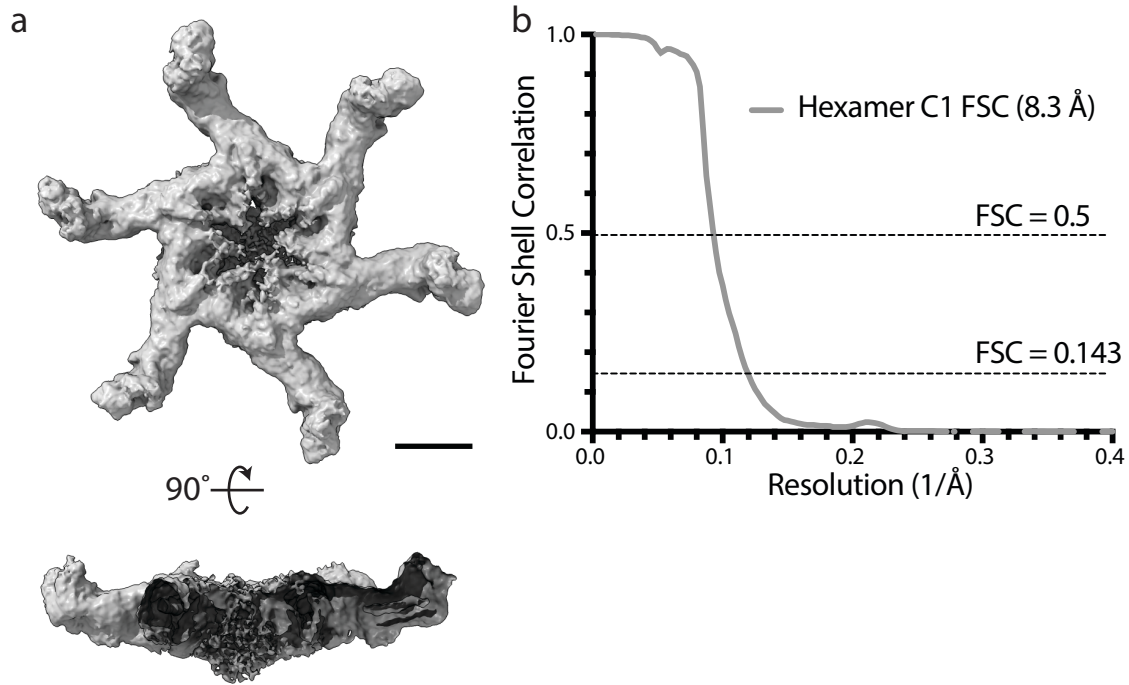
Figure 2.4. 3D reconstruction of VacA hexamer with no applied symmetry. (a) Cryo-EM density map of the VacA hexamer with no applied symmetry (C1) at 8.3 Å resolution. *En face* and central slice view after rotating 90° on the x-axis. Scale bar, 50 Å. (b) FSC curve of the VacA hexamer with no applied symmetry (C1).
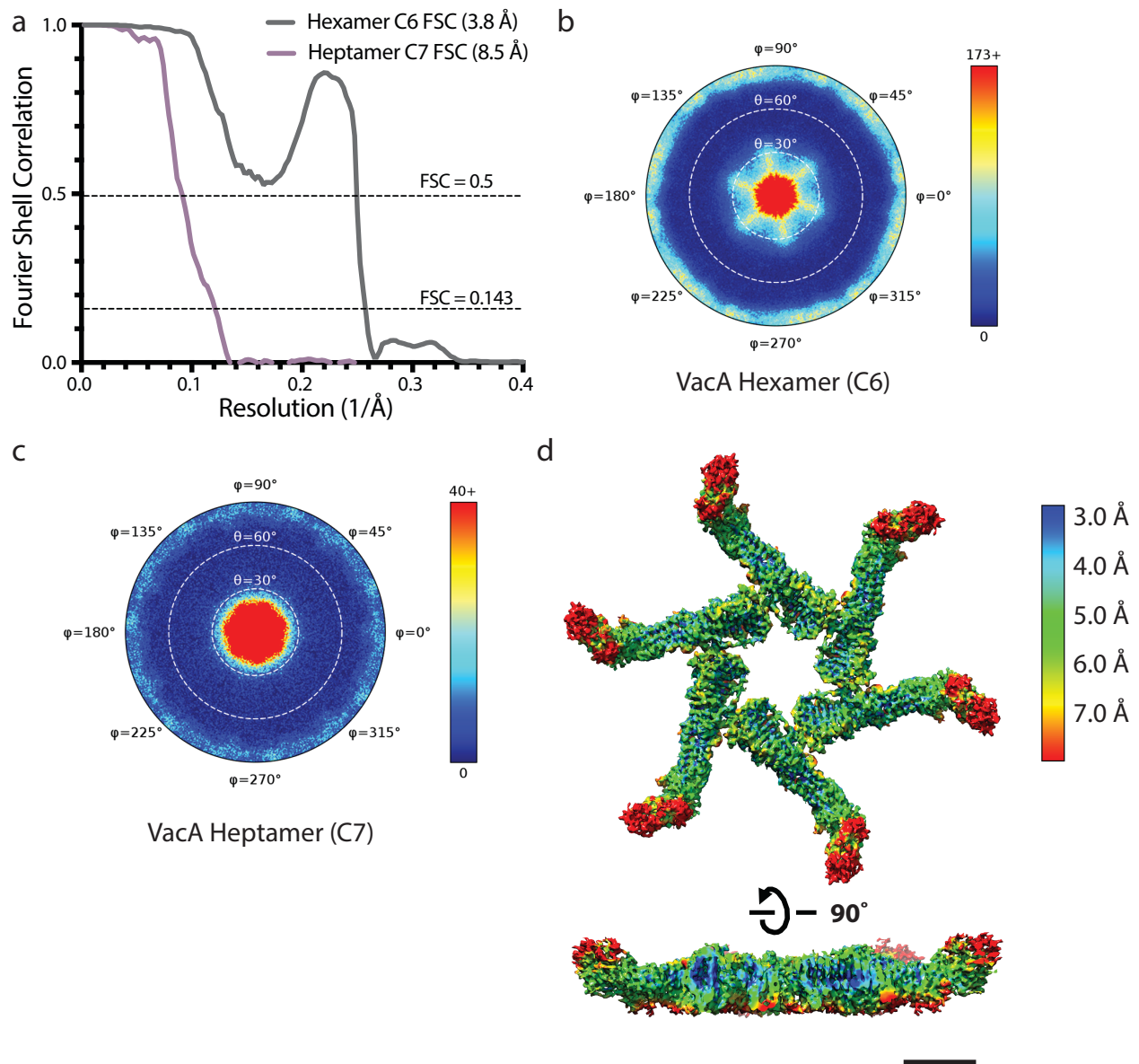
Figure 2.5. 3D reconstruction of VacA hexamer and heptamer with symmetry. (a) FSC curve between half-maps of VacA hexamer with C6 symmetry (grey curve) and heptamer with C7 symmetry (purple curve). (b) Angular distribution plot for the 3.8 Å VacA hexamer. (c) Angular distribution plot for 8.5 Å VacA heptamer. (d) Local resolution map of the 3.8 Å VacA hexamer cryo-EM map.
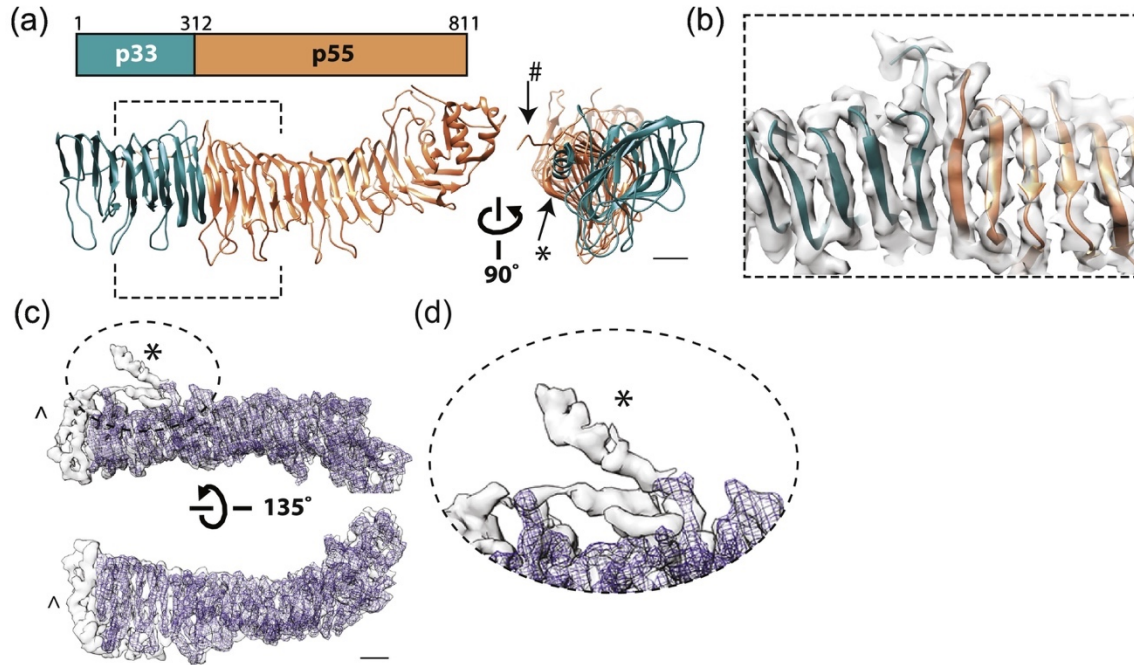
Fig. 2.6. Structure and interactions of VacA p88. (a) Schematic diagram of VacA p88 monomer subdivided into the p33 (dark cyan) and p55 (coral) domains. Ribbon representation of VacA p88 model. Dashed brackets show the region of the map shown in panel b. Arrow shows direction and degrees of rotation. Scale bar, 10 Å. (b) Closer view of a portion of the p88 model and the density map shows the clear beta-strand separation. (c) Overlay of the density of one p88 protomer extracted from the hexamer density (grey) and the crystal structure of VacA Δ346-347 (blue mesh) [39]. Arrow shows direction and degrees of rotation. Dashed oval shows the regions of the maps shown in panel d. Scale bar, 10 Å. (d) Closer view of a portion of the p88 model (grey) and VacA Δ346-347 crystal structure (blue mesh). (c-d) ^, marks the p33 region with additional beta-strands, #, marks position of alpha-helix.*, marks the position of density that extends from the p55 domain and interacts with an adjacent protomer.
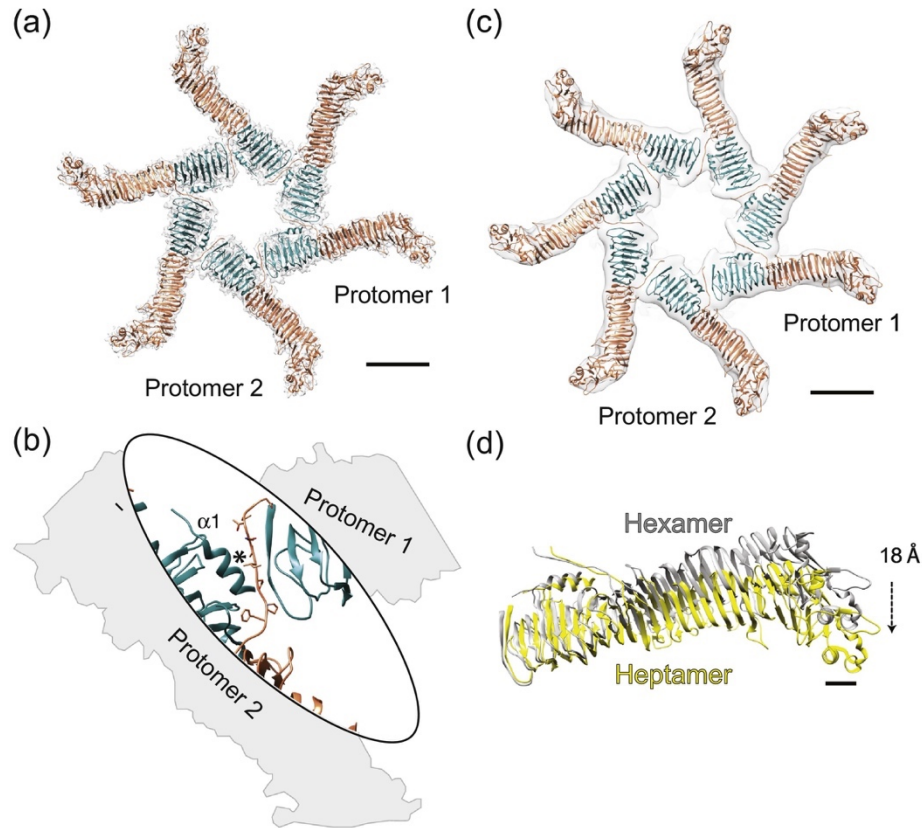
Fig. 2.7. Comparison of VacA hexamer and heptamer. (a) Envelope of the density of the 3.8 Å hexamer structure (grey) fitted with six p88 ribbon models. Dark cyan, p33 domain; Coral, p33 domain. Scale bar, 50 Å (b) Enlarged view of the interaction interface between protomer 1 and 2. The continuation of the cryo-EM density is traced in grey. The position of residues 346 and 347 in protomer 2 is marked with an *. These residues are part of the p55 extension that contacts the adjacent protomer. (c) Envelope of the density of the 8.5 Å structure (grey) fitted with seven p88 ribbon models. Dark cyan, p33 domain; Coral, p55 domain. Scale bar, 50 Å. (d) The superposition of the p88 map in protomer 1 of the hexamer (grey) with the p88 map in protomer 1 of the heptamer (yellow) shows the change of angle at the C-terminus of p88. The arrow indicates the direction of the 18 Å shift between protomers in the hexamer and heptamer. Scale bar, 10 Å.
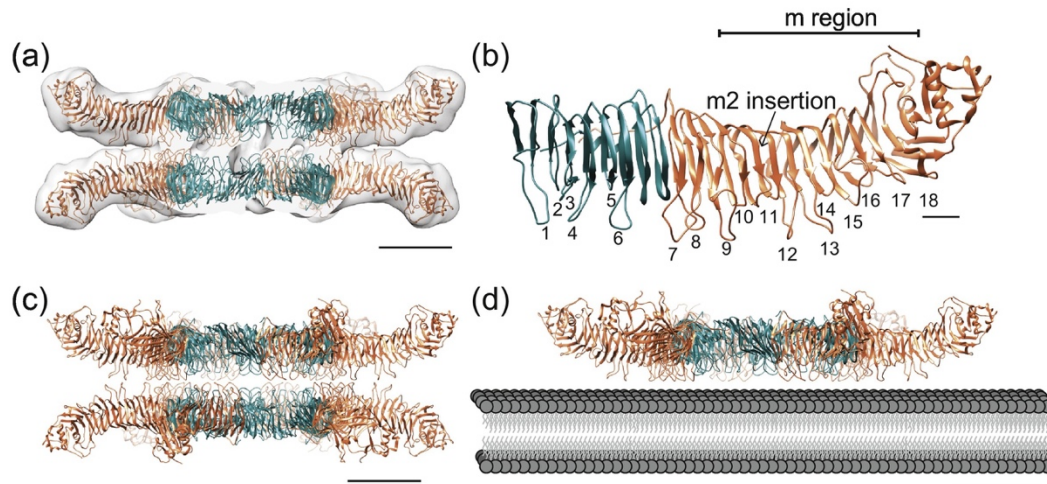
Figure 2.8. Model for VacA interactions in dodecamers and on membrane. (a) Central slice view of the envelope of the cryo-EM density map of a VacA dodecamer (grey) fitted with 2 models of VacA hexamers (total of 12 p88 models). Dark cyan, p33 domain; Coral, p55 domain. Scale bar, 50 Å. (b) Ribbon diagram of p88 model. Loops that interact between dodecamer layers are numbered. The general position of the m-region is marked by brackets. p33 domain in dark cyan, and p55 domain in coral. An arrow marks the position where ~20 m2-specific residues would be inserted [6, 15]. Scale bar, 10 Å. (c) Similar view as in panel a, but without the envelope of the cryo-EM density. (d) Model of how a VacA hexamer would sit on a host cell membrane. The same loop regions that interact in the double layer oligomer are predicted to interact with host cell membranes. Dark cyan, p33 domain; Coral, p55 domain. Scale bar, 50 Å.
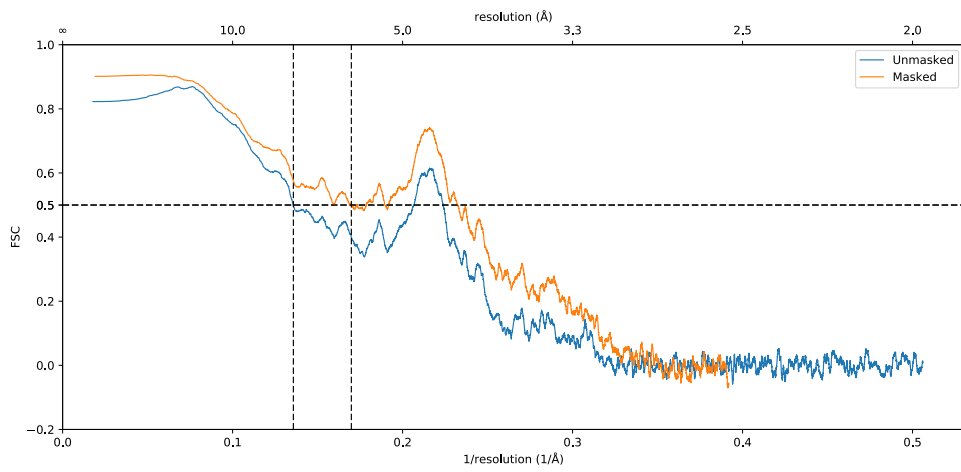
Figure 2.9. VacA hexamer map refinement and model validation. Model-map FSC curve for VacA hexamer.

References

[1] Marshall BJ, Goodwin CS, Warren JR, Murray R, Blincow ED, Blackbourn SJ, et al. Prospective double-blind trial of duodenal ulcer relapse after eradication of Campylobacter pylori. Lancet. 1988;2:1437-42.

[2] Suerbaum S, Michetti P. *Helicobacter pylori* infection. The New England journal of medicine. 2002;347:1175-86.

[3] Fuchs CS, Mayer RJ. Gastric carcinoma. The New England journal of medicine. 1995;333:32-41.

[4] de Martel C, Ferlay J, Franceschi S, Vignat J, Bray F, Forman D, et al. Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. The Lancet Oncology. 2012;13:607-15.

[5] Anonymous. Schistosomes, liver flukes and *Helicobacter pylori* IARC monographs on the evaluation of carcinogenic risks to humans. International Agency for Research on Cancer. Lyons, France1994.

[6] Atherton JC, Cao P, Peek RM, Jr., Tummuru MK, Blaser MJ, Cover TL. Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori*. Association of specific vacA types with cytotoxin production and peptic ulceration. The Journal of biological chemistry. 1995;270:17771-7.

[7] Cover TL. Helicobacter pylori Diversity and Gastric Cancer Risk. mBio. 2016;7.

[8] Figueiredo C, Machado JC, Pharoah P, Seruca R, Sousa S, Carvalho R, et al. *Helicobacter pylori* and interleukin 1 genotyping: an opportunity to identify high-risk individuals for gastric carcinoma. Journal of the National Cancer Institute. 2002;94:1680-7.

[9] Memon AA, Hussein NR, Miendje Deyi VY, Burette A, Atherton JC. Vacuolating cytotoxin genotypes are strong markers of gastric cancer and duodenal ulcer-associated Helicobacter pylori strains: a matched case-control study. Journal of clinical microbiology. 2014;52:2984-9.

[10] Rhead JL, Letley DP, Mohammadi M, Hussein N, Mohagheghi MA, Eshagh Hosseini M, et al. A new *Helicobacter pylori* vacuolating cytotoxin determinant, the intermediate region, is associated with gastric cancer. Gastroenterology. 2007;133:926-36.

[11] Cover TL, Blaser MJ. Purification and characterization of the vacuolating toxin from *Helicobacter pylori*. The Journal of biological chemistry. 1992;267:10570-5.

[12] Foegeding NJ, Caston RR, McClain MS, Ohi MD, Cover TL. An Overview of Helicobacter pylori VacA Toxin Biology. Toxins. 2016;8.

[13] Telford JL, Ghiara P, Dell'Orco M, Comanducci M, Burroni D, Bugnoli M, et al. Gene structure of the Helicobacter pylori cytotoxin and evidence of its key role in gastric disease. The Journal of experimental medicine. 1994;179:1653-58.

[14] Pagliaccia C, de Bernard M, Lupetti P, Ji X, Burroni D, Cover TL, et al. The m2 form of the *Helicobacter pylori* cytotoxin has cell type-specific vacuolating activity. Proceedings of the National Academy of Sciences of the United States of America. 1998;95:10212-7.

[15] Ji X, Fernandez T, Burroni D, Pagliaccia C, Atherton JC, Reyrat JM, et al. Cell specificity of *Helicobacter pylori* cytotoxin is determined by a short region in the polymorphic midregion. Infection and immunity. 2000;68:3754-7.

[16] Wang WC, Wang HJ, Kuo CH. Two distinctive cell binding patterns by vacuolating toxin fused with glutathione S-transferase: one high-affinity m1-specific binding and the other lower-affinity binding for variant m forms. Biochemistry. 2001;40:11887-96.

[17] Torres VJ, Ivie SE, McClain MS, Cover TL. Functional properties of the p33 and p55 domains of the *Helicobacter pylori* vacuolating cytotoxin. The Journal of biological chemistry. 2005;280:21107-14.

[18] Gonzalez-Rivera C, Algood HM, Radin JN, McClain MS, Cover TL. The Intermediate Region of *Helicobacter pylori* VacA Is a Determinant of Toxin Potency in a Jurkat T Cell Assay. Infection and immunity. 2012;80:2578-88.

[19] Ye D, Willhite DC, Blanke SR. Identification of the minimal intracellular vacuolating domain of the *Helicobacter pylori* vacuolating toxin. The Journal of biological chemistry. 1999;274:9277-82.

[20] Letley DP, Atherton JC. Natural diversity in the N terminus of the mature vacuolating cytotoxin of *Helicobacter pylori* determines cytotoxin activity. Journal of bacteriology. 2000;182:3278-80.

[21] McClain MS, Cao P, Cover TL. Amino-terminal hydrophobic region of Helicobacter pylori vacuolating cytotoxin (VacA) mediates transmembrane protein dimerization. Infection and immunity. 2001;69:1181-4.

[22] Czajkowsky DM, Iwamoto H, Cover TL, Shao Z. The vacuolating toxin from *Helicobacter pylori* forms hexameric pores in lipid bilayers at low pH. Proceedings of the National Academy of Sciences of the United States of America. 1999;96:2001-6.

[23] Szabo I, Brutsche S, Tombola F, Moschioni M, Satin B, Telford JL, et al. Formation of anion-selective channels in the cell plasma membrane by the toxin VacA of *Helicobacter pylori* is required for its biological activity. The EMBO journal. 1999;18:5517-27.

[24] Iwamoto H, Czajkowsky DM, Cover TL, Szabo G, Shao Z. VacA from *Helicobacter pylori*: a hexameric chloride channel. FEBS letters. 1999;450:101-4.

[25] Tombola F, Oregna F, Brutsche S, Szabo I, Del Giudice G, Rappuoli R, et al. Inhibition of the vacuolating and anion channel activities of the VacA toxin of *Helicobacter pylori*. FEBS letters. 1999;460:221-5.

[26] Tombola F, Carlesso C, Szabo I, de Bernard M, Reyrat JM, Telford JL, et al. *Helicobacter pylori* vacuolating toxin forms anion-selective channels in planar lipid bilayers: possible implications for the mechanism of cellular vacuolation. Biophysical journal. 1999;76:1401-9.

[27] Vinion-Dubiel AD, McClain MS, Czajkowsky DM, Iwamoto H, Ye D, Cao P, et al. A dominant negative mutant of *Helicobacter pylori* vacuolating toxin (VacA) inhibits VacA-induced cell vacuolation. The Journal of biological chemistry. 1999;274:37736-42.

[28] McClain MS, Iwamoto H, Cao P, Vinion-Dubiel AD, Li Y, Szabo G, et al. Essential role of a GXXXG motif for membrane channel formation by *Helicobacter pylori* vacuolating toxin. The Journal of biological chemistry. 2003;278:12101-8.

[29] Ivie SE, McClain MS, Torres VJ, Algood HM, Lacy DB, Yang R, et al. *Helicobacter pylori* VacA subdomain required for intracellular toxin activity and assembly of functional oligomeric complexes. Infection and immunity. 2008;76:2843-51.

[30] Genisset C, Galeotti CL, Lupetti P, Mercati D, Skibinski DA, Barone S, et al. A *Helicobacter pylori* vacuolating toxin mutant that fails to oligomerize has a dominant negative phenotype. Infection and immunity. 2006;74:1786-94.

[31] Lupetti P, Heuser JE, Manetti R, Massari P, Lanzavecchia S, Bellon PL, et al. Oligomeric and subunit structure of the *Helicobacter pylori* vacuolating cytotoxin. The Journal of cell biology. 1996;133:801-7.

[32] Lanzavecchia S, Bellon PL, Lupetti P, Dallai R, Rappuoli R, Telford JL. Three-dimensional reconstruction of metal replicas of the *Helicobacter pylori* vacuolating cytotoxin. Journal of structural biology. 1998;121:9-18.

[33] Adrian M, Cover TL, Dubochet J, Heuser JE. Multiple oligomeric states of the *Helicobacter pylori* vacuolating toxin demonstrated by cryo-electron microscopy. Journal of molecular biology. 2002;318:121-33.

[34] El-Bez C, Adrian M, Dubochet J, Cover TL. High resolution structural analysis of *Helicobacter pylori* VacA toxin oligomers by cryo-negative staining electron microscopy. Journal of structural biology. 2005;151:215-28.

[35] Chambers MG, Pyburn TM, Gonzalez-Rivera C, Collier SE, Eli I, Yip CK, et al. Structural analysis of the oligomeric states of *Helicobacter pylori* VacA toxin. Journal of molecular biology. 2013;425:524-35.

[36] Geisse NA, Cover TL, Henderson RM, Edwardson JM. Targeting of *Helicobacter pylori* vacuolating toxin to lipid raft membrane domains analysed by atomic force microscopy. The Biochemical journal. 2004;381:911-7.

[37] Pyburn TM, Foegeding NJ, Gonzalez-Rivera C, McDonald NA, Gould KL, Cover TL, et al. Structural organization of membrane-inserted hexamers formed by Helicobacter pylori VacA toxin. Molecular microbiology. 2016.

[38] Gangwer KA, Mushrush DJ, Stauff DL, Spiller B, McClain MS, Cover TL, et al. Crystal structure of the *Helicobacter pylori* vacuolating toxin p55 domain. Proceedings of the National Academy of Sciences of the United States of America. 2007;104:16293-8.

[39] Gonzalez-Rivera C, Campbell AM, Rutherford SA, Pyburn TM, Foegeding NJ, Barke TL, et al. A Nonoligomerizing Mutant Form of Helicobacter pylori VacA Allows Structural Analysis of the p33 Domain. Infection and immunity. 2016;84:2662-70.

[40] Scheres SH. RELION: implementation of a Bayesian approach to cryo-EM structure determination. Journal of structural biology. 2012;180:519-30.

[41] Grant T, Rohou A, Grigorieff N. cisTEM, user-friendly software for single-particle image processing. eLife. 2018;7.

[42] Afonine PV, Poon BK, Read RJ, Sobolev OV, Terwilliger TC, Urzhumtsev A, et al. Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr D Struct Biol. 2018;74:531-44.

[43] Kim S, Chamberlain AK, Bowie JU. Membrane channel structure of *Helicobacter pylori* vacuolating toxin: role of multiple GXXXG motifs in cylindrical channels. Proceedings of the National Academy of Sciences of the United States of America. 2004;101:5988-91.

[44] Skibinski DA, Genisset C, Barone S, Telford JL. The cell-specific phenotype of the polymorphic vacA midregion is independent of the appearance of the cell surface receptor protein tyrosine phosphatase beta. Infection and immunity. 2006;74:49-55.

Chapter III

Membrane Association Induces Conformational Changes in the *Helicobacter*

Pylori Pore Forming Toxin VacA

The work in this chapter is unpublished and includes contributions from the following authors: Erwin, A.L., Chang L., Torrez R.M., Hanks J.L., Campbell A.M., Caso G., Cover, T.L., and Ohi, M.D. A.L.E. designed and conducted the experiments. L.C. collected VacA s1m1 and Δ6-27 DDM negative stain datasets. R.M.T. optimized VacA s1/i1/m1 binding to SUVs. J.H conducted the VacA s1/i1/m1 detergent screen. A.M.C. and G.C. purified VacA s1/i1/m1 wild-type, Δ6-27, and Δ346-347 protein. T.L.C. provided feedback on all experiments. M.D.O designed and provided feedback on all experiments.

Abstract

   Helicobacter pylori infects ~50 percent of the human population and causes gastritis, peptic ulcer disease, and gastric cancer*. H. pylori* secretes an 88kDa pore forming toxin called VacA that is a key virulence factor implicated in these diseases. While recently determined structures of VacA oligomers have revealed key insights into how VacA can assemble in solution, there is limited structural information about VacA before and after membrane association. Here, we structurally analyze membrane associated VacA oligomers (detergent micelle and liposome bound) using negative stain and cryo electron microscopy. Comparison of these structures to VacA Δ6-27 mutant data and existing VacA soluble oligomer structures reveals differences in VacA oligomers upon membrane association. This work provides structural insights into the process of VacA pore formation and identifies regions of VacA that undergo conformational changes upon membrane association.

Introduction

   Pore-forming toxins are expressed by many pathogenic bacteria to kill other bacteria or as virulence factors that aid in the colonization and infection of hosts[1]. These proteins are produced by organisms as a soluble monomer that after binding to the plasma membrane of the target cell can oligomerize, insert, form active transmembrane oligomeric pores on the plasma membrane and/or membranes of intracellular organelles, and elicit a variety of downstream cellular effects[2]. This is a very heterogeneous group of proteins in terms of size, domain number, target membranes, cell receptors, oligomerization states, and mechanisms of action[3]. The two major classes of bacterial pore forming toxins, α and β, are distinguished by the

secondary structure of the transmembrane pore[4-6]. X-ray crystallography and EM experiments with soluble monomers, soluble oligomers, and oligomers in membrane-like environments (e.g. detergents, nanodiscs, liposomes) have shed light on how proteins from both classes undergo dynamic rearrangements from soluble monomers to intermediate oligomeric pre-pores to active oligomeric transmembrane pores [7-19].

The Gram-negative bacterium *Helicobacter pylori* that infects over half of the human population and causes gastritis, peptic ulcer disease, and gastric cancer secretes a pore forming toxin called vacuolating cytotoxin A (VacA) that is implicated in these diseases[20-24]. VacA is secreted by *H. pylori* as 88kDa monomers (residues 1-821) and shares very limited sequence similarity to any characterized proteins from other bacterial pore forming toxins. Upon binding to the surface of gastric epithelial cells, oligomerizing, and forming anion-selective transmembrane channels, VacA can cause vacuolation, membrane permeabilization, mitochondrial dysfunction, cell death, autophagy, and T cell inhibition[25-35]. These VacA-induced cellular alterations depend on the formation of active transmembrane channels that are predicted to form either in the plasma membrane, mitochondrial, or endosomal membrane[25, 26, 29, 36-42].

The mature, secreted VacA toxin consists of two functional regions, a 33kDa N-terminal region (p33 – residues 1-311) and a 55kDa C-terminal region (p55 – residues 312-821)[43]. p33 and p55 can vary between *H. pylori* strains in three polymorphic regions, the N-terminal signal region (s1 or s2), the intermediate region (i1 or i2) located near the C-terminus of p33, and the p55 mid- region (m1 or m2)[24, 44-46]. Humans infected with *H. pylori* containing *vacA* s1/i1/m1 alleles have an increased

risk of developing gastric adenocarcinoma compared to those infected with strains

containing *vacA* s2/i2/m2 alleles[24, 45, 47-49]. Additionally, the *vacA* genotype

influences the extent of VacA-induced cellular effects and activity *in vitro*. Sequence

diversity within s region affects pore activity, while variation within the i and m regions of

VacA has been functionally associated with differences in receptor binding and/or

cellular tropism[24, 44, 50-53].

Planar lipid bilayer and patch clamping experiments have shown that VacA can

oligomerize on lipid membranes and form chloride selective membrane channels[25,

36-38]. A VacA mutant with residues 6-27 deleted (VacA Δ6-27), containing the tandem

GXXXG motifs, does not form channels in planar lipid bilayers and lacks vacuolating

activity, showing that the GXXXG motifs (residues 14-26) are required for the

formation of chloride selective membrane channels[52, 54]. Computational modelling

of the VacA p33 GXXXG motifs suggests that the glycine residues in the GXXXG motifs

pack against small alanine or valine side chains to form a helical bundle as the anion-

selective transmembrane pore[55]. While previous work from our lab analyzed VacA

bound to membrane, showing that VacA can form hexamers and heptamers that are

inserted into the membrane, this work was not conclusive about the conformation of the

GXXXG motifs and central pore-forming region of VacA in a lipid environment[56].

Recently determined <4Å structures of soluble VacA hexamers show 88kDa

VacA protomers (p88) consist of a continuous right-handed β-helix that oligomerize

with neighboring protomers through salt bridges (K47:E338, K55:D346), side chain

hydrogen bonds (R50:T342, K75:Q343, K55, D346), and extensive main-chain

hydrogen bonds linking residues from both the p33 and p55 regions[17, 19].

However, density was not assigned for the N-terminus of p33 (residues 1-26) comprising the GXXXG motifs or residues (300-334) within the predicted flexible loop between the p33 and p55 regions, suggesting that these residue ranges are highly flexible in solution[17, 19]. The first resolved residues of the p33 region (residues 27-37) are a helical region buried in the oligomeric interface that has been suggested to swing down into the central density to facilitate GXXXG motif helical bundling in the context of membrane [17, 19].

While membrane association is essential for VacA function, the mechanism of how VacA undergoes a dynamic rearrangement from a soluble monomer to binding membrane and forming a transmembrane pore has not been fully characterized, limiting our understanding of this important virulence factor secreted by *H.* pylori. Here, we used negative stain and cryo electron microscopy (cryo-EM) to structurally analyze VacA in a liposome bound state and detergent micelle bound state resulting in snapshots of VacA associating with membrane. Together, this work provides structural insights into the process of VacA pore formation and identifies regions of VacA that undergo conformational changes upon membrane association.

Results

Partial insertion of VacA oligomers into small unilamellar vesicle lipid bilayers

To understand how VacA inserts into lipid bilayers, we stabilized a membrane associated conformation of VacA *in vitro* by incubating VacA wild-type s1/i1/m1 with eggPC/DOPS/Cholesterol (55/15/30 mol %) small unilamellar vesicles (SUVs). EM analysis of the VacA-bound SUVs revealed that the SUVs were 200-400Å in diameter,

wide enough to accommodate the ~140Å width of the predicted VacA pore forming region within the ~300Å wide single layer oligomers (Fig. 3.1A).

Cryo-EM image acquisition of VacA SUV particles was attempted on multiple grid types to increase the number of VacA associated liposome particles per frame. Similar to what other groups have observed with cryoEM imaging of membrane proteins embedded in liposomes, VacA-bound SUVs particles preferred the continuous ultra thin carbon coated grids compared to lacey carbon and Quantifoil grids with no coating[57]. 2-D classification of 1,148 particles selected manually from 3,800 micrographs revealed two main types of membrane associated VacA single layers (Fig. 3.1A). In some 2-D classes, the inner and outer leaflets of the bilayer were observed with single layer VacA oligomers positioned on top of the outer leaflet (Fig. 3.1A bottom panel). In other classes, a VacA pre-pore appeared to insert into the outer leaflet of the bilayer with membrane absent in the pore lumen (Fig. 3.1A bottom panel). More particles are needed to further characterize these different VacA pre-pore states. In addition, VacA Δ6-27 single layer oligomers were able to bind to small unilamellar vesicles for analysis by cryo-EM. This revealed that the mutant lacking the GXXXG motif region was still able to bind SUVs (Fig. 3.1B). More particles are needed to enable visualization of whether the deletion of this important pore forming region affects formation of a VacA pre-pore into the outer leaflet of the membrane bilayer.

Partial assembly of VacA oligomers incubated directly with LMNG detergent

To further visualize VacA pore formation, we attempted to induce pore formation *in vitro* by incubation of VacA monomers with detergent as a membrane mimetic and

analyze the structure of the complex by single particle cryo-EM. We first incubated acid activated soluble VacA monomers with a variety of detergents and analyzed the particles formed by negative stain EM (Fig. 3.2A-C). We were interested in the detergent conditions that produced the most single layered oligomers, indicative of detergent micelles forming around the pore forming region of VacA and thereby preventing double layer formation. While we observed that VacA appeared to form a mixture of both double layer and single layer oligomers in the presence of LMNG, Cymal-6, and Cymal-7, from negative stain EM micrographs it appeared that 1.1x the CMC of LMNG (0.0011%) resulted in the most single layer VacA pore complexes (Fig. 3.2A).

We further analyzed the VacA LMNG sample using single particle cryo-EM. Analyzing the VacA LMNG particles from micrographs using RELION 2D classification revealed 2D class averages representing *en face* views of intact and partially assembled single layer oligomers (Fig. 3.3A). To confirm the extent of VacA partial assembly, we collected 30˚-tilted data. 2D classification of this 30˚-tilted data also revealed class averages with tilted intact and partially assembled VacA oligomers (Fig. 3.3B). There did not appear to be a difference in the central region of the VacA LMNG micelle 2-D class averages from both untilted and 30˚-tilted data when compared to soluble VacA single layer oligomer 2-D class averages. While addition of LMNG to VacA was successful in preventing double layer formation, it did not appear to induce a conformational change in VacA when comparing VacA LMNG micelle complex 2-D class averages to previous soluble VacA 2-D class averages.

DDM detergent-solubilized VacA wild-type oligomers from LUVs form single layers with signal in central density that may be attributed to VacA N-terminal GXXXG motifs

Given the lack of clear single layer pore formation when acid-activated VacA was incubated with detergent directly, we next tested the ability of three detergents (DDM: 0.02%, LMNG: 0.002%, and Cymal-6: 0.4%) to extract VacA bound to large unilamellar vesicles (LUVs) comprised of eggPC/DOPS/Cholesterol (55/15/30 mol%) (Fig. 3.6). Negative stain images of these samples showed that solubilization of VacA-bound LUVs with DDM resulted in a majority of well-defined single-layer VacA particles (Fig. 3.4A), while solubilization with LMNG and Cymal-6 resulted in more double layer and aggregated particles (Fig. 3.4B-C). Further negative stain analysis of DDM-solubilized VacA s1/i1/m1 wild-type oligomers revealed more single-layer (55.6%) oligomers than double-layer (44.4%) oligomers, which differs from the majority double layer oligomer particle distribution observed in previous negative stain and cryoEM analyses of VacA wild-type soluble oligomers (Fig. 3.5A,C).

Since the VacA N-terminal GXXXG region is required for active pore formation and predicted to be the transmembrane domain that would be likely surrounded by the DDM micelle, we next tested whether VacA Δ6-27, a mutant with all three N-terminal GXXXG motifs deleted, is able to form single layer DDM-solubilized VacA oligomers (Fig. 3.5B)[55, 58]. Negative stain analysis revealed more single-layer (64.6%) than double layer (35.4%) VacA Δ6-27 oligomers (Fig. 3.5C), suggesting that similar to wildtype, VacA Δ6-27 can form single layer DDM-solubilized VacA oligomers. The deletion of the GXXXG motif region does not interfere with DDM acting as a membrane-mimetic, supporting a previous study which showed that the VacA GXXXG motif region

(residues 14-26) is not the only region of VacA that inserts into membrane [56]. Intriguingly, it appears that the central region of the DDM-solubilized VacA Δ6-27 oligomers have less signal when compared to DDM-solubilized VacA WT oligomers, as has been observed in previous LUV studies, suggesting that the VacA GXXXG motifs contribute to density in the central region.

DDM detergent-solubilized VacA oligomers from LUVs form hexameric and heptameric pre-pore complexes

To further investigate which regions of VacA are involved in membrane insertion, we analyzed the DDM-solubilized VacA wild-type s1/i1/m1 particles using single particle cryo-EM. Analysis of the single particles with 2D classification and sorting in cryoSPARC revealed VacA double layers and single layer hexamers and heptamers. 2D class averages of DDM-solubilized VacA hexamer and heptamer en face particles had clear density in the center of the oligomers (Fig. 3.6A, 3.7A). Defined central density was not observed in 2D class averages and 3D structures of soluble VacA wild-type s1/i1/m1 single layer oligomer cryo-EM data, suggesting that VacA oligomers undergo a dramatic conformational change in the presence of detergent/membrane[17, 19]. After separating the hexamer, heptamer, and double layer en face and slightly tilted particles using 2D classification, we determined 3D structures of the DDM-solubilized VacA hexamer and heptamer (Fig. 3.6B, 3.7B). With the lack of side view particle information, the 3D maps appeared anisotropic and particle orientation distribution maps revealed extreme preferred orientation for the VacA hexamer and heptamer particles (Fig. 3.6C, 3.7C).

Previous EM studies have observed that the majority of soluble VacA hexamer and heptamer particles orient as en face or double layer side views and consequently have used negative stain random conical tilt, extensive 3D classification, or signal subtraction of double layer side view particles to obtain isotropic 3D structures of soluble VacA hexamers[17, 19, 59]. To increase the range of particle orientations of VacA detergent micelle particles in vitrified ice, we collected micrographs of at a 40° tilt angle. While the tilting approach presents technical challenges (electron charging, less accurate drift correction of movies, decreased signal to noise, and less accurate CTF determination), we observed 2D class averages of hexamer, heptamer that had crisp features including the prominent density in the central region of the oligomers (Fig. 3.6D, 3.7D)[60]. The increase in the range of particle orientations was apparent in the 3D structures and particle orientation distribution maps (Fig. 3.6E-F, 3.7E-F). The symmetry applied maps displayed a basket-like density protruding ~30Å from N-terminal region of p88 on the membrane binding face of VacA (Fig. 3.6E, 3.7E). Given that typical membrane bilayers are ~40Å, it is likely that the VacA DDM micelle structure represents the membrane-associated pre-pore state in position to form the transmembrane channel in the correct context. This VacA DDM micelle pre-pore structure may correspond to the pre-pore state observed in the 2-D class averages of the VacA SUV sample (Fig. 3.1A).

Placement of the soluble VacA hexamer model (PDB 6NYF) into the 40°-tilted VacA DDM micelle hexamer map indicated that the majority of the soluble VacA model fit into the map (Fig. 3.8A)[19]. While the VacA DDM micelle map did not appear to have well-defined density at the oligomeric interface for residues 27-45, the overlay shows

that this helical region appears poised to move toward the spokes protruding from the membrane-binding face of each protomer (Fig. 3.8B). This hypothesized movement would bring the tandem GXXXG motif region (7-29) from each protomer together to form a helical bundle as the membrane spanning channel. Placement of the computational model for the GXXXG motif region (Model Archive File 1SEW) into the central density indicated that the density can accommodate the ~35Å x ~15Å predicted GXXXG motif helical bundle (Fig. 3.8A-C)[55]. Other residues located near the oligomeric interface that may contribute to the central density in the VacA DDM micelle map include residues 297-342, a predicted flexible loop that was not observed in the VacA soluble map (Fig. 3.8A-B)[43]. It is possible that upon binding to the appropriate receptor or entering the lower pH of the late endosomal compartments, residues 27-45 may move into the spokes protruding from each protomer, bringing the GXXXG motif bundle together in the central density. The predicted flexible loop residues 297-342 may act to stabilize the oligomeric interface in the transmembrane pore state while also serving to further anchor the complex into the membrane.

In an attempt to further improve the quality of the VacA DDM micelle structure and further pinpoint the exact regions of p88 forming the central density, 40˚-tilted dataset particles were combined with the untilted dataset particles in a 1:10 (untilted:40˚-tilted) ratio, increasing the range of particle orientations. However, this did not dramatically alter the resulting 3D structures (Fig. 3.9, 3.10). While combination of the 40˚-tilted particles with all of the untilted particles resulted in 3D structures with improved p88-p88 connectivity for the en face view, anisotropy was observed for the side view of the symmetry-applied heptamer 3D structure (Fig. 3.10E). Addition of more

40˚-tilted data and further 3D classification and focused refinement steps may be needed to further improve the resolution of these VacA DDM micelle structures and clearly determine regions of VacA p88 involved in the formation of the pre pore state.

Discussion

While several structures α-pore forming toxins associated with membrane have been determined to date, the structural organization of the *H. pylori VacA* membrane channel has not been clearly defined[8, 10, 12-14, 16]. Only recently was structural basis of VacA oligomerization revealed with cryo-EM structures of soluble VacA hexamers, heptamers, dodecamers, and tetradecamers[17, 19]. However, none of these structures had well-defined central density or assigned residues for the N-terminal GXXXG motif residues (14-26) shown to be critical for VacA channel activity. Our studies analyzing VacA bound to small unilamellar vesicles and VacA extracted from large unilamellar vesicles with DDM detergent reveal that VacA wild-type oligomers that form in the presence of lipids adopt a conformationally different state compared to VacA oligomers in solution (Fig. 3.1, 3.6-3.8). While our previous study showed negative stain EM 2D class averages of VacA hexamers on LUVs, there has not been a definitive observation of the structural organization of the central region within membrane associated VacA oligomers[56]. Here, we show that VacA forms both hexamers and heptamers with defined central density when extracted from LUVs with DDM at pH 7.2(Fig. 3.5-3.8). Defined central density was not observed in 2D classification and 3D structures of soluble VacA wild-type s1/i1/m1 oligomer cryo-EM data, suggesting that VacA oligomers undergo a dramatic conformational change in the presence of

detergent/membrane[17, 19]. 3-D reconstructions of the VacA DDM micelle complexes show that the central density appears to extend ~30Å from the membrane-associated face of the VacA oligomers (Fig. 3.6E,3.8B). Further, 2-D class averages of VacA bound to SUVs show that VacA appears to insert into the outer leaflet of the SUV lipid bilayer (Fig. 3.1A). Together, these data reveal previously unseen features in the central region of membrane associated VacA oligomers and suggest that these VacA DDM complexes reflect a pre-pore state of the *H. pylori* VacA anion channel.

The soluble VacA hexamer models (PDB 6ODY and 6NYF) and the computational model for the VacA GXXXG motif (residues 7-29) helical bundle (Model Archive File 1SEW) fit into the 40°-tilted VacA DDM micelle hexamer map arms and central density, respectively (Fig. 3.8)[17, 19, 55]. The central density volume accommodating the dimensions of the VacA GXXXG computation model aligns with data shown of 2-D class averages of DDM solubilized VacA Δ6-27 oligomers that lack signal in the center of the oligomers when compared to DDM solubilized VacA WT oligomers (Fig. 3.5). We hypothesize that helical oligomer interface residues 27-45 may undergo a greater than 90° counter-clockwise movement into the spokes protruding from each protomer, bringing the GXXXG motif helical bundle together in the central density. It is not clear which direction the GXXXG bundle would face, with the N-terminus (residue 6) buried within the outer leaflet of the membrane bilayer or facing into the endosome lumen (or outside the cell). It is plausible that changes in the tilt of the p88 arms around the plane parallel to the membrane would be required to facilitate these conformational changes of forming the pre-pore state. An improved resolution, in combination with more cryo-EM data and improved 2-D class averages and 3-D

reconstructions of VacA Δ6-27 bound to SUVs and VacA Δ6-27 DDM micelle complexes will be required to investigate these open questions and examine more subtle changes in VacA during the process of pore formation.

It is unclear whether the predicted loop (300-334) that was not observed in the soluble VacA oligomer structures contributes to the central density in the VacA DDM micelle map (Fig. 3.8A-B)[17, 19, 43]. This flexible linker could serve different functions depending on the VacA pore state. It is possible it stabilizes the oligomeric interface in the transmembrane pore state or anchors the complex into the membrane. The pore forming toxins TcdA and YenA2 both have linkers similar in length to the VacA predicted loop (300-334) that lack secondary structure in the pre-pore states and are stabilized and form secondary structure in the pore-inserted states[10, 13]. In the both cases, the collapse of the linker is the most likely driver of membrane insertion[13, 61]. Channel activity assays with the VacA Δ294-331 mutant, a mutant with the linker region deleted, and high resolution structural information of VacA in the pre-pore and pore-formed states will be required to understand the role of the VacA predicted loop and the overall dynamics of VacA pore formation[62].

In summary, we have shown that VacA can assemble into hexamer and heptamers pre-pore complexes in the presence of membrane. It remains to be determined how VacA transitions from this pre-pore state to forming a transmembrane pore and whether there are multiple required steps to this process. More data to improve VacA wild-type and mutant structural analysis will help define the regions involved in pre-pore formation. How VacA transmembrane pore formation is initiated *in vivo* remains unknown. Given the localization of VacA to late endosome compartments,

it has been hypothesized that pH will play an important role in triggering pore

formation[39, 63]. Planar lipid bilayer assays with VacA that have been conducted at pH

4 have shown that VacA forms anion selective pores in membrane [37, 54]. Additional

channel activity experiments with VacA wild-type and mutants associated with

membrane at varying pH values will help understand the role of pH and the different

pore states that exist during the process of pore formation. The use of gastric epithelial

cell membranes for experiments may also trigger VacA pore formation due to the

presence of unidentified receptors. Cryo-electron tomography of vacuoles isolated from

cells or extraction of VacA from cell membranes using styrene copolymers in

conjunction with mass spectrometry may help determine important elements for

triggering VacA transmembrane pore formation[64]. Further structural and functional

analyses of VacA wild-type and mutants in the context of native membrane will be

important to fully define how VacA forms an active transmembrane pore.
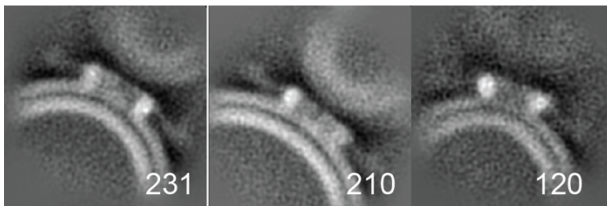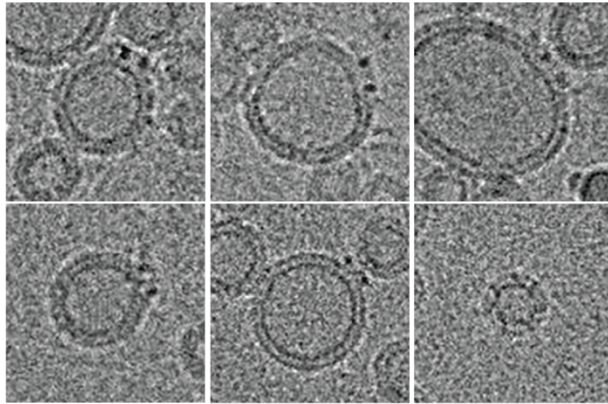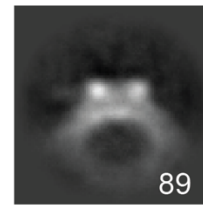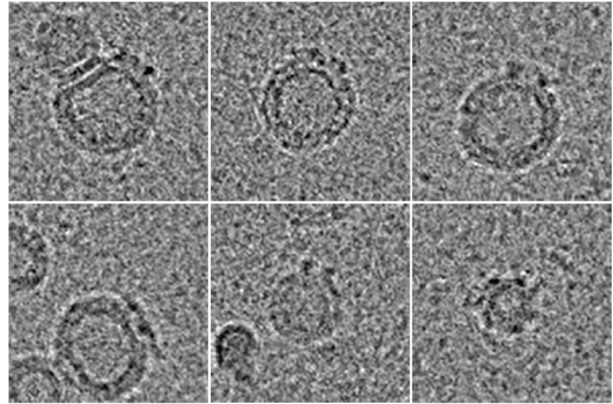
Figure 3.1. Cryo-EM analysis of VacA s1/i1/m1 oligomers bound to SUVs. a) Representative VacA s1/i1/m1 wild-type bound to SUV (55/15/30 mol % eggPC/DOPS/Cholesterol) particles from cryo-EM micrographs (top) and 2-D class averages (bottom). 2D class average box size, 364Å. B) Representative VacA s1/i1/m1 Δ6-27 bound to SUV (55/15/30 mol % eggPC/DOPS/Cholesterol) particles from cryo-EM micrographs (top) and 2-D class averages (bottom). 2-D class average box size, 490Å. All raw particle box sizes, 400Å. Number of particles shown in bottom right corner of each 2-D class average.
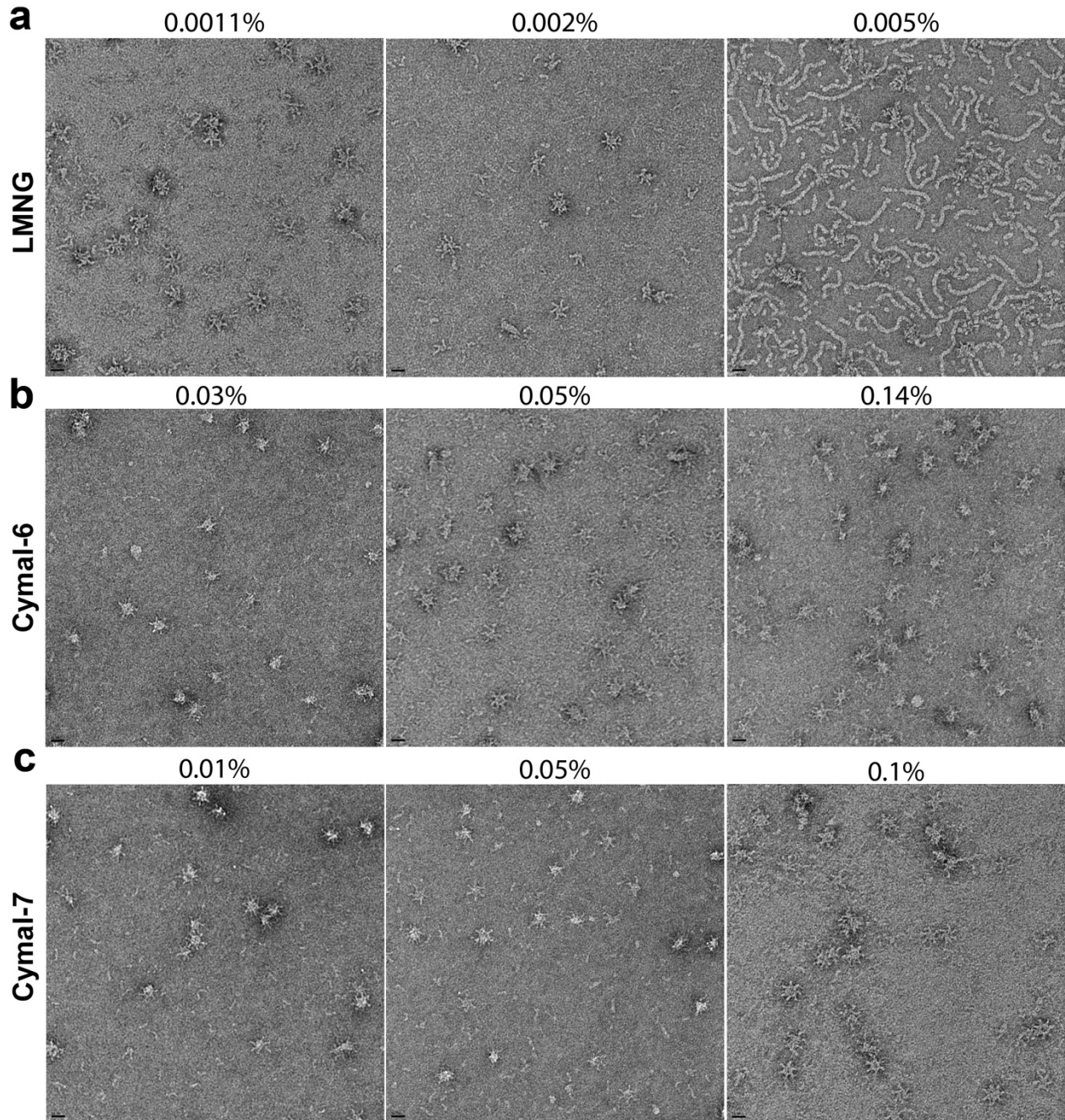
Figure 3.2. Negative stain EM analysis of VacA s1/i1/m1 detergent micelle complexes. a) VacA s1/i1/m1 wild-type incubated with increasing amounts of LMNG above the CMC (0.0011%, 0.002%, 0.005%). b) VacA s1/i1/m1 wild-type incubated with increasing amounts of Cymal-6 above the CMC (0.03%, 0.005%, 0.14%). C) VacA s1/i1/m1 wild-type incubated with increasing amounts of Cymal-7 above the CMC (0.01%, 0.05%, 0.1%). Scale bar (lower left of each image), 20 nm.

a untilted
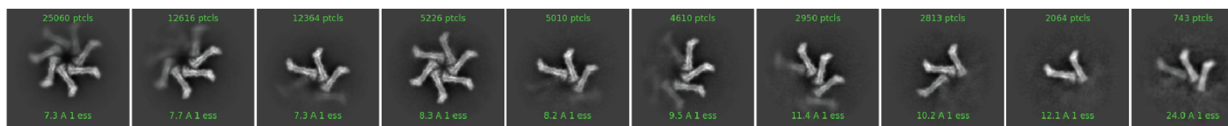
b 30˚-tilted
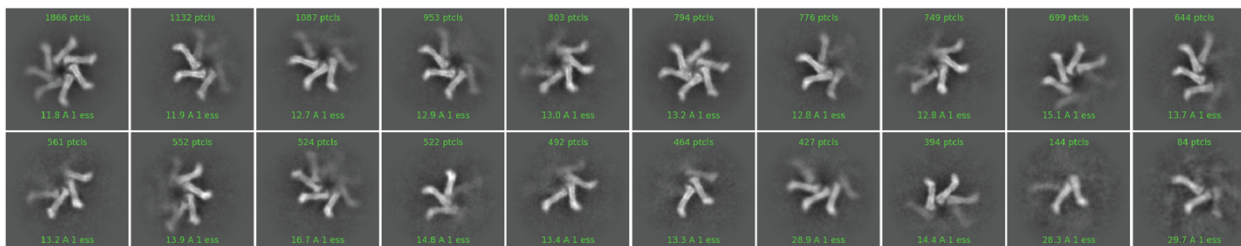
Figure 3.3. Cryo-EM analysis of VacA s1/i1/m1 LMNG (0.0011%) detergent micelle complexes. a) 2-D class averages of VacA s1/i1/m1 LMNG (0.0011%) complexes from untilted data. b) 2-D class averages of VacA s1/i1/m1 LMNG (0.0011%) complexes from 30˚-tilted data. Box sizes for 2D class averages, 480 Å.
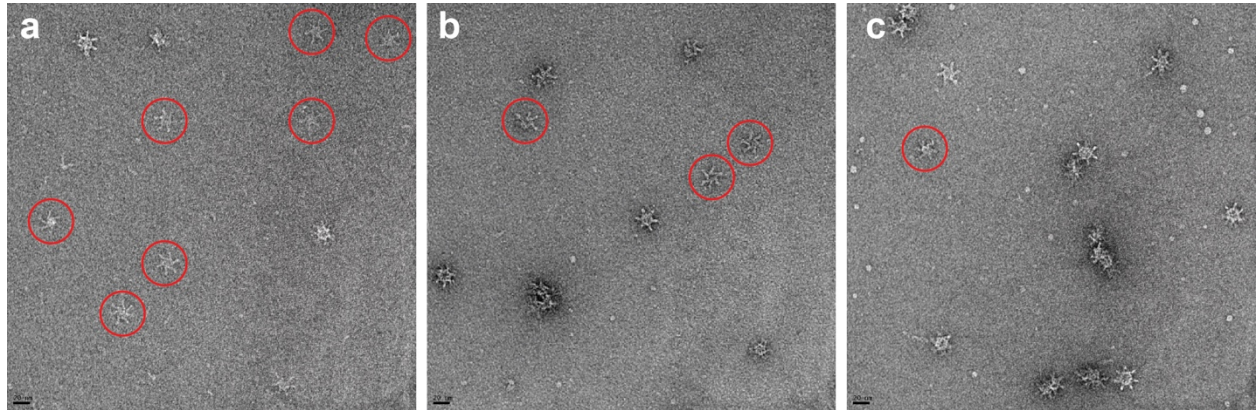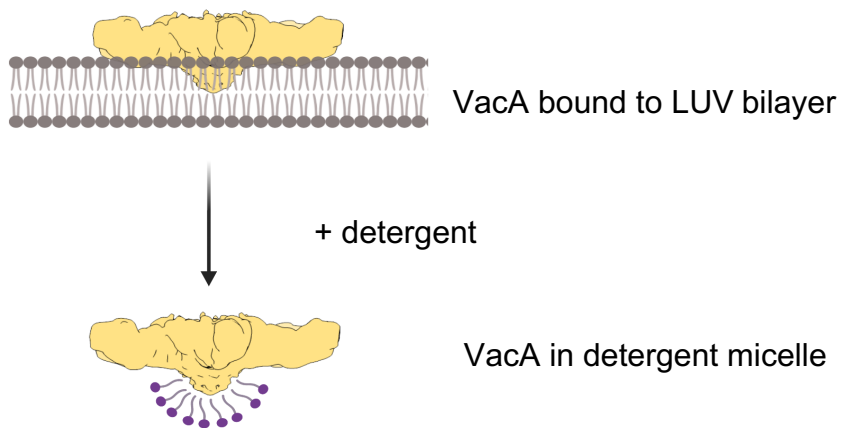
Figure 3.4. Negative stain micrographs of detergent solubilized VacA wild-type s1/i1/m1 oligomers. VacA wild-type s1/i1/m1 oligomers were bound to LUVs (55/15/30 mol % eggPC/DOPS/Cholesterol) and solubilized with a) DDM (0.002%), b) LMNG (0.002%), and c) Cymal-6 (0.4%). Representative single layer VacA oligomers are circled in red. All scale bars, 20nm. d) Cartoon schematic of detergent solubilization of VacA LUV-bound oligomers. VacA oligomer shown in orange. Detergent molecules shown in purple. Made with BioRender.
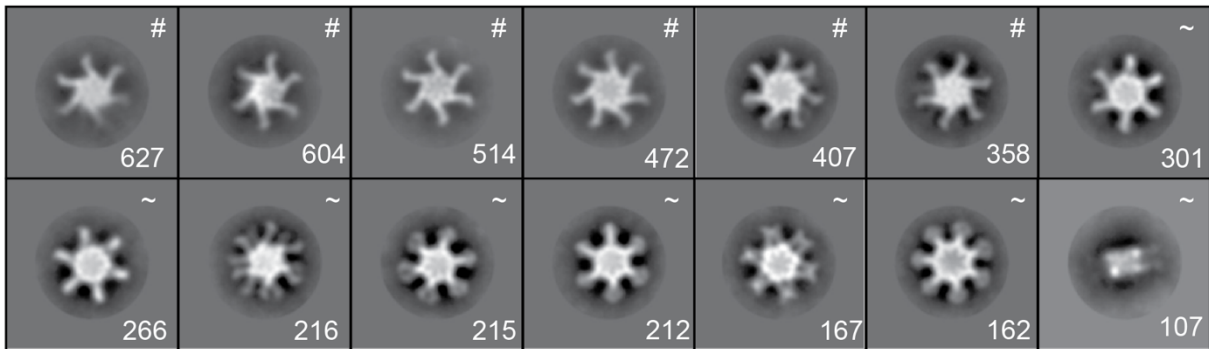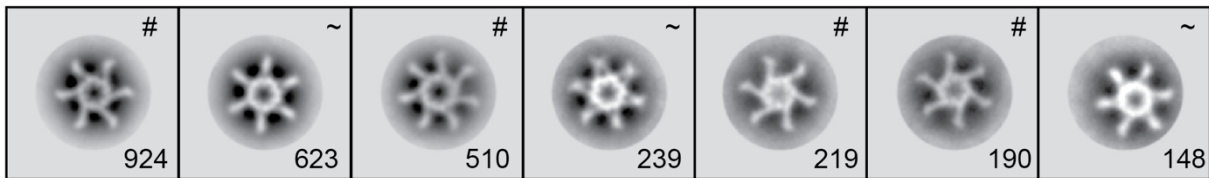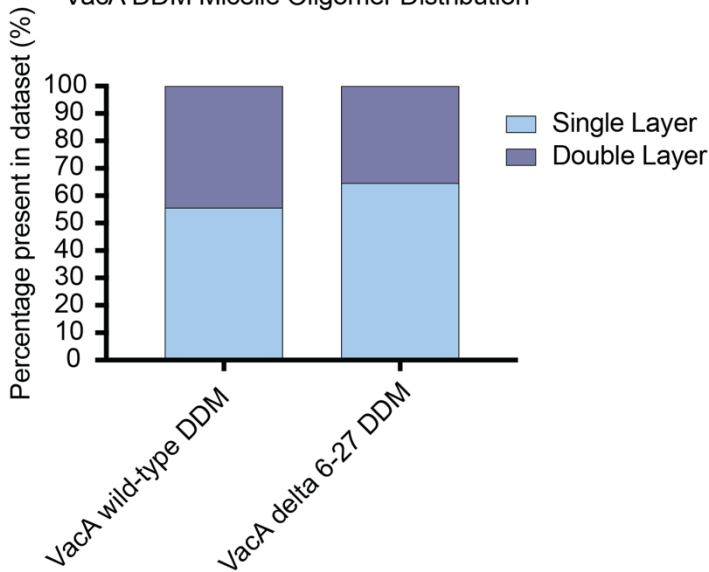
Figure 3.5. Characterization of negatively stained DDM-solubilized VacA s1/i1/m1 oligomers.
Representative 2-D class averages of a) VacA s1/i1/m1 wild-type and b) VacA s1/i1/m1 Δ6-27 solubilized from large unilamellar vesicles (55/15/30 mol % eggPC/DOPS/Cholesterol) with n-Dodecyl-ß-D-Maltoside (DDM) detergent. Number of particles shown in bottom right corner of each class. Box size, 618Å. Mask size, 400Å. ~ indicates double layer VacA; # indicates single layer VacA. c) Distribution of VacA DDM micelle single layer (blue) and double layer (purple) particles shown as percentage of total particles in dataset.

a VacA WT DDM untilted

b

C1   C6

c

C1

C6

d VacA WT DDM 40°-tilted

e

C1   C6

f

C1

C6

GSFSC Resolution: 8.36Å

No Mask (12Å)
Spherical (11Å)
Loose (8.8Å)
Tight (8.4Å)
Corrected (8.4Å)
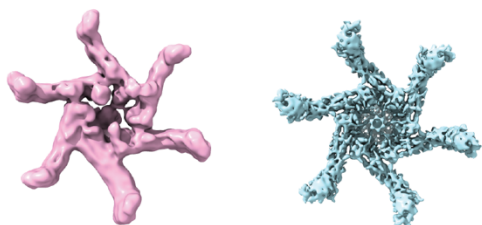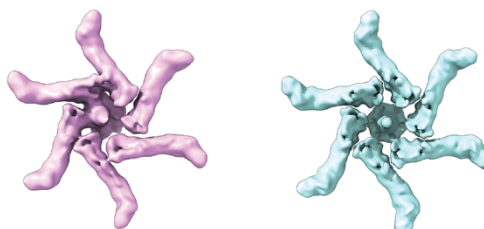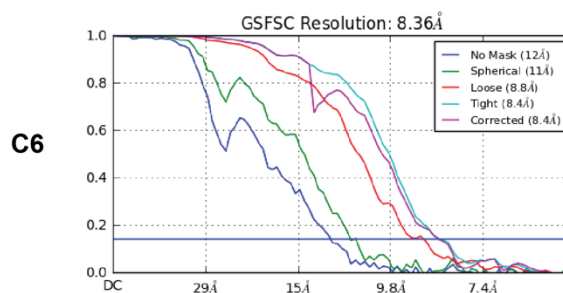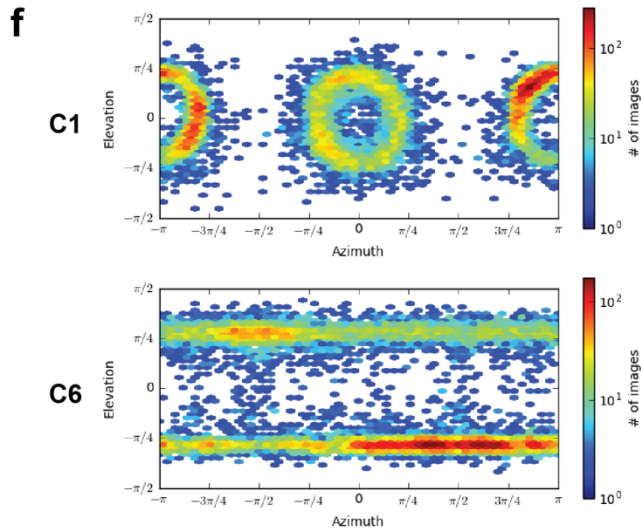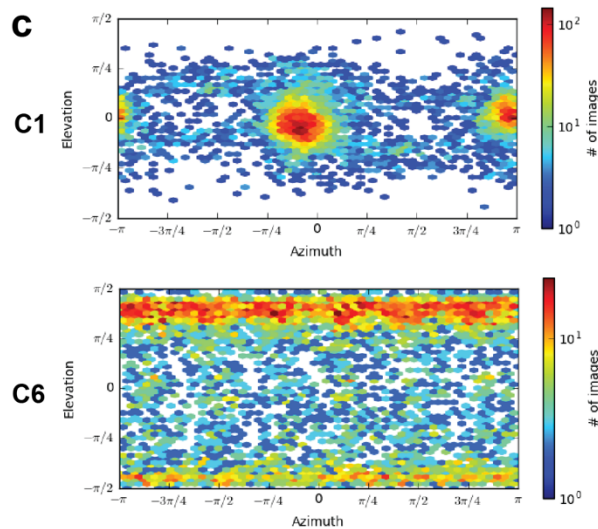
Figure 3.6. Cryo-EM analysis of DDM-solubilized VacA s1/i1/m1 wild-type hexamers.

a) Representative 2-D class averages of untilted VacA s1/i1/m1 wild-type hexamer particles solubilized from LUVs (55/15/30 mol % eggPC/DOPS/Cholesterol) with DDM. Number of particles is shown in bottom right corner of each class. Box size, 588Å. b) 3-D reconstruction of untilted VacA s1/i1/m1 wild-type hexamers refined with C1 (left) and C6 symmetry (right). Top, *en face* views. Bottom, views rotated around the x-axis -135˚. c) Particle orientation distribution maps for the C1 and C6 refinement maps from (b). d) Representative 2-D class averages of 40˚-tilted VacA s1/i1/m1 wild-type hexamer particles solubilized from LUVs (55/15/30 mol % eggPC/DOPS/Cholesterol) with DDM. Number of particles is shown in bottom right corner of each class. Box size, 588Å. e) 3-D reconstruction of 40˚-tilted VacA s1/i1/m1 wild-type hexamers refined with C1 (left) and C6 symmetry (right). Top, *en face* views. Bottom, views rotated around the x-axis -135˚. f) Particle orientation distribution maps for the C1 and C6 refinement maps from (e). FSC curves for the C6 refinement map from (e).

a VacA WT DDM untilted

1178 796 455

b

135° 135°

C1 C7

c

C1

C7

d VacA WT DDM 40°-tilted

2487 2464 2355

e

135° 135°

C1 C7

f

C1

C7

GSFSC Resolution: 10.08Å

No Mask (13Å)
Spherical (13Å)
Loose (11Å)
Tight (10Å)
Corrected (10Å)

C7
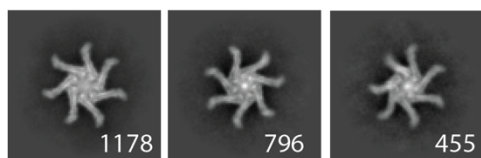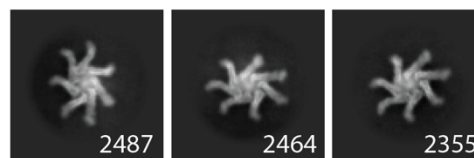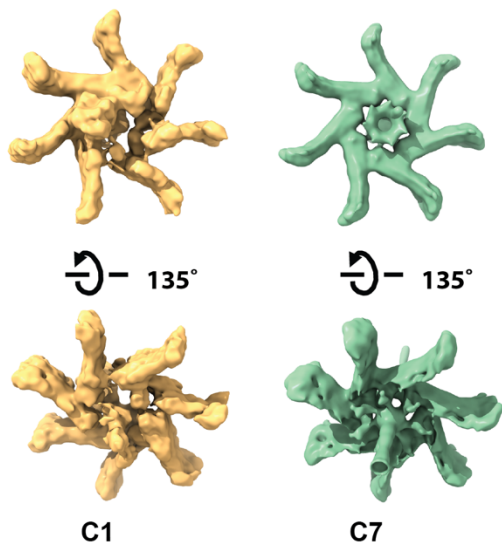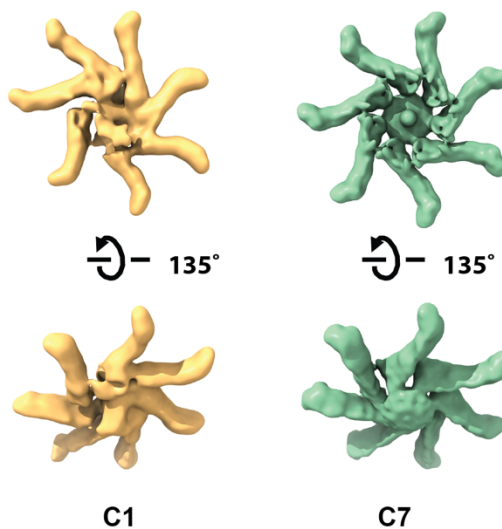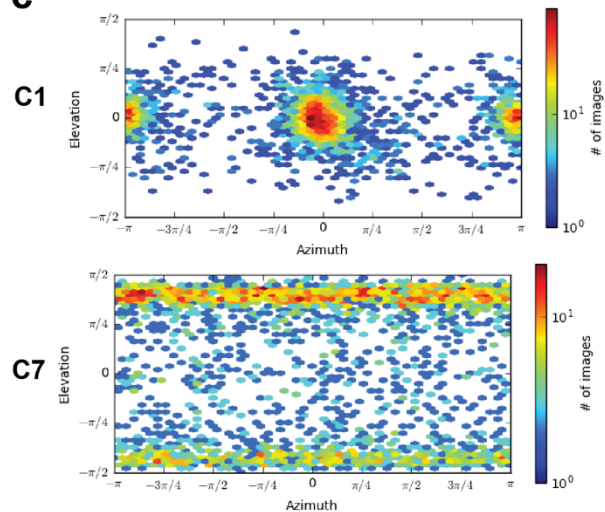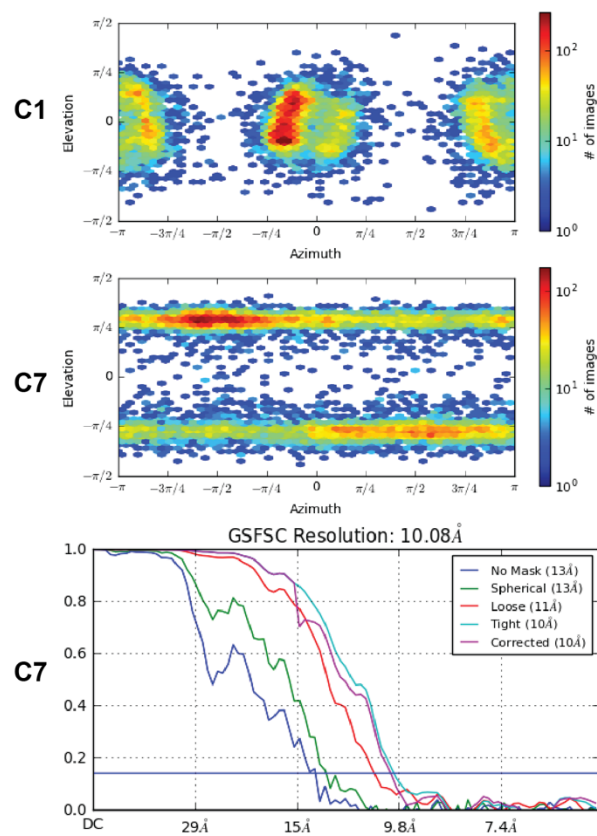
Figure 3.7. Cryo-EM analysis of DDM-solubilized VacA wild-type s1/i1/m1 heptamers.

a) Representative 2-D class averages of untilted VacA s1/i1/m1 wild-type heptamer particles solubilized from LUVs (55/15/30 mol % eggPC/DOPS/Cholesterol) with DDM. Number of particles is shown in bottom right corner of each class. Box size, 588Å. b) 3-D reconstruction of untilted VacA s1/i1/m1 wild-type hexamers refined with C1 (left) and C7 symmetry (right). Top, *en face* views. Bottom, views rotated around the x-axis -135˚. c) Particle orientation distribution maps for the C1 and C7 refinement maps from (b). d) Representative 2-D class averages of 40˚-tilted VacA s1/i1/m1 wild-type heptamer particles solubilized from LUVs (55/15/30 mol % eggPC/DOPS/Cholesterol) with DDM. Number of particles is shown in bottom right corner of each class. Box size, 588Å. e) 3-D reconstruction of 40˚-tilted VacA s1/i1/m1 wild-type heptamers refined with C1 (left) and C7 symmetry (right). Top, *en face* views. Bottom, views rotated around the x-axis -135˚. f) Particle orientation distribution maps for the C1 and 76 refinement maps from (e). FSC curves for the C7 refinement map from (e).

a

p55

p33

protomer interface

en face view

GXXXG motifs Model Archive File (7-29)
Upstream of GXXXGs (27-45)
Predicted Loop (300-334)

b

endosome
lumen

endosome
membrane

~40Å

cytosol

endosome
lumen

~30Å

endosome
membrane

~40Å

cytosol

GXXXG motifs (7-29)
Model Archive File

side view
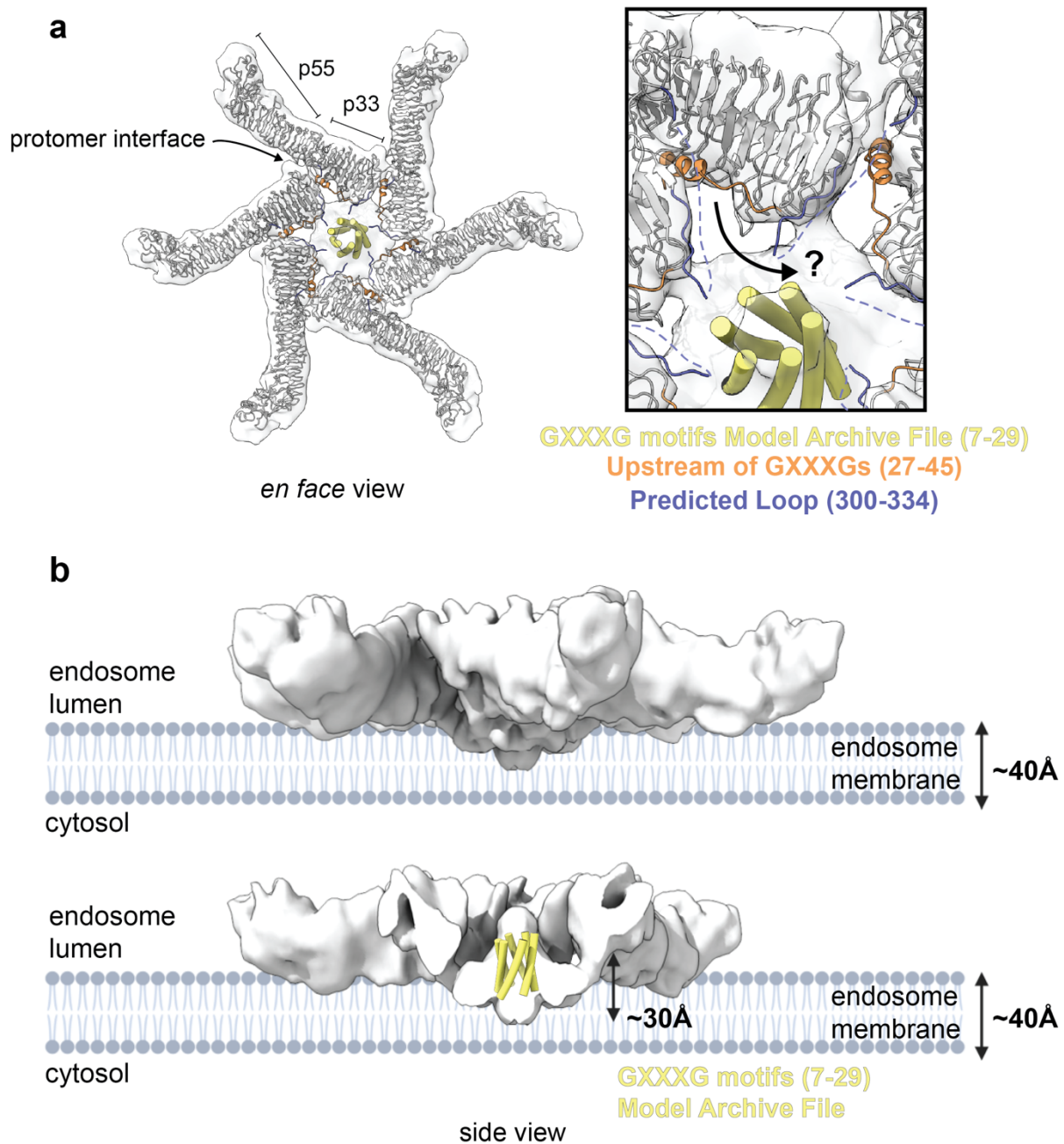
Figure 3.8. Comparison of soluble VacA wild-type s1/i1/m1 oligomer with DDM-solubilized VacA wild-type s1/i1/m1 oligomer. a) Left: En face view of soluble VacA wild-type s1/i1/m1 oligomer model PDB 6NYF and VacA GXXXG motifs computational model archive file 1SEW placed into DDM-solubilized VacA wild-type s1/i1/m1 oligomer 3-D reconstruction (from 40˚-tilted data). Right: Closer view of residues at the protomer interface that are poised to be involved in membrane pore formation. Residues 7-29 (Model Archive File 1SEW) are shown in yellow, 27-45 (PDB 6NYF) are shown in orange, and 297-342 (PDB 6NYF) are shown in purple. b) Top: Model of how DDM-solubilized

68

VacA wild-type s1/i1/m1 oligomer (from 40˚-tilted data) could partially insert into a lipid bilayer on the inner leaflet of the ~40Å endosomal membrane bilayer. Bottom: Slice view through side view of DDM-solubilized VacA oligomer to indicate length of density protruding toward center of oligomer. Residues 7-29 (Model Archive File 1SEW, shown in yellow) placed into central density with residue 7 and 29 at the bottom and top of each helix, respectively.
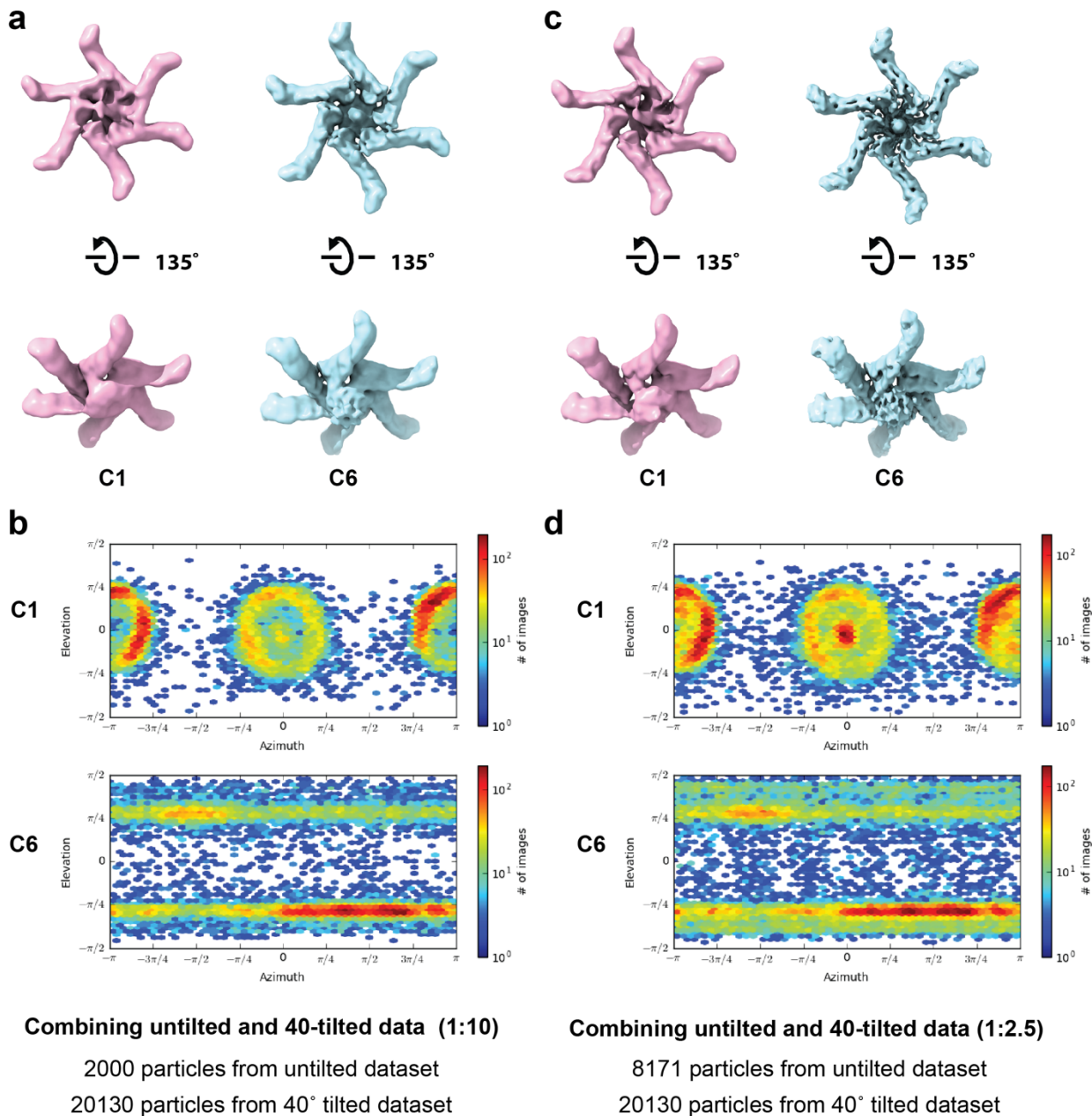
Figure 3.9. Combining untilted and 40˚-tilted cryo-EM datasets of DDM-solubilized VacA wild-type s1/i1/m1 hexamers

a) 3-D reconstruction of untilted and 40˚-tilted (1:10 ratio) VacA s1/i1/m1 wild-type hexamer particles solubilized from LUVs (55/15/30 mol % eggPC/DOPS/Cholesterol) with DDM refined with C1 (left) and C6 symmetry (right). Top, *en face* views. Bottom, views rotated around the x-axis -135˚. b) Particle orientation distribution maps for the C1 and C6 refinement maps from (a). c) 3-D reconstruction of untilted and 40˚-tilted (1:2.5 ratio) VacA s1/i1/m1 wild-type hexamer particles solubilized from LUVs (55/15/30 mol % eggPC/DOPS/Cholesterol) with DDM refined with C1 (left) and C6 symmetry (right). Top, *en face* views. Bottom, views rotated around the x-axis -135˚. d) Particle orientation distribution maps for the C1 and C6 refinement maps from (c).

Figure 3.10. Combining untilted and 40˚-tilted cryo-EM datasets of DDM-solubilized VacA wild-type s1/i1/m1 heptamers.
a) 3-D reconstruction of untilted and 40˚-tilted (1:10 ratio) VacA s1/i1/m1 wild-type heptamer particles solubilized from LUVs (55/15/30 mol % eggPC/DOPS/Cholesterol) with DDM refined with C1 (left) and C6 symmetry (right). Top, *en face* views. Bottom, views rotated around the x-axis -135˚. b) Particle orientation distribution maps for the C1 and C7 refinement maps from (a). c) 3-D reconstruction of untilted and 40˚-tilted (1:2.5 ratio) VacA s1/i1/m1 wild-type heptamer particles solubilized from LUVs (55/15/30 mol % eggPC/DOPS/Cholesterol) with DDM refined with C1 (left) and C7 symmetry

(right). Top, *en face* views. Bottom, views rotated around the x-axis -135°. d) Particle orientation distribution maps for the C1 and C7 refinement maps from (c).

References

[1] Los FC, Randis TM, Aroian RV, Ratner AJ. Role of pore-forming toxins in bacterial infectious diseases. Microbiol Mol Biol Rev. 2013;77:173-207.

[2] Bischofberger M, Iacovache I, van der Goot FG. Pathogenic pore-forming proteins: function and host response. Cell Host Microbe. 2012;12:266-75.

[3] Dal Peraro M, van der Goot FG. Pore-forming toxins: ancient, but never really out of fashion. Nat Rev Microbiol. 2016;14:77-92.

[4] Gouaux E. Channel-forming toxins: tales of transformation. Curr Opin Struct Biol. 1997;7:566-73.

[5] Lesieur C, Vecsey-Semjen B, Abrami L, Fivaz M, Gisou van der Goot F. Membrane insertion: The strategies of toxins (review). Mol Membr Biol. 1997;14:45-64.

[6] Parker MW, Feil SC. Pore-forming protein toxins: from structure to function. Prog Biophys Mol Biol. 2005;88:91-142.

[7] Bokori-Brown M, Martin TG, Naylor CE, Basak AK, Titball RW, Savva CG. Cryo-EM structure of lysenin pore elucidates membrane insertion by an aerolysin family protein. Nat Commun. 2016;7:11293.

[8] Brauning B, Bertosin E, Praetorius F, Ihling C, Schatt A, Adler A, et al. Structure and mechanism of the two-component alpha-helical pore-forming toxin YaxAB. Nat Commun. 2018;9:1806.

[9] Gatsogiannis C, Lang AE, Meusch D, Pfaumann V, Hofnagel O, Benz R, et al. A syringe-like injection mechanism in Photorhabdus luminescens toxins. Nature. 2013;495:520-3.

[10] Gatsogiannis C, Merino F, Prumbaum D, Roderer D, Leidreiter F, Meusch D, et al. Membrane insertion of a Tc toxin in near-atomic detail. Nat Struct Mol Biol. 2016;23:884-90.

[11] Gatsogiannis C, Merino F, Roderer D, Balchin D, Schubert E, Kuhlee A, et al. Tc toxin activation requires unfolding and refolding of a beta-propeller. Nature. 2018;563:209-13.

[12] Peng W, de Souza Santos M, Li Y, Tomchick DR, Orth K. High-resolution cryo-EM structures of the E. coli hemolysin ClyA oligomers. PLoS One. 2019;14:e0213423.

[13] Piper SJ, Brillault L, Rothnagel R, Croll TI, Box JK, Chassagnon I, et al. Cryo-EM structures of the pore-forming A subunit from the Yersinia entomophaga ABC toxin. Nat Commun. 2019;10:1952.

[14] Roderer D, Hofnagel O, Benz R, Raunser S. Structure of a Tc holotoxin pore provides insights into the translocation mechanism. Proc Natl Acad Sci U S A. 2019;116:23083-90.

[15] Savva CG, Clark AR, Naylor CE, Popoff MR, Moss DS, Basak AK, et al. The pore structure of Clostridium perfringens epsilon toxin. Nat Commun. 2019;10:2641.

[16] Schubert E, Vetter IR, Prumbaum D, Penczek PA, Raunser S. Membrane insertion of alpha-xenorhabdolysin in near-atomic detail. Elife. 2018;7.

[17] Su M, Erwin AL, Campbell AM, Pyburn TM, Salay LE, Hanks JL, et al. Cryo-EM Analysis Reveals Structural Basis of Helicobacter pylori VacA Toxin Oligomerization. J Mol Biol. 2019;431:1956-65.

[18] van Pee K, Neuhaus A, D'Imprima E, Mills DJ, Kuhlbrandt W, Yildiz O. CryoEM structures of membrane pore and prepore complex reveal cytolytic mechanism of Pneumolysin. Elife. 2017;6.

[19] Zhang K, Zhang H, Li S, Pintilie GD, Mou TC, Gao Y, et al. Cryo-EM structures of Helicobacter pylori vacuolating cytotoxin A oligomeric assemblies at near-atomic resolution. Proc Natl Acad Sci U S A. 2019;116:6800-5.

[20] Marshall BJ, Warren JR. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. Lancet. 1984;1:1311-5.

[21] Suerbaum S, Michetti P. Helicobacter pylori infection. N Engl J Med. 2002;347:1175-86.

[22] Atherton JC, Blaser MJ. Coadaptation of Helicobacter pylori and humans: ancient history, modern implications. J Clin Invest. 2009;119:2475-87.

[23] Cover TL, Blaser MJ. Purification and characterization of the vacuolating toxin from Helicobacter pylori. J Biol Chem. 1992;267:10570-5.

[24] Atherton JC, Cao P, Peek RM, Jr., Tummuru MK, Blaser MJ, Cover TL. Mosaicism in vacuolating cytotoxin alleles of Helicobacter pylori. Association of specific vacA types with cytotoxin production and peptic ulceration. J Biol Chem. 1995;270:17771-7.

[25] Szabo I, Brutsche S, Tombola F, Moschioni M, Satin B, Telford JL, et al. Formation of anion-selective channels in the cell plasma membrane by the toxin VacA of Helicobacter pylori is required for its biological activity. EMBO J. 1999;18:5517-27.

[26] Galmiche A, Rassow J, Doye A, Cagnol S, Chambard JC, Contamin S, et al. The N-terminal 34 kDa fragment of Helicobacter pylori vacuolating cytotoxin targets mitochondria and induces cytochrome c release. EMBO J. 2000;19:6361-70.

[27] Gebert B, Fischer W, Weiss E, Hoffmann R, Haas R. Helicobacter pylori vacuolating cytotoxin inhibits T lymphocyte activation. Science. 2003;301:1099-102.

[28] Sundrud MS, Torres VJ, Unutmaz D, Cover TL. Inhibition of primary human T cell proliferation by Helicobacter pylori vacuolating toxin (VacA) is independent of VacA effects on IL-2 secretion. Proc Natl Acad Sci U S A. 2004;101:7727-32.

[29] Willhite DC, Blanke SR. Helicobacter pylori vacuolating cytotoxin enters cells, localizes to the mitochondria, and induces mitochondrial membrane permeability changes correlated to toxin channel activity. Cell Microbiol. 2004;6:143-54.

[30] Cover TL, Blanke SR. Helicobacter pylori VacA, a paradigm for toxin multifunctionality. Nat Rev Microbiol. 2005;3:320-32.

[31] Terebiznik MR, Raju D, Vazquez CL, Torbricki K, Kulkarni R, Blanke SR, et al. Effect of Helicobacter pylori's vacuolating cytotoxin on the autophagy pathway in gastric epithelial cells. Autophagy. 2009;5:370-9.

[32] Calore F, Genisset C, Casellato A, Rossato M, Codolo G, Esposti MD, et al. Endosome-mitochondria juxtaposition during apoptosis induced by H. pylori VacA. Cell Death Differ. 2010;17:1707-16.

[33] Domanska G, Motz C, Meinecke M, Harsman A, Papatheodorou P, Reljic B, et al. Helicobacter pylori VacA toxin/subunit p34: targeting of an anion channel to the inner mitochondrial membrane. PLoS Pathog. 2010;6:e1000878.

[34] Radin JN, Gonzalez-Rivera C, Ivie SE, McClain MS, Cover TL. Helicobacter pylori VacA induces programmed necrosis in gastric epithelial cells. Infect Immun. 2011;79:2535-43.

[35] Jain P, Luo ZQ, Blanke SR. Helicobacter pylori vacuolating cytotoxin A (VacA) engages the mitochondrial fission machinery to induce host cell death. Proc Natl Acad Sci U S A. 2011;108:16032-7.

[36] Tombola F, Carlesso C, Szabo I, de Bernard M, Reyrat JM, Telford JL, et al. Helicobacter pylori vacuolating toxin forms anion-selective channels in planar lipid bilayers: possible implications for the mechanism of cellular vacuolation. Biophys J. 1999;76:1401-9.

[37] Czajkowsky DM, Iwamoto H, Cover TL, Shao Z. The vacuolating toxin from Helicobacter pylori forms hexameric pores in lipid bilayers at low pH. Proc Natl Acad Sci U S A. 1999;96:2001-6.

[38] Iwamoto H, Czajkowsky DM, Cover TL, Szabo G, Shao Z. VacA from Helicobacter pylori: a hexameric chloride channel. FEBS Lett. 1999;450:101-4.

[39] Papini E, Satin B, de Bernard M, Molinari M, Arico B, Galli C, et al. Action site and cellular effects of cytotoxin VacA produced by Helicobacter pylori. Folia Microbiol (Praha). 1998;43:279-84.

[40] Debellis L, Papini E, Caroppo R, Montecucco C, Curci S. Helicobacter pylori cytotoxin VacA increases alkaline secretion in gastric epithelial cells. Am J Physiol Gastrointest Liver Physiol. 2001;281:G1440-8.

[41] Campello S, Tombola F, Cabrini G, Zoratti M. The vacuolating toxin of Helicobacter pylori mimicks the CFTR-mediated chloride conductance. FEBS Lett. 2002;532:237-40.

[42] Willhite DC, Cover TL, Blanke SR. Cellular vacuolation and mitochondrial cytochrome c release are independent outcomes of Helicobacter pylori vacuolating cytotoxin activity that are each dependent on membrane channel formation. J Biol Chem. 2003;278:48204-9.

[43] Telford JL, Ghiara P, Dell'Orco M, Comanducci M, Burroni D, Bugnoli M, et al. Gene structure of the Helicobacter pylori cytotoxin and evidence of its key role in gastric disease. J Exp Med. 1994;179:1653-58.

[44] Pagliaccia C, de Bernard M, Lupetti P, Ji X, Burroni D, Cover TL, et al. The m2 form of the Helicobacter pylori cytotoxin has cell type-specific vacuolating activity. Proc Natl Acad Sci U S A. 1998;95:10212-7.

[45] Rhead JL, Letley DP, Mohammadi M, Hussein N, Mohagheghi MA, Eshagh Hosseini M, et al. A new Helicobacter pylori vacuolating cytotoxin determinant, the intermediate region, is associated with gastric cancer. Gastroenterology. 2007;133:926-36.

[46] Gangwer KA, Shaffer CL, Suerbaum S, Lacy DB, Cover TL, Bordenstein SR. Molecular evolution of the Helicobacter pylori vacuolating toxin gene vacA. J Bacteriol. 2010;192:6126-35.

[47] Cover TL. Helicobacter pylori Diversity and Gastric Cancer Risk. mBio. 2016;7:e01869-15.

[48] Figueiredo C, Machado JC, Pharoah P, Seruca R, Sousa S, Carvalho R, et al. Helicobacter pylori and interleukin 1 genotyping: an opportunity to identify high-risk individuals for gastric carcinoma. J Natl Cancer Inst. 2002;94:1680-7.

[49] Memon AA, Hussein NR, Miendje Deyi VY, Burette A, Atherton JC. Vacuolating cytotoxin genotypes are strong markers of gastric cancer and duodenal ulcer-associated Helicobacter pylori strains: a matched case-control study. J Clin Microbiol. 2014;52:2984-9.

[50] Ji X, Fernandez T, Burroni D, Pagliaccia C, Atherton JC, Reyrat JM, et al. Cell specificity of Helicobacter pylori cytotoxin is determined by a short region in the polymorphic midregion. Infect Immun. 2000;68:3754-7.

[51] Letley DP, Atherton JC. Natural diversity in the N terminus of the mature vacuolating cytotoxin of Helicobacter pylori determines cytotoxin activity. J Bacteriol. 2000;182:3278-80.

[52] McClain MS, Cao P, Iwamoto H, Vinion-Dubiel AD, Szabo G, Shao Z, et al. A 12-amino-acid segment, present in type s2 but not type s1 Helicobacter pylori VacA proteins, abolishes cytotoxin activity and alters membrane channel formation. J Bacteriol. 2001;183:6499-508.

[53] Wang WC, Wang HJ, Kuo CH. Two distinctive cell binding patterns by vacuolating toxin fused with glutathione S-transferase: one high-affinity m1-specific binding and the other lower-affinity binding for variant m forms. Biochemistry. 2001;40:11887-96.

[54] Vinion-Dubiel AD, McClain MS, Czajkowsky DM, Iwamoto H, Ye D, Cao P, et al. A dominant negative mutant of Helicobacter pylori vacuolating toxin (VacA) inhibits VacA-induced cell vacuolation. J Biol Chem. 1999;274:37736-42.

[55] Kim S, Chamberlain AK, Bowie JU. Membrane channel structure of Helicobacter pylori vacuolating toxin: role of multiple GXXXG motifs in cylindrical channels. Proc Natl Acad Sci U S A. 2004;101:5988-91.

[56] Pyburn TM, Foegeding NJ, Gonzalez-Rivera C, McDonald NA, Gould KL, Cover TL, et al. Structural organization of membrane-inserted hexamers formed by Helicobacter pylori VacA toxin. Mol Microbiol. 2016;102:22-36.

[57] Yao X, Fan X, Yan N. Cryo-EM analysis of a membrane protein embedded in the liposome. Proc Natl Acad Sci U S A. 2020;117:18497-503.

[58] McClain MS, Iwamoto H, Cao P, Vinion-Dubiel AD, Li Y, Szabo G, et al. Essential role of a GXXXG motif for membrane channel formation by Helicobacter pylori vacuolating toxin. J Biol Chem. 2003;278:12101-8.

[59] Chambers MG, Pyburn TM, Gonzalez-Rivera C, Collier SE, Eli I, Yip CK, et al. Structural analysis of the oligomeric states of Helicobacter pylori VacA toxin. J Mol Biol. 2013;425:524-35.

[60] Su M. goCTF: Geometrically optimized CTF determination for single-particle cryo-EM. J Struct Biol. 2019;205:22-9.

[61] Leunk RD, Johnson PT, David BC, Kraft WG, Morgan DR. Cytotoxic activity in broth-culture filtrates of Campylobacter pylori. J Med Microbiol. 1988;26:93-9.

[62] Burroni D, Lupetti P, Pagliaccia C, Reyrat JM, Dallai R, Rappuoli R, et al. Deletion of the major proteolytic site of the Helicobacter pylori cytotoxin does not influence toxin activity but favors assembly of the toxin into hexameric structures. Infect Immun. 1998;66:5547-50.

[63] Molinari M, Galli C, Norais N, Telford JL, Rappuoli R, Luzio JP, et al. Vacuoles induced by Helicobacter pylori toxin contain both late endosomal and lysosomal markers. J Biol Chem. 1997;272:25339-44.

[64] Autzen HE, Julius D, Cheng Y. Membrane mimetic systems in CryoEM: keeping membrane proteins in their native environment. Curr Opin Struct Biol. 2019;58:259-68.

[65] Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012;9:676-82.

Chapter IV

N-terminal Transmembrane-Helix Epitope Tag for Electron Microscopy and X-Ray Crystallography of Small Membrane Proteins

The work in this chapter is unpublished and in the submission process with authors as listed: McIlwain, B.C.*, Erwin, A.L.*, Davis, A.R., Koff B.B., Bylund, T., Chuang G.-Y., Kwong, P., Ohi, M., Lai, Y.-T., and Stockbridge, R.B. *indicates equal contribution. B.C.M. conceived this work; performed experiments: protein purification, biochemical analysis, protein crystallization, and negative stain grid preparation; analyzed data: crystal structure solution, EM image processing and analysis; prepared figures; wrote the original draft of the manuscript. A.L.E. performed experiments: negative stain grid preparation and EM data collection; analyzed data: EM image processing and analysis; reviewed and edited the manuscript. A.R.D. performed experiments: protein purification and crystal tray setup; analyzed data: crystal structure refinement. B.B.K. performed experiments: protein purification. T.B. performed experiments: bioinformatic analysis. G.-Y.C. performed experiments: bioinformatic analysis. P.D.K. supervised NIH-VRC team members; reviewed and edited the manuscript. M.D.O. supervised EM data collection and image processing; reviewed and edited the manuscript. Y.-T.L. conceived and directed this work; reviewed and edited the manuscript. R.B.S. conceived, directed, and supervised all aspects of this work; wrote the original draft of the manuscript.

Abstract

Structural studies of membrane proteins, especially small membrane proteins, are associated with well-known experimental challenges. Complexation with monoclonal antibody fragments is a common strategy to augment such proteins; however, generating antibody fragments that specifically bind a target protein is not trivial. Here we identify a helical epitope, from the membrane-proximal external region (MPER) of the gp41-transmembrane subunit of the HIV envelope protein, that is recognized by several well-characterized antibodies and that can be fused as a contiguous extension of the N-terminal transmembrane helix of a broad range of membrane proteins. To analyze whether this MPER-epitope tag might aid structural studies of small membrane proteins, we determined an X-ray crystal structure of a membrane protein target that is structurally intractable without crystallization chaperones, the Fluc fluoride channel, fused to the MPER epitope and in complex with antibody. We also demonstrate the utility of this approach for single-particle electron microscopy with Fluc and two additional small membrane proteins that represent different membrane protein folds, AdiC and GlpF. These studies show that the MPER epitope provides a structurally defined, rigid docking site for antibody fragments that is transferable among diverse membrane proteins and can be engineered without prior structural information. Antibodies that bind to the MPER epitope serve as effective crystallization chaperones and electron microscopy fiducial markers for small membrane proteins.

Introduction

Structural analysis of small (<100-kDa) membrane proteins can be challenging. A biochemically tractable protein might not crystallize due to lack of lattice-forming crystal contacts. At the same time, such proteins may be too small and indistinct to be visualized with electron microscopy (EM) and, when visible in vitrified ice, often suffer from low signal-to-noise, leading to misalignment in disordered detergent micelles[1, 2]. A number of strategies have been used to overcome these challenges. One common approach that has been used for decades in X-ray crystallography is the addition of soluble chaperone proteins such as antibody fragments. Antibody fragments have also proven to be useful tools in high-resolution cryo-electron microscopy (cryo-EM)[3-6]. Two such fragments are the single-chain variable-domain fragment (scFv) and fragment antigen-binding (Fab). scFv fragments are composed of a single 25-kDa unit, the variable domain of an antibody joined by a linker; scFvs are often extremely rigid, leading to highly ordered crystals. Fab fragments consist of two 25-kDa units, the constant and variable domains, which are arranged as an open clamshell through two elbow regions. The low-density area at the center of a Fab fragment appears as a hole – a feature particularly useful for particle alignment from EM images of particles in either vitrified ice or negative stain. Additionally, the 50-kDa Fab fragments effectively increase the size of complexed particles, and can overcome problems with preferred particle orientation, reducing anisotropy of datasets by improving the distribution of Euler angles of the particles in single particle cryo-EM analysis[7, 8]. Antibody fragments that bind targets specifically could also be used as localization tags, which are useful for

interpreting low-resolution EM density maps in order to unambiguously localize regions

of the protein[9] and map macromolecule topology[10].

Unfortunately, several non-trivial limitations accompany the use of antibody

fragments for structural studies. Antibodies with binding specificities to a target protein

are generally discovered by immunization of the target protein in small laboratory

animals. The requisite immunization and antibody discovery campaign can take several

months, antibody fragments discovered by this method sometimes lack stability or

biochemical tractability, flexible loops with limited utility for structural structures are often

preferentially recognized, and it can be difficult to generate antibodies against small

membrane proteins, which can be poorly immunogenic. Additional complications arise if

antibodies are desired against a structural target in a particular conformation or a

substrate-occupied state. Identification of "plug-and-play" chaperones or fiducial

markers that can be used for many different protein targets has been a recent focus of

protein engineering[11-13].

A general catch-22 faced by the plug-and-play approach is that chaperone

markers need to be fixed relative to the target protein to be useful as fiducials for

structural determination, but it has not been clear how to fix such chaperone markers in

the absence of structural information. Thus, one cannot generate a useful chaperone

marker by merely appending a tag to the target membrane protein, as the introduced

tag will be structurally flexible relative to the target membrane protein. To harness the

utility of Fab fragments while minimizing their limitations, we hypothesized that

sequence information could be exploited to identify suitable sites for the introduction of

a structurally specific probe. In particular, the location of transmembrane helices can

often be identified by sequence information, and we propose to exploit the N-terminal transmembrane helix as a general acceptor site for a helical epitope tag for structural studies of membrane proteins.

Here, we identify a helical epitope from the membrane-proximal external region (MPER) of the gp41-transmembrane subunit of the HIV envelope glycoprotein and show that this MPER epitope can be genetically fused to unrelated small membrane proteins and bound by MPER-targeting antibody fragments. We demonstrate the utility of this approach for structural studies by both X-ray crystallography and single particle electron microscopy.

Results

An HIV-1-MPER helical epitope, its recognition by monoclonal antibodies, and its suitability as a grafted tag

The MPER of gp41 consists of ~20 amino acids forming two short helical segments (one more distal and the other more proximal to the membrane; Uniprot ID: Q70626). The more proximal helix of the MPER extends relatively perpendicular from the plane of the membrane[14, 15] (Fig. 4.1A, B) and is recognized by multiple broadly neutralizing antibodies isolated from chronically infected patients. Due to the importance of this epitope in structural immunology, several MPER-Fab complexes have been characterized in biochemical and structural detail[16-22] (Fig. 4.1B, C). These antibody fragments bind in various poses, and usually have long, hydrophobic CDR loops to facilitate binding[17]. Although representative antibodies are able to bind to the isolated MPER peptide, in the full biological context, the membrane itself also comprises part of

the epitope[15, 23] and contributes to the binding affinity, which is typically in the low nanomolar range.

To assess the suitability of the most membrane-proximal helix of the MPER as a helical tag, we first assessed small membrane proteins in the PDB for their ability to accept the MPER epitope tag as graft for the N-terminus of the 1st transmembrane helix in order to create a chimera between the epitope tag and membrane protein. We focused on small membrane proteins that lack extracellular domains, since these are particularly difficult to characterize using standard structural techniques. We identified 304 unique structures of membrane proteins smaller than 100 kDa, with at least two transmembrane helices and at least half of residues embedded in the membrane. Of these proteins, we found 89% to be sterically compatible with such an N-terminal helix graft. For those membrane proteins structurally compatible with the MPER epitope tag, we further analyzed whether antibody fragments 10E8 and 4E10 could bind to the tag in one of several registers without clashing with the target protein or the membrane. Our analysis suggests that over half of unique small membrane proteins in the PDB, representing diverse folds, are likely to be structurally compatible with the N-terminal helix tag and its recognition by an MPER-binding antibody fragment (Fig. 4.1D, Table 3 and 4).

Tagged-membrane protein design and epitope positioning

To experimentally assess binding of antibody fragments to a membrane protein hosting the N-terminal MPER epitope tag, we designed such a construct using a fluoride channel from the Fluc family as the membrane protein scaffold (Fig. 4.2A,B). Flucs are

small (30-kDa) structurally characterized membrane proteins involved in bacterial resistance to fluoride ions[24, 25]. Homologues from *Bordetella pertussis* (Fluc-Bpe) and *E. coli* (Fluc-Ec2) have been structurally characterized[25, 26]. The channels are assembled as antiparallel dimers, with each monomer comprised of four transmembrane helices. Based on the known structure of the Fluc channels, we genetically fused the MPER epitope to the N-terminal end of the Fluc-Bpe coding sequence and incubated the resultant fused chimera with the MPER-binding antibody Fab fragment 10E8v4, a variant of the human-isolated 10E8 antibody that has been engineered for greater stability and solubility[27]. Gel filtration chromatography revealed a 2-mL shift in the retention time of the protein fraction, confirming the formation of a stable MPER-Fluc-Bpe/10E8v4 Fab complex (Fig. 4.2C).

The positioning of the epitope sequence along TMH1 is an important experimental consideration, since the face of the epitope, and antibody binding surface, rotates by 100° for each amino acid shift along TMH1. Figure 4.2C shows models of the complexes with the helical epitope shifted N-terminally along TMH1 (n+1, n+2, n+3), assuming a rigid-body rotation/translation of the Fab together with its binding site. To explicitly test the effect of changing the helical register, we expressed and purified these three additional constructs (n+1, n+2, n+3). When purified individually, these MPER-Fluc-Bpe fusion proteins were monodisperse. However, upon complexation with 10E8v4 Fab fragments, the complexes of these variants did not elute as a single peak, indicating that these constructs would not be suitable for further structural work (Fig. 4.2C). Thus, the MPER epitope must be engineered into the proper position with respect to the three-dimensional structure of the membrane protein to facilitate antibody

83

fragment binding, and the correct positioning can be experimentally determined by systematically shifting the register of the fusion.

Structural analysis of MPER-Fluc-Ec2/10E8v4 complex

For structural analysis, we prepared a SEC-purified stable complex between MPER-Fluc-Ec2 and 10E8v4 for vapor diffusion crystallography. For Fluc channels, structure determination required complexation with soluble crystallography chaperones called monobodies[25, 26]. Alone, the channels, which are almost entirely membrane-embedded, do not even form crystals[25]. Strikingly, the MPER-Fluc-Ec2 /10E8v4 complex began to form crystals within 24 h which reached a maximum size in approximately one week. These crystals diffracted to 3.3 Å Bragg spacing. The structure was solved by molecular replacement, with one Fluc-Ec2 dimer and two 10E8v4 Fab fragments per asymmetric unit (Fig. 4.3A, Table 2). Initial maps showed unambiguous Fo-Fc density into which we modelled the MPER peptide (Fig. 4.3A). Notably, MPER-Fluc-Ec2 made no contribution to the crystal lattice, with crystal contacts instead formed exclusively by the 10E8v4 Fab fragments bound to each side of the antiparallel Fluc-Ec2 dimer (Fig. 4.3B).

We wondered whether complexation with the 50-kDa Fab fragments would enable observation of the 30-kDa channels by single particle electron microscopy, a technique often limited by target protein size. MPER-Fluc-Ec2/10E8v4 sample was applied to a carbon-coated grid and negatively stained, as previously described[28]. The raw micrographs of MPER-Fluc-Ec2 bound to 10E8v4 antibody fragments showed elongated particles that measured approximately 180Å in the longest dimension (Fig.

4.3C, left panel). 2-D class averages revealed particles with strong central density flanked by Fab fragment clamshells at each end (Fig. 4.3C, middle panel). The reconstructed 3-D volume closely corresponded to the expected particle shape and size based on the crystal structure (Fig. 4.3C, right panel).

Fab fragment binding to additional MPER-membrane protein fusions.

To test whether this MPER-epitope tag-based approach would be generally useful for studying small membrane proteins, we designed MPER fusions to two additional structurally characterized targets representing different protein folds: AdiC, a bacterial arginine/agmatine exchanger that possesses the LeuT-fold[29], and GlpF, a bacterial glycerol channel that possesses the aquaporin fold[30].

The AdiC protein from *Salmonella enterica* is a biological dimer with ten TMHs per subunit. Each subunit is approximately 47-kDa in size. Based on the crystal structure[29], we modeled the placement of the helical MPER epitope so that Fab fragment binding to this epitope would not clash with the extracellular portions of the transporter (Fig. 4.4A). The MPER-AdiC /10E8v4 complex elutes 0.5 mL earlier than AdiC alone, and SDS-PAGE analysis of this fraction confirmed that both MPER-AdiC and the 10E8v4 Fab were present in this peak (Fig. 4.4A). Further analysis with negative stain electron microscopy revealed particles consistent with a globular membrane protein bound to a Fab fragment; in some of the 2-D class averages the 'hole' between constant and variable domains is readily visible (Fig. 4.4B, for example, top left average).

GlpF is a 28-kDa aquaglyceroporin that permits diffusion of glycerol and some linear polyalcohols in *E. coli*[30]. The channel is composed of six membrane-spanning TMHs and two re-entrant hairpins. GlpF crystallized as a tetramer[30], but biochemical studies have suggested that the oligomer is weakly associated and that monomeric proteins are both frequently observed[31, 32] and functional[33, 34]. As with the MPER-Fluc-Ec2 and MPER-AdiC complexes, the MPER-GlpF /10E8v4 complex elutes from the size exclusion column as a single monodisperse peak that contains both protein components (Fig. 4.4C). Analysis with negative stain electron microscopy showed well-dispersed particles, with characteristic Fab density readily visible in both raw micrographs and 2-D class averages (Fig. 4.4D). The particle size is consistent with a GlpF monomer bound to a single Fab fragment.

Design of a single-chain variable fragment (scFv) for bacterial expression.

Thus far, we have shown that by grafting the gp41 MPER epitope to the N-terminus of small membrane proteins, Fab fragments that recognize this motif can be used for structural studies. To expand the accessibility of the MPER epitope approach, we engineered a scFv based on the 10E8v4 Fab fragment, which could be periplasmically expressed in *E. coli* (Fig. 4.5). As only the variable domain is responsible for antigen binding, the heavy and light chains of the variable domain are concatenated via a $(Gly_3Ser)_4$ linker; the resulting protein could be expressed (Fig. 4.5B) with a yield of ~3 mg/L *E. coli* culture. Purified 10E8v4-scFv was soluble and monodisperse, and bound stably to MPER-Fluc-Ec2 as shown by size exclusion chromatography (Fig. 4.5C). Like the full-length 10E8v4 Fab fragment, the 10E8v4-scFv

promoted crystallization of the Fluc proteins, which do not crystallize in the absence of soluble chaperones[25]. Crystals formed after approximately three days in several conditions from a commercial screening block. Initial hits diffracted to 6.5Å Bragg spacing (Fig. 4.5D); although further optimization would be required to solve the structure of this complex by X-ray crystallography, these results show that the variable domain alone can act as a crystallization chaperone for MPER-membrane protein fusions. In addition, we imaged these samples by negative stain electron microscopy. The particles observed were in agreement with the expected architecture of the MPER-Fluc-Ec2 dimer bound to two molecules of scFv: discrete 3-unit particles (Figure 4.5E) that appeared shorter than the MPER-Fluc-Ec2/10E8v4 Fab particles observed previously (Fig. 4.3C).

Discussion

Here we present a strategy to engineer a high-affinity antibody-binding site to membrane proteins in a structurally specific manner. We have demonstrated the usefulness of this method by complexing antibody fragments to three small (<100-kDa) membrane proteins, each representing a different membrane protein fold. The method described here exploits the opportunity for 'off-the-shelf' antibodies that recognize an epitope that is transferable between diverse membrane proteins.

As with most structural biology approaches, our approach has both limitations and advantages. Since MPER must be appended to the N-terminus, the approach demands that the N-terminal helix be continuous and relatively vertical with respect to the membrane, and precludes the presence of N-terminal soluble domains. On the other

87

hand, this approach is most useful for proteins that are among the most difficult to study using current structural biology techniques: small helical bundles that are almost entirely embedded within the membrane. In the proof-of-principle experiments described here, we were aided in the epitope placement by high resolution structures of the target proteins. For the analysis of novel proteins, additional experimental troubleshooting would be required to determine the proper positioning of the epitope along the native TM1 helix. However, for targets with homology to structurally characterized proteins, or for small membrane proteins with straightforward hydrophobicity profiles, epitope placement can be initially predicted, and then experimentally confirmed by shifting the epitope along the TMH1 sequence and testing binding using gel filtration chromatography.

We also note that for some targets, clashes between the bound antibodies might influence the oligomeric state of the membrane protein. The examples described here exhibit different behaviors in this regard. For Fluc, two Fabs bind to the dimer without conflict. For AdiC, an obligate dimer, only one Fab fragment is accommodated per dimer. (In contrast, the smaller scFv binds with a stoichiometry of two per AdiC dimer, so that each epitope is occupied). And for GlpF, a protein for which the monomeric state is both frequent and functional, Fab binding to the N-terminal epitope favors the monomeric form over the tetrameric form.

The MPER-based strategy for antibody complexation has a number of clear advantages. As a contiguous extension of the N-terminal transmembrane helix of a target protein, the epitope is a structurally defined and rigid docking site for the antibody fragments. Since the antibodies that recognize the MPER epitope have innate affinity

for the membrane, this approach is therefore compatible – and perhaps would be improved – when membrane mimetics like nanodiscs are used. Moreover, the helical epitope is entirely discrete from the target protein and requires no alteration of native protein sequence. Thus, introducing this epitope is a milder maneuver than other strategies that depend on, for example, replacement of one of the helices of the target protein with a transmembrane epitope[11]. The substrate occupancy of the binding site should not influence the binding of the membrane protein to antibody fragments, allowing multiple protein conformations and ion-occupied states to be interrogated using the same target-protein construct and Fab pair.

Second, this strategy leverages the combined knowledge of the robust community performing structural and biochemical studies on MPER-targeting antibody fragments. Multiple, structurally distinct antibodies that bind to the MPER epitope in varying orientations have been identified[16-21] and some of these made available in public repositories. In some cases, the human-isolated antibody fragments have been engineered to increase solubility and biochemical tractability, which are beneficial for structural applications[27]. While we have focused on 10E8v4 here, other known and characterized antibody fragments could be tested for target binding and effect on oligomerization, behavior in vitreous ice, and ability to form stable crystal contacts.

And finally, the approach designed here allows the experimenter to choose between larger, more featureful Fabs or smaller, easily produced scFvs. We have demonstrated the utility of the scFvs for X-ray crystallography; in addition, we anticipate that despite the absence of the 'hole' between the constant and variable domains of the antibody fragment, the scFvs will be useful as an electron microscopy label for

interpreting low-resolution EM density maps as a localization tag[13]. This scFv is

recombinantly expressed in bacteria, significantly reducing the costs associated with

using MPER-binding antibody fragments as crystallization chaperones, cryo-EM fiducial

markers, or localization tags.

In closing, we would also like to note that although this manuscript has focused

on the application of the MPER epitope to structural biology of membrane proteins,

MPER is also an important target for HIV vaccine development. An antigen design that

employs small membrane proteins as scaffolds to display MPER in close proximity to

the membrane might be an effective vaccine strategy to elicit broadly neutralizing

antibodies against HIV.
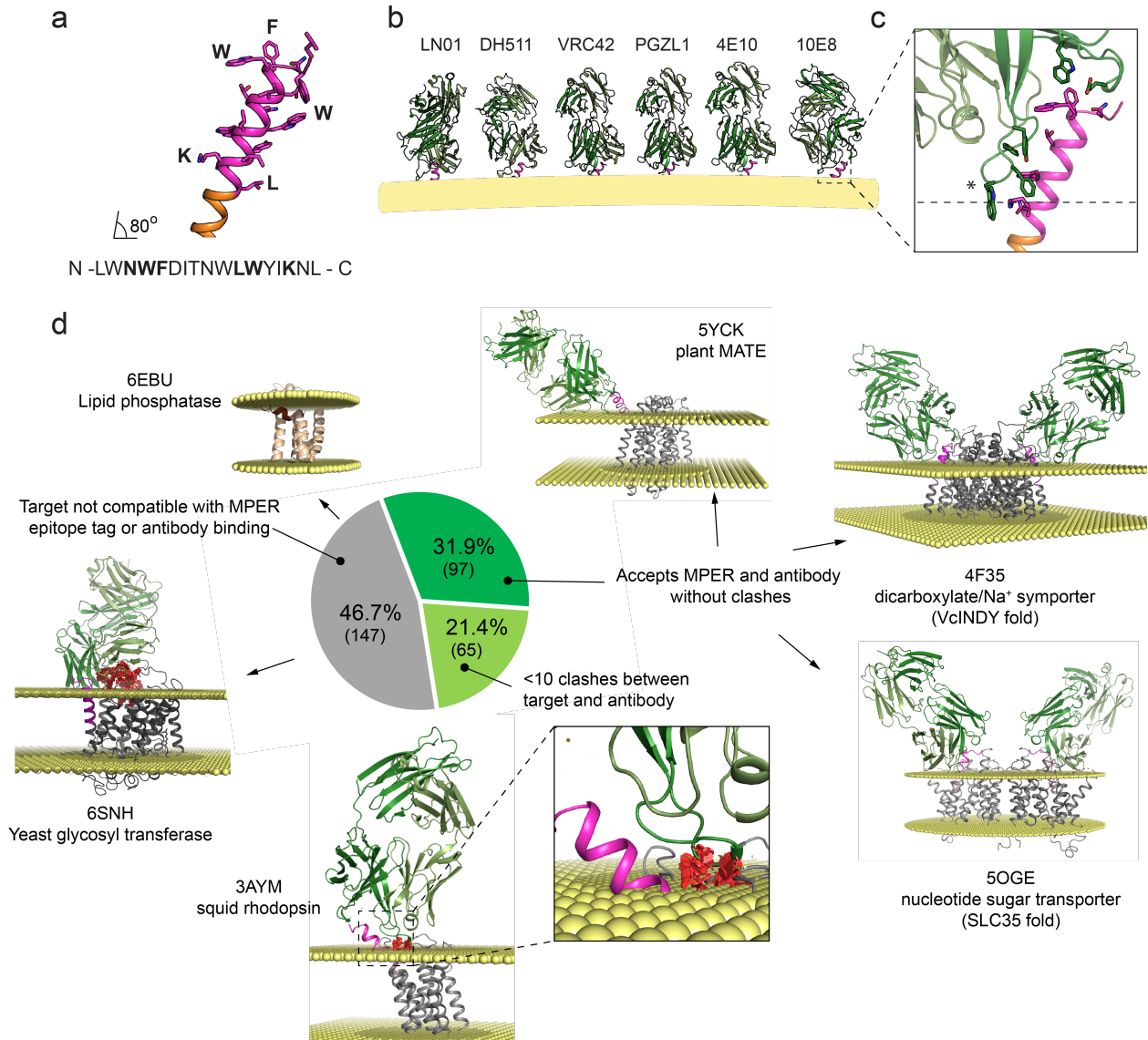
Figure 4.1. The HIV-1 MPER epitope, antibody recognition, and suitability as a structural tag. a) MPER epitope sequence and structure shown relative to the plane of the membrane. Residues critical for binding the antibody 10E8 are shown as sticks, and displayed in bold in the sequence. b) Structures of Fab fragments bound to the MPER epitope (LN01 pdb:6SNC[21]; DH501 pdb:6P3B[16]; 10E8 pdb:5JNY[37]; VRC42 pdb:6MTQ[18]; PGZL1 pdb:6O3G[17]; 4E10 pdb:1TZG[36]). Heavy and light chains shown in shades of green, and MPER in magenta. MPER helices are oriented in the same way for each structure, with the approximate position of the viral membrane shown as a yellow slab. c) View of MPER epitope (magenta) with bound 10E8v4 (green). Residues involved in antibody binding are shown as sticks, and * denotes Trp (W) residue that interacts with membrane lipids. Approximate position of viral membrane shown with dashed line. d) Bioinformatic analysis of MPER fusion and antibody binding to small membrane proteins (see Methods for details). For targets in dataset, pie chart shows the percentage (total number in parentheses) that is (i)

compatible with MPER fusion and 10E8 or 4E10 antibody fragment binding; (ii) compatible with MPER fusion and antibody binding with <10 clashes between target and antibody fragment; (iii) not compatible with this strategy because the target's first TM helix is at an oblique angle to the membrane or bound Fab clashes with target. For each category, example models are shown. MPER, antibody fragments, and membrane colored as in panels (a) and (b). Clashes between target and antibody shown in red. For 6ebu, oblique TM1 is shown in chocolate.
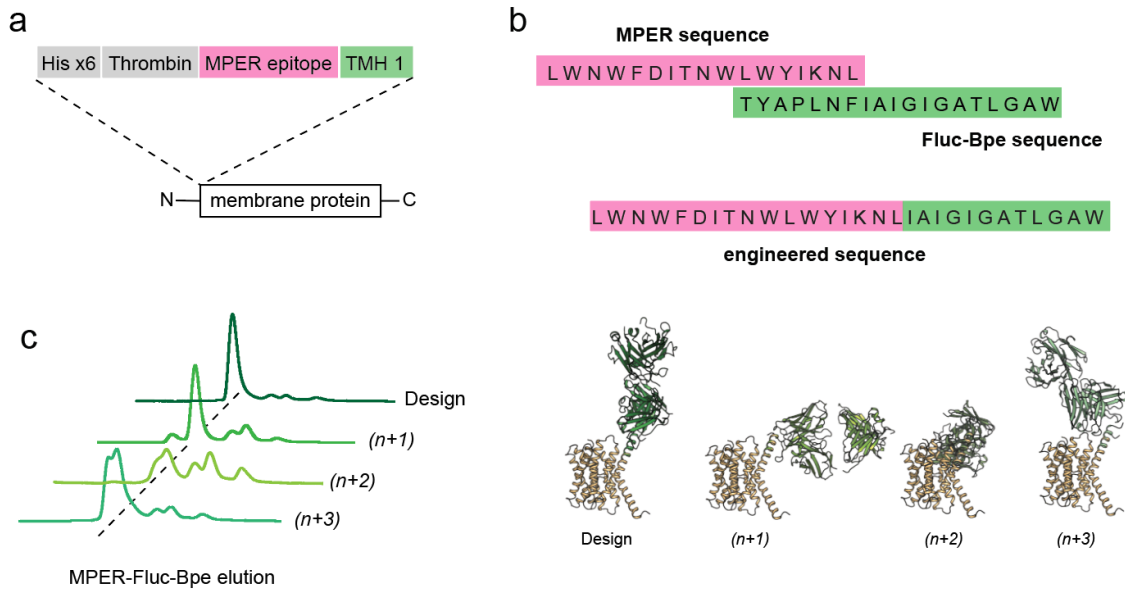
Figure 4.2. Design of Fluc fluoride channel with N-terminal MPER epitope tag. a) Construct design. b) Alignment of MPER sequence (magenta) to Fluc-Bpe TMH1 sequence (green), starting at residue 2 (N-terminal methionine not shown). c) Gel filtration chromatogram of MPER-Fluc-Bpe/10E8v4 complex. Left, comparison of gel filtration profiles of four MPER-Bpe fusion proteins with the position of the MPER epitope shifted N-terminally by the indicated number of amino acids. Right, models showing how the Fluc-Bpe/10E8v4 interface would be expected to change for each construct.
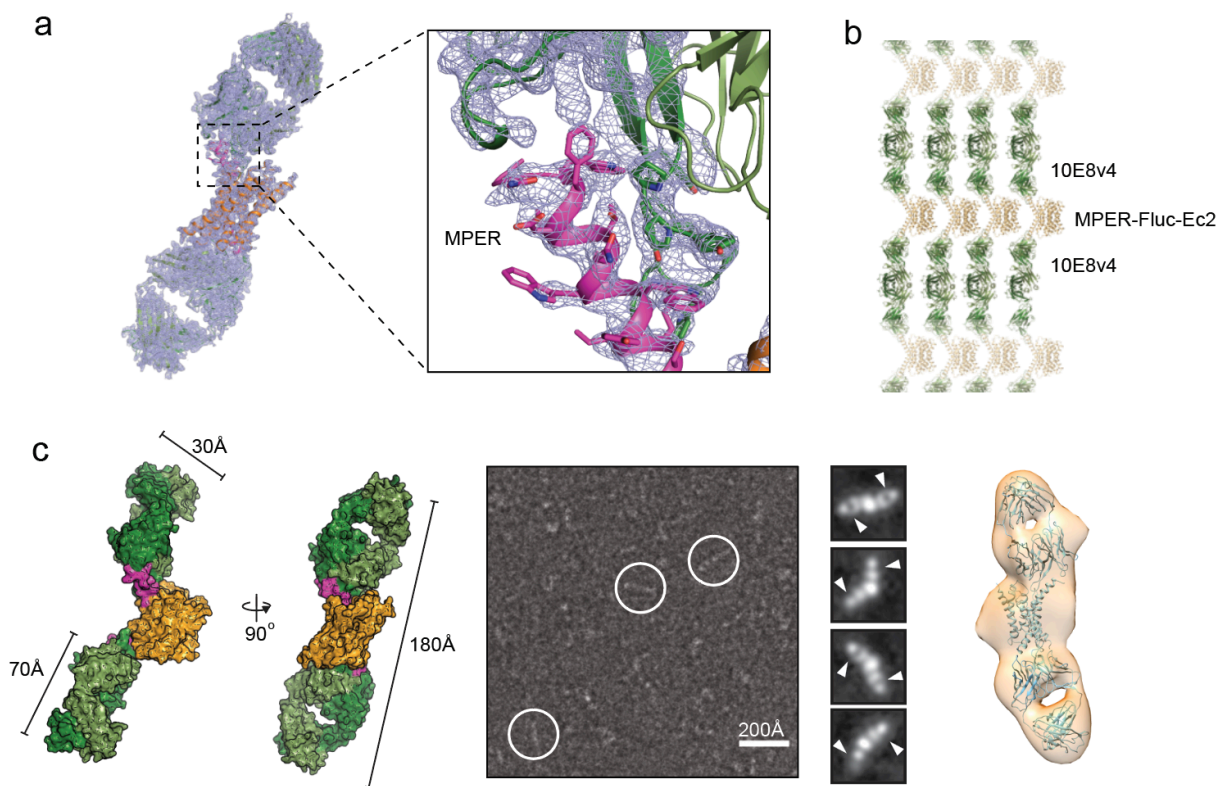
Figure 4.3. Structural characterization of the Fluc-Ec2/10E8v4 complex. a) Model of MPER-Fluc-Ec2/10E8v4 complex with 2Fo-Fc electron density map of contoured at 1.0 σ. MPER colored magenta, Fluc-Ec2 orange, and 10E8v4 green. b) MPER-Fluc-Ec2/10E8v4 crystal lattice. MPER-Fluc-Ec2 in yellow, and 10E8v4 in green. c) Left panel: MPER-Fluc-Ec2/10E8v4 crystal structure shown as surface representation. The dimensions of the expected particles are shown. Middle: Representative negative stain micrograph and 2-D class averages of MPER-Fluc-Ec2/10E8v4 complex. White arrows indicate bound antibody fragments with a characteristic "hole" between the heavy and light domains. Scale bar, 300Å. Box size, 398Å. Right, 3-D reconstruction of particles. Density envelope has been fitted with the crystal structure from (a).
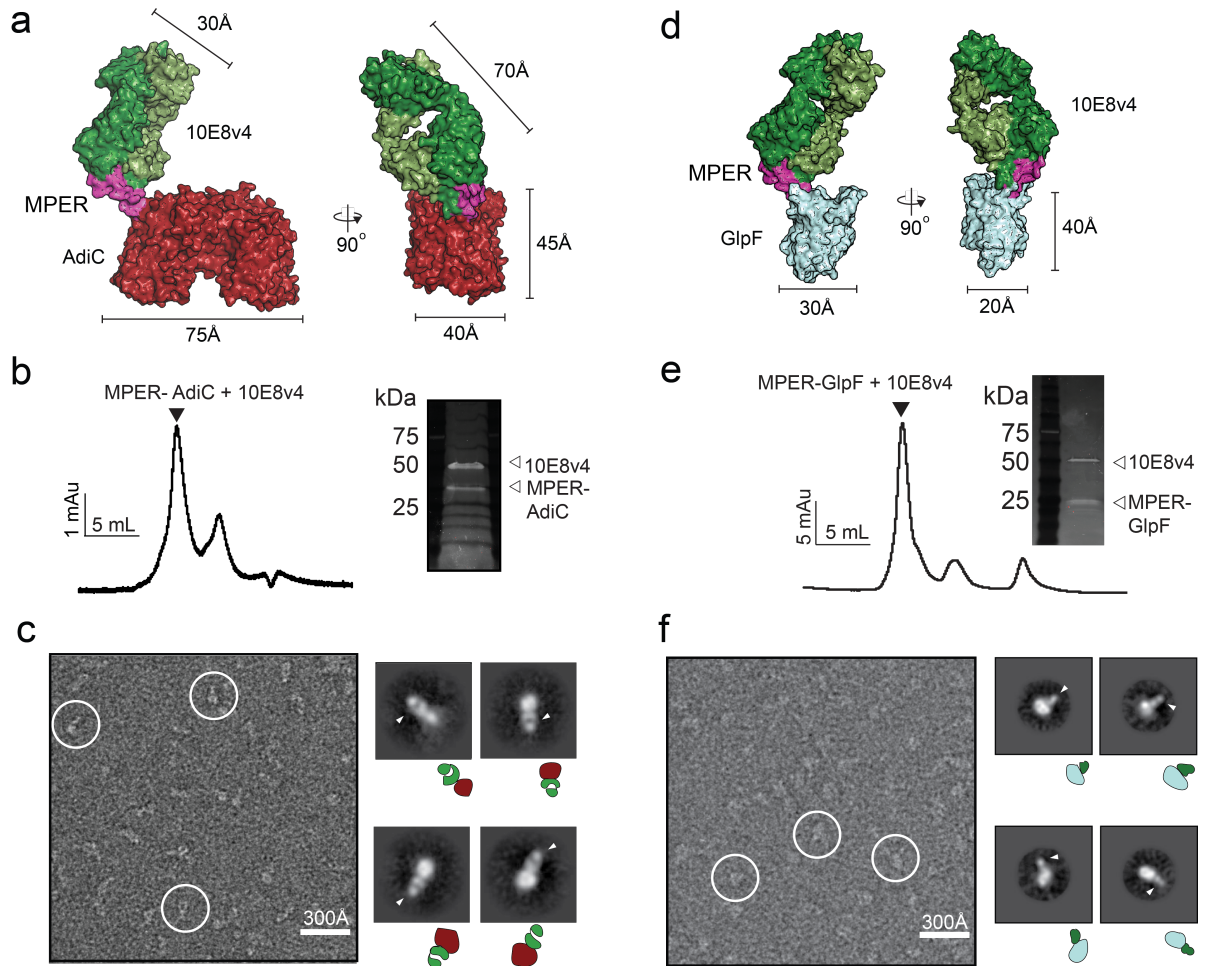
Figure 4.4. 10E8v4 binding to AdiC and GlpF bearing MPER epitope tag. a) Top:
Model of AdiC (dark red) fused with the MPER epitope (magenta) and bound to 10E8v4
Fab (green). Model based on known structures of AdiC (pdb: 3NCY)[29] and 10E8v4.
Protein shown as surface representation, with approximate dimensions shown. Bottom:
Gel filtration chromatogram of AdiC:10E8v4 complex. Right, SDS-PAGE gel of the
indicated fraction with major components labeled. b) Representative negative stain
micrograph and 2-D class averages of MPER-AdiC/10E8v4. Cartoon representation of
the AdiC dimer (dark red) and 10E8v4 Fab (green) shown below each average. White
arrows indicate 10E8v4 Fab. Scale bar, 300Å. Box size, 319Å. c) Model of GlpF
monomer (light blue) fused with the MPER epitope (magenta) and bound to 10E8v4 Fab
(green). Model based on known structures of GlpF (pdb: 1FX8)[30] and 10E8v4. Protein
shown as surface representation, with approximate dimensions shown. Bottom: Gel
filtration chromatogram of the MPER-GlpF/10E8v4 complex. Right, SDS-PAGE gel of
the indicated fraction with major components labeled. d) Representative negative stain
micrograph and 2-D class averages of MPER-GlpF/10E8v4. Cartoon representation of
GlpF (light blue) bound to 10E8v4 Fab (green) shown below each average. White
arrows indicate 10E8v4 Fab. Scale bar, 300Å. Box size, 468Å.
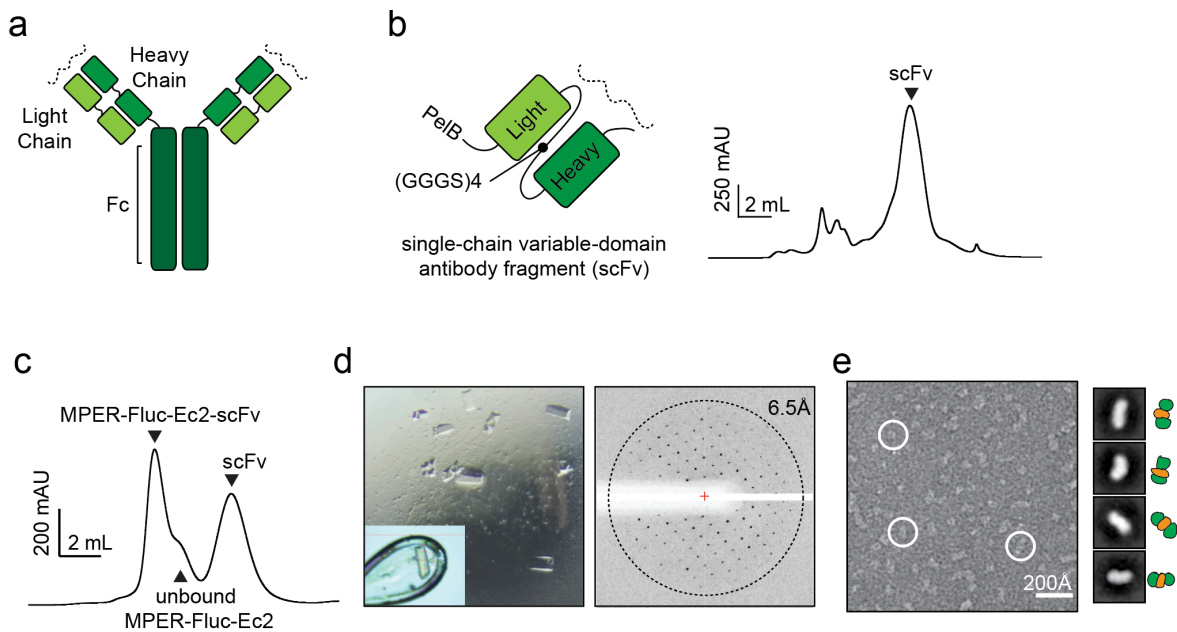
Figure 4.5. Design of a single-chain variable-domain antibody fragment (scFv) based on 10E8v4. a) Left, cartoon of full-length antibody (IgG) with the heavy and light chains that comprise the Fab fragment labeled. Epitope binding surface is represented by dashed line. b) Left, scFv design. Right, gel elution profile of bacterial periplasmic supernatant with scFv peak indicated. c) Gel elution chromatogram of MPER-Fluc-Ec2 incubated with scFv. d) Crystals of MPER-Fluc-Ec2/scFv grown in sitting drop format, and resulting diffraction pattern. e) Representative negative stain micrograph and 2-D class averages of MPER-Fluc-Ec2/scFv complex. Scale bar, 300Å. Box size, 398Å. Cartoons are provided as interpretations of the orientation of each class average.

References

[1] Herzik MA, Wu M, Lander GC. High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM. Nature communications. 2019;10:1-9.

[2] Nygaard R, Kim J, Mancia F. Cryo-electron microscopy analysis of small membrane proteins. Curr Opin Struct Biol. 2020;64:26-33.

[3] Wu S, Avila-Sakar A, Kim J, Booth DS, Greenberg CH, Rossi A, et al. Fabs enable single particle cryoEM studies of small proteins. Structure. 2012;20:582-92.

[4] Kim J, Tan YZ, Wicht KJ, Erramilli SK, Dhingra SK, Okombo J, et al. Structure and drug resistance of the Plasmodium falciparum transporter PfCRT. Nature. 2019;576:315-20.

[5] Bloch JS, Pesciullesi G, Boilevin J, Nosol K, Irobalieva RN, Darbre T, et al. Structure and mechanism of the ER-based glucosyltransferase ALG6. Nature. 2020;579:443-7.

[6] Coleman JA, Yang D, Zhao Z, Wen PC, Yoshioka C, Tajkhorshid E, et al. Serotonin transporter-ibogaine complexes illuminate mechanisms of inhibition and transport. Nature. 2019;569:141-5.

[7] Dang S, Feng S, Tien J, Peters CJ, Bulkley D, Lolicato M, et al. Cryo-EM structures of the TMEM16A calcium-activated chloride channel. Nature. 2017;552:426-9.

[8] Butterwick JA, del Mármol J, Kim KH, Kahlson MA, Rogow JA, Walz T, et al. Cryo-EM structure of the insect olfactory receptor Orco. Nature. 2018;560:447-52.

[9] Gubellini F, Fronzes R. Labeling of Membrane Complexes for Electron Microscopy. Membrane Protein Structure and Function Characterization: Springer; 2017. p. 125-38.

[10] Flemming D, Thierbach K, Stelter P, Böttcher B, Hurt E. Precise mapping of subunits in multiprotein complexes by a versatile electron microscopy label. Nature structural & molecular biology. 2010;17:775-8.

[11] Kim JW, Kim S, Lee H, Cho G, Kim SC, Lee H, et al. Application of antihelix antibodies in protein structure determination. Proceedings of the National Academy of Sciences. 2019;116:17786-91.

[12] Dutka P, Mukherjee S, Gao X, Kang Y, de Waal PW, Wang L, et al. Development of "Plug and Play" Fiducial Marks for Structural Studies of GPCR Signaling Complexes by Single-Particle Cryo-EM. Structure. 2019;27:1862-74. e7.

[13] Yeates TO, Agdanowski MP, Liu Y. Development of imaging scaffolds for cryo-electron microscopy. Curr Opin Struct Biol. 2020;60:142-9.

[14] Fu Q, Shaik MM, Cai Y, Ghantous F, Piai A, Peng H, et al. Structure of the membrane proximal external region of HIV-1 envelope glycoprotein. Proc Natl Acad Sci U S A. 2018;115:E8892-E9.

[15] Irimia A, Serra AM, Sarkar A, Jacak R, Kalyuzhniy O, Sok D, et al. Lipid interactions and angle of approach to the HIV-1 viral membrane of broadly neutralizing antibody 10E8: Insights for vaccine and therapeutic design. PLoS pathogens. 2017;13:e1006212.

[16] Nicely NI, Dennison SM, Spicer L, Scearce RM, Kelsoe G, Ueda Y, et al. Crystal structure of a non-neutralizing antibody to the HIV-1 gp41 membrane-proximal external region. Nature structural & molecular biology. 2010;17:1492.

[17] Zhang L, Irimia A, He L, Landais E, Rantalainen K, Leaman DP, et al. An MPER antibody neutralizes HIV-1 using germline features shared among donors. Nature communications. 2019;10:1-16.

[18] Krebs SJ, Kwon YD, Schramm CA, Law WH, Donofrio G, Zhou KH, et al. Longitudinal analysis reveals early development of three MPER-directed neutralizing antibody lineages from an HIV-1-infected individual. Immunity. 2019;50:677-91. e13.

[19] Williams LD, Ofek G, Schätzle S, McDaniel JR, Lu X, Nicely NI, et al. Potent and broad HIV-neutralizing antibodies in memory B cells and plasma. Science immunology. 2017;2.

[20] Rujas E, Insausti S, García-Porras M, Sánchez-Eugenia R, Tsumoto K, Nieva JL, et al. Functional contacts between MPER and the anti-HIV-1 broadly neutralizing antibody 4E10 extend into the core of the membrane. Journal of molecular biology. 2017;429:1213-26.

[21] Pinto D, Fenwick C, Caillat C, Silacci C, Guseva S, Dehez F, et al. Structural Basis for Broad HIV-1 Neutralization by the MPER-Specific Human Broadly Neutralizing Antibody LN01. Cell host & microbe. 2019;26:623-37. e8.

[22] Huang J, Ofek G, Laub L, Louder MK, Doria-Rose NA, Longo NS, et al. Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. Nature. 2012;491:406-12.

[23] Irimia A, Sarkar A, Stanfield RL, Wilson IA. Crystallographic identification of lipid as an integral component of the epitope of HIV broadly neutralizing antibody 4E10. Immunity. 2016;44:21-31.

[24] Stockbridge RB, Robertson JL, Kolmakova-Partensky L, Miller C. A family of fluoride-specific ion channels with dual-topology architecture. Elife. 2013;2:e01084.

[25] Stockbridge RB, Kolmakova-Partensky L, Shane T, Koide A, Koide S, Miller C, et al. Crystal structures of a double-barrelled fluoride ion channel. Nature. 2015;525:548-51.

[26] McIlwain BC, Newstead S, Stockbridge RB. Cork-in-Bottle Occlusion of Fluoride Ion Channels by Crystallization Chaperones. Structure. 2018;26:635-9 e1.

[27] Kwon YD, Georgiev IS, Ofek G, Zhang B, Asokan M, Bailer RT, et al. Optimization of the Solubility of HIV-1-Neutralizing Antibody 10E8 through Somatic Variation and Structure-Based Design. J Virol. 2016;90:5899-914.

[28] Ohi M, Li Y, Cheng Y, Walz T. Negative staining and image classification—powerful tools in modern electron microscopy. Biological procedures online. 2004;6:23-34.

[29] Fang Y, Jayaram H, Shane T, Kolmakova-Partensky L, Wu F, Williams C, et al. Structure of a prokaryotic virtual proton pump at 3.2 A resolution. Nature. 2009;460:1040-3.

[30] Fu D, Libson A, Miercke LJ, Weitzman C, Nollert P, Krucinski J, et al. Structure of a glycerol-conducting channel and the basis for its selectivity. Science. 2000;290:481-6.

[31] Manley DM, McComb ME, Perreault H, Donald LJ, Duckworth HW, O'Neil JD. Secondary structure and oligomerization of the E. coli glycerol facilitator. Biochemistry. 2000;39:12303-11.

[32] Borgnia MJ, Agre P. Reconstitution and functional comparison of purified GlpF and AqpZ, the glycerol and water channels from Escherichia coli. Proc Natl Acad Sci U S A. 2001;98:2888-93.

[33] Trefz M, Keller R, Vogt M, Schneider D. The GlpF residue Trp219 is part of an amino-acid cluster crucial for aquaglyceroporin oligomerization and function. Biochim Biophys Acta Biomembr. 2018;1860:887-94.

[34] Cymer F, Schneider D. A single glutamate residue controls the oligomerization, function, and stability of the aquaglyceroporin GlpF. Biochemistry. 2010;49:279-86.

Chapter V

Future Directions

Pore forming toxins are united by their dynamic assembly mechanism: secretion as a monomer from the bacterium, binding to target host cells, oligomerizing, and forming a pore in a lipid bilayer[1]. While these steps have been visualized using a number of techniques over the last several decades, X-ray crystallography and cryo-electron microscopy have been crucial for determining the structural basis of oligomerization and pore formation for a wide variety of pore forming toxins[2-15]. Several groups have observed pore forming toxins as soluble oligomers and oligomeric pre-pores, where the transmembrane pore has not completely formed, revealing insights into how pore formation occurs[6, 9, 10, 14, 15]. In this dissertation, I discussed my research aimed at understanding the dynamic assembly mechanism of vacuolating cytotoxin A (VacA), an 88kDa secreted monomer from *H. pylori* that oligomerizes into hexamers and heptamers in a membrane environment to form transmembrane anion channels in gastric epithelial cells[14]. Additionally, I contributed to the characterization of novel methods for the visualization of small membrane proteins with cryo-EM. Determination of small membrane protein structures using single particle cryo-EM is currently a challenge, and fiducial markers like the MPER-epitope tag will be used to make structure determination of these proteins easier by addressing 3-D alignment and preferred orientation issues. Innovative strategies like fiducial markers will be required to address remaining questions

surrounding the dynamic mechanisms of VacA and bacterial pore forming toxins in general to understand how these unique membrane proteins elicit diverse effects on a wide variety of cell types across the animal kingdom. In the following chapter, I discuss several of these remaining questions and indicate future experimental approaches that begin to address them.

Does p88 conformationally change upon oligomerization and membrane binding?

VacA p88 is secreted as a soluble 88 kDa monomer from *H. pylori* that subsequently forms oligomers on the membrane of host cells[16]. While a previous study determined a crystal structure of a majority of the p55 region of VacA (355-811), there was not a full-length high resolution structure of monomeric VacA p88 to enable insights into the conformational changes VacA undergoes from its monomeric state to oligomeric state prior to pore formation[17]. Wild-type VacA p88 assembles into a heterogeneous mixture of oligomer types at neutral pH and for this reason was not amenable for X-ray crystallography studies of WT VacA p88[18-20]. Gonzalez-Rivera et al. used the non-oligomerizing mutant VacA Δ346-347 to determine a 4.2Å crystal structure of monomeric VacA[21]. This crystal structure contained density for the p55 region that was determined previously by X-ray crystallography (355-811) and additional β-helical density extending from the p55 region that contributes to part of the p33 region of VacA[17]. Placing their model into a 15Å negative stain EM map of the wild-type VacA hexamer showed p55 localized to the peripheral arms of the hexamer and additional β-helical density extending from the p55 region localizing closer to the center of the hexamer. It was uncertain whether the VacA Δ346-347 p88 monomer model undergoes conformational changes

101

during the process of oligomerization due to the low-resolution nature of the negative stain EM map, which was the only 3D information for VacA known at the time of publication. Docking the Δ346-347 crystal structure into recently published high-resolution cryo-EM VacA hexamer structures revealed that the Δ346-347 p88 monomer structure lacked density for the GXXXG motif region (residues 1-26), alpha helical density upstream of the GXXXG motif region (residues 27-45), β-strands extending into the p33 region (residues 46-109), predicted flexible loop (residues 293-334), and p55 β-strand extension (residues 335-345) that interact with a p33 β-strand (residues 47-75) from the neighboring protomer in the oligomer structure[14, 15, 21]. It is likely that residues involved in oligomerization are not stabilized in the p88 monomeric form (residues 47-75, 335-345) or in the absence of membrane (residues 1-26) and thus unable to be visualized in the Δ346-347 crystal structure.

To further define conformational changes VacA undergoes during oligomerization, we aimed to use single-particle cryo-EM to determine a 3D structure of a full-length VacA p88 monomer in a closer to native state compared to crystallography. First, we plunged WT VacA acid-activated to pH 3.0 and analyzed the particles using 2D classification (Fig. 5.1A,B). All particles appeared to be *en face*, situated in the vitreous ice with the longest dimension of the monomer parallel to the plane of the ice (Fig. 5.1B). 2D class averages of *en face* views had crisp features and demonstrated the characteristic hollow nature of β-helical core observed in the 2.4Å p55 crystal structure (PDB: 2QV3), 4.2 Å Δ346-347 p88 monomer crystal structure, and VacA hexamer EM density maps (PDB-6ODY/EMD-20024; PDB-6NYF/EMD-0542)[14, 15, 17, 21]. Neighboring protomers were observed in 2-D class averages of acid-activated VacA particles, suggesting that VacA is able to

engage in weak dimeric interactions under acidic conditions. However, this limited further 3-D structure determination of full-length monomeric VacA p88 with the acid-activated VacA sample.

We next analyzed the Δ346-347 VacA non-oligomerizing mutant with single particle cryo-EM at neutral pH, as this mutant would result in a completely monomeric sample compared to acid-activated WT VacA. Compared to the 4.5Å crystal structure of VacA Δ346-347, cryo-EM analysis may enable structure determination of dynamic monomeric states of VacA Δ346-347 in solution that may not have been crystallizable. 2-D analysis of the Δ346-347 VacA mutant at pH 7 revealed class averages of VacA monomers, with no neighboring protomers (Fig. 5.1C). While there were not 2-D class averages representing views of the longest dimension perpendicular in the ice, we hoped that we had enough p88 monomer particles (with the longest dimension parallel to the plane of the ice) at all rotations around the x-axis. A 3D reconstruction of the 2-D class averages resulted in an 8Å EM density map. The particle orientation map indicated strong preferred orientation that we had suspected from the class averages. Overlaying this Δ346-347 EM density map onto the VacA hexamer EM density map revealed no major conformational changes in VacA p88 from the soluble monomer to soluble oligomer state (Fig. 5.1D)[14]. Overcoming the preferred orientation of this sample and improving the resolution will enable definition of local conformational changes that may occur in regions that contribute to the oligomeric interface. Further structural studies of VacA monomers in the presence of detergent or lipids may enable more insight into which regions of VacA p88 insert into membrane prior to VacA oligomer pore formation as has been suggested by biochemical insertion assays of VacA Δ6-27 and VacA Δ346-347[22].

Which residues comprise the VacA oligomeric interface in solution and in membrane?

The <4Å cryo-EM structures of soluble VacA hexamers determined that VacA p88 protomers oligomerize with neighboring protomers through salt bridges (K47:E338, K55:D346), side chain hydrogen bonds (R50:T342, K75:Q343, K55, D346), and several main-chain hydrogen bonds linking residues from the p33 and p55 regions[14, 15]. Additionally, mass spectrometry of VacA crosslinked with EDC (0Å linker) and BS3 (11.4Å linker) to understand the proximity of residues within VacA oligomers resulted in several crosslinks between residues in the p33 and p55 regions, including K44 and E338 (unpublished results from the Cover lab). To investigate the importance of these residues for VacA oligomerization in solution, we introduced cysteine mutations to induce disulfide bond formation between the salt bridge residues determined from the soluble hexamer cryo-EM structure (K47:E338, K55:D346) and the residue pair identified from mass spectrometry (K44:E338)[15]. Negative stain EM analysis of the resulting disulfide mutants revealed that while VacA wild-type oligomers are oligomers at pH 8 and dissociate at pH 3, VacA 44C-338C, 47C-338C, and 55C-346C remain oligomerized at pH 3 (Fig. 5.2, 5.3). This shows that introducing a disulfide bond at these locations prevents VacA from dissociating. Further, when these samples are treated with DTT at pH 8 the oligomers remain associated, while at pH 3 in the presence of reducing agent DTT, these mutants dissociate, indicating that 1) interactions in addition to the one disulfide bond are responsible for VacA oligomerization (and these other interactions are pH dependent), and 2) each of these disulfide bonds alone is sufficient to hold protomers together under dissociating conditions (to varying degrees).

Further investigation of the residues involved in oligomerization could be conducted by generating alanine and charge-reversal point mutations for each of the residues pinpointed in the soluble VacA hexamer structure to be important for oligomerization in solution (alone and in combination) followed by analyses with SDS-PAGE gel separation and negative stain EM. Additionally, alanine and charge-reversal point mutations could be made for all of the interactions identified to be important for VacA double layer formation in solution. These point mutations could be used to generate soluble VacA hexamers and heptamers and test their capacity to bind to membrane at neutral pH and subsequently form a functional pore. The temporal sequence of VacA membrane binding and oligomerization prior to pore formation remains unknown and dodecamer interaction residue point mutations (or VacA disulfide mutants at pH 3) could be used to help define this. It is possible that the VacA dodecamer interacting residues are critical for membrane binding and insertion, and cell vacuolation assays or VacA insertion assays (sodium carbonate treatment of VacA-bound liposome samples, followed by analysis of the sample fractionated with a discontinuous sucrose gradient) to assess the effect of these point mutations (at pH 7) on membrane insertion could be conducted to determine this. Additionally, point mutations for the double layer interaction residues may prove useful for further biochemical and structural experiments, as double layer formation can decrease the amount of VacA that binds to membrane and result in heterogeneous samples for cryo-EM analysis.

While VacA Δ346-347 results in VacA p88 that is unable to oligomerize, examination of VacA mutants with the deletion of residues Δ334-343, a stretch of residues comprising a portion of the protruding p55 segment from one protomer that

105

forms β-strand contacts with the neighboring protomer and contains critical soluble hexamer interaction residue E338, does not affect the ability of VacA to oligomerize into hexamers, heptamers, dodecamers, and tetradecamers (Fig. 5.5B,C). This observation, in combination with the negative stain EM results that show VacA 44C-338C, 47C-338C, 55C-346C disulfide mutants treated with DTT remain associated at pH 8 and dissociate at pH 3, indicate that multiple pH-dependent interactions are required for VacA oligomerization.

Intriguingly, while most of the disulfide mutants and the Δ334-343 mutant form hexamers, heptamers, dodecamers, and tetradecamers in solution at pH 8, 2-D class averages of ~24,000 VacA 44C-338C particles show only hexamer and dodecamer classes at pH 8 (Fig 5.3B). The negative stain analysis of VacA 44C 338C (identified from mass spectrometry rather than the soluble VacA hexamer cryo-EM structure) indicates that this residue pair may be responsible for hexamer vs. heptamer formation in solution. It remains unclear which oligomer type forms anion-selective channels in gastric epithelial cells and whether the different stoichiometries have functional significance. Identification of the structural and cellular environment requirements for hexamers and heptamers, future experiments with mutants that only form hexamers or heptamers, and in situ cryo-EM visualization of VacA are required to understand the functional significance of VacA hexamers and heptamers.

What is the structure and function of the VacA pre-pore state?

While we have further improved our understanding of VacA oligomerization in solution, the structural basis of oligomerization of VacA in a membrane environment

remains to be confirmed. In Chapter 3, we determined an intermediate resolution structure of VacA in a pre-pore state in the context of membrane. Improvement of the resolution is required to determine the interactions critical for protomer association and pre-pore formation when VacA is bound to membrane. Given that over 100,000 untilted particles contributed to the 3.2Å and 3.8Å soluble VacA hexamer cryo-EM structures, incorporation of 4 times more 40˚-tilted data (80,000 particles) is likely required to improve the resolution of this structure to define the regions contributing to the VacA pre-pore state and investigate the role of the GXXXG motif region (1-45) and dynamic flexible loop (300-334) in VacA pore formation. Other strategies that may be useful to consider for improving the resolution include combatting preferred orientation through addition of detergents, changing the grid-type, or use of blot-free plunging to minimize air-water interface interactions that affect particle behavior[23]. Another strategy to consider is truncating VacA p88 to decrease the longest dimension and number of hydrophobic loops, perhaps enabling a wider distribution of particle views in vitrified ice. Additionally, the use of a tag or antibody fragment, isolated from patient serum or phage display, could be employed to alter VacA particle behavior in vitrified ice[24, 25].

Structural and functional analysis of VacA deletion mutants could be used to further understand the role of the GXXXG motifs and flexible loop in VacA pore formation. Preliminary 2-D class averages of 30˚-tilted VacA ∆6-27 solubilized from large unilamellar vesicles show oligomers that appear to lack the strong central density observed in VacA wild-type 2-D class averages (Fig. 5.4). Addition of more data and subsequent 3-D analysis is required to confirm that the GXXXG motifs contribute to the pre-pore state central density. Another deletion mutant that may be interesting to investigate is VacA

107

Δ294-331, which has the predicted flexible loop region at the p33/p55 junction deleted. Negative stain EM analysis shows that this mutant is able to oligomerize (Fig. 5.5A). Further analysis is required to determine whether this VacA mutant is able to conduct chloride across a planar lipid bilayer and understand the contribution of this loop to VacA pore formation.

What is required for VacA pore formation?

How VacA pore formation is triggered *in vivo* once VacA oligomers form on lipid bilayers of the plasma membrane or late endosome remains to be determined. While we tested whether various detergents and conditions could induce VacA pore formation, perhaps a major factor required to induce pore formation is a pH 5.5 environment similar to what may be found within the late endosome[26]. There are several examples of endocytosed bacterial toxins that undergo pH-induced conformational changes[27-29]. It would be interesting to test whether VacA incubated with LUVs at pH 5.5 and subsequently extracted with DDM results in a VacA pore formed state. Planar lipid bilayer assays conducted at pH 4 showed that VacA can form anion selective pores, while at pH 7, no current was detected[30]. Future electrophysiology analysis of VacA channel-forming activity could be conducted with VacA wild-type and mutants at varying pH values to understand the regions of VacA that may be important for pH-dependent pore formation.

Specific lipid interactions with VacA may be required for VacA pore formation. Preliminary 2D class averages suggest that VacA partially inserts into the lipid bilayer of SUVs. As shown in recent papers, thorough 2D classification and 3D analysis of these noisy, low contrast liposome-embedded particles could be used to shed insight on key

lipid interactions [10, 31]. Another strategy is to use nanodiscs that can accommodate the pore-forming region of VacA. Unpublished results from our lab indicate that VacA does not bind efficiently to MSP2N2 nanodiscs (15-16.5nm in diameter)[32]. Thus, next steps include binding VacA to large circularized nanodiscs (cNW50)[33]. Our preliminary data suggest that VacA p88 can bind cNW50 at pH 3 and pH 7, but only form oligomers on cNW50 nanodiscs at pH 7 (Fig. 5.6). VacA binding to cNW50 nanodiscs must optimized to determine a VacA pore formed state using single particle cryo-EM. For a more native approach, VacA transmembrane pores could be isolated from enlarged late endosomes from cultured gastric epithelial cells treated with VacA to analyze VacA inserted into the late endosome membrane with cryo-electron tomography.

VacA may need to bind a cell surface component or receptor to trigger pore formation. It has been proposed that VacA is able to bind sphingomyelin and several epithelial cell receptors, including receptor protein tyrosine phosphatases (RPTP) α and β, epidermal growth factor receptor (EGFR), heparan sulfate, and low-density lipoprotein receptor-related protein 1 (LRP1)[34-41]. We did a pilot experiment to test whether gamma-secretase, a protein shown to be involved in the membrane insertion of HPV capsid protein, is involved in VacA membrane insertion [42]. Initial experiments showed that treatment of VacA treated AGS cells with the gamma-secretase inhibitor XX1 did not prevent vacuolation, a readout for VacA membrane insertion and pore forming activity (Fig. 5.7). To more broadly screen for receptors important for VacA pore formation, one approach would be to extract VacA from gastric epithelial cell membranes (plasma membrane and enlarged late endosome membranes) using styrene copolymers and analyze the structure of the VacA particles[43]. Mass spectrometry could be used to

identify the receptors that associate with VacA extracted from the gastric epithelial cell plasma membrane and enlarged late endosome membranes.

What are future directions of cryoEM to visualize membrane proteins?

Structural analysis of membrane proteins can be challenging, as observed with the efforts to determine cryo-EM structures of ~528 kDa VacA hexamer in solution and solubilized with detergent as a membrane mimetic (Chapters 2 and 3)[14]. The structural analysis of <100 kDa membrane proteins with single particle cryo-EM can be even more difficult due to the low signal-to-noise data which can lead to poor 2-D and 3-D alignment of particles[44]. Antibody fragments have been used to enable structural analysis of membrane proteins with cryo-EM by increasing the size of particles, aiding in 2-D and 3-D alignment of particles, and overcoming preferred particle orientation[45-50]. However, the discovery of antibody fragments suitable for structural studies of the protein of interest can be difficult and time consuming. In Chapter 4, we identified and characterized the helical MPER-epitope tag as a strategy to complex small membrane proteins with existing, easy to produce MPER-targeting antibody fragments for structural analysis with X-ray crystallography and single-particle electron microscopy. 50 kDa 10E8 antibody fragments bound to MPER-tagged small membrane proteins (for example, the 30 kDa *E. coli* fluoride channel MPER-Ec2-FluC) were clearly visualized with negative stain EM, demonstrating the utility of this method for studying membrane-embedded helical bundles (Fig. 5.8A).

While this technique has the potential to facilitate structure determination of small membrane proteins, preliminary cryo-EM experiments with *E. coli* fluoride channel FluC and *S. enterica* arginine/agmatine exchanger AdiC underscored the challenge of imaging

small membrane proteins in the presence of detergents. Cryo-EM micrographs and 2-D class averages of MPER-Ec2-FluC complexed with 10E8 fab fragment in the presence of 2mM DDM revealed particles that were ~60Å x ~60Å (Fig. 5.8B). Cryo-EM micrographs and 2-D class averages of MPER-AdiC complexed with 10E8 fab fragment in the presence of 4mM DM revealed a mixture of particles that were 40Å x 50Å and ~25Å x 80Å (Fig. 5.8C). It is not clear from these imaging sessions whether the classified particles are detergent micelles, dissociated 10E8 fab fragments, or MPER-Ec2-FluC:10E8/MPER-AdiC:10E8 complexes. DDM and DM micelle particles have been measured with SAXS to be ~20Å x 36Å and 18Å x 30Å, respectively[51]. 10E8 antibody fragments (~70Å) were not discernable in raw cryo-EM micrographs or 2-D class averages of MPER-Ec2-FluC:10E8 and MPER-AdiC-10E8 (Fig 5.8B,C). It is possible that these are detergent micelles (alone and overlapping), given that 1 and 2 fab bound MPER-Ec2-FluC:10E8 measures 110Å and 180Å in the longest dimension, and MPER-AdiC:10E8 measures ~100Å in the longest dimension. Another possibility is that these are dissociated or denatured complexes. If this is the case, rapid-plunge freezing methods, crosslinking, or addition of detergent before freezing could be used to prevent air-water interface issues[44].

The similarity in the size of detergent micelles and small membrane proteins can make it a challenge to determine structures of small membrane proteins with cryo-EM or know how to optimize grid preparation. Future use of the MPER-epitope tag for small membrane protein characterization with cryo-EM will require using close to CMC detergent concentration and obtaining thin vitreous ice using blot-free methods[23]. It will be useful to collect images of a buffer only control, antibody fragment alone, and complex

at different concentrations for comparison to understand the best way forward for small membrane protein structure determination with this technique. In addition to innovations and sample-specific optimization of antibody fragment technology, improving the ease of stabilizing membrane proteins in the absence of detergent (i.e. nanodiscs) and determining small membrane protein structures incorporated into liposomes will be important for investigating the molecular mechanisms of membrane proteins in a native environment and open up the possibility of studying the effect of pH or ionic gradients on membrane protein structure[31, 52].

Open questions remaining

The target and function of VacA remain unclear. Determination of the membrane site (i.e. plasma membrane, late endosome, mitochondria) at which VacA forms transmembrane pores will be critical for understanding the requirements for pore formation (specific receptor, pH, substrate requirements). One possibility is that VacA is secreted directly into endosomes containing internalized *H. pylori*. This may provide the right environment and VacA concentration with other *H. pylori* factors present to induce VacA pore formation and downstream cellular effects of VacA that aid *H. pylori* colonization of the human gastric epithelium. While VacA is able to conduct chloride across planar lipid bilayers, structural information about the pore forming region is required to illuminate how VacA is able to transport chloride and/or other ions or molecules. Future experiments aimed at determining the structure and possible cofactors of the VacA transmembrane pore extracted from a more native cell membrane environment will improve our understanding of VacA activity and *H. pylori* pathogenesis.
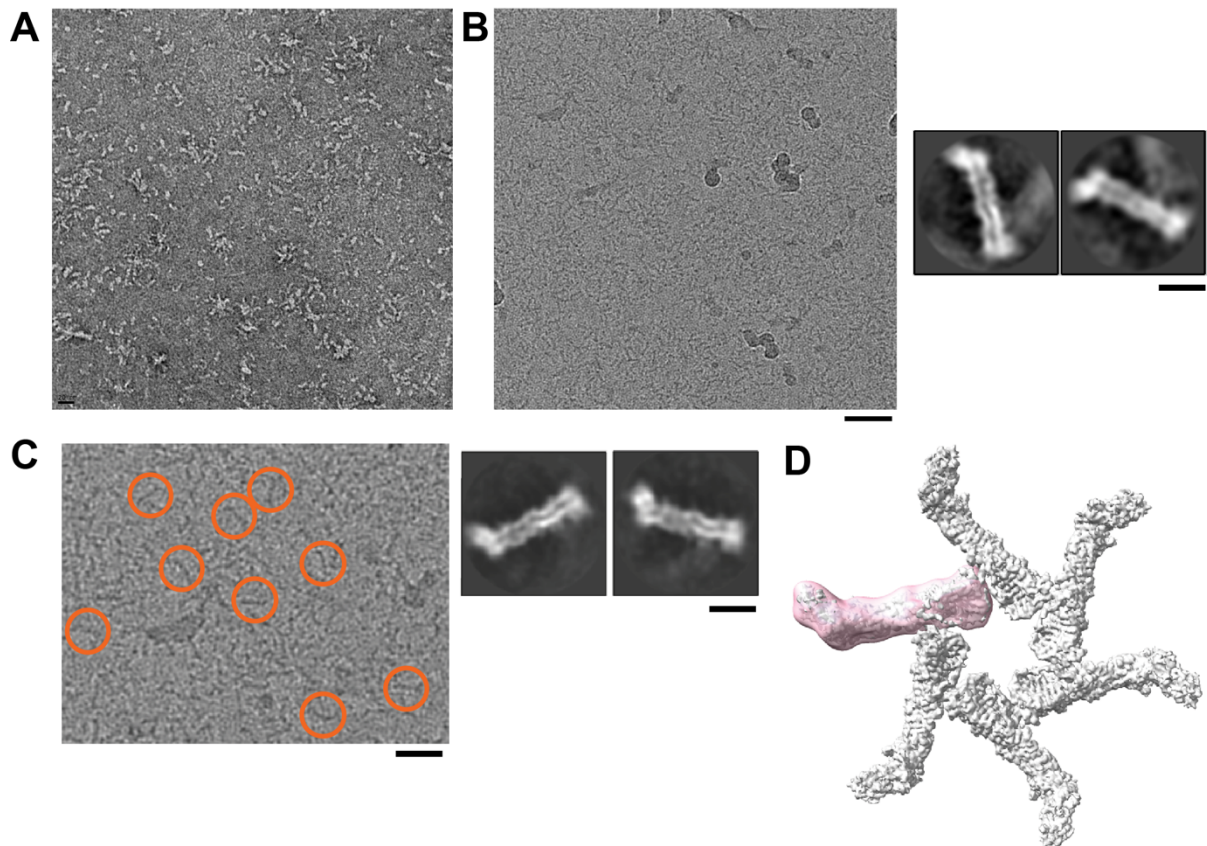
Figure 5.1. Cryo-EM analysis of VacA p88 monomers. A) Negative stain EM image of acid-activated VacA. Scale bar, 20nm. B) Left, representative micrograph of acid-activated VacA in vitreous ice. Scale bar, 40nm. Right, 2-D class averages of acid-activated VacA particles. Scale bar, 5nm. C) Left, representative micrograph of Δ346-347 VacA at pH 7. Example particles circled. Scale bar, 30nm. Right, 2-D class averages of Δ346-347 VacA at pH 7. Scale bar, 5nm. D) 3D reconstruction of Δ346-347 VacA particles shown in pink placed into VacA hexamer map EMD-20024.
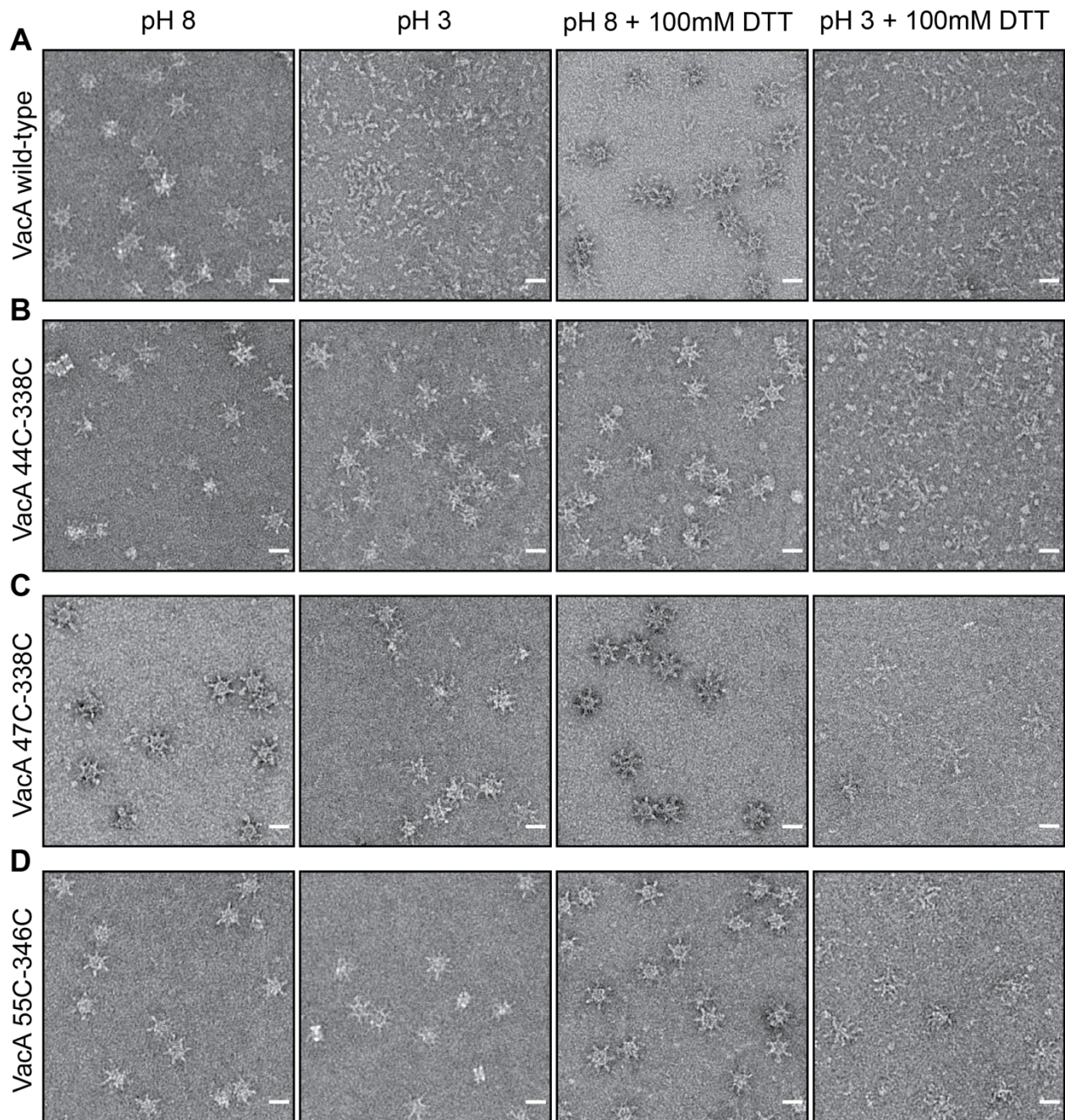
Figure 5.2. Negative stain EM of VacA disulfide mutant reveal that the oligomers remain intact at low pH and upon 100mM DTT treatment at neutral pH.
Negative stain EM micrographs of A) VacA wild-type, B) VacA 44C-338C, C) VacA 47C-338C, and D) VacA 55C-346C at pH 8, pH 3, pH 8 + 100mM DTT, pH 3 + 100mM DTT (left to right). Scale bars, 200Å.
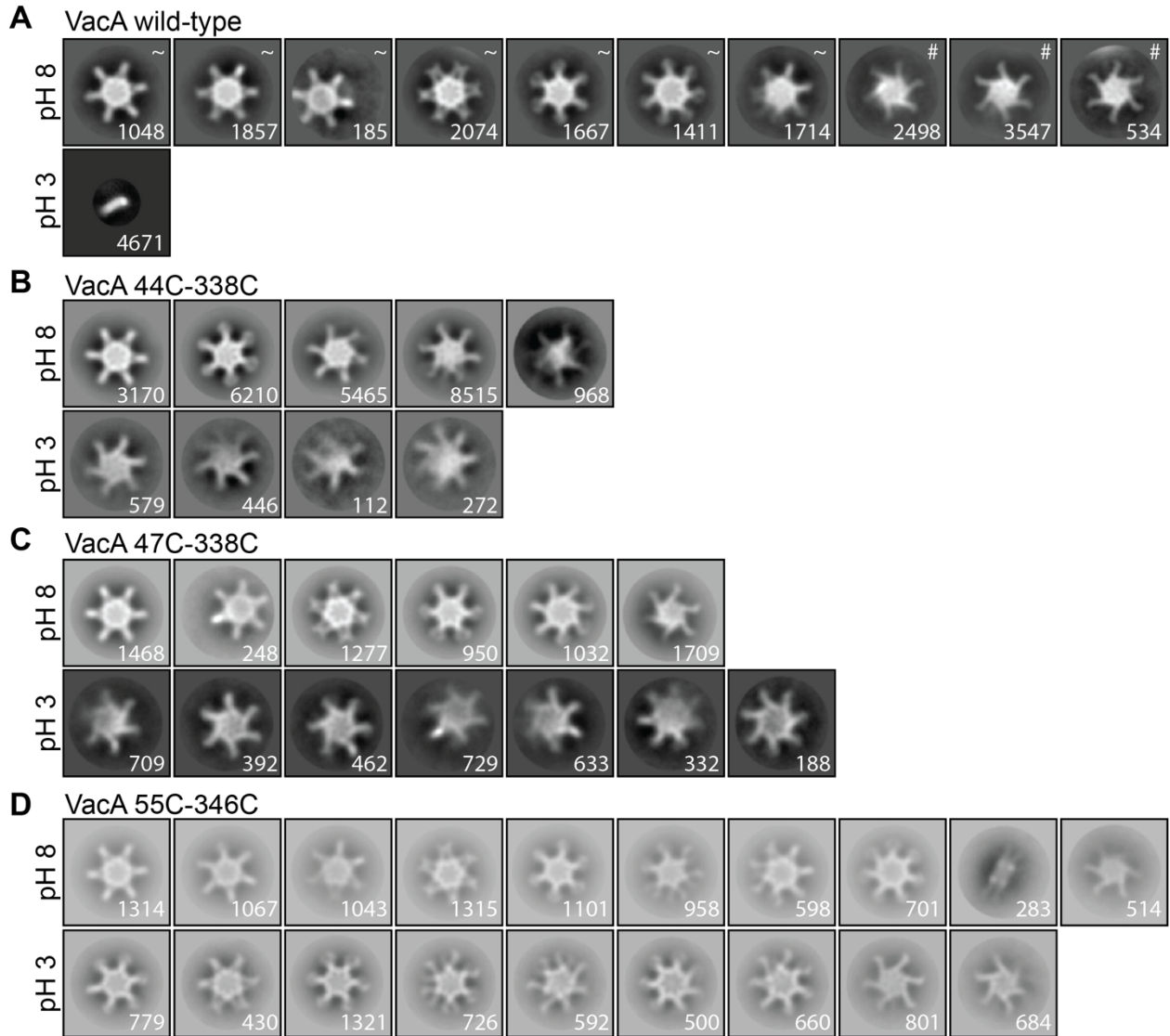
Figure 5.3. Characterization of negatively stained VacA disulfide mutant oligomers. Representative class averages of A) VacA wild-type (~, double layer VacA; # single layer VacA), B) VacA 44C-338C, c) VacA 47C-338C, and D) VacA 55C-346C at neutral and low pH. Number of particles is shown in bottom right corner of each class. Box size, 468Å.

A
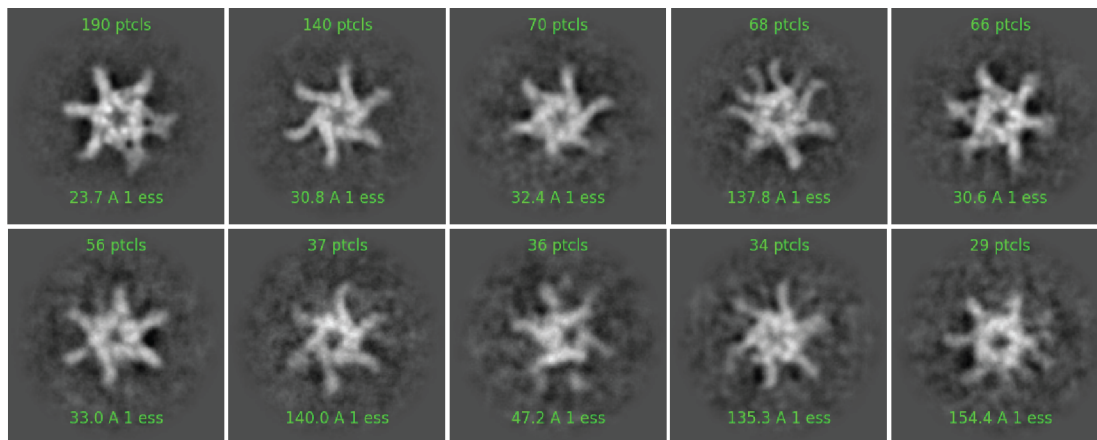
Figure 5.4. Preliminary cryo-EM analysis of 30˚-tilted VacA Δ6-27 DDM micelle oligomers. A) 2-D class averages of VacA Δ6-27 solubilized from large unilamellar vesicles (55/15/30 mol % eggPC/DOPS/Cholesterol) with n-Dodecyl-ß-D-Maltoside (DDM) detergent. Number of particles shown at top of each class. Box size, 588Å.

A VacA Δ294-331

B VacA Δ334-343

C VacA Δ334-343
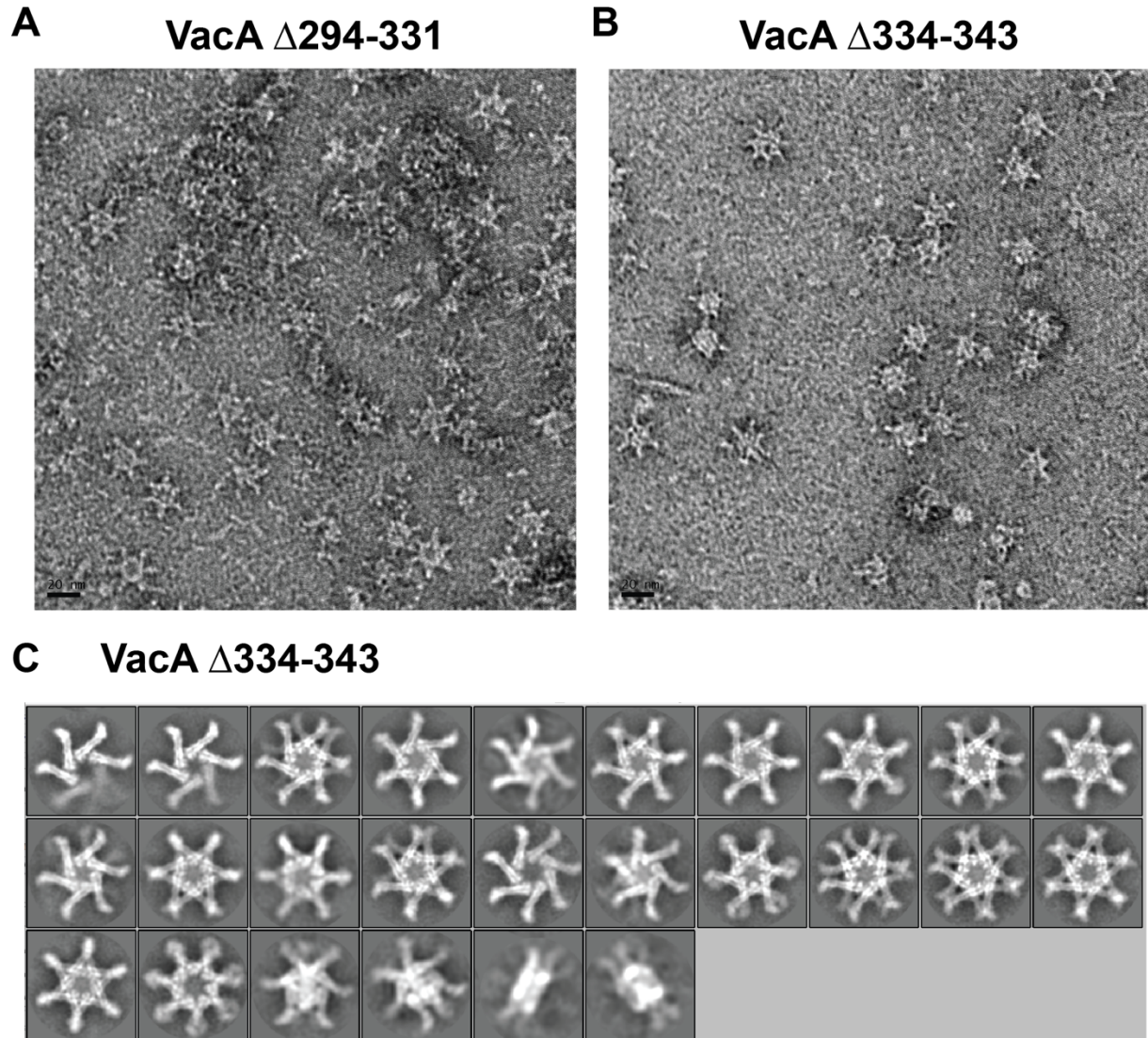
Figure 5.5. EM analysis of VacA oligomers with deletions in the oligomeric interface and flexible loop region able to form multiple oligomeric types in solution. Negative stain EM micrograph of soluble A) VacA Δ294-331 and B) soluble VacA Δ334-343. C) 2-D class averages of VacA Δ334-343 particles in vitrified ice. Scale bar, 20nm. Box size, 323.2Å.
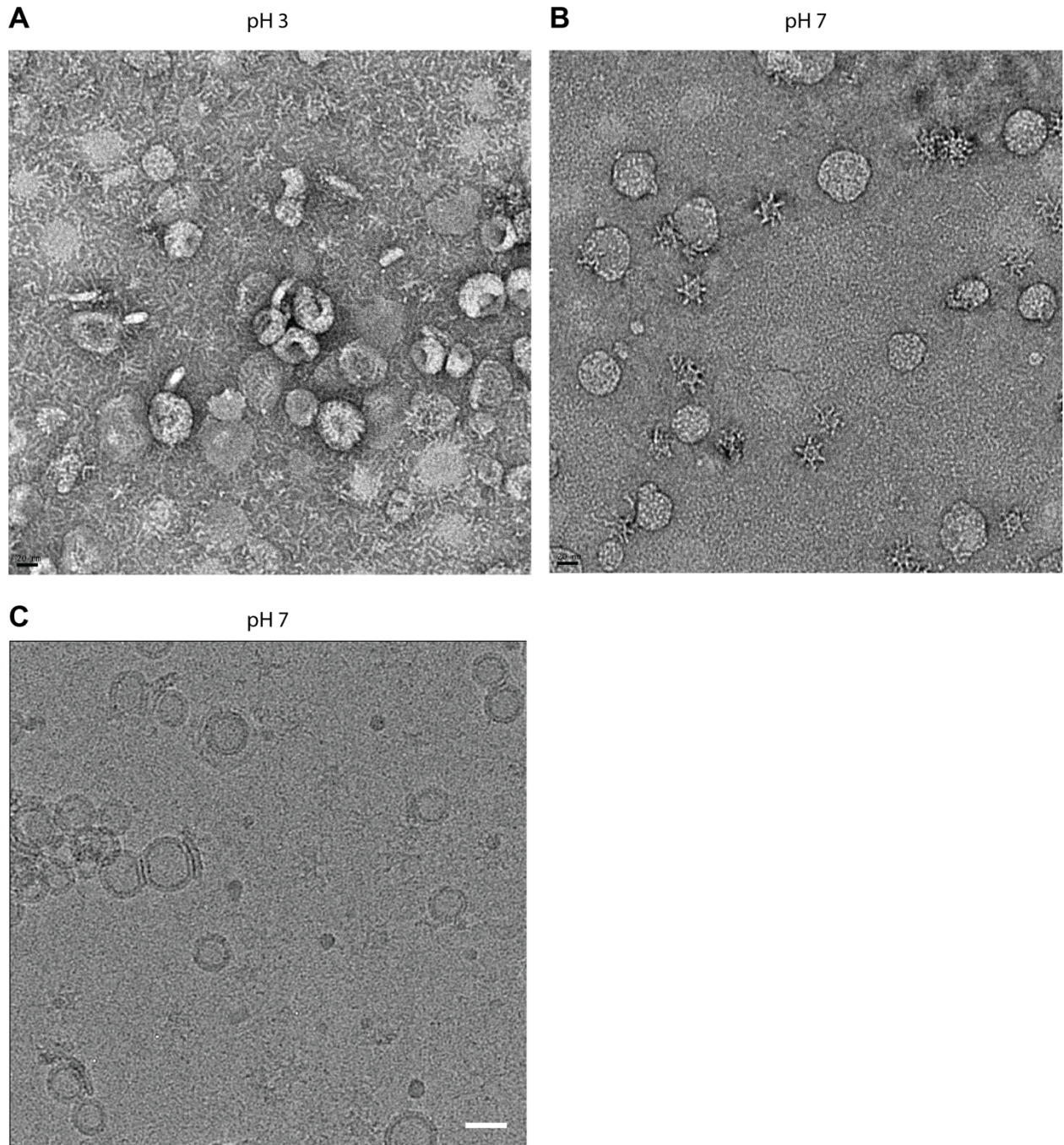
Figure 5.6. EM analysis of VacA wild-type incubated with cNW50 nanodiscs.
Negative stain EM micrograph of VacA s1/i1/m1 wild-type incubated with cNW50 nanodiscs in an A) pH 3 buffer and B) pH 7 buffer. Scale bar, 20nm. C) Cryo-EM micrograph of VacA s1/i1/m1 wild-type incubated with cNW50 nanodiscs. Scale bar, 30nm.
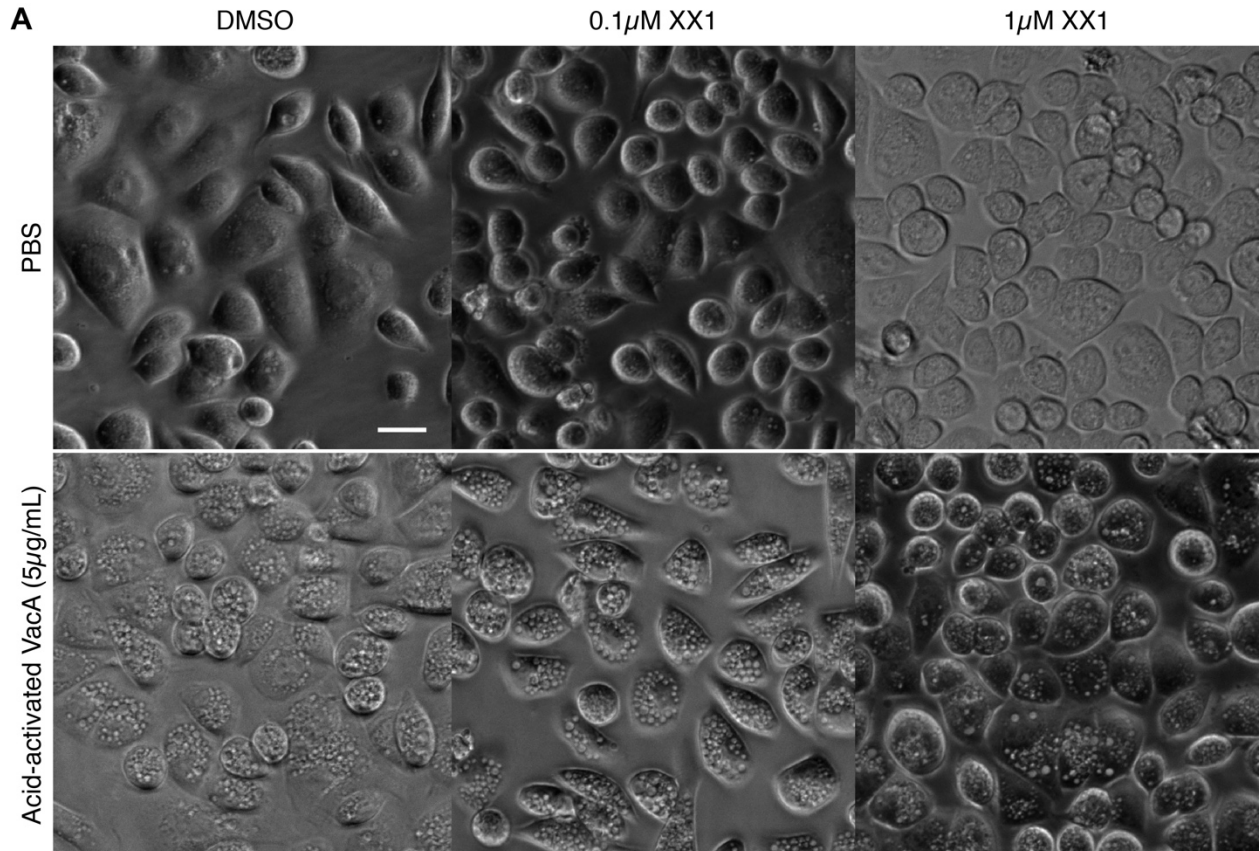
Figure 5.7. AGS cells treated with acid-activated VacA in the presence of 5mM NH₄Cl remain vacuolated upon gamma secretase inhibitor. A) Top, AGS cells were treated with (from left to right) DMSO, 0.1μM XX1 gamma secretase inhibitor, 1μM XX1 gamma secretase inhibitor for 4 hrs in the presence of 5mM NH₄Cl. Bottom, AGS cells were co-treated with 5μg/mL acid-activated VacA wild-type and (from left to right) DMSO, 0.1μM XX1, and 1μM XX1 in the presence of 5mM NH₄Cl. Scale bar, 20μm.

A

31,465 particles
(mixture of single/double fab bound) → 2D sorting → 4,441 particles
(best 2 fab bound classes)

↓ ab initio (1 class)   ↓ ab initio (1 class)

1 fab bound ab initio   2 fab bound ab initio

Heterogeneous Refinement (31,465 particles; 2 classes)

15,958 particles   15,507 particles

B MPER-Ec2-FluC:10E8

38737 ptcls | 30387 ptcls | 28246 ptcls
3.1 A 2 ess | 13.1 A 2 ess | 7.5 A 2 ess

27227 ptcls | 25334 ptcls | 23736 ptcls
9.7 A 2 ess | 3.1 A 3 ess | 8.1 A 2 ess

C MPER-AdiC:10E8

4091 ptcls | 2510 ptcls | 2424 ptcls
10.2 A 1 ess | 10.2 A 2 ess | 10.8 A 1 ess

2305 ptcls | 2294 ptcls | 2202 ptcls
10.9 A 3 ess | 10.6 A 3 ess | 9.7 A 1 ess

120

Figure 5.8. EM analysis of small membrane proteins complexed with 10E8 fab fragment. A) Heterogeneous refinement of negatively stained MPER-Ec2-FluC:10E8 particles reveals two Ec2-FluC bound to 1 or 2 10E8 fab fragments. Cryo-EM micrograph and 2-D class averages of B) MPER-Ec2-FluC:10E8 (Box size, 182Å) and C) MPER-AdiC:10E8 (Box size, 182Å). Scale bars, 200Å.

References

[1] Dal Peraro M, van der Goot FG. Pore-forming toxins: ancient, but never really out of fashion. Nat Rev Microbiol. 2016;14:77-92.

[2] Gatsogiannis C, Lang AE, Meusch D, Pfaumann V, Hofnagel O, Benz R, et al. A syringe-like injection mechanism in Photorhabdus luminescens toxins. Nature. 2013;495:520-3.

[3] Gatsogiannis C, Merino F, Prumbaum D, Roderer D, Leidreiter F, Meusch D, et al. Membrane insertion of a Tc toxin in near-atomic detail. Nat Struct Mol Biol. 2016;23:884-90.

[4] Bokori-Brown M, Martin TG, Naylor CE, Basak AK, Titball RW, Savva CG. Cryo-EM structure of lysenin pore elucidates membrane insertion by an aerolysin family protein. Nat Commun. 2016;7:11293.

[5] Iacovache I, De Carlo S, Cirauqui N, Dal Peraro M, van der Goot FG, Zuber B. Cryo-EM structure of aerolysin variants reveals a novel protein fold and the pore-formation process. Nat Commun. 2016;7:12062.
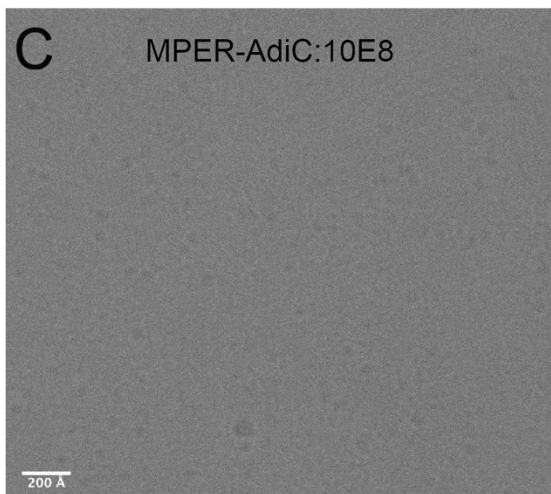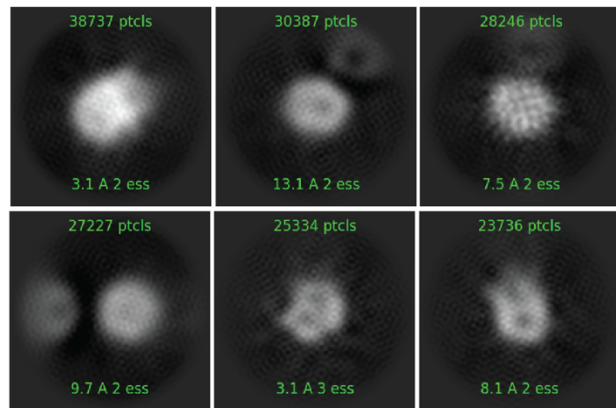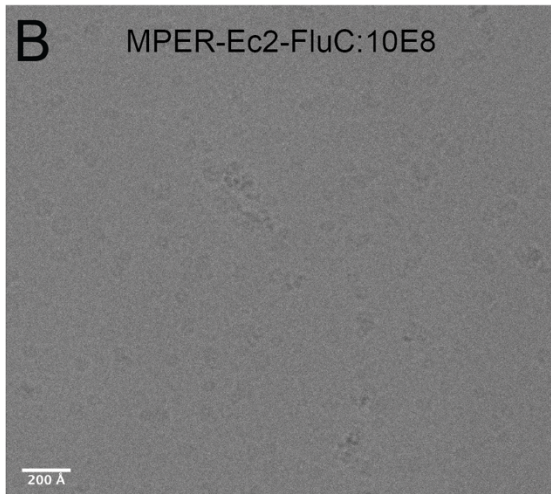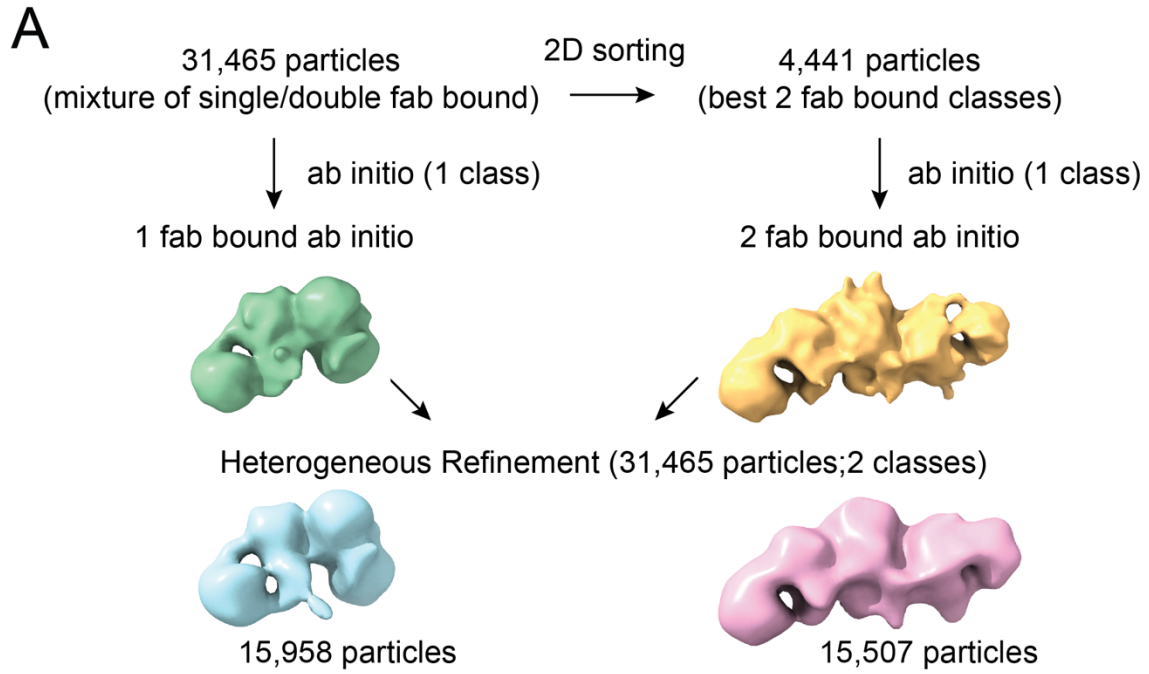
[6] van Pee K, Neuhaus A, D'Imprima E, Mills DJ, Kuhlbrandt W, Yildiz O. CryoEM structures of membrane pore and prepore complex reveal cytolytic mechanism of Pneumolysin. Elife. 2017;6.

[7] Brauning B, Bertosin E, Praetorius F, Ihling C, Schatt A, Adler A, et al. Structure and mechanism of the two-component alpha-helical pore-forming toxin YaxAB. Nat Commun. 2018;9:1806.

[8] Schubert E, Vetter IR, Prumbaum D, Penczek PA, Raunser S. Membrane insertion of alpha-xenorhabdolysin in near-atomic detail. Elife. 2018;7.

[9] Gatsogiannis C, Merino F, Roderer D, Balchin D, Schubert E, Kuhlee A, et al. Tc toxin activation requires unfolding and refolding of a beta-propeller. Nature. 2018;563:209-13.

[10] Piper SJ, Brillault L, Rothnagel R, Croll TI, Box JK, Chassagnon I, et al. Cryo-EM structures of the pore-forming A subunit from the Yersinia entomophaga ABC toxin. Nat Commun. 2019;10:1952.

[11] Peng W, de Souza Santos M, Li Y, Tomchick DR, Orth K. High-resolution cryo-EM structures of the E. coli hemolysin ClyA oligomers. PLoS One. 2019;14:e0213423.

[12] Roderer D, Hofnagel O, Benz R, Raunser S. Structure of a Tc holotoxin pore provides insights into the translocation mechanism. Proc Natl Acad Sci U S A. 2019;116:23083-90.

[13] Savva CG, Clark AR, Naylor CE, Popoff MR, Moss DS, Basak AK, et al. The pore structure of Clostridium perfringens epsilon toxin. Nat Commun. 2019;10:2641.

[14] Su M, Erwin AL, Campbell AM, Pyburn TM, Salay LE, Hanks JL, et al. Cryo-EM Analysis Reveals Structural Basis of Helicobacter pylori VacA Toxin Oligomerization. J Mol Biol. 2019;431:1956-65.

[15] Zhang K, Zhang H, Li S, Pintilie GD, Mou TC, Gao Y, et al. Cryo-EM structures of Helicobacter pylori vacuolating cytotoxin A oligomeric assemblies at near-atomic resolution. Proc Natl Acad Sci U S A. 2019;116:6800-5.

[16] Foegeding NJ, Caston RR, McClain MS, Ohi MD, Cover TL. An Overview of Helicobacter pylori VacA Toxin Biology. Toxins (Basel). 2016;8.

[17] Gangwer KA, Mushrush DJ, Stauff DL, Spiller B, McClain MS, Cover TL, et al. Crystal structure of the Helicobacter pylori vacuolating toxin p55 domain. Proc Natl Acad Sci U S A. 2007;104:16293-8.

[18] Lupetti P, Heuser JE, Manetti R, Massari P, Lanzavecchia S, Bellon PL, et al. Oligomeric and subunit structure of the Helicobacter pylori vacuolating cytotoxin. J Cell Biol. 1996;133:801-7.

[19] El-Bez C, Adrian M, Dubochet J, Cover TL. High resolution structural analysis of Helicobacter pylori VacA toxin oligomers by cryo-negative staining electron microscopy. J Struct Biol. 2005;151:215-28.

[20] Chambers MG, Pyburn TM, Gonzalez-Rivera C, Collier SE, Eli I, Yip CK, et al. Structural analysis of the oligomeric states of Helicobacter pylori VacA toxin. J Mol Biol. 2013;425:524-35.

[21] Gonzalez-Rivera C, Campbell AM, Rutherford SA, Pyburn TM, Foegeding NJ, Barke TL, et al. A Nonoligomerizing Mutant Form of Helicobacter pylori VacA Allows Structural Analysis of the p33 Domain. Infect Immun. 2016;84:2662-70.

[22] Pyburn TM, Foegeding NJ, Gonzalez-Rivera C, McDonald NA, Gould KL, Cover TL, et al. Structural organization of membrane-inserted hexamers formed by Helicobacter pylori VacA toxin. Mol Microbiol. 2016;102:22-36.

[23] Razinkov I, Dandey V, Wei H, Zhang Z, Melnekoff D, Rice WJ, et al. A new method for vitrifying samples for cryoEM. J Struct Biol. 2016;195:190-8.

[24] Donati M, Moreno S, Storni E, Tucci A, Poli L, Mazzoni C, et al. Detection of serum antibodies to CagA and VacA and of serum neutralizing activity for vacuolating cytotoxin in patients with Helicobacter pylori-induced gastritis. Clin Diagn Lab Immunol. 1997;4:478-82.

[25] Fahimi F, Sarhaddi S, Fouladi M, Samadi N, Sadeghi J, Golchin A, et al. Phage display-derived antibody fragments against conserved regions of VacA toxin of Helicobacter pylori. Appl Microbiol Biotechnol. 2018;102:6899-913.

[26] Geisow MJ, Evans WH. pH in the endosome. Measurements during pinocytosis and receptor-mediated endocytosis. Exp Cell Res. 1984;150:36-46.

[27] Choe S, Bennett MJ, Fujii G, Curmi PM, Kantardjieff KA, Collier RJ, et al. The crystal structure of diphtheria toxin. Nature. 1992;357:216-22.

[28] O'Keefe DO, Cabiaux V, Choe S, Eisenberg D, Collier RJ. pH-dependent insertion of proteins into membranes: B-chain mutation of diphtheria toxin that inhibits membrane translocation, Glu-349----Lys. Proc Natl Acad Sci U S A. 1992;89:6202-6.

[29] Pruitt RN, Chambers MG, Ng KK, Ohi MD, Lacy DB. Structural organization of the functional domains of Clostridium difficile toxins A and B. Proc Natl Acad Sci U S A. 2010;107:13467-72.

[30] Czajkowsky DM, Iwamoto H, Cover TL, Shao Z. The vacuolating toxin from Helicobacter pylori forms hexameric pores in lipid bilayers at low pH. Proc Natl Acad Sci U S A. 1999;96:2001-6.

[31] Yao X, Fan X, Yan N. Cryo-EM analysis of a membrane protein embedded in the liposome. Proc Natl Acad Sci U S A. 2020;117:18497-503.

[32] Ritchie TK, Grinkova YV, Bayburt TH, Denisov IG, Zolnerciks JK, Atkins WM, et al. Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. Methods Enzymol. 2009;464:211-31.

[33] Nasr ML, Baptista D, Strauss M, Sun ZJ, Grigoriu S, Huser S, et al. Covalently circularized nanodiscs for studying membrane proteins and viral entry. Nat Methods. 2017;14:49-52.

[34] Yahiro K, Niidome T, Kimura M, Hatakeyama T, Aoyagi H, Kurazono H, et al. Activation of Helicobacter pylori VacA toxin by alkaline or acid conditions increases its binding to a 250-kDa receptor protein-tyrosine phosphatase beta. J Biol Chem. 1999;274:36693-9.

[35] Utt M, Danielsson B, Wadstrom T. Helicobacter pylori vacuolating cytotoxin binding to a putative cell surface receptor, heparan sulfate, studied by surface plasmon resonance. FEMS Immunol Med Microbiol. 2001;30:109-13.

[36] Seto K, Hayashi-Kuwabara Y, Yoneta T, Suda H, Tamaki H. Vacuolation induced by cytotoxin from Helicobacter pylori is mediated by the EGF receptor in HeLa cells. FEBS Lett. 1998;431:347-50.

[37] Fujikawa A, Shirasaka D, Yamamoto S, Ota H, Yahiro K, Fukada M, et al. Mice deficient in protein tyrosine phosphatase receptor type Z are resistant to gastric ulcer induction by VacA of Helicobacter pylori. Nat Genet. 2003;33:375-81.

[38] Yahiro K, Wada A, Nakayama M, Kimura T, Ogushi K, Niidome T, et al. Protein-tyrosine phosphatase alpha, RPTP alpha, is a Helicobacter pylori VacA receptor. J Biol Chem. 2003;278:19183-9.

[39] Nakayama M, Hisatsune J, Yamasaki E, Nishi Y, Wada A, Kurazono H, et al. Clustering of Helicobacter pylori VacA in lipid rafts, mediated by its receptor, receptor-like protein tyrosine phosphatase beta, is required for intoxication in AZ-521 Cells. Infect Immun. 2006;74:6571-80.

[40] Gupta VR, Patel HK, Kostolansky SS, Ballivian RA, Eichberg J, Blanke SR. Sphingomyelin functions as a novel receptor for Helicobacter pylori VacA. PLoS Pathog. 2008;4:e1000073.

[41] Yahiro K, Satoh M, Nakano M, Hisatsune J, Isomoto H, Sap J, et al. Low-density lipoprotein receptor-related protein-1 (LRP1) mediates autophagy and apoptosis caused by Helicobacter pylori VacA. J Biol Chem. 2012;287:31104-15.

[42] Inoue T, Zhang P, Zhang W, Goodner-Bingham K, Dupzyk A, DiMaio D, et al. gamma-Secretase promotes membrane insertion of the human papillomavirus L2 capsid protein during virus infection. J Cell Biol. 2018;217:3545-59.

[43] Smith AAA, Autzen HE, Faust B, Mann JL, Muir BW, Howard S, et al. Lipid Nanodiscs via Ordered Copolymers. Chem 2020;6:1-14.

[44] Nygaard R, Kim J, Mancia F. Cryo-electron microscopy analysis of small membrane proteins. Curr Opin Struct Biol. 2020;64:26-33.

[45] Wu S, Avila-Sakar A, Kim J, Booth DS, Greenberg CH, Rossi A, et al. Fabs enable single particle cryoEM studies of small proteins. Structure. 2012;20:582-92.

[46] Dang S, Feng S, Tien J, Peters CJ, Bulkley D, Lolicato M, et al. Cryo-EM structures of the TMEM16A calcium-activated chloride channel. Nature. 2017;552:426-9.

[47] Butterwick JA, Del Marmol J, Kim KH, Kahlson MA, Rogow JA, Walz T, et al. Cryo-EM structure of the insect olfactory receptor Orco. Nature. 2018;560:447-52.

[48] Kim J, Tan YZ, Wicht KJ, Erramilli SK, Dhingra SK, Okombo J, et al. Structure and drug resistance of the Plasmodium falciparum transporter PfCRT. Nature. 2019;576:315-20.

[49] Coleman JA, Yang D, Zhao Z, Wen PC, Yoshioka C, Tajkhorshid E, et al. Serotonin transporter-ibogaine complexes illuminate mechanisms of inhibition and transport. Nature. 2019;569:141-5.

[50] Bloch JS, Pesciullesi G, Boilevin J, Nosol K, Irobalieva RN, Darbre T, et al. Structure and mechanism of the ER-based glucosyltransferase ALG6. Nature. 2020;579:443-7.

[51] Lipfert J, Columbus L, Chu VB, Lesley SA, Doniach S. Size and shape of detergent micelles determined by small-angle X-ray scattering. J Phys Chem B. 2007;111:12427-38.

[52] Pang SS, Bayly-Jones C, Radjainia M, Spicer BA, Law RHP, Hodel AW, et al. The cryo-EM structure of the acid activatable pore-forming immune effector Macrophage-expressed gene 1. Nat Commun. 2019;10:4288.

Chapter VI

Concluding Remarks

H. pylori is a human specific pathogen that colonizes the stomach and causes gastric inflammation, peptic ulcer disease, and gastric cancer[1]. The Gram-negative bacterium secretes virulence factors that aid in infection of the human stomach, including a pore forming toxin called vacuolating cytotoxin A (VacA)[2]. While VacA has important roles in *H. pylori* colonization of the human stomach and the development of *H. pylori*-related gastroduodenal diseases, the function of VacA has not been fully elucidated. Previous studies have suggested that VacA promotes *H. pylori* infection by oligomerizing into anionic channels/pores in various host cell membranes[3-5].

While VacA oligomerization has been shown to be required for the formation of active channels in planar lipid bilayer studies, the mechanism of VacA oligomerization remained unknown prior to my dissertation studies[3, 6, 7]. The determination of a 3.8Å cryo-EM structure of a VacA hexamer in the absence of a membrane environment in Chapter II revealed that residues from both the p33 and p55 regions of neighboring VacA protomers are required for oligomerization[8]. Structures of a VacA hexamer and VacA docecamers in the absence of a membrane environment determined in 2019 by Zhang*, Zhang*, Li* et al. identified two specific salt bridges from the p33 and p55 regions of neighboring VacA protomers that are important for oligomerization, K47:E338 and K55:D346[9]. Importantly, decades of biochemical analysis of VacA supported our

126

findings in Chapter II and the findings from Zhang*, Zhang*, Li* et al[10-12]. My work detailed in Chapter V further investigated the importance of these identified salt-bridge residues by introducing cysteine mutations to induce disulfide bond formation. Negative stain analysis of the resulting disulfide mutants showed that the disulfide bond at the salt-bridge residue locations are sufficient to prevent VacA from dissociating, validating the findings from the cryo-EM structures.

The oligomerization of VacA on membrane is required for active channel/pore formation. Chapter II suggested that the loop regions extending from the VacA beta-helix may be important for binding to host cell membranes. To further understand how VacA interacts with membranes, in Chapter III I incubated VacA with detergent, small unilamellar vesicles (SUVs), and large unilamellar vesicles (followed by detergent solubilization). In the context of membrane (SUVs) and detergent as a membrane mimetic, VacA formed hexamers and heptamers. Single-particle cryo-EM analysis of the VacA sample that was bound to large unilamellar vesicles followed by detergent solubilization with DDM revealed additional density in the center of the hexamer and heptamer particles which may correspond to the tandem GXXXG motif region predicted to form a helical bundle as the VacA transmembrane pore[3]. This work presents new structural information showing how VacA associates with membrane prior to pore formation.

In addition to the first two dissertation chapters focusing on structural analyses of VacA oligomerization and membrane association, Chapter IV presents the use of the membrane proximal external region (MPER) of the HIV envelope glycoprotein gp41 as an N-terminal transmembrane helix epitope tag for X-ray crystallography and single particle EM structural studies. Addition of the MPER tag to an N-terminal transmembrane

helix of a membrane protein enables complexation with well-characterized antibody fragments to MPER. Complexation with these antibody fragments can be used to facilitate crystallization and add to the molecular weight of a complex, improving contrast in EM studies and introducing features helpful for 2-D and 3-D alignment of particles. This technique is broadly applicable to the structural biology community, including to future structural studies of the VacA pore forming region. While there is an existing theoretical model of the VacA GXXXG motif region predicted to form a transmembrane helical bundle, it would be valuable to determine an experimental X-ray crystallography or electron microscopy structure of this N-terminal region of VacA using the MPER N-terminal transmembrane helix epitope tag. Altogether, the new insights from this dissertation inform our fundamental knowledge of VacA oligomerization and VacA membrane association, as well as characterize a new tool, the MPER N-terminal transmembrane helix epitope tag for the study of membrane proteins with confirmed and predicted transmembrane helices.

References

[1] Suerbaum S, Michetti P. Helicobacter pylori infection. N Engl J Med. 2002;347:1175-86.

[2] Foegeding NJ, Caston RR, McClain MS, Ohi MD, Cover TL. An Overview of Helicobacter pylori VacA Toxin Biology. Toxins (Basel). 2016;8.

[3] Kim S, Chamberlain AK, Bowie JU. Membrane channel structure of Helicobacter pylori vacuolating toxin: role of multiple GXXXG motifs in cylindrical channels. Proc Natl Acad Sci U S A. 2004;101:5988-91.

[4] McClain MS, Iwamoto H, Cao P, Vinion-Dubiel AD, Li Y, Szabo G, et al. Essential role of a GXXXG motif for membrane channel formation by Helicobacter pylori vacuolating toxin. J Biol Chem. 2003;278:12101-8.

[5] Pyburn TM, Foegeding NJ, Gonzalez-Rivera C, McDonald NA, Gould KL, Cover TL, et al. Structural organization of membrane-inserted hexamers formed by Helicobacter pylori VacA toxin. Mol Microbiol. 2016;102:22-36.

[6] Iwamoto H, Czajkowsky DM, Cover TL, Szabo G, Shao Z. VacA from Helicobacter pylori: a hexameric chloride channel. FEBS Lett. 1999;450:101-4.

[7] Vinion-Dubiel AD, McClain MS, Czajkowsky DM, Iwamoto H, Ye D, Cao P, et al. A dominant negative mutant of Helicobacter pylori vacuolating toxin (VacA) inhibits VacA-induced cell vacuolation. J Biol Chem. 1999;274:37736-42.

[8] Su M, Erwin AL, Campbell AM, Pyburn TM, Salay LE, Hanks JL, et al. Cryo-EM Analysis Reveals Structural Basis of Helicobacter pylori VacA Toxin Oligomerization. J Mol Biol. 2019;431:1956-65.

[9] Zhang K, Zhang H, Li S, Pintilie GD, Mou TC, Gao Y, et al. Cryo-EM structures of Helicobacter pylori vacuolating cytotoxin A oligomeric assemblies at near-atomic resolution. Proc Natl Acad Sci U S A. 2019;116:6800-5.

[10] Burroni D, Lupetti P, Pagliaccia C, Reyrat JM, Dallai R, Rappuoli R, et al. Deletion of the major proteolytic site of the Helicobacter pylori cytotoxin does not influence toxin activity but favors assembly of the toxin into hexameric structures. Infect Immun. 1998;66:5547-50.

[11] Genisset C, Galeotti CL, Lupetti P, Mercati D, Skibinski DA, Barone S, et al. A Helicobacter pylori vacuolating toxin mutant that fails to oligomerize has a dominant negative phenotype. Infect Immun. 2006;74:1786-94.

[12] Ivie SE, McClain MS, Torres VJ, Algood HM, Lacy DB, Yang R, et al. Helicobacter pylori VacA subdomain required for intracellular toxin activity and assembly of functional oligomeric complexes. Infect Immun. 2008;76:2843-51.

Chapter VII

Methods

Purification of VacA

H. pylori strains (from s1/i1/m1 parent strain 60190) were grown in brucella broth supplemented with lipid concentrate (BB-cholesterol) for 48 hours[1]. Cultures were centrifuged, followed by precipitation of proteins in the supernatant with 50% saturated solution of ammonium sulfate. Precipitated proteins were pelleted using centrifugation and subsequently resuspended in phosphate buffered saline (PBS), 1mM EDTA, and 0.02% sodium azide. Strep-tagged VacA (strep-tag incorporated at position 808) was incubated with Strep-Tactin resin (IBA), loaded into a gravity column, and washed with 50mM Tris, 150mM NaCl (pH 8.0). Lastly, VacA was eluted from the resin with 50mM Tris, 150mM NaCl, and 5mM D-desthiobiotin (pH 8.0).

VacA acid activation

Purified VacA was acid-activated to dissociate soluble oligomers by dropwise addition of 0.5M HCl to pH 3 or 4.0 depending on the experiment [2-4]. Samples were kept on ice until negative staining, addition to SUVs/LUVs/nanodiscs, vitrification within 30 minutes.

VacA SUV binding experiments

Egg PC (L-a-phosphatidylcholine 840051C), DOPS (1,2-dioleoyl-sn-glycero-3-phospho-L-serine sodium salt 840035C), and ovine wool cholesterol (7000000P) from Avanti Polar Lipids (Alabaster, AL) were mixed in chloroform in a 55/15/30 mol % ratio (eggPC/DOPS/Cholesterol). The chloroform was evaporated under N_2 stream and dried overnight in a vacuum desiccator. After fully dried, lipids were rehydrated to 2mg/mL final lipid concentration with 10mM HEPES pH 7.3, 100mM KCl. Lipids were vigorously vortexed for 2 minutes, and subsequently sonicated in a water bath sonicator for 50 minutes until the solution was clear. The resulting 20-50nm SUVs were used for VacA SUV binding experiments with 1-2 days of SUV preparation.

SUVs were prewarmed in a 37°C incubator for 15 minutes. VacA wild-type or Δ6-27 was diluted to 0.5mg/mL in 10mM HEPES pH 7.3, 100mM KCl, acid-activated, and then mixed slowly with prewarmed SUVs at a lipid to protein ratio (LPR) of 20:1. The mixture was incubated at 37°C for 15 mins.

VacA direct incubation with detergents

Acid-activated VacA wild-type was added to 10mM HEPES pH 7.3, 100mM KCl buffer with varying percentages of Anatrace detergents Lauryl Maltose Neopentyl Glycol (0.0011%, 0.002%, 0.005%; CMC: 0.001%), Cymal-6 (0.03%, 0.05%, 0.14%; CMC: 0.028%), and Cymal-7 (0.01%, 0.05%, 0.1%; CMC:0.0099%).

Extraction of VacA from LUVs with detergents

Egg PC (L-a-phosphatidylcholine 840051C), DOPS (1,2-dioleoyl-sn-glycero-3-phospho-L-serine sodium salt 840035C), and ovine wool cholesterol (7000000P) from Avanti Polar Lipids (Alabaster, AL) were mixed in chloroform in a 55/15/30 mol % ratio (eggPC/DOPS/Cholesterol). The chloroform was evaporated under N_2 stream and dried overnight in a vacuum desiccator. After fully dried, lipids were resuspended with 10mM HEPES pH 7.2, 100mM KCl to 1mg/mL, vortexed, and incubated at room temperature for 30 minutes. To generate the large unilamellar vesicles (LUVs), the resuspended lipids were frozen in a dry ice ethanol bath and thawed in a 37˚C water bath 3 times and then extruded 11 times through an 0.8µm polycarbonate membrane (610009 Millipore Sigma) using the Avanti mini extruder. The LUVs were mixed with acid activated strep-tagged VacA at a lipid to protein ratio (LPR) of 25:1 (w:w) for 30 minutes at room temperature.

The VacA LUV sample was pelleted at 50,000xg for 45 minutes at 7˚C using the Beckman TLA 100 rotor and TL100 centrifuge. The pellet containing VacA bound LUVs was washed gently 3 times with 10mM HEPES pH 7.2, 100mM KCl and then resuspended with 10mM HEPES pH 7.2, 100mM KCl, 2% DDM for 1 hour at room temperature. The sample was centrifuged at 50,000xg for 45 minutes at 7˚C using the Beckman TLA 100 rotor and TL 100 centrifuge. The supernatant containing solubilized VacA detergent micelle complexes was collected and bound to Strep-Tactin resin (IBA) for 30 minutes at 4˚C, washed 4 times with 10mM HEPES pH 7.2, 100mM KCl, 0.02% DDM, and then eluted with 10 mM HEPES pH 7.2, 100mM KCl, 0.02% DDM, 5mM D-Desthiobiotin.

132

Informatics-based screening of membrane proteins for compatibility with MPER-epitope tag

α-helical transmembrane proteins from the non-redundant PDB-TM database, excluding theoretical models, were obtained from https://www.rcsb.org (934 proteins). Transmembrane biological units were reconstructed from α-helical transmembrane chains identified in "biological unit" information from REMARK 350 of the PDB file. For complexes with a molecular weight of ≤100 kDa (457 proteins), the membrane was modelled using the PPM2.0 method (https://opm.phar.umich.edu), implemented using a program provided by Andrei Lomize. Based on the estimated membrane boundaries, the dataset was further filtered to include only complexes with two or more transmembrane helices and with more than 50% of residues inside the membrane. The proteins in this set (304 proteins) were analyzed for steric compatibility with the MPER epitope tag according to two criteria: a tilt angle for the first transmembrane helix of ≥45° relative to the plane of the membrane (294 proteins), and no clashes between the MPER epitope tag and the target protein. To assess clashes between the target protein and MPER, the last 8 residues (IKSLAAAA) of the α-helical MPER peptide (LWNWFDITNWLWYIKSLAAAA) were aligned in PyMol with the first TM helix of the target protein in four registers: *n* (the first eight residues of the first transmembrane helix of the target), *n*+1, *n*+2, and *n*+3. Each of these four models was assessed for clashes (C_α ≤3Å) between the MPER epitope tag and the biological unit. Models with an acceptable tilt angle for TM helix 1, and with no clashes between MPER and the target protein (1069 models representing 290 proteins) were further assessed for

133

complexation by antibody fragments 10E8 Fab, 10E8 scFv, 4E10 Fab, and 4E10 scFv. For each model, the antibody fragment was positioned relative to the MPER epitope based on structures of antibody-MPER complexes (PDB: 1TZG and PDB: 5JNY) and assessed for clashes ($C_\alpha \leq 5$Å) with the target protein, the membrane, or, for targets with oligomeric construction, other antibody fragments. For these calculations, the residues N-terminal to the twelfth residue of the target's first transmembrane helix were not considered. To assess membrane clashes, the membrane boundary was extended to the length of the antibody complex, and the membrane-penetrating CDR loops of the antibody fragments were excluded from the calculation. Different cutoffs for MPER and antibody clashes calculations were chosen because the MPER clash reflects intra-protein interactions, whereas the antibody clash reflects inter-protein interactions. All models that could accommodate at least one of the four antibody fragments with <10 clashes were manually inspected.

Antibody preparation

10E8v4 antibody expression vector DNA (heavy chain and light chain) was obtained through the AIDS Reagent Program, Division of AIDS, NIAID, NIH from Dr. Peter Kwong (cat. No. 12866 and 12877) and used to transfect Expi293 HEK cells. Antibodies were harvested from media 5 days post-transfection using Protein A agarose (Pierce). Antibodies were processed to antibody fragments by papain digestion (1:200 antibody: papain), followed by filtration on an immobilized Protein A agarose column (Pierce) to separate the Fc domain from Fab fragments. Fab fragments were further purified by gel filtration chromatography.

MPER-tagged protein preparation

Synthetic constructs with a hexahistidine tag, a thrombin recognition site, the MPER epitope sequence (LWNWFDITNWLWYIKNL), and a sequence encoding the fluoride channel from *B. pertussis* (Bpe) or *E. coli* (Ec2), the arginine/agmatine antiporter from *S. entericus* (AdiC), or the glycerol channel from *E. coli* (GlpF) were designed so that the MPER sequence was positioned at the membrane boundary (sequences in Supplementary Table 5). Synthetic gene constructs were cloned into the pET21a protein expression vector and transformed into *E. coli* C41 cells. Cells were grown at 37 ºC to an O.D$_{600}$ of ~1 before induction. The Fluc-Ec2 and Fluc-Bpe variants were induced for 1 hour with 4 mM isopropyl β-1-thiogalactopyranoside (IPTG); the GlpF variant was induced for 2 hours with 1 mM IPTG; the AdiC variant was induced for 1 hour with 0.2 mM IPTG. Cells were harvested by centrifugation, resuspended in cell breaking buffer (100 mM NaCl, 20 mM Tris pH 8.0) with lysozyme (1 mg/mL), protease inhibitor cocktail (Roche), PMSF (200 mM), and lysed by sonication on ice. After detergent extraction with 2% n-Decyl-β-D-Maltoside (DM) for 2 hours, cell debris were removed by centrifugation, and His-tagged proteins were isolated using cobalt affinity resin (Takara). Bound protein was eluted with 400 mM imidazole, and applied to a desalting column to remove imidazole. His tags were cleaved by overnight thrombin digestion (1.58U thrombin per mg protein, 4 ºC). For antibody complexation, protein and 10E8v4 Fab were combined in a 1:1 stoichiometric ratio and incubated for 2 hours at room temperature prior to gel filtration (Superdex200 equilibrated in 100 mM NaCl, 10

mM HEPES pH 7.5, 4 mM DM). Fractions corresponding to the protein:10E8v4 Fab

complex were isolated and used for further studies.

Ec2/10E8v4 antibody fragment crystallography

Samples containing stable MPER-Fluc-Ec2/10E8v4 complex were concentrated to 10

mg/mL and mixed with an equal volume of crystallization solution in 24-well sitting-drop

vapor diffusion trays. After optimization of initial hits, MPER-Fluc-Ec2/10E8v4 crystals

grew in 100 mM NaCl, 34-39% PEG 300, 100 mM MES pH 6.2-7.0. For MPER-Fluc-

Ec2/scFv crystallization, the stable complex was formed as for MPER-Fluc-Ec2/10E8v4,

and protein samples were mixed with an equal volume of crystallization solution from

commercial screening blocks (MemGold and MemGold2, Molecular Dimensions) in 96

well plates. The best-diffracting crystals were grown in 100 mM calcium acetate, 100

mM HEPES pH 7.5, 33.8% PEG 600. Crystals were frozen in liquid nitrogen prior to

data collection at the Life Sciences Collaborative Access Team beamline 21-ID-D at the

Advanced Photon Source, Argonne National Laboratory. Phases were calculated by

molecular replacement with Phaser using Fluc-Ec2 (pdb:5A43) and the constant and

variable domains of 10E8v4 (pdb:5IQ9) as search models, followed by iterative rounds

of refinement with Refmac and model building in real space with Coot. Figures were

prepared with PyMol.

**Design and purification of a single-chain, variable-domain antibody fragment
(scFv)**

136

Design of scFv was aided by the crystal structure of 10E8v4 bound to the helical epitope. The construct was designed as: Heavy Chain$_{variable}$-(Gly$_3$Ser)$_4$ linker-Light Chain$_{variable}$-TEV cleavage site (ENLYFQS)-His x6 (sequence in Supplementary Table 4) and sub-cloned into the periplasmic expression vector pET26b. *E. coli* (Origami B(DE3), Novagen) bearing this construct was grown at 37 ºC until reaching an O.D$_{600}$ of ~ 0.8, and induced with 50 μM IPTG overnight at 16 ºC. Cells were harvested and resuspended in osmotic shock buffer (200 mM Tris pH 8, 5 mM EDTA, 20% sucrose) and incubated for 1 hour at room temperature with occasional mixing. Cell suspension was centrifuged (16,000 r.p.m) to pellet the bacterial cytoplasm and cellular debris. Supernatant containing scFv was purified using gel filtration (Superdex 200 equilibrated in 100 mM NaCl, 10 mM HEPES pH 7.5, 4 mM DM).

VacA treatment with DTT

100mM DTT was added to VacA (30μg/mL) and incubated for 1 hour at 37˚C. For analysis of DTT-treated samples at pH 3, 0.2-0.4μL of 0.5M HCl was added to pH 3 (tested with litmus paper). Samples were kept at room temperature and negatively stained within 30 minutes.

cNW50 nanodisc preparation and incubation with VacA

POPC/POPG in chloroform was mixed at a 3:2 ratio (61.33μL of 25mg/mL POPC, 41.47μL of 25mg/mL POPG). Chloroform was dried off with N$_2$ gas and left in a vacuum desiccator for 1 hour. The lipid mixture was resolubilized in 225μL 20mM Tris-HCl, pH 7.5, 100mM NaCl, 50mM sodium cholate for 10 minutes until lipid solution was clear. 0.5mg of lyophilized cNW50 (purchased from NOW Scientific) was resuspended

with 472µL 20mM Tris-HCl, pH 7.5, 100mM NaCl (to 10µM). 75µL of 10µM cNW50 was

thawed and added to the resolubilized lipid mixture and then incubated for 1 hour on

ice. 0.5g of wet biobeads was added to the cNW50 lipid mixture for an additional 1 hour

incubation on ice. Next, the mixture was rocked on an orbital shaker for 3 hours in a

cold room. After removing the biobeads, the sample was run over a Superose 6 10/300

increase column. The peak collected around 9mL was negatively stained revealing

large ~50 nm nanodiscs and liposomes. Acid activated VacA in 50mM Tris, 150mM

NaCl pH 7 was added to lipid particles from the 9mL peak and diluted further in 50mM

Tris, 150mM NaCl pH 7 or pH 3 citrate acid buffer. 3µL was negatively stained and

imaged on the Morgagni. 3.5µL was plunged (1s blot time, humidity 100%, 20˚C) onto

glow discharged (30s 5mA) R1.2/1.3 200 mesh quantifoil grids coated with thin carbon

(homemade) for cryo-EM screening on the 200keV Arctica microscope.

AGS cell vacuolation assay in the presence of gamma secretase inhibitor

15×10^4 AGS cells plated into four well chamber slides (1.7cm2/well) in RPMI

complete media were treated with 01µM or 1µM XX1 gamma-secretase inhibitor (gift

from Dr. Billy Tsai) or mock DMSO for 30 minutes at 37˚C. Acid-activated VacA

(5µg/mL) or mock PBS was added to the cells treated with inhibitor or DMSO in the

presence of 5mM NH_4Cl. Treated cells were imaged with a 40x/0.60 objective lens and

CCD camera on a DeltaVision at 2 hour and 4 hour timepoints to visualize the extent of

vacuolation. Images for figures were prepared using Fiji[5].

Negative stain EM analysis

Protein samples (3 µL of 10-100µg/ml) were adsorbed for 1 minute to a glow discharged (30s at 5mA) 400-mesh copper grid (Structure Probe, Inc.) that was coated with colloidion film followed by carbon in a carbon evaporator. The grids were washed twice and then negatively stained in 0.07% uranyl formate solution. Transmission electron microscopy (TEM) images were obtained on a FEI Morgagni electron microscope run at 100 keV at a magnification of 22,000x (2.1Å/pixel) and then recorded on a Gatan Orius charge-coupled device camera.

Negative stain EM data collection and processing

Membrane-extracted VacA DDM micelle samples

The negatively stained VacA DDM micelle complex samples were imaged with a FEI Tecnai T12 microscope operated at 120keV. Images (107 for wild-type, 272 for Δ6-27) were acquired with Leginon[6]. Images were recorded at a pixel size of 2.342Å/px. CTF estimation was performed on the images using gCTF[7]. Particles were picking using the crYOLO general negative stain model and analyzed with the crYOLO box manager using SBGrid[8, 9]. Particle extraction (with a 618Å box size) and 2D classification (400Å mask size) were performed using RELION 3.0.8[10]. 2-D class averages were generated using 4,662 VacA wild-type and 2,973 VacA Δ6-27 DDM particles. Stacked bar graph was generated using GraphPad Prism 7.0c.

Disulfide mutants

VacA wild-type controls and VacA disulfide mutant (VacA 44C 338C, VacA 47C 338 C, and VacA 55C 346C) samples at pH 3 and pH 8 were negatively stained and

imaged with a FEI Tecnai T12 microscope operated at 120keV. Images (141 for VacA

wild-type pH 8, 65 for VacA wild-type pH 3, 119 for VacA 44C 338C pH 8, 97 for VacA

44C 338C pH 3, 245 for VacA 47C 338C pH 8, 248 for VacA 47C 338C pH 3, 175 for

VacA 55C 346C pH 8, and 132 for VacA 55C 346C pH 3) were acquired with

Leginon[6]. Images were recorded at a pixel size of 2.342Å/px. CTF estimation was

performed on the images using CTFFIND4[11]. Particles (41156 for VacA wild-type pH

8, 23884 for VacA wild-type pH 3, 37939 for VacA 44C 338C pH 8, 20172 for VacA 44C

338C pH 3, 16526 for VacA 47C 338C pH 8, 27562 for VacA 47C 338C pH 3, 18682 for

VacA 55C 346C pH 8, and 17101 for VacA 55C 346C pH 3) were picked using the

crYOLO general negative stain model and analyzed with the crYOLO box manager

using SBGrid[8, 9]. Particle extraction (618Å box size for all samples except VacA wild-

type pH 3: 244Å) and 2D classification (400Å mask size for all samples except VacA

wild-type pH 3: 200Å) were performed using RELION 3.0.8[10]. 2-D class averages

were generated using 16,535 VacA wild-type pH 8, 11,150 VacA wild-type pH 3, 24,325

VacA 44C 338C pH 8, 1409 VacA 44C 338C pH 3, 6,684 VacA 47C 338C pH 8, 4,922

VacA 47C 338C pH 3, 8,894 VacA 55C 346C pH 8, and 6,493 VacA 55C 346C pH 3

particles.

MPER-Ec2-FluC:10E8, MPER-AdiC:10E8, MPER-GlpF:10E8, MPER-Ec2-FluC:scFv

The negatively-stained target/antibody samples were imaged with a FEI Tecnai

T12 microscope operated at 120 keV (MPER-Fluc-Ec2/10E8, MPER-AdiC/10E8,

MPER-GlpF/10E8) or a FEI Tecnai T20 microscope operated at 200 keV (MPER-Fluc-

Ec2/scFv). Images were acquired with Leginon[6]. Images were recorded at a pixel size

of 2.34A/px (MPER-Fluc-Ec2, MPER-GlpF), 1.45A/px (MPER-AdiC), or 1.37A/px (MPER-Fluc-Ec2/scFv). CTF estimation was performed on the images using CTFFIND4[11]. Particles were picked using the crYOLO general negative stain model and analyzed with the crYOLO box manager using SBGrid[8, 9]. Extracted particles were then imported into cryoSPARC for 2-D class averaging[12].. Averages were generated using 9,582 (MPER-Fluc-Ec2), 15,290 (MPER-AdiC), 7,594 (MPER-GlpF), and 6,781 (MPER-Fluc-Ec2/scFv) particles. Ab initio models and subsequent 3-D refinements were generated in cryoSPARC[12]. For MPER-FluC-Ec2 specifically, 118 images were acquired with Leginon[6]. Images were recorded at a pixel size of 2.342Å/px. CTF estimation was performed on the images using CTFFIND4[11]. 43,169 particles were picked using the crYOLO general negative stain model and analyzed with the crYOLO box manager using SBGrid[8, 9]. Particle extraction (with a 398Å box size) was performed using cryoSPARC v2[12]. An ab initio 3D volume that appeared as MPER-Ec2-FluC:10E8 with one fab fragment bound was generated using 31,465 single and double fab fragment bound particles. Another ab initio 3D volume was generated using 4,441 double fab fragment bound particles. Both of these volumes were used for a cryoSPARC v2 heterogeneous refinement job with 2 classes to sort MPER-Ec2-FluC:10E8 particles into single fab fragment and double fab fragment bound complexes and generate 3-D volumes representing single fab bound (15,958 particles) and double fab bound (15,507 particles) MPER-Ec2-FluC[12]. Maps were analyzed in ChimeraX[13].

Cryo-EM grid preparation, data collection, and processing

VacA s1/i1/m1 soluble oligomers,

3 µL of wild-type VacA s1m1 was loaded onto glow-discharged QUANTIFOIL R2/2 200 mesh grids (EMS) and vitrified using a Vitrobot (FEI, Mark IV). In the first session of data collection, the samples were visualized at liquid nitrogen temperature on a Polara electron microscope (Thermo Fisher Scientific) operating at 300 kV. Cryo-EM images were recorded at a nominal magnification of 40,109X using a K2 Summit direct electron detector (Gatan) in counted mode, corresponding to a pixel size of 1.25 Å per pixel with a dose rate of ~6.0 electrons $Å^{-2}s^{-1}$. In the second session of data collection, the samples were imaged on a Titan Krios electron microscope (Thermo Fisher Scientific) operating at 300 kV. Images were recorded at a nominal magnification of 50,000X using a K2 Summit direct electron detector in counted mode, corresponding to a pixel size of 1.01 Å per pixel with a dose rate of ~6.0 electrons $Å^{-2}s^{-1}$.

Movie frames were first dose-weighted and aligned using Motioncor2 [14]. The contrast transfer function (CTF) values were determined using CTFFIND4 [11]. Image processing was carried out using CryoSPARC [12], RELION (v1.4 and 2.0) and cisTEM [10, 15-17]. A total of 16,462 and 5,471 cryo-EM images were recorded using a Titan Krios and Polara (Thermo Fisher Scientific) respectively. The 3D VacA dodecamer was calculated using data collected on the Polara using RELION for 2D classification, *ab initio* model calculation, and 3D refinement starting with C1 symmetry and ending the final rounds with C6 symmetry. The final resolution of the VacA dodecamer was 12 Å.

Using automated particle picking in RELION, particles were identified and extracted from dose weighted micrographs. For the hexamer and heptamer 3D

142

reconstructions, the images from the Polara and Titan Krios were combined for all data processing steps. In order to combine the images, particle images extracted from Titan Krios images were down sampled to a pixel size of 1.25 Å to match the sampling of Polara for data combination. A data set of ~500,000 particles were extracted. Two-dimensional (2D) classifications were performed using RELION-2. Initial 3D models of the hexamer and the heptamer were generated as follows. Classes representing the *en face* views of hexamers and heptamers, and side views of double-layer oligomers were selected and combined (133,827 particles). These particles were moved to CryoSPARC for *ab initio* model calculation of either a hexamer or heptamer using no applied symmetry (C1). A 3D hexamer and heptamer were chosen for use as the initial model for homogenous refinement in CryoSPARC was then done using C6 for the hexamer refinement or C7 for the heptamer refinement. Hexamer, heptamer, and double-layer side views were combined and a round of heterogenous refinement with no applied symmetry (C1) was done in CryoSPARC.

The best-defined 3D hexamer and heptamer models were chosen for further refinement and signal subtraction in RELION. 77% of the particles were found in the 3D model of the hexamer and 23% of the particles were found in the 3D model of the heptamer. The models were then used in RELION for 3D refinement with either C6 or C7 applied symmetry, after which signal subtraction was used on each layer of the double layer side views to subtract one layer and leave one layer remaining. Signal subtracted images of the side views were then combined with either the *en face* views of the hexamers or heptamers. Next supervised refinement in RELION was done with no applied symmetry (C1) and then the final 3D refinement was done using cisTEM.

The final resolution of the VacA hexamer and heptamer was 3.8 Å and 8.5 Å respectively. Reported resolutions are based on the gold standard Fourier shell correlation (FSC) using the 0.143 criterion [18]. The FSC curve for the hexamer oscillates in the range of 5 – 7 Å resolution with a correlation score dropping to around 0.55. The typical spacing of beta strands in the VacA structure is around 6 Å. It is likely that the FSC curve oscillation seen is related to p88 being dominated by beta-strands. High-resolution noise substitution was used to correct for the effects of soft masking on the FSC curves. Local resolution was determined using ResMap [19].

VacA SUV samples

5 μL of VacA SUV samples was applied to glow discharged (5mA 30s) ultra thin carbon-coated Quantifoil R2/2 200 mesh copper grids (Electron Microscopy Sciences Q250CR2-2nm) with a wait time of 30 seconds. The sample was blotted 1 time for 11 seconds with a blot force 5 and then plunged into liquid ethane using a Vitrobot Mark IV set at 100% humidity/4˚C. 3,800 VacA wild-type SUV movies were recorded (0.91Å/px) using Leginon on the 200keV Arctica microscope equipped with a Gatan K2 detector with a total dose of ~60e-/Å^2 with a -2 to -3.5μm defocus range[6]. 753 VacA Δ6-27 SUV movies were recorded using Leginon on the 200 keV Glacios microscope equipped with a Gatan K2 detector with a total dose of ~60e-/Å^2 at 0.98Å/px[6]. All frames were aligned using MotionCor2[14]. CTFFIND4 (for wild-type) or gCTF (for Δ6-27) were used to estimate CTF values[7, 11]. 1158 VacA wild-type SUV particles and 2643 VacA Δ6-27 SUV particles were selected manually in RELION 3.0.8, followed by particle extraction (VacA wild-type box size 364Å, VacA Δ6-27 box size 490Å)[10]. 2D

classification of VacA wild-type SUV particles was performed using cisTEM[17]. 2D classification of VacA Δ6-27 SUV was performed using cryoSPARC v2[12].

VacA LMNG micelle

3.5 µL of 75µg/mL VacA LMNG micelle sample was applied to glow discharged (5mA, 30s) Quantifoil R1.2/1.3 200 mesh copper grids (Electron Microscopy Sciences Q250-CR1.3). The sample was blotted 1 time for 1 second with a blot force 1 and then plunged into liquid ethane using a Vitrobot Mark IV set at 100% humidity/22˚C. 2,611 VacA LMNG micelle untilted and 816 30˚-tilted movies were recorded (1.01Å/px) on the 300keV Titan Krios microscope equipped with a Gatan K2 detector with a total dose of ~60e-/Å^2 and -2 to -3.5µm defocus range. All frames were aligned using MotionCor2[14]. CTFFIND4 was used to estimate CTF values[11]. 161,996 untilted VacA LMNG micelle particles and 69,643 30˚-tilted VacA LMNG micelle particles were picked with template-based picking in RELION 3.0.8, followed by particle extraction (VacA LMNG micelle particle box size 484.8Å)[10]. 2D classification of untilted (final particle count 73,576) and 30˚-tilted (final particle count 13,843) VacA wild-type LMNG particles was performed using cryoSPARC v2[12].

Membrane-extracted VacA DDM micelle samples

3.5µL of eluted VacA DDM micelle (wild-type or Δ6-27) sample (final DDM concentration 0.02%) was applied 2 times to glow discharged (5mA, 30s) ultra-thin carbon coated Quantifoil R 1.2/1.3 200 mesh copper grids (Electron Microscopy Sciences Q250CR1.3-2nm) with a wait time of 30 seconds for each application. The

sample was blotted 1 time for 4.5 or 5.5 seconds with a blot force 1 and then plunged into liquid ethane using a Vitrobot Mark IV set at 100% humidity/4°C. 739 untilted VacA wild-type DDM micelle, 6192 40°-tilted VacA wild-type DDM micelle, and 737 30°-tilted VacA Δ6-27 DDM micelle sample movies were recorded using Leginon on the 200keV Glacios microscope equipped with a Gatan K2 detector (0.98Å/px) with a total dose of ~60e-/Å^2 and a -2 to -3.5μm defocus range[6]. All frames were aligned using MotionCor2 (VacA wild-type: untilted/40°-tilted; VacA Δ6-27 30°-tilted) or cryoSPARC v2 Patch Motion (VacA wild-type:40°-tilted)[12, 14]. gCTF (VacA wild-type: untilted/40°-tilted; VacA Δ6-27 30°-tilted), goCTF (VacA wild-type 40°-tilted), or Patch CTF (VacA wild-type 40°-tilted) were used to estimate CTF values[7, 12, 20].

91,311 untilted VacA DDM micelle particles, 69,643 40°-tilted VacA DDM micelle, and 44,694 30°-tilted VacA Δ6-27 DDM micelle particles were picked with template based picking (wild-type) or Blob picker (Δ6-27) in cryoSPARC v2, followed by particle extraction (box size 588Å)[12]. 2D classification of untilted VacA wild-type DDM micelle particles (final particle count hexamer: 8171 heptamer: 5357), 40°-tilted VacA wild-type DDM micelle particles (final particle count hexamer: 20130 heptamer: 20006), and 30°-tilted VacA Δ6-27 DDM micelle particles (726) was performed using cryoSPARC v2[12]. Ab initio models were generated with cryoSPARC v2 with sorted VacA wild-type DDM micelle hexamer and heptamer particles, followed by C1 and C6/C7 non-uniform 3-D refinement for untilted and 40°-tilted datasets[12]. For the VacA wild-type DDM micelle 40°-tilted particles, ab initio models were generated using multiple classes to generate good and junk models for further use of heterogeneous refinement in cryoSPARC v2 to sort particles in 3D, resulting in less biased particle sorting for further C1 and C6 non-

146

uniform refinement[12]. VacA wild-type DDM micelle untilted and tilted particles were combined in different ratios (with "Minimize over per-particle scale" and "Use scales from current alignment in reconstruction turned on in all refinements due to different contrast between the tilted and untilted data) to test whether this helped improve the quality of the maps. Maps were analyzed in ChimeraX[13, 21].

Acid-activated VacA wild-type/VacA Δ346-347 p88 monomers/VacA Δ334-343 soluble oligomers

3 μL of 100μg/mL acid-activated wild-type VacA s1m1, VacA Δ346-347, or VacA Δ334-343 was applied to glow discharged (5mA, 30s) Quantifoil R1.2/1.3 200 mesh copper grids (Electron Microscopy Sciences Q250-CR1.3). The sample was blotted 1 time for 1 second with a blot force 1 and then plunged into liquid ethane using a Vitrobot Mark IV set at 100% humidity/4˚C.

253 acid-activated wild-type VacA s1m1, 1,059 VacA Δ346-347, and 1,561 VacA Δ334-343 movies were recorded (1.01Å/px) using SerialEM on the 300keV Titan Krios microscope equipped with a Gatan K2 detector with a total dose of ~60e-/Å^2 and -2 to -3.5μm defocus range[22]. All frames were aligned using MotionCor2[14]. CTF estimation was performed on the images using gCTF[7]. 51,180 acid-activated wild-type VacA s1m1, 686,301 VacA Δ346-347, and 13,074 VacA Δ334-343 particles were picked with template-based picking in RELION 3.0.8, followed by particle extraction (150Å for acid activated wild-type VacA s1m1, 162Å for VacA Δ346-347, ~320Å for VacA Δ334-343) and 2D classification was performed using RELION and cryoSPARC

[10, 12]. Subsequent ab initio 3D reconstruction and homogeneous C1 3D refinement for acid-activated wild-type VacA s1m1 and VacA Δ346-347 was done in cryoSPARC.

MPER-Ec2-FluC:10E8

3.5µL of 0.5mg/mL MPER-Ec2-FluC:10E8 in 100mM NaCl, 10mM HEPES pH 7.5, 2mM DDM was applied to glow discharged (5mA 30s) Quantifoil R1.2/1.3 200 mesh copper grids (Electron Microscopy Sciences Q250-CR1.3). The sample was blotted 1 time for 3 seconds with blot force 1 and then plunged into liquid ethane using a Vitrobot Mark IV set at 100% humidity/22˚C. 1,979 movies were recorded (0.91Å/px) using Leginon on the 200keV Arctica microscope equipped with a Gatan K2 detector with a total dose of ~60e/$Å^2$ with a -1.0µm to -2.0µm defocus range[6]. All frames were aligned using MotionCor2[14]. CTFFIND4 was used to estimate CTF values for aligned micrographs[11]. 980,921 particles were picked using the crYOLO general model and analyzed with the crYOLO box manager using SBGrid[8, 9]. Particle extraction (with a 182Å box size) and 2D classification was performed using cryoSPARC v2[12]. 2-D class averages were generated using 878,537 particles. Particle dimension measurements done in Fiji[5].

MPER-AdiC:10E8

3.5µL of 0.1mg/mL MPER-AdiC-FluC:10E8 in 100mM NaCl, 10mM HEPES pH 7.5, 4mM DM was applied 2 times to glow discharged (5mA 30s) Quantifoil R1.2/1.3 200 mesh copper grids (Electron Microscopy Sciences Q250-CR1.3). The sample was blotted 1 time for 3 seconds with blot force 1 and then plunged into liquid ethane using a

Vitrobot Mark IV set at 100% humidity/22°C. 886 movies were recorded (0.91Å/px) using Leginon on the 200keV Arctica microscope equipped with a Gatan K2 detector with a total dose of ~60e/Å^2 with a -2.0μm to -3.0μm defocus range[6]. All frames were aligned using MotionCor2[14]. CTFFIND4 was used to estimate CTF values for aligned micrographs[11]. 300,568 particles were picked using the cryoSPARC v2 Blob picker[12]. Particle extraction (with a 182Å box size) and 2D classification was performed using cryoSPARC v2[12]. 2-D class averages were generated using 81,676 particles. Particle dimension measurements done in Fiji[5].

Model building, structure refinement and validation of soluble VacA hexamer.

The models of the VacA p55 domain (2QV3 [23]) and Δ346-347 VacA non-oligomerizing mutant [1] were docked into one protomer of the 3.8 Å VacA hexamer cryo-EM density map using the Phenix Dock in map tool and coordinates were exported to COOT [24]. The model was then extended in COOT [24]. The final model was obtained by iterative cycles of manual model refinement in COOT and refinement with Real-space Refinement in the Phenix suite of programs [25, 26]. During refinement, the PDB 2QV3 was used as a reference model. Secondary structure was predicted from the poly-alanine model using Phenix.secondary_structure_restraints with the CaBLAM search method [27] and used as refinement restraints. Phenix Map Symmetry was used to determine the symmetry of the map. The Phenix Apply NCS operators tool was used to transform the p88 model into the six arms of the hexamer to generate a complete structure. A final Real-space Refinement was completed on the hexamer model. Validation of the model was performed using the Comprehensive Validation (cryo-EM)

tool in Phenix. FSC curves were calculated for the model in relation to half-map 1 and compared with that for the summed map. Programs used for structure determination and refinement were accessed through SBGrid [8]. Figures were prepared with Chimera, and ChimeraX [13, 28].

References

[1] Gonzalez-Rivera C, Campbell AM, Rutherford SA, Pyburn TM, Foegeding NJ, Barke TL, et al. A Nonoligomerizing Mutant Form of Helicobacter pylori VacA Allows Structural Analysis of the p33 Domain. Infect Immun. 2016;84:2662-70.

[2] Cover TL, Hanson PI, Heuser JE. Acid-induced dissociation of VacA, the Helicobacter pylori vacuolating cytotoxin, reveals its pattern of assembly. J Cell Biol. 1997;138:759-69.

[3] de Bernard M, Papini E, de Filippis V, Gottardi E, Telford J, Manetti R, et al. Low pH activates the vacuolating toxin of Helicobacter pylori, which becomes acid and pepsin resistant. J Biol Chem. 1995;270:23937-40.

[4] McClain MS, Schraw W, Ricci V, Boquet P, Cover TL. Acid activation of Helicobacter pylori vacuolating cytotoxin (VacA) results in toxin internalization by eukaryotic cells. Mol Microbiol. 2000;37:433-42.

[5] Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012;9:676-82.

[6] Suloway C, Pulokas J, Fellmann D, Cheng A, Guerra F, Quispe J, et al. Automated molecular microscopy: the new Leginon system. J Struct Biol. 2005;151:41-60.

[7] Zhang K. Gctf: Real-time CTF determination and correction. J Struct Biol. 2016;193:1-12.

[8] Morin A, Eisenbraun B, Key J, Sanschagrin PC, Timony MA, Ottaviano M, et al. Collaboration gets the most out of software. Elife. 2013;2:e01456.

[9] Wagner T, Merino F, Stabrin M, Moriya T, Antoni C, Apelbaum A, et al. SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. Commun Biol. 2019;2:218.

[10] Scheres SH. RELION: implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol. 2012;180:519-30.

[11] Rohou A, Grigorieff N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J Struct Biol. 2015;192:216-21.

[12] Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods. 2017;14:290-6.

[13] Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci. 2018;27:14-25.

[14] Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, Agard DA. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods. 2017;14:331-2.

[15] Scheres SH. Semi-automated selection of cryo-EM particles in RELION-1.3. J Struct Biol. 2015;189:114-22.

[16] Kimanius D, Forsberg BO, Scheres SH, Lindahl E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. Elife. 2016;5.

[17] Grant T, Rohou A, Grigorieff N. cisTEM, user-friendly software for single-particle image processing. Elife. 2018;7.

[18] Scheres SH. Processing of Structurally Heterogeneous Cryo-EM Data in RELION. Methods Enzymol. 2016;579:125-57.

[19] Swint-Kruse L, Brown CS. Resmap: automated representation of macromolecular interfaces as two-dimensional networks. Bioinformatics. 2005;21:3327-8.

[20] Su M. goCTF: Geometrically optimized CTF determination for single-particle cryo-EM. J Struct Biol. 2019;205:22-9.

[21] Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, et al. UCSF ChimeraX: Structure Visualization for Researchers, Educators, and Developers. Protein Sci. 2020.

[22] Mastronarde DN. Automated electron microscope tomography using robust prediction of specimen movements. J Struct Biol. 2005;152:36-51.

[23] Gangwer KA, Mushrush DJ, Stauff DL, Spiller B, McClain MS, Cover TL, et al. Crystal structure of the Helicobacter pylori vacuolating toxin p55 domain. Proc Natl Acad Sci U S A. 2007;104:16293-8.

[24] Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004;60:2126-32.

[25] Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, et al. PHENIX: building new software for automated crystallographic structure determination. Acta crystallographica Section D, Biological crystallography. 2002;58:1948-54.

[26] Afonine PV, Poon BK, Read RJ, Sobolev OV, Terwilliger TC, Urzhumtsev A, et al. Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr D Struct Biol. 2018;74:531-44.

[27] Sobolev OV, Afonine PV, Adams PD, Urzhumtsev A. Programming new geometry restraints: parallelity of atomic groups. J Appl Crystallogr. 2015;48:1130-41.

[28] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004;25:1605-12.

Tables

Table 1. VacA hexamer data collection, refinement and validation statistics

Data collection and processing

Magnification	40109 (Polara)/ 50000 (Titan Krios)
Voltage (kV)	300 (Polara) /300 (Titan Krios)
Electron exposure (e⁻/Å²)	48
Defocus range (μm)	-1.5 to -3.5
Pixel size (Å)	1.25
Symmetry imposed	C6
Initial particle images (no.)	~500,000
Final particle images (no.)	~133,827
Map resolution (Å)	3.8
FSC threshold	0.143

Refinement

Initial model used (PDB code)	2QV3
Model resolution (Å)	3.2
FSC threshold	0.143
Model Composition	
Chains	6
Non-hydrogen atoms	28116
Protein residues	4158
Bonds (RMSD)	
Length (Å) (# > 4σ)	0.007 (0)
Angles (°) (# > 4σ)	1.255 (45)

Validation
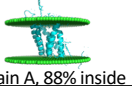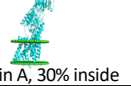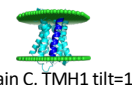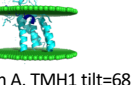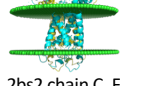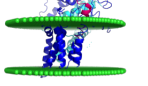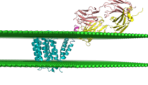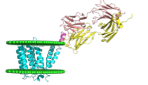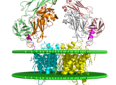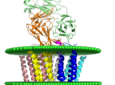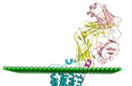
EMRinger Score	1.40
MolProbity score	2.67
Clash score	16.55
Ramachandran plot (%)	
Outliers	0.00
Allowed	21.54
Favored	78.46
Rotamer outliers (%)	1.79
Cβ outliers (%)	0.00
Peptide plane (%)	

Cis proline/general	0.0/0.1
Twisted proline/general	0.0/1.2
CaBLAM outliers (%)	12.92
ADP (B-factors)	
Iso/Aniso (#)	28116/0
Protein (Min/max/mean)	14.91/150.72/69.76
Occupancy	
Mean	1.00
occ = 1 (%)	100.00
0 < occ < 1 (%)	0.00
occ > 1 (%)	0.00

Table 2. MPER-Ec2-10E8v4 data collection and refinement statistics

	MPER-Ec2-10E8v4
Data collection	
Space group	P32
Cell dimensions	
a, b, c (Å)	99.1, 99.1, 167.6
α, β, γ (°)	90, 90, 120
Resolution (Å)	54.14-3.2(3.43-3.2)
Data Completeness	99.9 (100)
R_{merge}	0.144 (1.08)
R_{meas}	0.152 (1.14)
Mn $I / \sigma I$	14.8 (3.5)
Multiplicity	19.3 (20.4)
Refinement	
Resolution (Å)	34.0-3.2
No. reflections	26, 788
R_{work} / R_{free}	24.0 / 30.5
Ramachandran Favored	89.3
Ramachandran Outliers	2.0
Clashscore	23.1
R.m.s. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	1.01

Table 3. Pipeline for modeling MPER-membrane protein fusions and antibody fragment binding.

Analysis phase	Step	Number of proteins	%	Example of proteins that	
				satisfy conditions	do not satisfy conditions
Selection of dataset	PDB-TM database	1094			
	Alpha-helical TM proteins minus theoretical models	934			
	Molecular weight of TM biounit* <=100 kDa	457			
	>= 2 TM helices	373			
Analysis of MPER tag acceptance	50% of biounit residues inside the membrane	304	100%	6ffv chain A, 88% inside	4bbj chain A, 30% inside
	Tilt of TMH1 <=45°	294	96.7%	4cad chain C, TMH1 tilt=11°	6ebu chain A, TMH1 tilt=68°
	Align 10e8/4e10+ MPER tag to TMH1 (n, n+1,n+2,n+3), add tag to identical proteins in biounit	294	96.7%	1xio align MPER to **residue 3** residue 4	residue 5 residue 6
	Can accept MPER tag (no clash with biounit 3A Ca-Ca;)	290	95.4%	2bs2 chain C, F	4huq chain S, T
Analysis of Antibody acceptance	Membrane extended to include the length of antibody	290	95.4%	3v5s chain A, 4e10	3v5s chain A, 4e10
	Tolerate 4e10 or 10e8 Fv/Fab (no clash 5A Ca-Ca)	97	31.9%	5y78 chains A, B , 4e10	1oed chain A
	Tolerate 4e10 or 10e8 Fv/Fab (minor clash)	65	21.4%	3aym chain A, 4e10	4pgr chain A

* 49 proteins had only one chain in pdb, but n-homo-mer in biological unit. They were examined manually. Five of them (3zux, 6nwd, 1fx8, 5guf,5xls) could be in monomeric state and were considered monomer in this study.

Table 4. Models for MPER-membrane protein fusions and antibody binding analyzed according to the sequence position of the first residue of TM helix 1.

Step	TMH1 start at residue position		
	1-25	26-50	>50
50% protein inside membrane	280	20	4
Tilt of TMH1 <=45°	270	20	4
Can accept MPER tag	266	20	**4**
Tolerate 4e10 or 10e8 Fv/Fab	87	9	1
Tolerate 4e10 or 10e8 Fv/Fab (minor clash)	61	4	0

Table 5. Sequences of MPER constructs

His-Tag   MPER epitope   Thrombin-site

MPER-Fluc-Bpe	GHHHHHHGGLVPRGSASLWNWFDITNWLWYIKN LIAIGIGATLGAWLRWVLGLRLNGAGWPWGTLTAN LVGGYLIGVMVALIASHPEWPAWIRLAAVTGFLGG LTTFSTFSAETVDMLERGVYATAAAYAGASLAGSL AMTGLGLATVRLLLR
MPER-Fluc-Ec2	GHHHHHHGGLVPRGSASLWNWFDITNWLWYIKN LIAVIIGGSVGCTLRWLLSTRFNSLFPNLPPGTLVV NLLAGLIIGTALAYFLRQPHLDPFWKLMITTGLCGG LSTFSTFSVEVFALLQAGNYIWALTSVLVHVIGSLIM TALGFFIITILFA
MPER-AdiC	MHHHHHHSGGLVPRGSGSASLWNWFDITNWLWY IKNLTLMVSGNIMGSGVFLLPANLAATGGIAIYGWL VTIIGALALSMVYAKMSSLDPSPGGSYAYARRCFG PFLGYQTNVLYWLACWIGNIAMVVIGVGYLSYFFPI LKDPLVLTLTCVAVLWIFVLLNIVGPKMITRVQAVAT VLALVPIVGIAVFGWFWFKGETYMAAWNVSGMNT FGAIQSTLNVTPMVIHRGRKRLRCGRCGKKTRNAT
MPER-GlpF	GHHHHHHGGLVPRGSASLWNWFDITNWLWYIKN LIAEFLGTGLLIFFGVGCVAALKVAGASFGQWEISVI WGLGVAMAIYLTAGVSGAHLNPAVTIALWLFACFD KRKVIPFIVSQVAGAFCAAALVYGLYYNLFFDFEQT HHIVRGSVESVDLAGTFSTYPNPHINFVQAFAVEM VITAILMGLILALTDDGNGVPRGPLAPLLIGLLIAVIG ASMGPLTGFAMNPARDFGPKVFAWLAGWGNVAF TGGRDIPYFLVPLFGPIVGAIVGAFAYRKLIGRHLPC DICVVEEKETTTPSEQKASL
10E8-scFv	GWSCIILFLVATATGVHSEVRLVESGGGLVKPGGS LRLSCSASGFDFDNAWMTWVRQPPGKGLEWVGR ITGPGEGWSVDYAESVKGRFTISRDNTKNTLYLEM NNVRTEDTGYYFCARTGKYYDFWSGYPPGEEY FQDWGQGTLVIVSSASGGGGSGGGGSGGGGGSG GSGWSCIILFLVATATGSVTASELTQDPAVSVALKG TVTITCRGDSLRSHYASWYQKKPGQAPVLLFYGK NNRPSGIPDRFSGSASGNRASLTITGAQAEDEA DYYCSSRDKSGSRLSVFGGGTKLTVLSQPKAAEN LYFQSHHHHHH

Variable-Heavy Chain   Variable-Light Chain Linker   TEV |