Structure determination of membrane-located complexes: Aquaporin 8 and YscC secretin

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Julia Kowal

aus Kraków, Polen

Basel, 2011

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Andreas Engel und Prof. Dr. Henning Stahlberg

Basel, den 14.12.2010

<div align="right">

Prof. Dr. Martin Spiess

Dekan

</div>

creative commons

Attribution-Noncommercial-No Derivative Works 2.5 Switzerland

You are free:

to Share — to copy, distribute and transmit the work

Under the following conditions:

Attribution. You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).

Noncommercial. You may not use this work for commercial purposes.

No Derivative Works. You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work. The best way to do this is with a link to this web page.

- Any of the above conditions can be waived if you get permission from the copyright holder.

- Nothing in this license impairs or restricts the author's moral rights.

Quelle: http://creativecommons.org/licenses/by-nc-nd/2.5/ch/deed.en   Datum: 3.4.2009

Table of contents

List of figures

List of tables

1 Introduction

1.1 Biological membrane and membrane proteins

(...) the chromocytes [erythrocytes] *of different animals are covered by a layer of lipoids* [lipids] *just two molecules thick. (...) We therefore suspect that every chromocyte is surrounded by a layer of lipoids, of which the polar groups are directed to the inside and to the outside* [of the cell], *in much the same way as Bragg (1) supposes the molecules to be oriented in a 'crystal' of a fatty acid, and as the molecules of a soap bubble are according to Perrin (2). On the boundary of the two phases, one being the watery solution of hemoglobin, and the other the plasma, such an orientation seems a priori to be the most probable one.*

[Gorter, E. and F. Grendel, On Bimolecular Layers of Lipoids on the Chromocytes of the Blood. 31 Mar 1925, [1]]

In 1925, two Dutchmen, Gorter and Grendel extracted lipids from red blood cells, calculated the total area of a membrane they had purified, created the monolayer out of isolated membranes and finally measured the monolayer area. They observed that area of the monolayer was twice the estimated area of original membranes and thus concluded that the cell membrane is composed of two layers of lipids. Additionally, they correctly predicted the phospholipid molecules orientation in the bilayer.

The biological membrane is key to the life of the cell. Membranes are the sheet-like structures made mainly of lipids and proteins. All cells are surrounded by plasma membrane that is 6 to 10 nm thick. Moreover, eukaryotic cells possess additional internal membranes around several organelles. Proteins are crucial for both membrane stability and function.

In 1972, Singer and Nicolson proposed "the fluid mosaic model" describing the structure of biological membrane [2]. The model included aspects such as organization of lipid bilayers, diffusion of proteins and lipids freely within the plane of bilayer (membranes are two-dimensional fluids), embedding proteins into the lipid layers (mosaic formation) and an asymmetric distribution of proteins and lipids in the bilayer. However, years of membrane studies changed the general view in the membrane architecture [3]. It was verified that the membrane, which composition is not random, is rather mosaic than fluid. Besides, it was observed that membranes are patchy, with segregated regions of

structure and function. The next correction of Singer-Nicolson model concerned lipid regions that vary in composition and thickness. Within the bilayer, the clustering lipids form rafts that move around inside the cell and function as platforms for the attached or integrated proteins [4]. Additionally, the crowdedness of biological membranes seems to be much higher than assumed and the proteins inserted into the bilayer often have small-spanning domains but huge ectodomains, which are limiting exposure of lipid to the aqueous regions (Figure 1.1). To summarize, the mobility of proteins in a lipid bilayer depends on lipid composition, thickness, protein crowding, and interactions with cytoskeleton and other proteins.

Plasma membrane has many functions such as responding to external signals and transmitting them into the cell, providing a barrier to the water soluble molecules, transport of molecules via channels, cell-to-cell communication, separating cellular reactions by compartments and creating ion gradients between them, which are used for converting an energy and signal transduction.

The biological membranes, depending on the cell or organelle specialization, vary in composition of the proteins that are scattered within lipid bilayer (Table 1.1), which are constantly in motion and perform most of the membrane functions.

TYPE OF MEMBRANE COMPONENT	SHORT DESCRIPTION	EXAMPLE
Transmembrane proteins (Integral proteins)	Proteins have hydrophobic membrane-spanning domain in-between the lipids and hydrophilic cytosolic and extracellular domains that interact with internal and external molecules. The membrane-spanning domain is made of α-helices and/or β-sheets motives.	Aquaporins, G protein-coupled receptors, ion channels, proton pumps
Lipid anchored proteins	Proteins covalently bound to the lipid molecules, which are inserted into the membrane; function as molecular switches.	G proteins (guanine nucleotide-binding proteins)
Peripheral proteins	Proteins attached temporarily to the transmembrane proteins or associated with peripheral regions of membrane. After dissociation, function in cytoplasm.	Some hormones and enzymes
Phospholipids	Class of lipids that is a major component of lipid bilayer. Typically contain a phosphate group, diglyceride, and a simple organic molecule (i.e. choline). Zwitterionic.	Phosphatidylcholine (lecithin) (PC) Phosphatidylethanolamine (cephalin) (PE) Phosphatidic acid (PA) Phosphatidylserine (PS) Sphingomyelin

Glycolipids	Membrane component composed of lipids that are covalently bonded to monosaccharide or polysaccharides. They function as energy sources and cellular markers.	Glyceroglycolipids Glycosphingolipids - Cerebrosides - Gangliosides
Cholesterol	Sterol that stiffens the bilayer, decreases its permeability, regulates the activity of some integral proteins.	In parallel, component of bile acids, steroid hormones (glucocorticoids, estrogens, androgens and progestagens) and fat-soluble vitamins (A,D,E,K)

Table 1.1: Main membrane components.
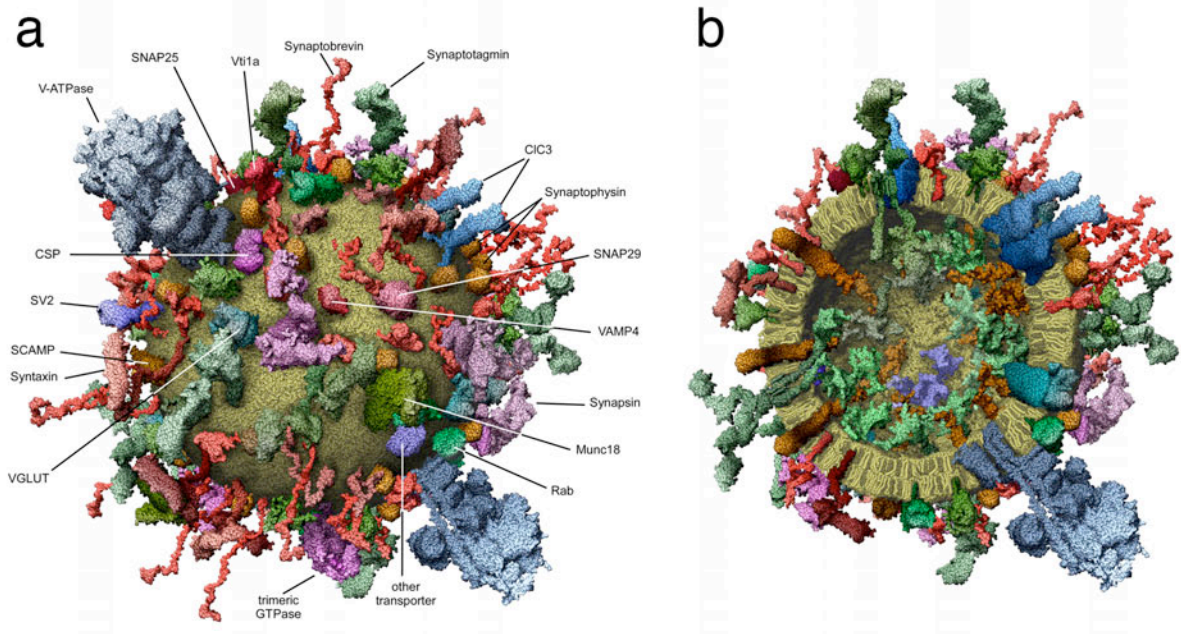
Figure 1.1: Molecular model of the synaptic vesicle with inserted proteins. The macromolecules are presented at near atomic resolution. (a) Overview of the vesicle. (b) View of the vesicle cross-section. Lipids are colored in *green* and darker fragments of the bilayer correspond to the cholesterol molecules. Images were taken from [5].

1.2   Aquaporins

The New York Times (January 26, 2009) [6]: *YOU WON THE 2003 NOBEL CHEMISTRY PRIZE FOR YOUR DISCOVERY OF AQUAPORINS. WHAT EXACTLY ARE THEY?*

Peter Agre: The plumbing system for cells. Every cell in our body is primarily water. But the water doesn't just sit in the cell, it moves through it in a very organized way. The process occurs rapidly in tissues that have these aquaporins or water channels.
Until 1985, when my lab found the protein they are made of, aquaporins hadn't yet been identified. There had been a controversy in biology for more than 100 years about how water moved through cells. The assumption was that it somehow leaked through the cell membrane. And indeed some water moves that way. But the very rapid movement of water through some cells was not explained by this theory.

The homeostasis of water, the most abundant molecule in living organisms, is crucial for physiology of all cells. Biological membranes have a defined lipid composition. To pass the lipid bilayer barrier water molecule needs activation energy of 10 - 20 kcal/mol. Nevertheless, in red blood cells, renal tubules or secretory glands incomprehensibly high-rates of water transport were observed. The first proof of the existence of a water channel, which reduce the energy barrier in a cell to <5 kcal/mol, was given in 1984 by the observation that organic mercurials decrease water permeability in red blood cells. Additionally, the low channel permeability of H^+ ions was explained by a specific fixed charge in the channel or the fact that water molecules within the channel do not form a continuum [7]. In 1992, it was discovered that oocytes from *Xenopus laevis* injected with in vitro-transcribed CHIP28 RNA exhibited increased osmotic properties due to water permeability. Experiments showed that the volume of the oocytes was increasing due to water uptake when incubated in hypoosmotic solution and proved that CHIP28, later renamed to Aquaporin-1 (AQP1), is the water channel protein in red blood cells [8]. AQP1, occurring at a level of ~120000-160000 copies per cell, facilitates the water transport at a rate of 3 billion water molecules per second. Such transport is bidirectional accordingly to the predominant osmotic gradient. The Cys189 was identified as the Hg^{2+} inhibitory site [9].

Taking into account the myriad of physiological processes connected with water transport and diseases (nephrogenic diabetes insipidus [AQP2], congenital cataract [AQP0], onset obesity [AQP7]), the discovery of aquaporins was a breakthrough in physiology, biology and medicine. Peter Agre was awarded the 2003 Nobel Prize in Chemistry for his discovery of aquaporins.

1.2.1 Function and structure of Aquaporins

Aquaporins were found in animal, plant, fungi and bacteria kingdoms of life. Ubiquitous process of aquaporin-mediated water transport across biological membranes is possible because of the special protein structure created by the fusion and duplication of 3 transmembrane helices [10]. Aquaporins control flux of water out and into the cells. Several functions are attributed to aquaporins. Some aquaporin family members, called aquaglyceroporins, facilitate transport of glycerol, urea, carbon dioxide and arsenite in addition to water [11-13]. In prokaryotic organisms water channels are believed to protect against osmotic shock and sudden freezing, in eukaryotes aquaporins perform a number of physiological functions including concentrating the urine in kidneys, keeping water homeostasis in brain, excreting of moisture from sweat, tear and salivary glands, controlling glycerol uptake into the cell, supporting the buoyancy of fish-eggs in seawater and controlling opening and closing of flower petals [14].
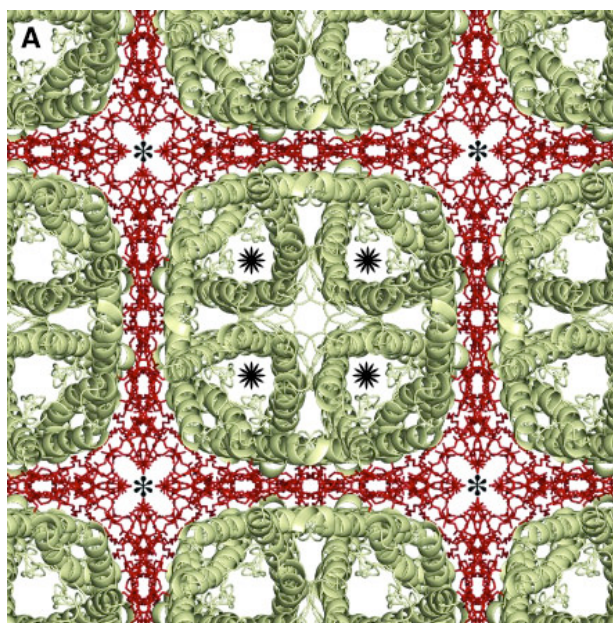
Figure 1.2: Image shows top view of the AQP0 2D crystal in the *E. coli* polar lipid bilayer at 2.3 Å resolution [15]. AQP0 tetramers are colored in *green* and the surrounding *E. coli* polar lipids in *red*. The AQP0 channels were marked with stars. The lipids in the central area between four adjacent tetramers marked with asterisks (*) were not defined.

Aquaporins share the characteristic basic structure and features. In the biological membrane they form tetramers with each monomer functioning as a separate pore (Figure

1.2). The water is transported via pore using the osmotic gradient. Several studies focused on water transport through aquaporins, without passage of protons, were performed up to now. The high-resolution structure of AQP1 solved by electron crystallography in 1997 [16, 17] revealed the "hourglass fold" predicted 5 years earlier by Jung *et al.* [18], which corresponds to the six transmembrane α-helices and seventh helix consisting of two half-helices formed by loops B and E (Figure 1.3 A, B). Both loops B and E, which fold into the channel from opposite sites of the membrane, contain the highly conserved NPA (Asn-Pro-Ala) sequence motif.

Proton conduction in a bulk water proceeds due to the Grotthuss mechanism, which explains that protons are transferred between water molecules via hydrogen bonds and transient hydronium ions. Surprisingly, the study of water channel structure revealed the interrupted hydrogen bonds in the channel water chain and strictly defined orientations of water molecules in the pore. The mechanism of proton exclusion in aquaporins was reported and confirmed by molecular dynamics simulations that made it possible to observe the transport of water molecules through the pore.

Since the α-helices have a natural dipole moment, the two half-helices in aquaporin oriented in opposite directions both position their positive dipole at the central NPA region of the pore. This creates the electrostatic barrier for the proton transfer, prevents water molecules from hydrogen bonding to adjacent water molecules, and induces alignment of the water molecules dipole moments (a rotation by 180°) as they pass the NPA aquaporin motif [19, 20].

At the extracellular side of the pore the constriction region is located, which acts as the water filter. This region consisting of aromatic/arginine (ar/R) motif is a selectivity filter (Figure 1.3 C) conserved in all aquaporins. The diameter and polarity of this region is decisive for transport of molecules other than water.
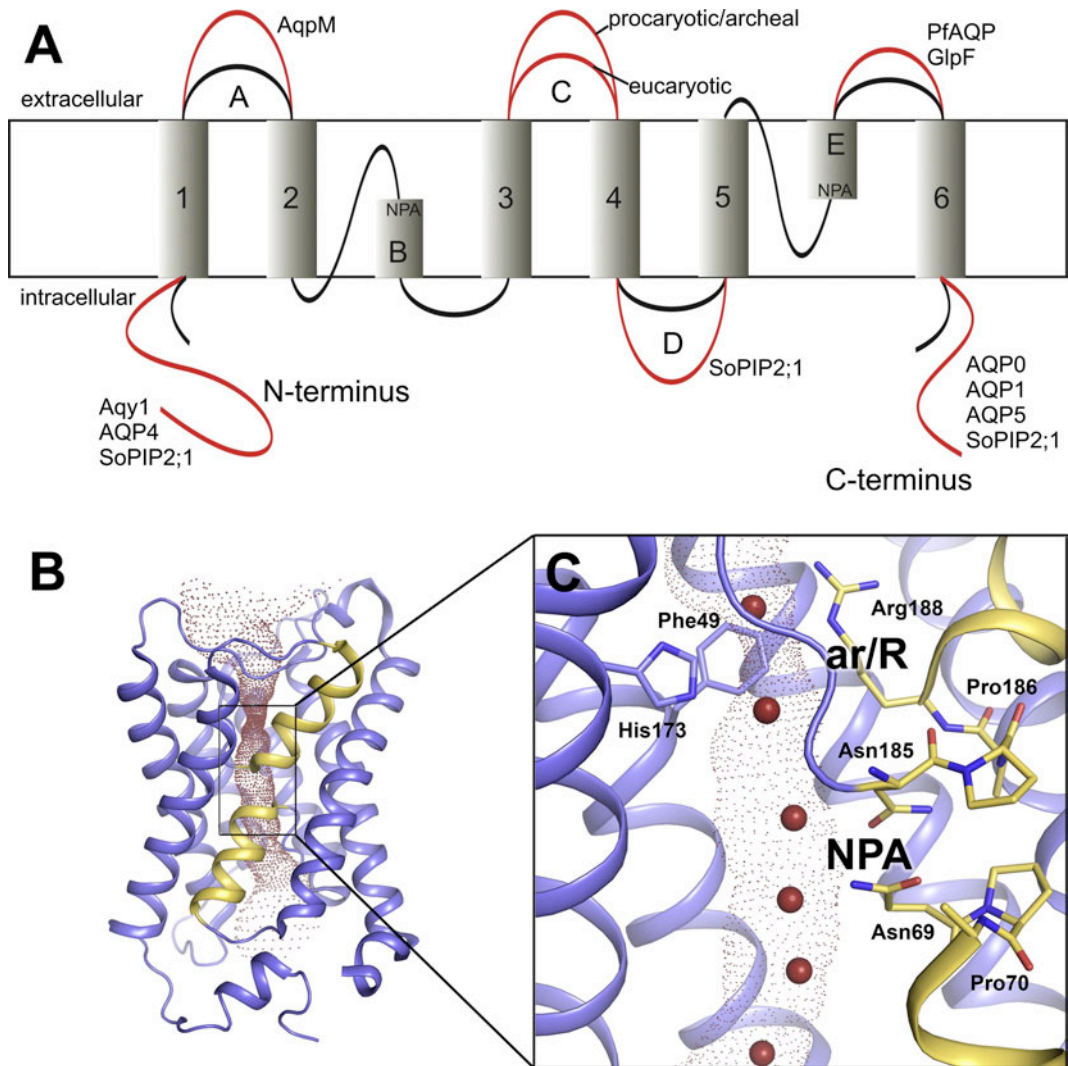
Figure 1.3: Structural overview of aquaporins [14]. (A) Topographic representation of aquaporin shows the basic fold and highlights unusually long intra- and extracellular regions (*red*). (B) Water permeation through structure of human AQP5 [21]. Six trans-membrane helices and two half-helices from loops B and E were highlighted in *blue* and *yellow*, respectively. (C) The aquaporin signature NPA-motif (asparagine, proline, alanine) and ar/R motif, which acts as selectivity filter, are shown in detail.

Aquaporins are regulated posttranslationally by gating, whereby the flux rate through the pore is regulated, or by trafficking, whereby proteins are transported from the cell to the membranes. The phosphorylation at serine and threonine residues [22] plays the role both in gating and trafficking. Besides, changes in pH, osmolality and divalent cations concentrations can also influence the water transport in different cells [23]. Two mechanisms of aquaporin gating were proposed: *capping* and *pinching* (Figure 1.4, [24]).

The *capping* was reported to occur i.e. in SoPIP2;1 aquaporin present in spinach leaf plasma membranes, which gates when loop D and N-terminus cap the pore from the cytoplasmic side and moves residues to block the channel [25]. The *pinching* relies on movement of few residues that pinches upon ar/R region and restricts the passage of water molecules, and was suggested for i.e. mammalian AQP0 from eye lens fibre cells [26]. Additionally, two forms of AQP0 were observed: closed, forming highly ordered stacks of lens junctions upon proteolytic cleavage of the cytoplasmic termini, and opened non-junctional forming the open water channel. Activity of AQP0, which have two histidine residues within pore, can be also regulated by changes of pH [23].

In plants it was reported that there are at least two forms of the *capping* gating mechanism. As a response to drought the dephosphorylation of two serine residues (Ser 115 and Ser 274) of SoPIP2;1 regulates the gating, whereas during flooding, channel is closed in response to the drop of cytosolic pH and thus protonation of a fully conserved histidine (His 193) in the cytosolic D-loop [27].

Human AQP2 is located in renal conducting duct. When the V2 receptor localized on the basolateral cell membrane binds hormone vasopressin, the AQP2 is phosphorylated by the cAMP dependent protein kinase A and transported to the apical plasma membrane [28].
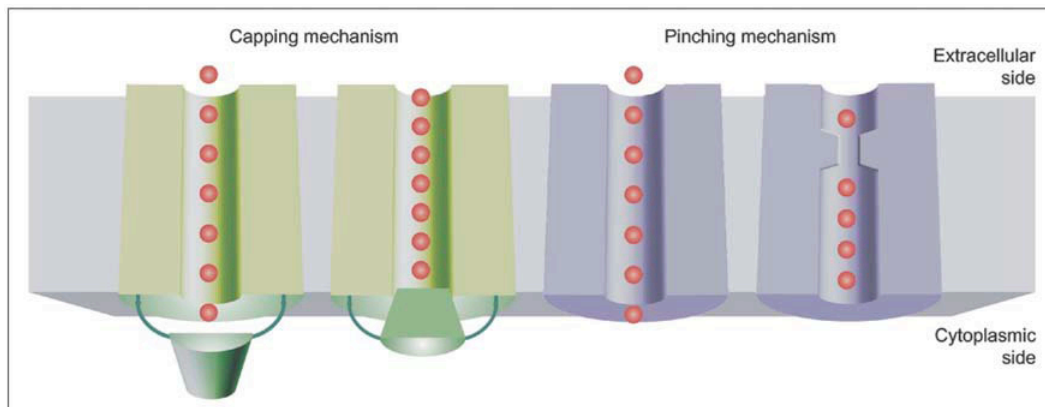
Figure 1.4: Images of the *capping* and the *pinching* gating mechanisms in SoPIP2;1 (*green*) and AQP0 (*purple*), respectively [24].

Thirteen aquaporin crystal structures solved by electron and X-ray diffraction at resolution better than 3.5 Å have been reported. These proteins are: ovine [26] and bovine [29] **AQP0**, human [30] and bovine [31] **AQP1**, human [32] and rat [33] **AQP4**, human

[21] **AQP5**, plant **SoPIP2;1** from spinach [25], bacterial [34] **AQPZ**, **GlpF** from *Escherichia coli* [35], **Aqy1** from yeast *Pichia pastoris* [36], archeal **AQPM** from *Methanothermobacter marburgensis* [37], and **PfAQP** from parasite *Plasmodium falciparum* [38].

To date, the thirteen mammalian aquaporins (AQP0-AQP12) were found, unique in their localization, regulation and trafficking. Channels AQP0, AQP1, AQP2, AQP4, AQP5, AQP6 and AQP8 belong to the group of aquaporins, while AQP3, AQP7, AQP9 and AQP10 are included in the group of aquaglyceroporins, permeable to small polar solutes such as glycerol and urea in addition to water (Figure 1.5). AQP11 and AQP12 belong to the unclassified aquaporin subfamily, not assigned to neither aquaporins nor aquaglyceroporins family. AQP6 and AQP8 are in aquaporin group only on the basis of the sequence analysis. The human aquaporins, AQP3, AQP7, AQP9 are permeable to urea, and AQP3, AQP7, AQP8, AQP9, and possibly AQP10, are permeable to ammonia. The new subgroup of NH_3-transporting aquaporins called aquaammoniaporins was recently suggested [39, 40]. The characteristic of AQP0-AQP12 is shown in Table 1.2.

AQUAPORIN	MAJOR SITES OF EXPRESSION	FUNCTION/COMMENTS	REFERENCES
AQP0	Eye: lens fiber cells	**Low water permeability** Fluid balance within the lens Forms membrane junctions upon proteolytic cleavage, acts as adhesion protein At neutral pH its permeability is 40 times lower than that of AQP1 Mutations cause congenital cataract pH, Ca^{2+}/calmodulin regulated	[26]
AQP1	Red blood cells, kidney: proximal tubule, eye: ciliary epithelium, brain, lung	**High water permeability** Carbon dioxide permeability Osmotic protection Concentration of urine Production of aqueous humor and cerebrospinal fluid Alveolar hydration state Readsorbs 90 % of the 180 l of water filtered by the kidney per day Permeates water molecules across the membrane at a rate of 3×10^9 s^{-1} per channel	[19, 41]
AQP2	Kidney: collecting ducts, vas deferens	**High water permeability** Stored in intracellular vesicles located throughout the cell Only aquaporin regulated by vasopressin Mediates antidiuretic hormone activity Mutations cause nephrogenic diabetes insipidus	[11] [28]
AQP3*	Kidney: collecting ducts, trachea: epithelial cells,	**High water and glycerol permeability, moderate urea and ammonia permeability**	[42]

	colon, skin, eye, lung	Reabsorbtion of water into blood Secretion of water into trachea pH regulated	
AQP4	Kidney: collecting ducts, brain, muscle, stomach, lung: bronchial epithelium	**High water permeability** Reabsorbtion of water Fluid balance Osmosensing function Bronchial fluid secretion Functions both as water channel and cell adhesion protein, forms junctions	[33]
AQP5	Stomach, duodenum, pancreas, airways, lungs, salivary glands, sweat glands, eyes, lacrimal glands, inner ear	**High water permeability** Production of saliva Production of tears	[21]
AQP6	Kidney	**High water and anions permeability** Localized intracellularly rather than at the plasma membrane Function as an ion channel pH regulated Water permeability is accompanied by an increased ion conductance with the following halide permeability sequence: $NO_3^->I^->Br^->Cl^->SO_4^{2-}$	[43, 44]
AQP7*	Fat cells, kidney, testis, sperm	**High water, glycerol, urea, ammonia and arsenite permeability** Responsible for glycerol release during starvation Has been proposed as a central agent in fat metabolism Mutations cause obesity	[45]
AQP8	Pancreas, testis, liver, kidney, gastrointestinal tract, airways, heart, salivary glands	**High water, ammonia and hydrogen peroxide permeability** Localized intracellularly rather than in the plasma membrane Neither urea nor glycerol can penetrate	[43, 46, 47]
AQP9*	Liver, leukocytes, brain, testis	**High glycerol, ammonia, urea and arsenite permeability, and low water permeability** Permits glycerol uptake for gluconeogenesis Permeable to olyols (e.g. glycerol and mannitol), purines and pyrimidines	[48]
AQP10*	Intestinal epithelial cells, duodenum, jejunum	**High glycerol and urea permeability, and low water permeability** Exists as two isoforms derived from two different transcripts, one is highly permeable to water, urea and glycerol, second shows only poor water permeability	[49, 50]
AQP11#	Kidney, liver, testes, brain	Does not transport water, urea, glycerol or ions	[51]
AQP12#	Pancreas	Localized in intracellular sites in acinar cells Unknown function	[52]

Table 1.2: Characterization of mammalian aquaporins. Aquaglyceroporins are highlighted with asterisks; pores not assigned to any of the groups are marked with a hash. Aquaporins with known structure are underlined.
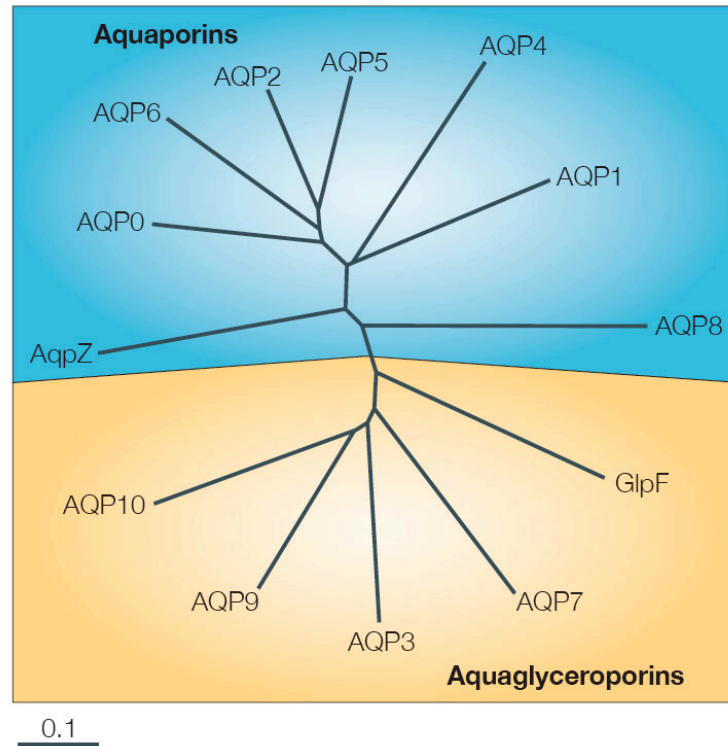
Figure 1.5: The aquaporin family tree [53]. Phylogenetic tree illustrates the relationship of the AQP0–AQP10 human aquaporins and the *E. coli* aquaporin homologues, AqpZ and GlpF. The aquaporins are typically permeated only by water and the aquaglyceroporins are permeated by water and small polar solutes. The scale bar represents the evolutionary distance: 0.1 equals 10 substitutions per 100 amino-acid residues.

1.2.2   The human Aquaporin 8

The human AQP8 cDNA has been cloned in 1998 [54] after its identification one year earlier in mouse [55] and rat [54, 56]. Human AQP8 shares only 30-35% of sequence homology to other aquaporins and a sequence identity of 75 and 76% with, respectively, rat and mouse orthologs. AQP8 sequence is the most similar to the wheat AQP TaTIP2;1, which is able to transport ammonium (Figure 1.6). Both channels have unusually long N-termini and shorter C-termini and their putative pores share most of the key residues

(His72, Ile198, Gly207, Cys208 and Arg213) [57]. By sequence alignment to AQP1 the supposed mercury-sensitive cysteine residue of AQP8 was determined: Cys208 in human, Cys209 in rat and Cys207 in mouse AQP8. On the phylogenetic tree of mammalian aquaporins AQP8 is at an equal distance from both groups of aquaporins and aquaglyceroporins (Figure 1.5). This divergence reveals uniqueness of AQP8 and suggests that it may belong to the new subgroup in mammalian aquaporins.

AQP8 is expressed in the inner mitochondrial membrane of several mammalian tissues as well as in liver and kidney where the urea and glycerol fluxes are essential for organism metabolism. In rat, the AQP8 was identified in the gastrointestinal tract, testis, airways and kidney cells, mainly in intracellular vesicles. In addition to AQP8, the only mammalian aquaporins reported to occur in intracellular vesicles are AQP2, AQP6 [43] and AQP12 [52]. This localization would suggest that AQP8 is not primarily involved in the water transport but rather in an osmoeqilibrium between vesicular and cytoplasmic compartments in a cell. Upon stimulation with the hormone glucagon AQP8 is redistributed to the canalicular membrane via a microtubule-depending, cAMP-stimulated mechanism [58].

In a liver, AQP8 appeared as glycosylated 34-kDa protein in intracellular vesicles and the apical plasma membrane domain [59], whereas in the mitochondrial inner membrane 28-kDa non-glycosylated form was found [47]. Interestingly, when AQP8 was expressed in yeast not only the 28-kDa AQP8 band appeared but an additional 26-kDa band was also identified suggesting that protein import into the yeast mitochondria is accompanied by a proteolysis of AQP8 N-terminus polypeptide [60].

In comparison to other aquaporins the less restrictive ar/R region of AQP8 allows polar solutes such as ammonia [40] and H_2O_2 [61] to permeate through the pore in addition to water, whereas neither urea nor glycerol can pass [54]. In the study of ammonia (NH_3) and ammonium (NH_4^+) permeability in aquaporin-expressing *Xenopus oocytes* it was observed that AQP8 support significant fluxes of ammonia and has physiological implications for liver and kidney function [62]. In parallel, ammonia transport was determined in stopped flow light scattering experiments using formamide as ammonia analog. It was observed that formamide transport in mitochondria increased 3 times in the presence of AQP8 [60]. Additionally, it was reported that the mitochondrial AQP8 has preference for ammonia over water molecules. The ammonia permeability through the pore was determined to be 8 x 10^5 NH_3 molecules/second/channel, twice higher compared to the permeability of water [63].

Ammonia transport facilitated by AQP8 might have significant physiological roles. AQP8 found in mitochondrial inner membrane and liver is speculated to be involved in cell-to-cell signaling, generation of reactive oxygen species [61] and urea cycle [62]. Moreover, it seems most likely that the diffusional transport of ammonia, which eliminates the severe neurotoxicity caused by hyperammonemia, could be a central function for AQP8 in mitochondria.

In addition to above observations, many AQP8 studies gave contradictory results. The evidence against functionally significant expression of AQP8 in rat brain, liver and kidney mitochondria was published [64]. Therefore, it seems clear that there are still uncertainties regarding the physiological role and characterization of AQP8 within aquaporin family.
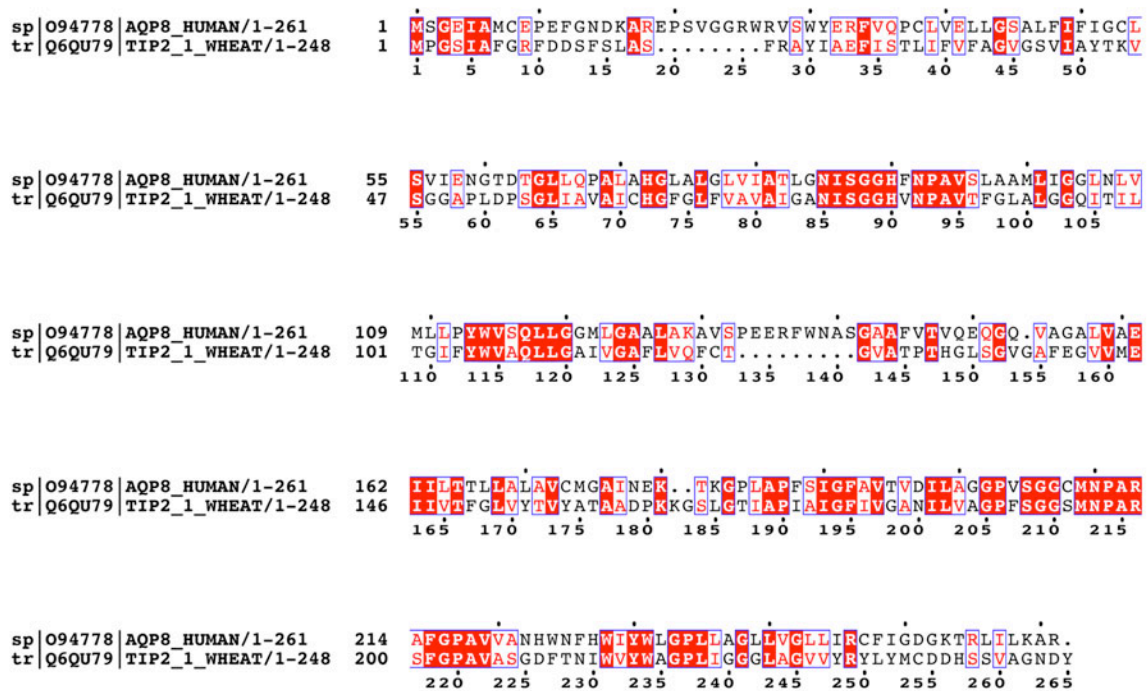
Figure 1.6: Alignment of two aquaporins, 27.4-kDa human AQP8 to 25.2-kDa wheat *Triticum aestivum* TIP2;1. The sequence alignment was done using ClustalW2 [65] (default settings, matrix Blosum30, number of iterations 10) and graphically illustrated by ESPript 2.2 [66]. Identical residues are highlighted in red boxes and similar residues are framed in blue.

1.3   Type III secretion system

The outer membrane of Gram-negative bacteria, besides the protection against harmful compounds like digestive enzymes, detergents and antibiotics in the extracellular environment, functions also as a barrier to uptake nutrients and secrete proteins. The transport of the hydrophilic molecules across the bacteria outer membrane is facilitated by a large variety of protein channels formed by substrate-specific porins and receptors. For the secretion of proteins Gram-negative bacteria have developed specific pathways that involve outer membrane proteins. In case of the type II and III secretion systems these outer membrane proteins are related and belong to the family of secretins, which includes also proteins involved in secretion of filamentous phages and assembly of type IV pili.

The type III secretion system (T3SS) is special nanoscale machinery by which Gram-negative bacteria mediate protein translocation into eukaryotic cells. It has been identified in more than 25 different species of pathogenic organisms including *Yersinia, Pseudomonas, Shigella, Salmonella, EPEC (enteropathogenic E. coli)* and is essential for their virulence [67]. In parallel, it was reported that *rhizobia*, a diverse group of soil bacteria which induce the formation of nitrogen-fixing nodules on the roots of legumes, clearly exploit pathogenic secretion system and effectors for symbiotic purposes [68].

T3SS is composed of a cylindrical basal structure spanning the inner and outer bacterial membranes and the peptidoglycan, into which a hollow external needle is embedded and an array of effector proteins. The T3S apparatus, closely related to the flagellum, is a functionally and structurally conserved macromolecular machine consisting of ~25 different proteins. Bacteria docked to the surface of target eukaryotic cells or enclosed in the vacuole, deliver during infection the effector proteins across the membrane to the host cytoplasm or at the cytosolic face of the membrane [69-71]. Effectors injected to the cell have an influence on the biochemical and regulatory processes conducive to infection. They interrupt signaling cascades leading to phagocytosis, disturb the dynamics

of the cytoskeleton, control pro-inflammatory responses, induce apoptosis, prevent autophagy and influence the intracellular trafficking [72, 73].

The *Yersinia pestis* bacterium has in the past caused the bubonic plague, called Black Death, which in years 1348-1350 spread throughout Central Asia, Mediterranean Region and Europe by fleas that made use of black rat (*Rattus rattus*). The pathogenicity of *Yersinia* resulted in unmatched devastation of people population in Europe (Figure 1.7). *Y. pestis* has imposing ability to overcome the mammalian host cell immune system and intensively multiply because of its pathogenic arsenal. Two another closely related food-borne pathogens *Yersinia enterocolitica* and *Yersinia pseudotuberculosis* responsible for diarrhea and tuberculosis-like symptoms in humans, respectively, share with *Y. pestis* the core of the pathogenicity: the Yop virulon, which allows extracellular *Yersinia* to attach to the host cell surface and inject effector proteins across the membrane [74].

Figure 1.7: The figure shows images that correspond to the plague caused by the Gram-negative bacterium *Yersinia pestis.* (a) *Doktor Schnabel von Rom* (*Doctor Beak from Rome*) engraved by Paul Fürst, Rome; 1656, [75]. A medical doctor equipped with wide-brimmed black hat, long waxed coat, wooden cane and a bird-like mask with red glass

eyepieces and a long beak with pleasantly smelling herbs and spices visited patients afflicted with the bubonic plague. (b) *Omne Bonum,* a Fourteenth-Century Encyclopedia of Universal Knowledge, compiled by Cistercian James le Palmer, London; 1360-1375, British Library. A scene showing infected monks being blessed by a priest. (c) Plaque in Weymouth commemorates the arrival of plague in 1348.

1.3.1   The needle complex

The T3S apparatus is made up of 3 major parts: a cytoplasmic region (C ring), a region spanning both bacterial membranes and the peptidoglycan and the large external needle. The `needle complex` (or injectisome basal body plus needle), corresponding to the transmembrane and extracellular parts, was first isolated from *Salmonella typhimurium* [76] and visualized by electron microscopy in 1998. In 2001 the three-dimensional reconstruction of *Shigella flexneri* needle complex (Figure 1.8) with resolution of 17 Å from negatively stained particles was performed [77]. It was found that the complex is made of two sets of different diameter rings and the needle is a helical polymer of diameter of 7 nm and inner channel of 3 nm. Besides, four major components of the *S. flexneri* basal body's transmembrane region were identified: two proteins in the inner membrane and a secretin in the outer membrane with attached periplasmic part of the needle [78, 79]. Cryo EM analysis and symmetry particle averaging of *S. typhimurium* needle complex [80, 81] extended the resolution, showed the inner rod with a central channel of 2-3 nm diameter connected with the hollow transmembrane chamber by the socket-like structure. The studies revealed that the secretion channel markedly narrows at the entry point to the needle. The rotational symmetry of the lower rings of the needle complex was determined to be 19-22-fold, in contrast to *S. flexneri* where side views of the needle complex correlated the best with 24-fold symmetry. The rotational symmetry of outer membrane secretins was reported to be 12-14-fold [82]. It was found that extracted needle complex structures have neither C ring, nor ATPase and transmembrane proteins that form the export apparatus, most probably because they are broken off during the purification processes.
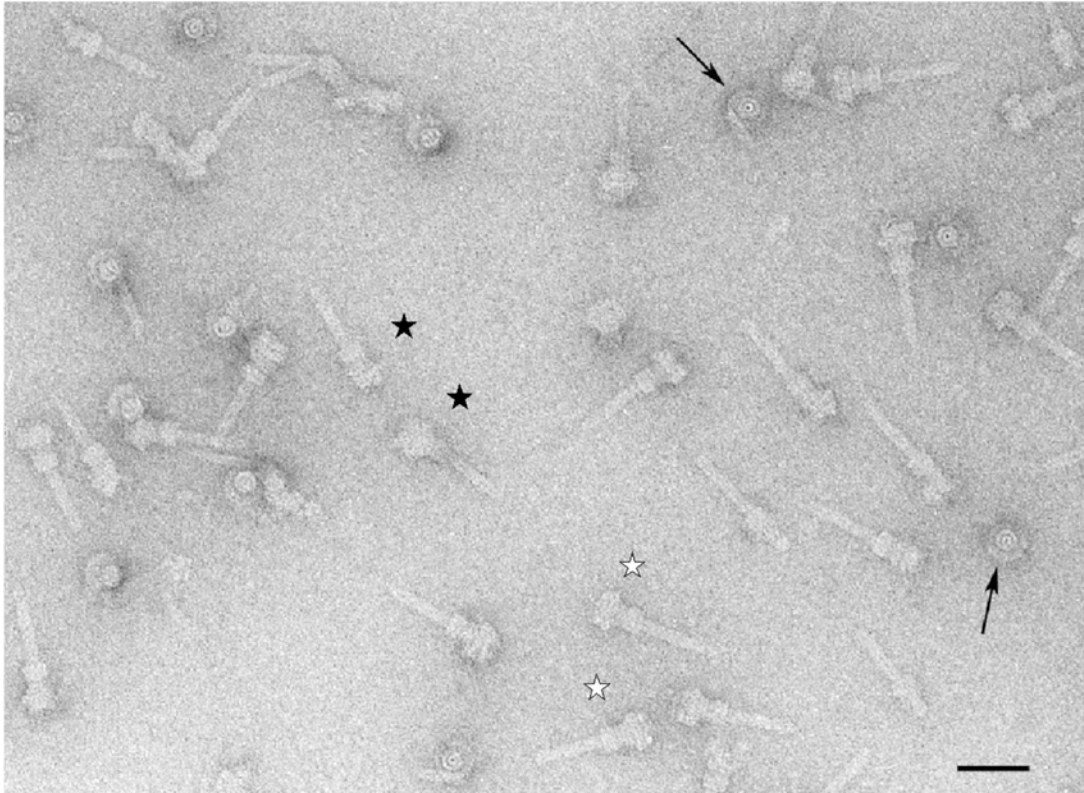
Figure 1.8: Uranyl-acetate stained *Shigella flexneri* needle complexes taken from [83]. Well-preserved needles with the outer and inner membrane rings are indicated with white stars. Damaged particles are highlighted with black stars and top views of bases without needles with arrows. Scale bar corresponds to 50 nm.

1.3.2   The YscC oligomer as part of *Yersinia enterocolitica* weaponry

The injectisome acts as the syringe through which the Yop (*Yersinia* outer proteins) effectors travel unfolded from the bacterial cytoplasm across the both membranes and peptidoglycan, via needle to the host cell cytoplasm (Figure 1.9).

In *Yersinia spp.,* the T3SS is encoded on a virulence plasmid. However, the secretion of Yops can be artificially controlled and induced by incubation of the bacteria at 37°C in media without Ca^{2+}. Then, the secreted Yops have no cytotoxic effect on cultured cells, suggesting that physical contact of *Yersinia* with the host cell and translocation of Yops into the host cytosol is obligatory for *Yersinia* pathogenesis *in vivo*. Several *Yersinia* Yop effectors were identified, i.e. YopH (phosphotyrosine phosphatase which dephosphorylate adhesion proteins), YopP/YopJ (which have an anti-inflammatory role and induce apoptosis of macrophages) or YopE and YopT (which inactivates GTPases that

control cytoskeleton dynamics) [74]. To inject the Yop effectors the physical contact of the needle with a host cell is necessary.

Approximately 25 types of proteins are needed to build the injectisome. They are structural components, molecules involved in the assembly process and function of the injectisome, i.e. molecular ruler and chaperones. In *Yersinia*, these proteins are called from YscA to YscY (**Y**op **sec**retion). The highly conserved injectisome, made of Ysc proteins, consists of cylindrical basal body submerged in both bacterial membranes joined by a central rod and connected to the hollow needle (Figure 1.10).

INJECTISOME COMPONENTS HOMOLOGUES	BACTERIA			
	Yersinia enterocolitica	*Shigella flexneri*	*Salmonella typhimurium*	*EPEC (enteropathogenic E. coli)*
	YscC (secretin)	MxiD	InvG	EscC
	YscW (pilotin)	MxiM	InvH	
	YscJ (inner ring)	MxiJ	PrgK	EscJ
	YscD (connector)	MxiC	PrgH	EscD
	YscF (needle)	MxiH	PrgI	EscF

Table 1.3: Injectisome components homologues.

Needle (needle: YscF, tip: LcrV, translocation pore: YopB, YopD, length control/molecular ruler: YscP)

The *Y. enterocolitica* injectisome needle is a hollow stiff tube 65±10 nm long and 7 nm thick, with an inner diameter of about 2.5 nm. The needle, which is formed by polymerization of 139±19 subunits of the 9-kDa-protein YscF with fixed twist of 5.6 subunits per turn, terminates with a bell shaped tip structure constructed by a homopentamer of the protein LcrV [84-87].

The needle does not puncture the target cell but upon the contact with the host membrane builds a translocation pore using two hydrophobic proteins YopB and YopD. One of the translocon proteins (YopB) interacts with the tip protein using its N-terminal domain, which influences a proper YopB insertion into the host membrane. The translocon, embedded in a membrane, is made of 6-8 subunits of both proteins as was revealed by EM and AFM [69, 88]. It was shown that in the absence of any of the

translocon components, Yops are secreted out of the bacteria but are not able to enter the host cell.

The structure of the needle-tip acts as an assembly platform for the translocation pore [89]. After assembly of the basal body, which is the core of an export machine, the needle subunits are exported. The same machinery changes the substrate specificity over the time and transfer sequentially translocation pore proteins and effectors.

To reveal the needle structure several studies were performed using the homologues needle subunits, MxiH (*S. flexneri*), BsaL (*Burkholderia pseudomallei*) and PrgI (*S. typhimurium*). They shed light onto the atomic resolution structure of 'building blocks' of the needle. All the needle components structures have the two-helix bundle coiled-coiled motif, which was docked into 16 Å EM density reconstruction of *S. flexneri* needle and provided atomic predictions of the helical needle structure [90].

The YscP protein acts as a molecular ruler that determines the length of the *Yersinia* injectisome needle. The ruler concept was taken from few known examples when one elongated protein determines the number of subunits of another protein polymerizing into a tubular structure, i.e. tail of bacteriophagues [91]. Similar mechanisms were proposed for the bacterial injectisome [92] and flagellar hook [93]. The protein YscP is the early substrate of the *Yersinia* export machinery, and is driven by two N-terminal export signal sequences. Since there is a linear dependence between lengths of the YscP helical domain and the formed needle, it was found that when the molecular ruler detects that needle has reached the final length its C-terminal domain switches the export machine substrate specificity to stop export of the YscF needle subunits [72, 92, 94].

Outer rings (ring: YscC, chaperone: YscW)

Secretins exist in the outer membrane as stable oligomeric complexes and were found in T3S and T2S system, type IV pilus and filamentous phage. The outer membrane of *Yersinia* is spanned by a triple ring of the YscC secretin, a unique membrane protein that forms the 12-14-meric ring structures of diameter ~110 Å [83, 95, 96]. Top view of the oligomer show large ring-like structures with an internal cavity that might function as an export channel. The C-terminus of YscC forms the trypsin-resistant domain, called secretin homology region, embedded in the outer membrane. The flexible N-terminus protrudes into the periplasmic space where it recognizes and binds inner membrane ring. Though C-terminal protease-resistant domain, predicted to be a β-barrel, is highly

conservative in all secretins, the N-terminal flexible domain shows much lower sequence homology. The low-resolution structure of PulD from *K. oxytoca* obtained in cryo EM revealed the approximate dimensions and localization of both domains [97]. Recently, the structure of N-terminal part of EscC, homolog of YscC, was solved [98].

For correct localization and assembly, the YscC oligomer requires a specific lipoprotein, pilotin YscW [99]. In 2005, the crystal structure of MxiM (YscW homologue) shed light on the assistance in proper insertion and assembly of the oligomer. It was revealed that pilotin binds to the C-terminus helix of the secretin [100], protects the secretin against proteolysis and mediates its stable insertion into the bacteria outer membrane. In a PulD secretin, the PulS pilot protein remains associated with the outer membrane protein after its insertion into the lipid bilayer. Without the pilotin the PulD secretin is assembled in the inner membrane what highlights the importance of the lipoprotein.

Inner rings (ring: YscJ, connector: YscD)

The inner membrane ring, called also MS (**m**embrane **s**panning) ring, constituted by the YscJ and YscD proteins (homologues: Table 1.3) exhibits a bigger diameter than the rings in the outer membrane. It was reported that the large periplasmic homologue lipoprotein MxiJ is embedded in the inner membrane with its lipidated N-terminus and single C-terminal transmembrane helix [101]. So far, only the structure of EscJ was solved. In the crystal structure the individual proteins formed 24-subunits helical assembly [102]. Besides, the structure was fitted to the *S. typhimurium* basal body density map indicating that inner ring is made of 24 EscJ. Additionally, the rotational symmetry experiments of *S. typhimurium* injectisome suggested for the inner ring (PrgK) a 19-22-fold symmetry, in contrast to the *S. flexneri* (MxiJ) where side views of the needle complex correlated the best with the 24-fold symmetry. The stoichiometry of PrgK : PrgH : secretin (InvG) in *S. typhimurium* basal body has been proposed to be 22±1.7 : 17.4±1.4 : 13±1 assuming the secretin with 12-14 subunits as a reference [80]. On the other hand, when the MxiJ structure was docked into the cryo EM density map of the *S. flexneri* basal body it was suggested that the inner ring consists of 12 proteins [83].

The YscD protein, much less conserved than YscJ, probably takes part in inner membrane ring formation and connects the YscJ with the YscC secretin [98].

C ring (YscQ)

In the cryo EM reconstructions of the *S. flexneri* and *S. enterica SPI-1* needle complexes the C ring was not observed [83]. However, it is known that YscQ protein has a high similarity to proteins FliN and FliM forming the 45-50 nm C ring in the flagellum and it is suspected to be an essential component in injectisome C ring. Data obtained from YscQ orthologs suggest the presence of a C ring in the lower part of the injectisome which would act as the platform for the ATPase complex [103].

Export apparatus (transmembrane α-helical proteins: YscR, YscS, YscT, YscU, YscV, associated ATPase complex: YscN, YscK, YscL)

After formation of a *Yersinia* injectisome basal body five highly conserved proteins (YscR, YscS, YscT, YscU and YscV) are thought to form the translocation channel integrated in the inner and outer membrane rings. ATPase (YscN), associated with the transmembrane α-helical proteins was shown to interact with C ring (YscQ) and YscK and YscL proteins. ATPase, which is the conserved and essential inner membrane component in T3SS, is responsible for energizing the export apparatus. The recent structures of the flagellar ATPase and *EPEC* ATPase (EscN) pointed to the similarity of YscN with F1 ATPase [104]. It is believed that the hexameric YscN is attached to the C ring of T3S pore with the central cavity filled with the protein that links it to the injectisome basal body.
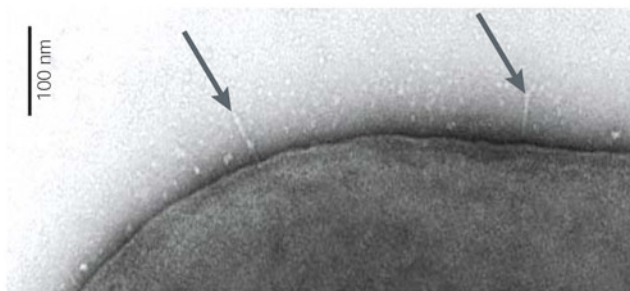
Figure 1.9: Electron micrograph with marked injectisome needles protruding from *Yersinia enterocolitica* [105].
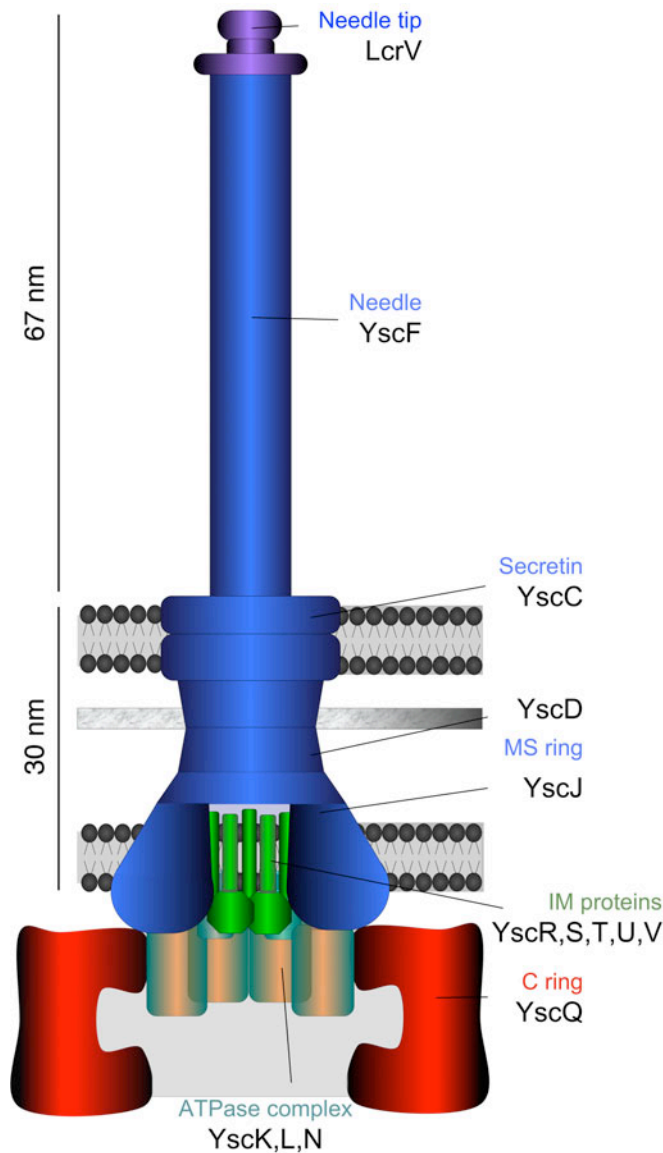
Figure 1.10: Model of the *Yersinia* Type III Secretion injectisome. Needle tip, needle, secretin YscC, MS ring, **i**nner **m**embrane (IM) proteins and ATPase with adequate proteins in the **c**ytosolic ring (C ring) are highlighted. Model was based on [106].

1.3.3   Assembly of *Yersinia* injectisome

Assembly of the *Yersinia* injectisome is a complex process that involves more than 20 types of proteins, which form the structures in the outer and inner membranes (OM and IM), C ring and an external needle. Up to 2002, when it was proposed that the IM ring is formed first [107], little was known about the order of assembly of the complex. Recently, the new assembly model was confirmed [108], based on experiments with injectisome constituents fused to fluorescent proteins. It was found that the secretin ring in the outer

Yersinia membrane is formed first (Figure 1.11). It was also observed that the membrane ring-forming proteins (YscC, YscD, and YscJ) were required for the assembly of any cytosolic structure and that YscD was required for YscJ to become attached to YscC. After the outer and inner rings formation, the cytosolic components assemble onto the structure. Then the huge cytosolic complex made of ATPase (YscN, inactive), interacting proteins YscK and YscL and the component of C ring (YscQ) is formed independently on the export apparatus. However, it was confirmed that the C ring (YscQ) with its components is the integral part of the injectisome.
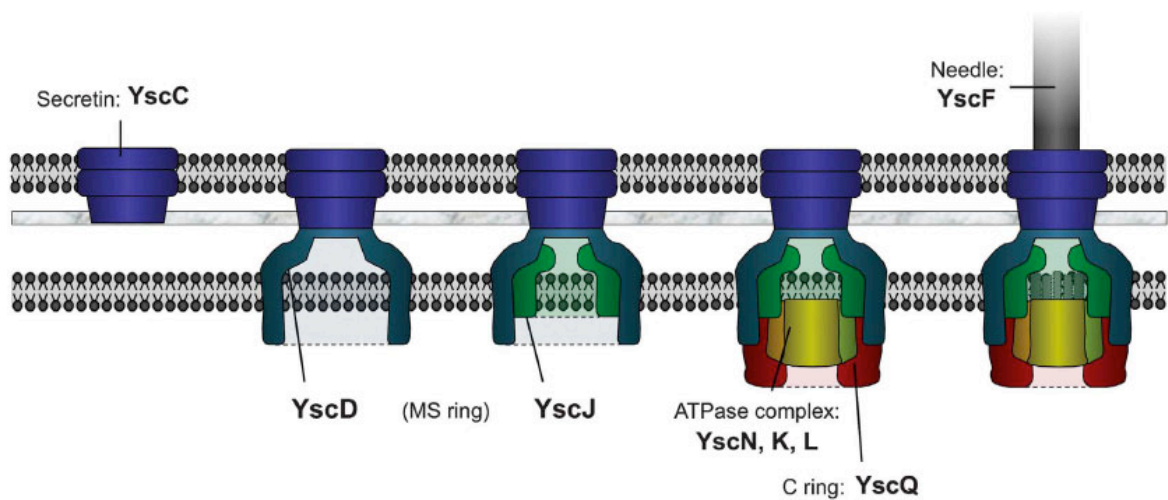
Figure 1.11: Model of the assembly of the Yersinia injectisome taken from [108]. The formation of the injectisome starts at the YscC secretin ring in the outer membrane, followed by YscD and YscJ MS ring, and then the ATPase - C ring complex (YscN, K, L, and Q) assembled at the cytoplasmic side of the inner membrane. The assembly of IM proteins (YscR, S, T, U, and V) is not clear. At the very end the needle made of YscF subunits and tip (LcrV) is formed.

2 An approach to solve the structure of human Aquaporin 8

Abstract

Among the thirteen human aquaporins (AQP0-12), AQP8 is unique in the primary structure. By sequence alignment it is evident that AQP8 creates a separate subfamily, which is apart from all the other mammalian aquaporins. The special constriction region of the pore, which determines the solute permeability, is unique in AQP8 and makes it permeable to both ammonia and H_2O_2 in addition to water. For mammalian aquaporins, the structures solved up to now, all belong to the water-permeable subfamily.

To better understand AQP8 selectivity and gating mechanism the high-resolution structure is necessary. To assess the structure, human AQP8 was overexpressed in methylotrophic yeast *Pichia pastoris* as a His-tagged protein. A wide screen of different detergents and detergent-lipid combinations for optimal protein purification and 2D-crystalization was essential to obtain well-ordered AQP8 crystal arrays. Removal of amino acids constituting affinity tags was necessary to achieve highly ordered crystals diffracting up to 3 Å. Atomic force microscopy, electron microscopy and gold labeling experiments revealed the double-layered nature of 2D crystals, with tetrameric organization of AQP8, which had termini exposed outwards of the 2D crystal. In parallel, alignments to AQP4 revealed a similar, extraordinary long N-terminal of AQP8. In analogy to AQP4, where only the short isoform is able to crystallize, 2D crystallization of the shorter AQP8 construct, with removed N-terminus, was initiated.

The protein expression, purification and 2D crystallization studies are part of the "Reconstitution of water channel function and crystallization of human aquaporin 8" manuscript of Maria Agemark Fellert[1], Julia Kowal, Wanda Kukulski, Kristina Nordén[1], Niklas Gustavsson[1], Urban Johanson[1], Andreas Engel and Per Kjellbom[1]. Described 2DX

---

[1] Center for Molecular Protein Science, Department of Biochemistry, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden

robot studies were presented in the publication "The 2DX robot: A membrane protein 2D crystallization Swiss Army knife" published in Journal of Structural Biology (see Appendix), 2010.

2.1 Materials and Methods

2.1.1 Expression of AQP8 in yeast *Pichia pastoris* and protein purification

Human AQP8 containing the c-myc epitope sequence and 6xHis-tag on the C-terminal was heterologously overexpressed in the methylotrophic yeast *Pichia pastoris*. The protein expression in *P. pastoris*, membranes preparation and protein purification protocols were optimized in the laboratory of Per Kjellbom in Lund, Sweden.

To grow *P. pastoris* culture expressing the human AQP8 the BMGY (1% yeast extract, 2% peptone, 100 mM potassium phosphate buffer pH 6.0, 1.34% yeast nitrogen base without amino acids, 4×10^{-5} % biotin, 1% glycerol) was inoculated with the glycerol stock and incubated at 28°C, 200 rpm, to an OD_{600} of 5. Exchanging the medium to BMMY (1% yeast extract, 2% peptone, 100 mM potassium phosphate buffer pH 6.0, 1.34% yeast nitrogen base without amino acids, 4×10^{-5} % biotin, 0.5% methanol) and increasing the volume 10 times induced the overexpression. Cells were pelleted, resuspended in BMMY and incubated in baffled flasks for 6 days (144 h) at 28°C, 200 rpm. Because of evaporation and consumption, 0.5% sterile methanol was added every 24 h. Cells were pelleted at 10000 g for 10 min and resuspended in breaking buffer (50 mM sodium phosphate buffer pH 7.4, 1 mM EDTA, 5% glycerol, 1 mM PMSF). Then cells were broken using the Bead Beater (BioSpec Products, USA) with a 50 ml chamber half-filled with glass beads (0.5 mm $\varnothing$), by 12 runs of 30 seconds each and 30 seconds of cooling time in between. Cell debris was pelleted at 4000 g for 10 min at 4°C. Total membrane fraction was collected by the supernatant ultracentrifugation at 41000 g for 2 h, at 4°C. Membrane pellets were resuspended and homogenized in a cold buffer A (20 mM HEPES-NaOH pH 7.8, 50 mM NaCl, 10% glycerol, 2 mM ß-Mercaptoethanol, 0.03 % NaN_3), aliquoted, frozen in the liquid nitrogen and stored at -80°C.

Crude *P. pastoris* membranes were diluted to a total protein concentration of 24 mg/ml in the buffer B (20 mM HEPES-NaOH, 300 mM NaCl, 10% glycerol, 2 mM ß-Mercaptoethanol, 0.03% NaN_3) and solubilized dropwise with different detergents

26

solutions (4 % DDM, 6 % DM, 4 % Fos-choline12, 4% Cymal-6 in the buffer B to a final concentration of 2 % DDM, 3 % DM, 2% Fos-choline12 or 2% Cymal-6, respectively) to a final total protein concentration of 12 mg/ml. Membranes were solubilized for 3 h at room temperature with gentle stirring. The material that did not solubilize was pelleted at 41000x g for 40 min at 4°C. Ni-NTA agarose (Qiagen) was equilibrated with buffer B with detergent (0.4% DM, 0.04% DDM, 0.15% Cymal-6 or 0.3% Fos-choline-12). 10 mM imidazole was added to the solubilized protein solution to minimize unspecific binding. For binding, the solution was incubated overnight with Ni-NTA agarose at 4°C. On the next day the slurry was placed into the PolyPrep column (BioRad) and washed with 50 bed volumes of buffer B with 70 mM imidazole and detergent. Aquaporin was eluted with 300 mM imidazole in buffer B with detergent. Protein concentration was determined by measuring absorbance at 280 nm (MW = 29934.9 Da, ε_{280nm} = 42970 $M^{-1}cm^{-1}$) [109], the purity and homogeneity of the sample were ascertained by 4-12% Bis-Tris NuPAGE (Invitrogen) and single particle EM.

2.1.2   Single particles and sample homogeneity

Single particle grids were prepared by adsorbing 3 µl of purified protein solution at concentrations of 0.1 - 0.001 mg/ml for 30 sec - 1 min on carbon-coated copper grids previously rendered hydrophilic by glow discharging. Grids were washed four times with nanopure water and stained twice for 12 seconds with 2% uranyl acetate. Single particles were imaged with a Phillips CM-100 or CM-10 electron microscopes operated at 80 kV.

2.1.3   Two-dimensional crystallization and endoprotease Lys-C treatment of 2D crystals

Recombinant AQP8 solubilized in 2% DM and isolated by Ni-affinity chromatography was reconstituted into 2D crystals by dialysis. Purified protein ready for 2D crystallization (1 mg/ml) was mixed with different lipids solubilized in 2% DM to achieve the final LPRs of 0.2, 0.3 and 0.4 (w/w). After one-hour preincubation at room temperature, 60 µl aliquots were dialyzed against a detergent-free buffer of varying composition in dialysis buttons using a dialysis membrane with a molecular cut-off of 10 kDa. Different temperature profiles were tested. After 12, 24 or 36 h at room temperature dialysis buttons were moved to 37°C for 48 h and thereafter back to room temperature for

2 days. Subsequently, the best crystallization procedure was repeated and tuned in the temperature-controlled dialysis machine [110]. For cleavage of the C-terminal fragment containing His-Tag and c-myc epitope sequence, to each 100 µl of AQP8-cmyc-His-tag sample 1 µg of Lys-C (Roche) was added after 36 h of the crystallization procedure. The serine protease Lys-C specifically cleaves peptide bonds C-terminally at lysine, at pH 7.0 – 9.0. Reaction was stopped after the end of crystallization. The 2D crystals were washed twice through centrifugation at 5000 rpm for 10 min at 4°C with subsequent removal of the supernatant and addition of fresh crystallization buffer solution.

An analog experiment was performed using endoprotease Glu-C (protease V8 from *Staphylococcus aureus*, Roche) that cleaves peptide bonds C-terminally at glutamic acid.

After the protease treatment, samples were analyzed by 4-12% Bis-Tris NuPAGE and western blot using TetraHis antibody (Qiagen). Briefly, after electrophoresis, gels were either stained by Coomassie brilliant blue R-250 or transferred to a polyvinylidene diflouride (PVDF) membrane at 100 V for 20 min. His-tags of AQP8 were detected by TetraHis antibody (Qiagen) diluted 1:2000 in Tris buffer saline and 3% bovine serum albumin. Anti-mouse IgG conjugated to alkaline phosphatase (GE Healthcare) was used as secondary antibody and the blots were developed with BCIP/NBT.

In parallel, the proper bands were cut out from the NuPAGE and send to the Mass Spectrometry Facility, Biozentrum, University of Basel [111] for analysis. Proteins isolated from the gel were digested with trypsin and resulting peptides were separated by capillary reverse phase chromatography coupled to the mass spectrometer. The masses of the peptides were measured and used for the protein fragments identification.

Additionally, gold labeling experiments using 1.8 nm Ni-NTA-bound gold beads (Nanoprobes, [112]) were performed.

2.1.4   Transmission electron microscopy and cryo electron microscopy

TEM was used for screening the crystallization trials. Grids were prepared as described in 2.2. Negatively stained specimens were imaged using a transmission electron microscope Phillips CM-100 or CM-10 operated at 80 kV. Electron micrographs were taken at magnifications of 7000x to 50000x and recorded on Kodak SO163 film or on a 2000x2000 pixel charge-coupled device (CCD) camera (Veleta, Olympus soft imaging solutions GmbH, Munster, Germany) at a nominal magnification of 50000x (130000x at the plane of the CCD).

Specimens of highly ordered AQP8 2D crystals were further analyzed by cryo EM. Selected samples were adsorbed on molybdenum grids coated with a carbon film, embedded in 7% trehalose [113] and frozen in liquid nitrogen. Electron micrographs were taken at low dose (<10 e$^-$ per Å^2) at magnification of 50000x with a Phillips CM-200 FEG microscope operated at 200 kV and at the temperature of -175ºC.

The electron diffraction patterns of 2D crystals were recorded at low electron doses (<5 e$^-$ per Å^2) using a Philips CM200 FEG, operated at 200 kV, at temperature -175ºC, equipped with a Gatan UltraScan 1000 2000x2000 pixel slow scan CCD camera.

2.1.5   Image Processing Library & Toolbox

Electron micrographs of 2D crystals were scanned with a Heidelberg Primescan D7100 drum scanner, at a step size of 5 μm (2000 lines/cm) resulting in a resolution of 1 Å/pixel. Digitized images were analyzed and processed with the IPLT [114, 115]. Briefly, the goal was to filter an image, of a negatively stained 2D crystal and further, to improve it by correlation averaging. First, two data viewers displayed the image of double-layered crystal and its Fourier transform. To improve the display of the Fourier transform image was normalized. In the displayed Fourier transform peaks corresponded to two identical crystal lattices which originated from the double-layered crystal. One of the two lattices was fitted manually modifying the default Lattice Overlay option values. Because all the information about the unit cell was on the points in the lattice and everything outside the spots was a noise the image was filtered with a Fourier Peak Filter. In the perfect crystal, the information is concentrated on a single pixel. In reality, due to crystalline imperfections (i.e. crystal bending, crystal size), the information was spread out over several pixels. In the next step, correlation averaging, the created Fourier filtered image was used as a reference for a cross-correlation with the original image. The cross-correlation image displayed peaks correlated to the areas that were matching the template. By optimizing the Peak Search Parameters correct positions of the most unit cells in the image were found. The correlation peak positions were refined using a 2D Gaussian fit. The position information was used afterwards for averaging the unit cells and generating the averaged image.
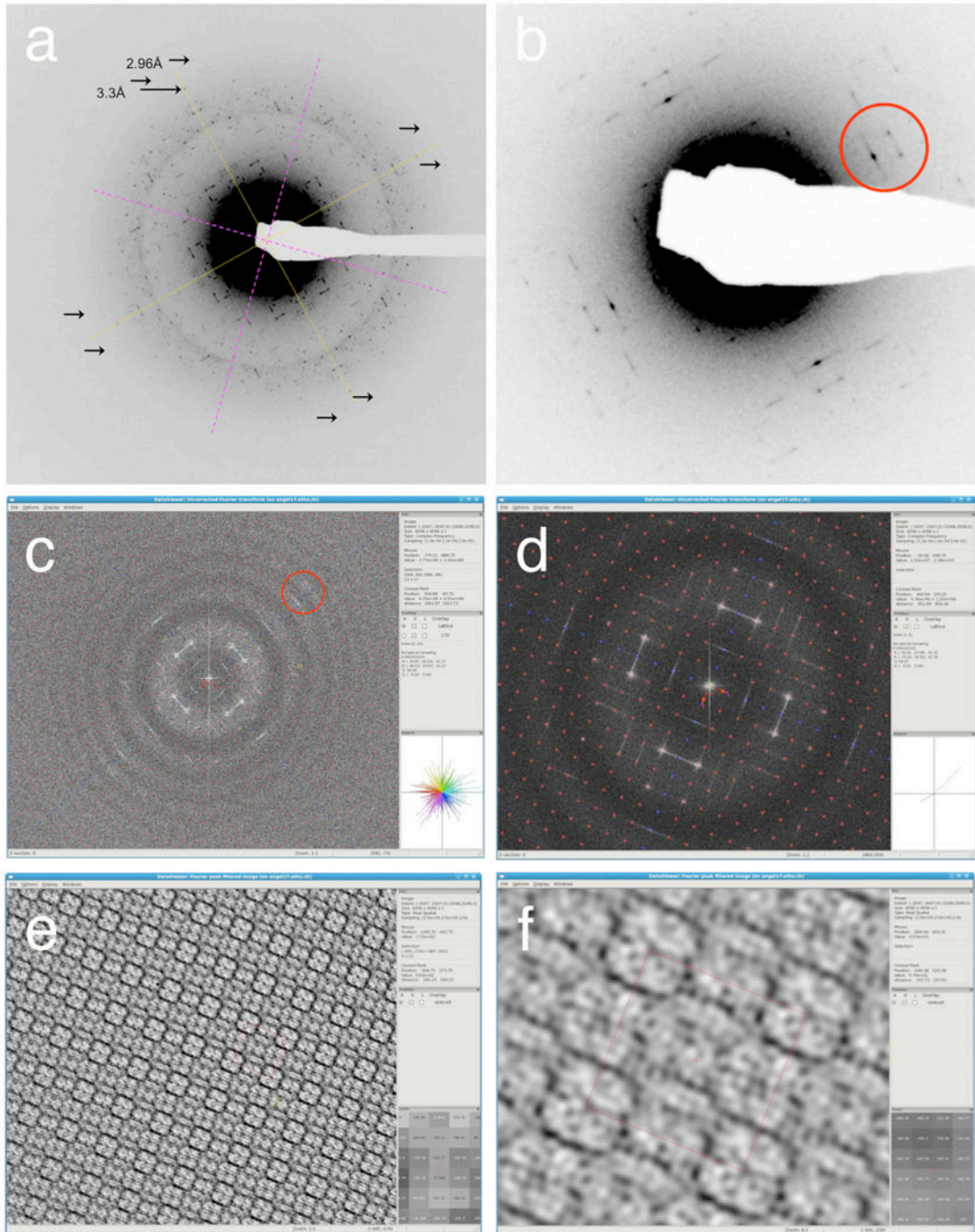
Figure 2.1: Difference between electron diffraction and cryo EM images. (a-b) Electron diffraction pattern of protease Lys-C treated crystal. (c-f) Image processing performed using IPLT software. Orders 12 and 14 were highlighted in (b) and (c) with the red dashed line circles; (c) The Fourier-transformed image; (d) Lattice fitting performed on (c). The unit cell vectors are: α=β=133.5 Å, γ=90°; (e) Filtered and backtransformed image; (f) Correlation average with the unit cell contour.

2.1.6   Gold labeling of AQP8 crystals

Intact His-tagged AQP8 2D crystals, as well as crystals digested previously with endoprotease Lys-C (Roche), were dialyzed for 2 h at 4ºC against the buffer (20 mM Tris pH 8.0, 150 mM NaCl) without chelating and reducing agents, EDTA and DTT, respectively. 20 µl of protein crystal solution (0.5 mg/ml) was mixed with equal amount of 1.8 nm Ni-NTA-bound gold beads (Nanoprobes [112]) of concentration 1 nmol/ml (Figure 2.2). 20 µl-drops were applied on the carbon-covered copper grids, previously rendered hydrophilic by glow discharging, and incubated for 1 h. In parallel, control experiment crystal solutions were mixed and incubated on the grids with Ni-NTA-bound gold beads in the buffer containing 300 mM imidazole. Grids were washed four times with buffer containing 20 mM imidazole to eliminate only unspecific binding and stained twice for 12 seconds with 2% uranyl acetate. Gold-labeled crystals were imaged with a Phillips CM-10 electron microscope operated at 80 kV.
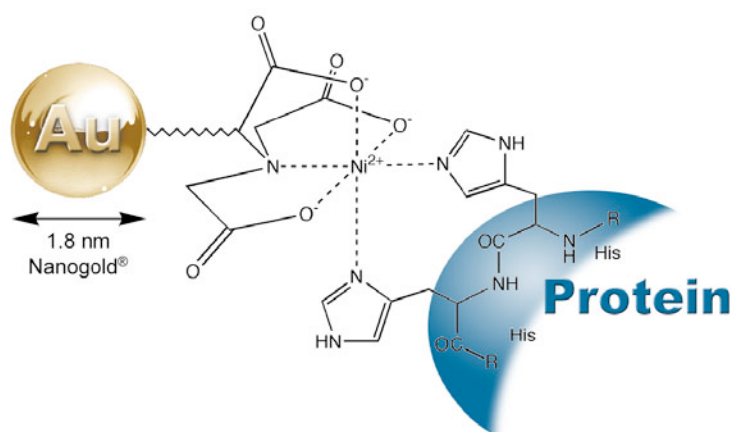
Figure 2.2: Interaction between a His-tagged protein and 1.8 nm Ni-NTA-Nanogold [112].

2.1.7   Atomic force microscopy

A commercial atomic force microscope (AFM) (Nanowizzard II Ultra, JPK Instruments, Berlin, Germany) was used for imaging two-dimensional crystals of AQP8. The AFM was protected from vibrations by placing it on an active damping table. Furthermore, the instrument was shielded from acoustic noise. For AFM experiment the mica prepared as described in [116], freshly cleaned with Scotch tape, was used as support.

The imaging was performed in buffer solution (100 mM NaCl, 20 mM HEPES-NaOH, pH 7) at room temperature. All buffers were prepared using ultrapure water. For all

experiments rectangular silicon nitride cantilevers (ORC8-PS, $k_{nominal}$ = 0.05 N·m^{-1}, Olympus Ltd., Tokyo, Japan) were used. Two-dimensional crystals of reconstituted AQP8 were adsorbed on a freshly cleaved mica surface in buffer solution (100 mM NaCl, 20 mM HEPES-NaOH, pH 7) for ~30 min at room temperature. Non-adsorbed material was removed by rinsing the sample several times with adsorption buffer. For imaging the adsorption buffer was used. Before imaging the membranes at forces <500 pN, the microscope was equilibrated to reduce cantilever drift.

Piezo height images were analyzed using commercial software (JPK Image Processing Software, JPK Instruments, Berlin, Germany).

2.1.8   Protein crystallization using 2DX robot

Crude membranes from *Pichia pastoris* expressing His-tagged AQP8 were solubilized with DDM and protein was purified as described in 2.2.1.

Purified protein (1 mg/ml) was dialyzed against crystallization buffer (20 mM Hepes pH 7.0, 100 mM NaCl, 5 mM EDTA, 10% glycerol, 2 mM DTT, 0.03% NaN$_3$) for 3 h at 4°C. The protein was then mixed with selected lipids to final LPRs ranging from 0.2 to 0.8 (w/w). In parallel to UV-spectroscopy measurements at 280 nm, the detergent concentration was measured with the DropBox [117].

In a fully automated 2DX robot using the cyclodextrin-driven method for detergent removal (Figure 2.3) the kinetics of the reconstitution processes and temperature are precisely controlled. Moreover, the method requires smaller volumes and lower protein concentrations in comparison to other 2D crystallization methods. The 2D crystallization process was monitored by the light scattering intensity measurement.

The total ternary mixture volume was 30 μl, DDM detergent concentration was 0.15% and the cyclodextrin titration rate ranged from 5 to 20 nl/(h μl). For crystallization, the following temperature profile was selected: 12 h at 23°C, 12 h increasing linearly to 37°C, and 48 h at 37°C. When proteins formed sheets and large vesicles, 0.3 μg of endoprotease Lys-C was added to remove a long, artificial C-terminus of protein.

Figure 2.3: Membrane protein crystallization robot [118]. (A) Contactless cyclodextrin dispenser. (B) Contactless volume measurement sensor. (C) Light scattering detector. (D) Contactless water dispenser maintaining the targeted sample volume. (E) An orbital microplate shaker allowing for simultaneous light scattering measurement provides homogenization of the crystallization mixture. (F) 96 wells commercial transparent microplate. A Peltier device regulates temperature.
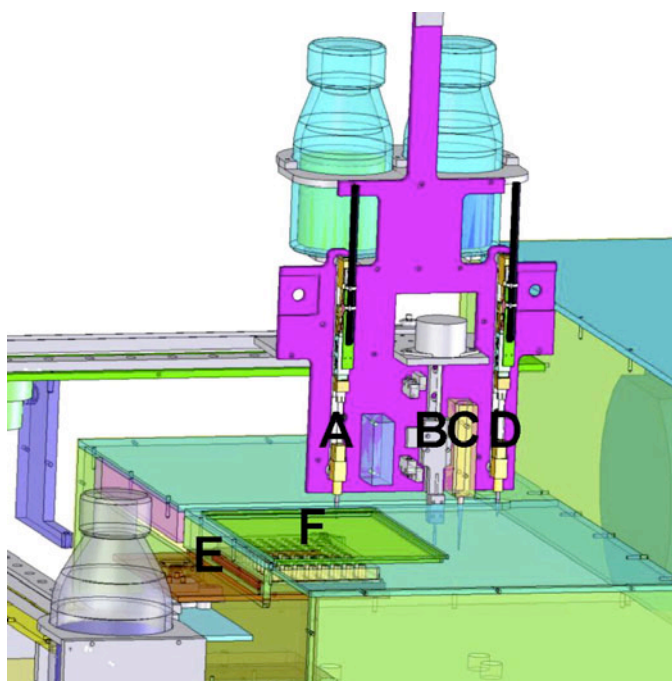
2.1.9 Cloning, transformation, expression and purification of new AQP8 constructs

Cloning

The laboratory of Per Kjellbom, Sweden, kindly provided the pPICZB vector with the subcloned construct AQP8-cmyc-His-tag[1]. All designed reverse and forward primers (Microsynth) used for cloning experiments of the new constructs were diluted with nanopure water to the end concentration of 10 μM.

Oligonucleotide primer sequences designed for 3 new constructs of AQP8, the template DNA and restriction enzymes used for cloning are shown in Table 2.1.

---

[1] In the AQP8 construct, provided by Per Kjellbom laboratory (Lund, Sweden) (AQP8-cmyc-His-tag) the first 5 amino acids from N-terminus are missing due to the used cDNA clone, which was the only one accessible on the market. The artificially added Ser2 was inserted for proper initiation of translation in *Pichia pastoris*. Thus, the original construct sequence starts with *MSCEP* while human AQP8 is *MSGEIAMCEP*. The amino acid numbering has been adjusted to the new construct - AQP8-cmyc-His-tag.

Table 2.1: New AQP8 constructs with physico-chemical parameters of a protein sequence [119], designed primers, the template DNA and restriction enzymes used for cloning.

NEW CONSTRUCT	PRIMERS	TEMPLATE DNA (MODIFIED CONSTRUCT)	RESTRICTION ENZYMES
AQP8-TEV-cmyc-His-tag Number of amino acids: 288 Molecular weight: 30566.6 Da Theoretical pI: 6.04	Jul_Not_Xbafor 5`-GGC CGC AGA AAA CCT GTA TTT TCA GGC T -3` Jul_Not_Xbarev 5` - CTA GAG CCT GAA AAT ACA GGT TTT CTG C - 3`	AQP8-cmyc-His-tag Number of amino acids: 283 Molecular weight: 29934.9 Da Theoretical pI: 6.18	XbaI / NotI
AQP8-His-tag Number of amino acids: 265 Molecular weight: 27915.7 Da Theoretical pI: 7.14	Not_His_Xbafor 5`- GGC CGC CAC CAT CAC CAT CAC CAT TGA CAT - 3` Not_His_Xbarev 5`- CTA GAT GTC AAT GGT GAT GGT GAT GGT GGC - 3`	AQP8-TEV-cmyc-His-tag	XbaI / NotI
MG$_{18}$-AQP8-His-tag Number of amino acids: 249 Molecular weight: 26168.8 Da Theoretical pI: 8.35	N_modfor 5` - CAG GGG AAT TCA AAA TGG GTG GCA GGT GGC GAG TGT C - 3` Not_His_Xbarev 5`- CTA GAT GTC AAT GGT GAT GGT GAT GGT GGC - 3`	AQP8-His-tag	XbaI / NotI

Three new AQP8 constructs were designed:

1. AQP8-TEV-cmyc-His-tag (AQP8 gene sequence with the inserted TEV protease cleavage site before the c-myc epitope sequence EQKLISEEDL and the 6xHis-tag on the C-terminus),
2. AQP8-His-tag (AQP8 gene sequence with the inserted 6xHis-tag on the C-terminus),
3. MG$_{18}$-AQP8-His-tag (AQP8 gene sequence starting at Gly18 with the inserted 6xHis-tag on the C-terminus).

The cloning strategy for AQP8-TEV-cmyc-His-tag and AQP8-His-tag constructs was based on the primers annealing, insertion of the modified DNA into the vector and an amplification of AQP8 gene with modified fragments in the pPICZB vector in *E. coli*. The procedure of annealing [120] involved the usage of a thermal cycler. 100 µl aliquots of the oligonucleotides mixed in PCR tubes were placed in the thermal cycler. The annealing had the following profile: heating to 95°C, 2 min at 95°C, ramp cooling to 25°C over a period of 45 minutes, a storage at temperature of 4°C. Template DNA digestion was performed with *XbaI* and *NotI* enzymes added at 37°C for 3h. The PCR product was purified using the QIAquick PCR Purification Kit, controlled by the agarose gel electrophoresis and stored at -20°C. The optimal ligation mix contained 12 µl of vector cut (10-50 ng), 5 µl of annealing product, 1 µl of T4 DNA ligase (Roche) and 2 µl of 10X T4 ligation buffer. The ligation was performed overnight at 16°C.

To the MG_{18}-AQP8-His-tag construct a shorter N-terminus was introduced by the polymerase chain reaction (PCR) using designed primers and the construct AQP8-His-tag as the template. The PCR mix contained 15 ng of template DNA, 50 nM of forward and reverse primer, 1 µl of Pfu DNA polymerase, 2.5 µl of Pfu DNA polymerase buffer and 1 µl of dNTP in a total reaction volume of 25 µl. The PCR was performed under the following conditions: 1 cycle of 30 sec at 95°C, then 20 cycles of 30 sec at 95°C, 30 sec at 55°C and 5 min at 72°C, and 1 cycle of 10 min at 72°C. The PCR product was purified using the QIAquick PCR Purification Kit, controlled by agarose gel electrophoresis and stored at -20°C.

Plasmid transformation into E. coli cells using Heat Shock

50 µl competent DH5α *E. coli* cells were thawed onto wet ice with 5 µl of PCR or ligation product added. The mixture was incubated 30 min on ice, followed by 45 sec heat shock at 42°C and 1 min incubation on the ice. 950 µl of SOC or LB medium was added to the bacteria and the mixture was kept at 37°C, at 225 rpm, for 1 h to express the antibiotic resistance gene. 100 µl, 200 µl and 500 µl of transformed cells were plated on low salt LB plates containing 25 µg/ml of Zeocin and grown at 37°C overnight.

Plasmid linearization and sequencing

The plasmid isolation from transformed *E. coli* cells grown at 37°C overnight in liquid low salt LB medium was done using the Qiagen Miniprep Kit. 170-350 ng of

isolated DNA was mixed with 1 μl of sequencing primer, 2 μl of BigDye-Premix Version 3.1 solution, 2 μl of 5X Sequencing buffer Version 3.1 in a total reaction volume of 10 μl and amplified in the thermal cycler under the following conditions: 30 cycles of 10 sec at 96°C and 4 min at 55°C, then cooled to 4°C. DNA, purified using the QIAquick PCR Purification Kit, was mixed with 1 μl of 125 mM EDTA and 2 μl of 3M sodium acetate, then with 50 μl of ethanol and incubated for 15 min on the ice. DNA was pelleted at 10000 rpm, at 4°C, for 10 min and the pellet was washed with 400 μl of 70% ethanol. Prepared dry DNA pellet was sent and sequenced in the Microsynth. The *E. coli* plasmid linearization was achieved by *PmeI* digestion at 37°C for 3h and controlled by agarose gel electrophoresis. The linearized DNA was purified using the QIAquick PCR Purification Kit. DNA was eluted in 100 μl of nanopure water to eliminate all salts and prevent an arching (electrical discharge), which reduces the viability of cells when pulsing the *Pichia pastoris* cells.

Transformation of Pichia pastoris cells

X-33 Pichia pastoris cells were grown on YPD plates at 30°C for 48 h. Colonies were scratched from plates and grown in 5 ml of YPD liquid medium overnight at 30°C, at 200 rpm. 200 ml YPD were inoculated with 2 ml of the overnight culture and grown to an OD_{600} of 1.6. Cells were pelleted at 4000 rpm for 10 min at 4°C, resuspended in 120 ml of the sterile, ice-cold water, pelleted, again resuspended in 60 ml of the sterile, ice-cold water, pelleted, resuspended in 5 ml of sterile, ice-cold 1M sorbitol, pelleted, resuspended in 0.3 ml of sterile, ice-cold 1M sorbitol and kept on ice. 100 μl of *P. pastoris* electrocompetent cells were mixed with 5 μl of purified DNA (0.5-1 mg/ml), incubated for 5 min on ice and transferred to the ice-cold 0.2 cm electroporation cuvette. The yeast electroporation was performed in the GenePulser apparatus (Bio-Rad) and the following parameters were applied: charging voltage = 1500 V, capacitance = 25 μF, resistance = 200 Ω. Immediately after the application of 1 pulse 1 ml of sterile, ice-cold 1M sorbitol was added to the cuvette with DNA/cells. Transformants were incubated for 2 h at 30°C, afterwards 150-200 μl of cells were plated on the YPD plates with 100, 200, 500 and 1000 μg/ml Zeocin and incubated at 30°C, in darkness, for 3-5 days. Yeast clones with the highest Zeocin resistency were grown on YPD plates containing 200 μg/ml Zeocin, at 30°C for 2 days. Yeast colonies were picked and cultured in 5 ml of liquid YPD medium at

30°C, overnight. 0.2 ml of sterile 80% glycerol was added to yeast culture aliquots of 0.8 ml in sterile eppendorfs. *P. pastoris* glycerol stocks were frozen and stored at -80°C.

2.1.10   Purification and 2D crystallization of new AQP8 constructs

All 3 new AQP8 constructs were designed to have 6xHis-Tag on the C-terminus, therefore the protein was purified according to the same protocol as AQP8-cmyc-His-tag construct as described in 2.1.1. An additional purification step occurred only for the construct with a TEV protease cleavage site.

The Tobacco Etch Virus (TEV) protease, 27 kDa, is a highly site-specific cysteine protease encoded by TEV and is often used for removing affinity tags from purified proteins. Protease recognizes a linear epitope of the general form of E-Xaa-Xaa-Y-Xaa-Q-(G/S) where the cleavage occurs between the Gln and Gly/Ser residues. The most efficient substrate for hydrolase is ENLYFQS [121], although it was shown that many different amino acids can be placed in the P1' position [122], what has relatively small impact on the enzyme efficiency. In the construct AQP8-TEV-cmyc-His-tag the protease recognition site is ENLYFQA and the proteolysis occurs between Gln and Ala residues. The pure His-tagged form of TEV protease was obtained from self-processing MBP-TEV fusion protein [123]. After the purification on Ni-NTA agarose 1 mg of the construct AQP8-TEV-cmyc-His-tag was mixed with 20 mg of the His-tagged TEV protease and dialyzed overnight in the Slide-A-Lyzer Cassette (Pierce) with 10 kDa molecular mass cut-off against the elution buffer, without imidazole, at 4°C. The His-tagged protease and removed His-tagged C-terminus of protein were bound to the Ni-NTA by vigorous shaking for 1 h at 4°C and the resin was collected on the PolyPrep column (Bio-Rad). The flow-through contained the protein which purity was determined by 4-12% Bis-Tris NuPAGE and the concentration by measuring absorbance at 280 nm.

The purified protein fractions of all constructs were kept for 2 days at room temperature and at 4°C to study the protein stability. Samples were ultracentrifuged at 150000 g and supernatants analyzed by NuPAGE. No visible pellet of possible precipitated protein was observed after ultracentrifugation and no protein degradation was detected using the gel. Single particles homogeneity and monodispersity were checked by negative stain EM.

Screening for 2D crystallization conditions and protein crystallization of new constructs was performed according to the protocol described in 2.1.3. Briefly, the pure

protein was mixed with DM-solubilized *E. coli* polar, DMPC or phosphatidyl lipids at protein end concentration of 0.5-1 mg/ml. The detergent removal from the protein-lipid-detergent ternary mixture was achieved by the 4-day long dialysis in dialysis buttons against crystallization buffers. The temperature profile was as follow: 36 h at room temperature, 48 h at 37°C and 24 h at room temperature. Crystal presence and quality were checked using the negative stain EM.

2.2   Results

2.2.1   Expression of AQP8 in yeast *Pichia pastoris* and protein purification

The clone of yeast *Pichia pastoris* heterologously expressing human AQP8 (construct AQP8-cmyc-His-tag), as well as the culturing, membrane preparation and protein purification protocols were established by M. Fellert from the laboratory of Per Kjellbom in Lund, Sweden.

144-hour long yeast expression was crucial to obtain the adequate total membrane protein content. Crude membranes were directly solubilized with selected detergent omitting the alkaine/urea stripping protocol that is used to remove membrane-associated proteins [124], which led to a huge loss of AQP8 protein. M. Fellert and W. Kukulski performed a wide screen of 9 different detergents and, after comparison of the purification efficiency in a quantitative manner, it turned out that the best-solubilizing AQP8 detergents are DM, DDM, Cymal-6 and Fos-choline-12 (unpublished results, Figure 2.4, Figure 2.5). Depending on the 2D crystallization method the protein was solubilized both in the low and high CMC detergents (Table 2.2). For the dialysis technique, typically, the high CMC detergent (DM) was used. For crystallization in an automated 2DX robot the protein was purified in DDM.

Table 2.2: CMC and concentrations of detergents used for solubilization and purification of the AQP8.

DETERGENT	CMC [%]	SOLUBILIZATION CONCENTRATION [%]	WASH AND ELUTION CONCENTRATIONS [%]
DM	0.087	3	0.4
DDM	0.0087	2	0.04
Cymal-6	0.028	2	0.15
Fos-choline-12	0.053	2	0.3

Extending the solubilization time from 2 to 3-4 hours at room temperature increased the AQP8 solubilization yield twofold. Efficient purification by Ni-NTA affinity chromatography routinely gave a single fraction of highly pure AQP8. Concentration of the eluted His-tagged AQP8 was typically 2 – 4 mg/ml making it possible to set up 2D crystallization trials at the protein concentration of 1 mg/ml. With the optimized protocol on average 1 – 2 mg of pure protein was obtained from 1 l of *P. pastoris* culture.
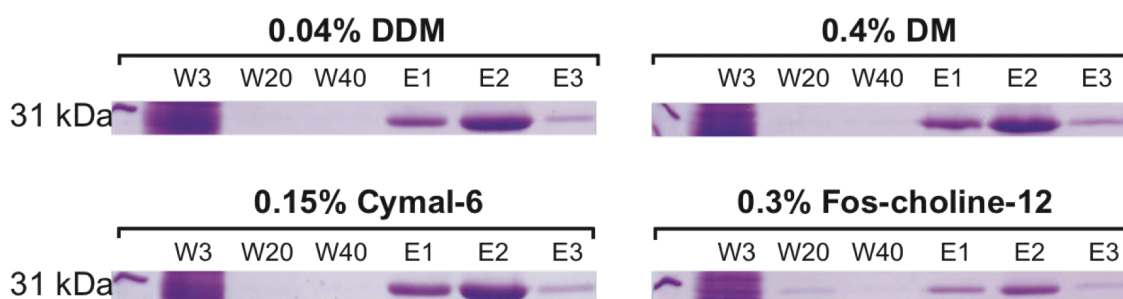
Figure 2.4: SDS-PAGE showing test purification with the four best detergents. Wash fractions collected after 3, 20 and 40 ml (W3, W20, W40) and elution fractions E1-3 are shown.

Single particles homogeneity was checked by negative stain EM. Electron micrographs were taken at a magnification of 50000x and a top view class average was calculated from 490 particles in EMAN software package. The evident 4-fold symmetry of the purified AQP8 complex confirms its tetrameric nature, which is typical for all known aquaporins.
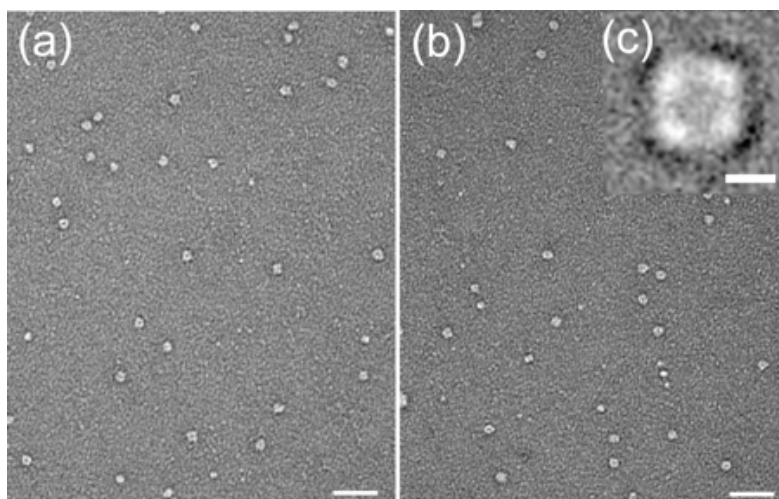
Figure 2.5: Electron micrographs of negatively stained single AQP8 particles purified in DM (a) and in DDM (b). The inset (provided by W. Kukulski) shows a top class average calculated from 490 DDM-purified particles. Scale bars correspond to 20 nm in (a-b) and 5 nm in the inset (c).

2.2.2   Two-dimensional crystallization of AQP8

Crystallization conditions optimization

To optimize crystallization, a variety of conditions were tested: a range of LPRs, different lipids, detergents, and buffers at various pH values as well as different salts with varying concentrations. Solubilized and highly purified AQP8 was reproducibly reconstituted into lipid bilayers using a mixture of POPE and POPC lipids (Figure 2.6) in the ratio of 3:7 (w/w). Mixtures of POPA:POPE:POPC lipids or DOPA:DOPE:DOPC lipids, both in ratio 1:2:7, gave equally good results. *E. coli* polar lipids were tested as well, but only small vesicles with a diameter in the order of 100 nm, were found. Neither POPS:POPC lipids (in ratio 3:7) nor pure DOPC turned out to be useful. In the first case, formation of small tubes, collapsed vesicles and aggregated structures were observed, whereas in the second case only aggregates and precipitations were noticed.

High protein concentrations during crystallization resulted in well-ordered 2D crystals. The optimal protein concentration was 1 mg/ml for which the best LPR was 0.25-0.3. Optimal NaCl concentration in the dialysis buffer was between 100 mM and 250 mM. Ordered crystals of AQP8 were obtained at pH 7.0. In the pH range between 7.4 and 8.4 the crystals were smaller and mixed with aggregates. At pH 4.0 and 9.0 the protein

precipitated, also at pH 5.0 and 6.0 a significant number of aggregates was observed. Therefore all further crystallization trials were performed at pH 7.0.

Addition of EDTA to the dialysis buffer reduced the aggregation observed in the presence of the divalent cations Mg^{2+} and Ca^{2+}. Nevertheless, aggregates occurred both in the absence and presence of divalent cations (Figure 2.7 c-d).

During the AQP8 crystallization process the temperature profile had a significant influence on the formation, size, and quality of the crystals. For instance, tubular crystals occurred when dialysis buttons were moved to the higher temperature during the first 36 h of crystallization. In contrast, incubation at room temperature for the first 36 h and subsequent temperature increase yielded large amounts of regularly packed sheets and vesicles-like structures of a few micrometers in size. Interestingly, the quality of AQP8 2D crystals improved during a two-week long storage at 4°C.

To precisely control the crystallization, most of the experiments were carried out in the temperature-controlled dialysis machine. The following temperature profile was selected: 12 h at 20ºC, 12 h linear increase to 37ºC, 48 h at 37ºC, and 24 h linear decline to 20ºC.

Optimal 2D crystallization conditions had the following parameters (Table 2.3): POPE/POPC lipids (in ratio 3:7), LPR of 0.25-0.3 and the dialyzing buffer containing 20mM HEPES-NaOH pH 7.0, 100 mM NaCl, 5 mM EDTA, 2mM DTT and 10% glycerol.
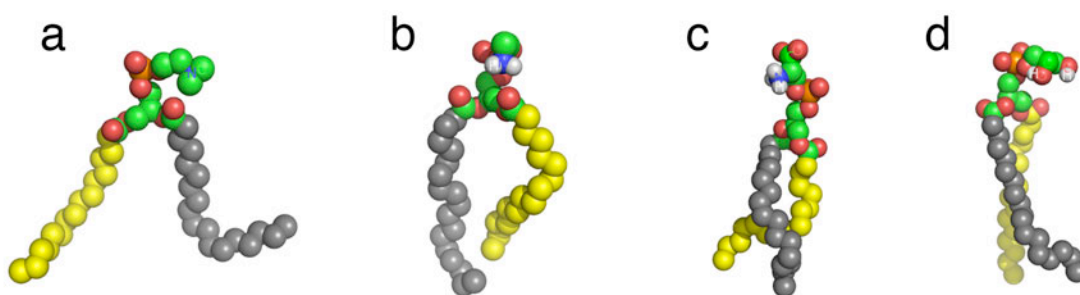
Figure 2.6: Models of (a) POPC, (b) POPE, (c) POPS and (d) POPG lipids. Colors correspond to: *yellow* – 1-palmitoyl chain, *gray* – 2-oleoyl chain, *green* – carbon, *red* – oxygen, *orange* - phosphorus, *blue* – nitrogen and *white* – hydrogen atoms in –PX head group.

Analysis of a projection map of intact AQP8

A large variety of different 2D crystallization conditions resulted in diverse protein packing arrangements in reconstituted lipid bilayers and number of 2D crystal lattice parameters. Electron micrographs of negatively stained reconstituted samples revealed long and rather thin tubular structures (Figure 2.7 a-b), vesicles and also well-ordered 2D crystalline sheets with a diameter of 2 μm and a length of several micrometers (Figure 2.7 e).

The unit cell vectors of tubular crystals, which were less ordered than sheets, were determined to be a = b = 100 Å and γ = 90° with 8 monomers per unit cell. The unit cell vectors of sheets grown in buffer containing 100 mM NaCl were a = b = 72 Å, γ = 90° (crystals exhibit a p4 symmetry and accommodate 1 tetramer per unit cell; Figure 2.7 g), and in 1 M NaCl were a = 68.5 Å, b = 133 Å, γ = 100° (two tetramers per unit cell, up–and–down orientation), respectively.

In the majority of cases, the power spectra showed two sets of diffraction spots indicating two crystalline layers or collapsed vesicles (Figure 2.7 f). The two lattices, which happen to be rotated by 65°, were analyzed separately with the IPLT software package [114]. The 2D projection map of AQP8 shows individual monomeric pores in a tetrameric organization with a well visible central depression (Figure 2.7 g).

Figure 2.7: 2D crystals of AQP8 were reconstituted by dialysis in the presence of lipid. (a) Crystals appear as well ordered tubes with a diameter of about 360 nm and a length of several micrometers, often attached to aggregates, vesicles or unrolled/open tubes. (b) Crystalline AQP8 tubes at a magnification 50000x. (c) Crystals after addition of 5 mM EDTA appear as large vesicles and sheets of regularly packed protein. (d) 2D crystals in the presence of Mg^{2+} appear as small vesicles or tubes that often clustered together. (e) Overview of a large crystal obtained under the optimized conditions. The double-layered crystals are large mono-crystalline protein-lipid arrays. (f) The power spectrum calculated from cryo EM image of crystalline sheet shown in (e) reveals two sets of diffraction spots indicating a double-layered crystal. Sharp diffraction spots marked by white arrow extend to the order (9,3) corresponding to a resolution of 8 Å. Point marked by an arrowhead corresponds to a resolution of 12 Å. (g) The 2D projection map calculated from a cryo electron micrograph. AQP8 packs into arrays with p4 symmetry, housing one tetramer per unit cell of size 72 Å. The dark areas between bright protein tetramers represent the lipid

bilayer. (h) Electron micrograph of a negatively stained Lys-C treated crystals. The large well-ordered rectangular double-layered crystals, with sizes of several micrometers in diameter are visible. (i) The electron diffraction pattern of a frozen–hydrated crystal showed in (h). On insert sharp diffraction orders marked by black arrow correspond to a resolution of 3 Å, by the arrowhead to 3.3 Å. Scale bars corresponds to 1 µm in (a, c-e), 200 nm in (b), 50 Å in (g) and 5 µm in (h).

Table 2.3: The summary of 2D-crystallization experiments.

PARAMETER	RANGE TESTED	OPTIMAL CONDITION
Detergent	DM, DDM, Cymal-6, Fos-choline12	DM
Lipids	*E. coli, E. coli* polar, soy, liver, brain, DMPC, DOPS, DOPE, DOPC, DOPA, POPS, POPC, POPE, POPA, POPE:POPC (1:1, 3:7), POPS:POPC (3:7), POPA:POPE:POPC (1:2:7), POPA:POPC (1:1, 3:7), DOPE:DOPC (3:7), DOPS:DOPC (3:7), DOPA:DOPC(1:1, 3:7), DOPA:DOPE:DOPC (1:2:7)	POPE:POPC (3:7)
LPR	0.1 - 1.3	0.25 - 0.3
Protein concentration	0.5 - 2 mg/ml	1 mg/ml
pH	4.0 - 9.0	7.0
Buffer	Citrate, MES, Hepes, Tris	Hepes
NaCl	50 mM - 1M	100 mM
$MgCl_2$	0 - 20 mM	-
$CaCl_2$	0 - 20 mM	-
EDTA	0 - 5 mM	5 mM
DTT	0 - 10 mM	2 mM
Glycerol	0 - 20%	10%
Temperature profile	4°C, 20°C, 37°C, 20°C→37°C, 20°C→4°C, 20°C→4°C→20°C, 20°C→37°C→20°C, 37°C→20°C	20°C→37°C→20°C
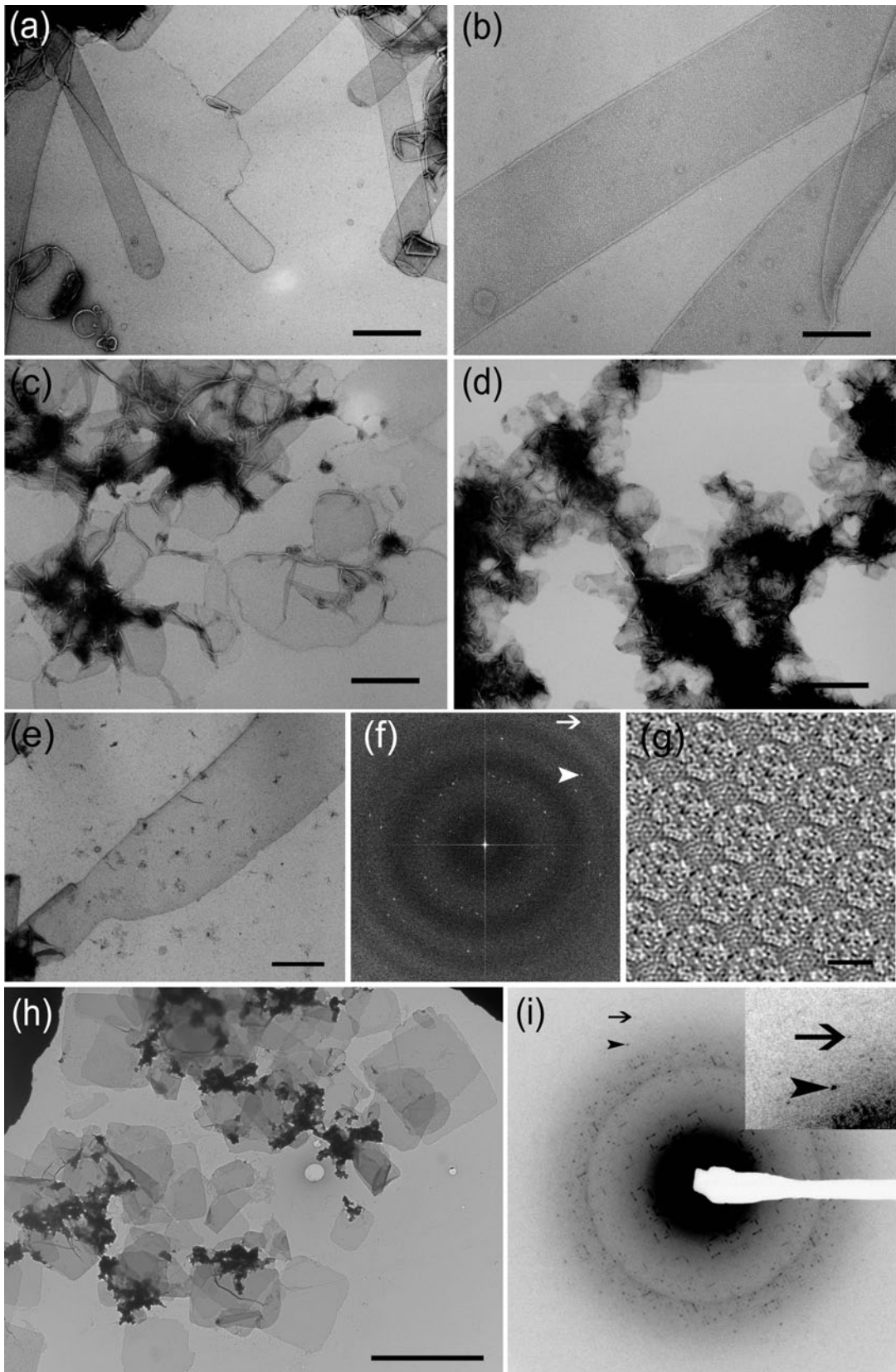
2.2.3 Atomic force microscopy of double-layered crystals

Intact 2D crystals of AQP8 adsorbed without wrinkles to the mica and were suitable for imaging by AFM. The overview in Figure 2.8 demonstrates the double-layered 2D crystal, whose thickness was measured to be 12 nm. Crystals often stacked and formed the multilayer structures made of 2 or 3 double-layers with heights of 24 or 36 nm, respectively.
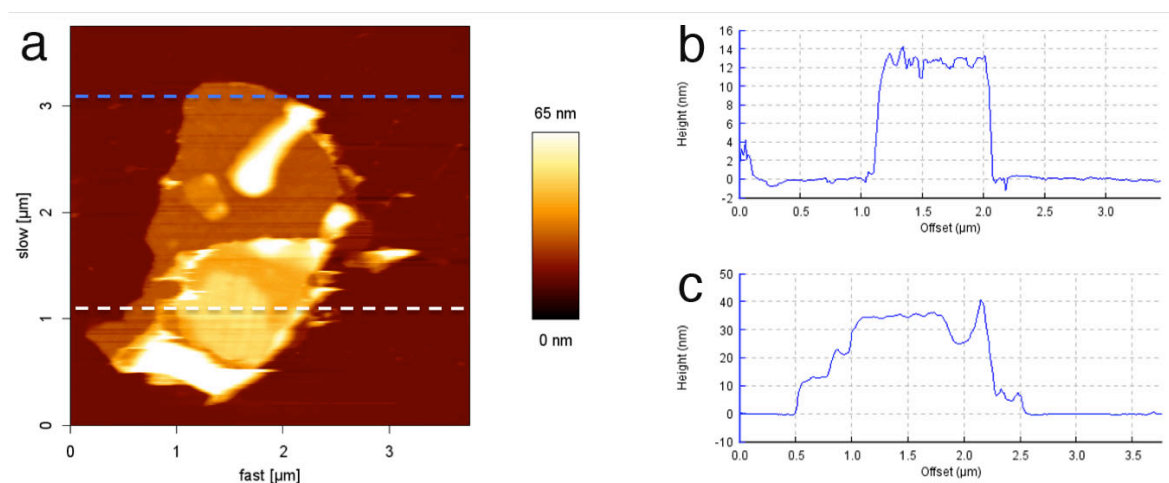
Figure 2.8: (a) AFM topograph of AQP8 2D crystals adsorbed to mica and recorded in buffer solution: 20 mM HEPES pH 7.0, 100 mM NaCl, 5 mM EDTA. Multi-layered areas can be clearly distinguished from double-layered crystals by their height above the mica surface. (b) Cross section analysis along the blue dashed line in the image (a). (c) Cross section analysis along the white dashed line in image (a).

2.2.4 Lys-C protease treated 2D crystals

2D crystallization with addition of the Lys-C protease

The AQP8 construct contains a 10 aa long c-myc epitope (EQKLISEEDL), 6xHis-tag and 11 aa long linkers at the C-terminus of the protein (Figure 2.14 a). It was suspected that this 27 aa sequence could possibly disturb crystallization. Therefore the fragment was removed enzymatically during the dialysis procedure. The 283 aa long AQP8 construct has 7 Lys-C cleavage sites (K11, K124, K175, K177, K248, K254 and K265). When the

protein was embedded in a lipid bilayer, theoretically all of these Lys were exposed to the enzyme except one located in the transmembrane helix 3 (K124). One Lys is positioned in the N-terminal region (K11), 2 Lys in the cytosolic D-loop (K175, K177) and 3 Lys in the C-terminal region of AQP8 (K248, K254, K265). However, it was assumed that lysines at the D-loop are hidden and inaccessible for the enzyme.

Addition of the Lys-C serine protease during the crystallization process resulted in formation of large rectangular protein-lipid arrays (Figure 2.7 h), instead of the round sheets and vesicle-like structures (Figure 2.7 c-e). The lattice parameters of Lys-C treated crystal were a = b = 133.5 Å, γ = 90°. The correlation average of Lys-C digested crystal was shown on Figure 2.9.

The Philips CM200 FEG was aligned according to the standard protocol for electron diffraction. Diffraction patterns of digested crystals recorded with a CCD camera at tilt angle 0˚, showed sharp spots at 3 Å, however, blurred diffraction spots were also visible at higher resolution. The two orders, marked by black arrow and the arrowhead in Figure 2.7 i, correspond to a resolution of 3 Å and 3.3 Å, respectively.
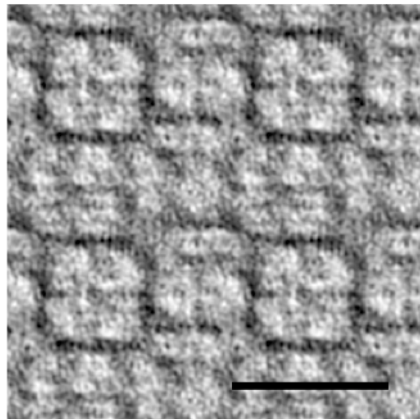
Figure 2.9: Correlation average of Lys-C digested 2D crystal. For cryo EM experiments crystals were embedded in 7% trehalose. Low-pass (10 Å) and high-pass (200 Å) thresholds were applied during IPLT processing. Unit cell vectors are a = b = 133.5 Å, γ = 90°. Scale bar corresponds to 100 Å.

Mass spectroscopy measurements of Lys-C protease digested crystals showed that removal of first 10 amino acids of N-terminus was achieved. However, neither His-tag nor any 11 aa or 18 aa peptide, corresponding to the potential fragments on the C-terminus,

could be identified. Therefore, two additional experiments were performed on 2D crystals: gold labeling using 1.8 nm Ni-NTA-bound gold beads (Nanoprobes), able to bind to the His-Tags, and Western blot with Tetra-His antibody (Qiagen).

Gold labeling of Lys-C treated crystals

To detect if His-tags are exposed on the surface of the double-layered crystal, 1.8 nm-Ni-NTA-bound gold beads were mixed with intact and Lys-C digested 2D crystals dialyzed previously against the binding buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl) without chelating (EDTA) and reducing (DTT) agents. After 1 h incubation of samples on the carbon-coated grids, each grid was washed with 20 mM imidazole to remove all unbounded gold particles. As shown on electron micrographs in Figure 2.10, only single gold particles sparsely appeared in the background. In control experiment, with 300 mM imidazole in the binding buffer, Ni-NTA-bound gold beads did not bind to the His-tagged crystals.

Surprisingly, Ni-NTA-bound gold particles had attached to the crystal surface, not only in the case of the intact but as well to the Lys-C treated crystals (Figure 2.10 a-d). Additionally, more gold particles were bound to the digested crystal surface in comparison to the intact ones. When the enzyme concentration had been increased from 0.005 mg/ml to 0.01 mg/ml (from ratio 1:200 to 1:100 (w/w)), number of gold particles bound to the 2D crystal even increased revealing densely packed gold-particles patches (Figure 2.10 e-f).

The results indicated that cytosolic AQP8 termini of both layers are exposed outwards of the double-layered crystal. Additionally, the result demonstrates that the removal of His-tags localized on the crystal surface, was not complete. Attachment of larger number of Ni-NTA-bound gold beads to the digested crystals suggests that at first 10 aa on the N-terminus of the protein are removed and then the His-tagged C-terminus, which is not easily accessible for protease, is being exposed.
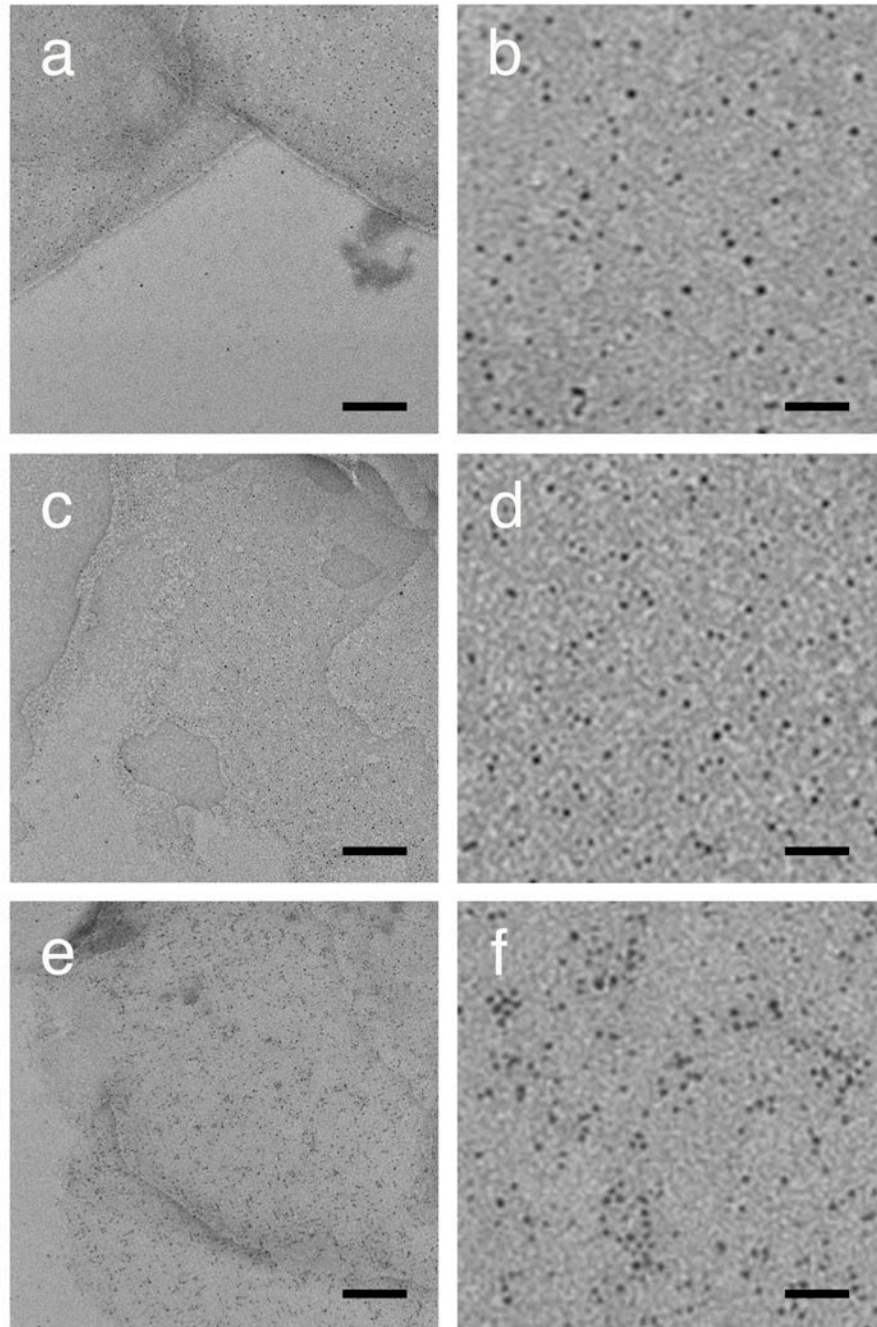
Figure 2.10: Gold labeling of AQP8 crystals. Black dots correspond to Ni-NTA-bound gold beads. (a-b) Intact crystals, (c-d) Crystals incubated with protease in ratio 1:200 (w/w), (e-f) Crystals treated with Lys-C in ratio 1:100 (w/w). Scale bars correspond to 100 nm in (a, c, e) and to 20 nm in (b, d, f).

Western blot with Tetra-His antibody

Western blot was performed in parallel to SDS PAGE. Intact, Lys-C and Glu-C treated AQP8 2D crystals were mixed with denaturating SDS-PAGE buffer and proteins were separated on the NuPAGE. After the protein transfer on the PVDF membrane, the Tetra-His antibody (Qiagen) for highly specific and sensitive detection of 6xHis-tag was attached to the C-terminal His-tagged proteins. The secondary antibody was conjugated to the alkaline phosphatase. The phosphatase hydrolyzed a substrate BCIP to form a blue intermediate that was then oxidized by NBT to produce a dimer of an intense insoluble purple dye. The western blot visualized His-tags present in protein or peptides (Figure 2.11).

In Figure 2.11, the image (c) which is the superimposed image (a) onto (b), revealed that a small percentage of protein in the sample with intact crystals was degradated from C-terminus (Figure 2.11, lines 1-2). Moreover, it turned out that after addition of the Lys-C protease during crystallization the first 10 aa from N-terminus were removed completely but some His-tags were inaccessible for enzyme (Figure 2.11, line 3). When the ratio enzyme:AQP8 was 1:100 (w/w) and the crystals diffracting to 3 Å formed, additional bands were observed on the SDS gel around 10 kDa and 15 kDa. This suggested that the cytosolic D-loop of 50% of all AQP8 (Figure 2.14 a) was cut by the protease on K175 or K177 (Figure 2.11, line 4). Because the enzyme was added to the protein reconstituted in the membrane, only exposed D-loops could have been fragmented. Amazingly, only ~75% of His-tags were accessible for Lys-C protease even after increasing enzyme:AQP8 ratio.

In crystals digested with a Glu-C protease, which hydrolyses peptide bonds on Glu, 100% of C-termini were removed (Figure 2.11, line 5) but when amount of enzyme was increased to a ratio of 1:100 (w/w) additional bands were visible on the gel (Figure 2.11, line 6) analogously to Lys-C treated crystals (Figure 2.11, line 4) suggesting that protein was cut by the enzyme on E174. Crystals formed during crystallization with Glu-C had much worse quality than those treated with Lys-C.
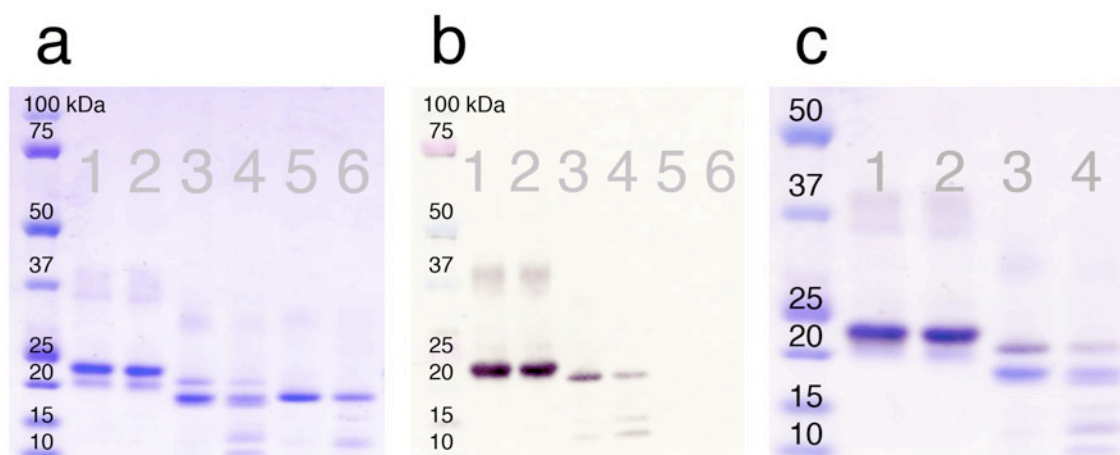
Figure 2.11: (a) NuPAGE and (b) Western blot of AQP8 2D crystals. Image (c) corresponds to superimposed image (a) onto image (b). In (b) Tetra-His antibody for detection of C-terminal 6xHis-tag was used. First line corresponds to protein molecular weight marker (Odyssey). Lines marked as 1-2 correspond to intact crystals, 3-4 to crystals treated with Lys-C protease in ratio 1:200 and 1:100 (w/w), respectively, 5-6 to crystals treated with Glu-C protease in the same manner.

IPLT

Symmetrically arranged stripes were observed on the diffraction patterns of AQP8 2D crystals that were treated with Lys-C protease. These lines, always 2-orders long and occuring both at low and high resolution, were examined using the IPLT software [114] (Figure 2.12).
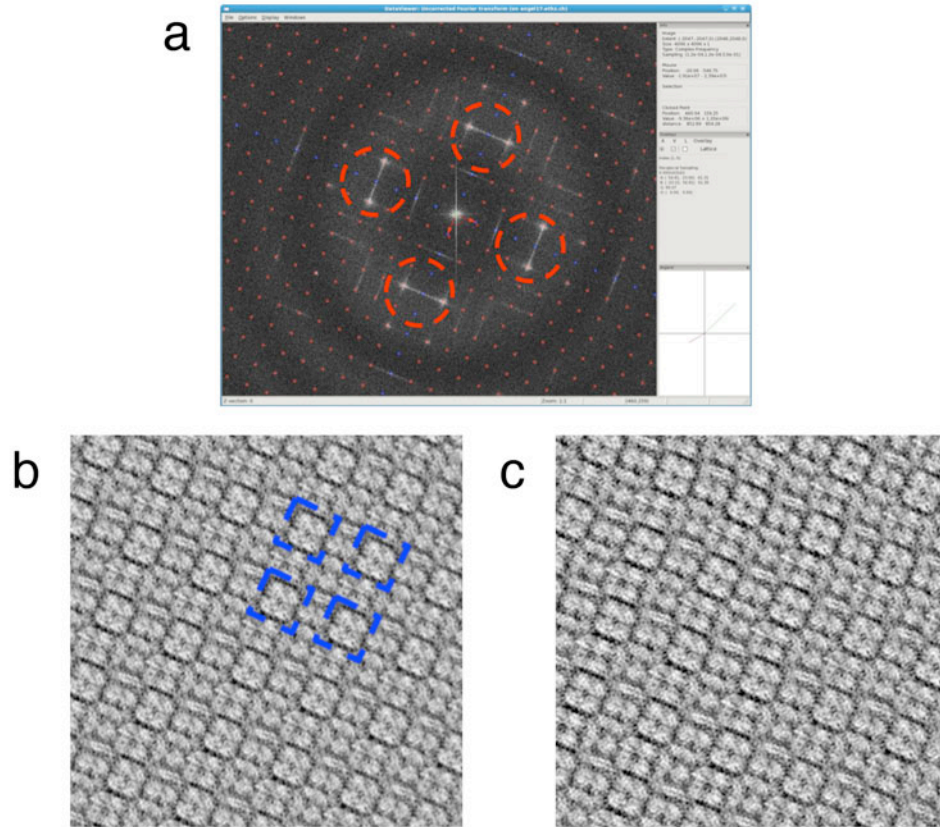
Figure 2.12: To observe what the stripes on the diffraction pattern (a) of Lys-C digested AQP8 crystal correspond to, the lines between orders (4, 1) and (4, -1) etc., highlighted by red dashed line circles, were selected in parallel to the peak information in a lattice and Fourier-backtransformed in IPLT. (b) Filtered and backtransformed image (peaks in the lattice). Four AQP8 overlapped from two crystal layers were highlighted with blue dashed line squares. (c) Filtered and backtransformed image (peaks plus lines from red dashed line circles in the lattice).

The Fourier-filtered and backtransformed image of AQP8 2D crystal represents two overlapped crystal layers adhering exactly to each other. The most conspicuous is the repetitive motive of 4 AQP8 overlapped from both layers and arranged at the corners of the square (Figure 2.12 b). In the backtransformed image where lines between spots were included to the lattice (Figure 2.12 c), the proteins located in between those 4 anchored aquaporins look more blurred or disordered in comparison to the normal backtransformed image. That indicates that the lines visible on the diffraction pattern originate from systematic crystal structure disorder, which could be connected with proteolysis of peptide bond at the D-loop.

The crystal disorders, which are likely the result of AQP8 proteolysis by Lys-C, led to distortions of diffraction pattern and made these crystals unsuitable for generating three-dimensional model at high resolution. For this reason, new AQP8 constructs were designed.

2.2.5 The 2DX robot: A membrane protein 2D crystallization Swiss Army knife

The crystallization conditions of the human aquaporin AQP8 solubilized in DDM were screened using the 2DX robot. AQP8 was reconstituted as double-layered crystalline sheets and packed vesicles of a few micrometers in diameter. As illustrated in Figure 2.13, the 2D crystals developed during gradual detergent neutralization by cyclodextrin from the ternary protein–lipid–detergent mixture over 60 h. The reconstitution process was monitored using the light scattering (Figure 2.13 F). By taking samples after 5, 11, 21 and 26 h (Figure 2.13 A, B, C and D, respectively), we could demonstrate how the crystallization kinetics influences crystal growth. To explore the effect of the reconstitution kinetics with AQP8, parallel experiments were run over 60 h, while the detergent removal rates were varied from 1 to 4 folds (5, 10, 15 and 20 nl/(h•μl)). When the 36 mM cyclodextrin solution was added at a rate of 20 nl/(h μl) to the crystallization mixture containing 3 mM detergent (DDM) and AQP8 at a concentration of 1 mg/ml, small proteoliposomes were formed. In contrast, large sheets were detected during reconstitutions at a rate of 5 nl/(h•μl) (Figure 2.13 E). Moreover, the effect of removing the 6xHis-tag of AQP8 with the protease Lys-C was explored at different stages of AQP8 reconstitution. Lys-C was added to improve the crystallinity of giant vesicles that formed when the reconstitution process was well advanced (Figure 2.13 D). If the protease was added too early, aggregation occurred. Best results with DDM solubilized AQP8 were obtained upon cleavage of the C- and N-terminal moiety by enzyme addition when proteins formed sheets and large vesicles. Diffraction patterns exhibited spots beyond 4 Å (highest observable Friedel pairs) (Figure 2.13 G). The temperature profile (12 h at 23˚C, followed by a slope reaching a plateau at 37˚C 12 h later) was the key to obtaining highly ordered AQP8 sheets.
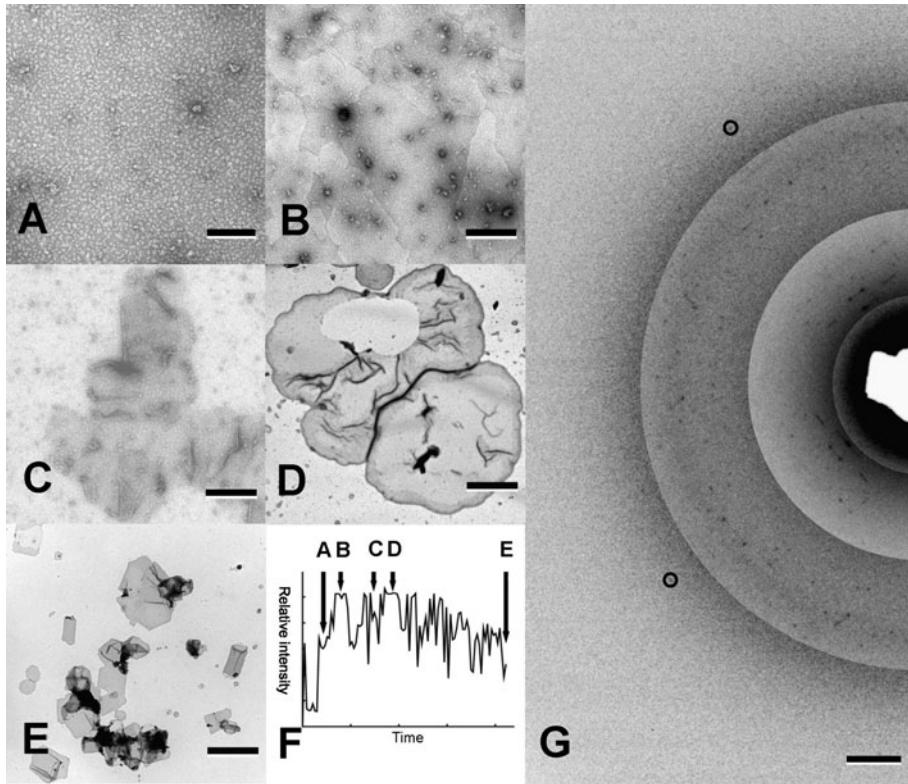
Figure 2.13: Electron micrographs of negatively stained samples taken at different stages during crystallization of AQP8. [118] (A) After 5 h, sufficient cyclodextrin has been added to neutralize all the free detergent molecules and small proteoliposomes started to assemble. (B) After 11 h, sheets and vesicles coexist. (C) After 21 h, sheets are more numerous while the number of vesicles has decreased. (D) After 26 h giant vesicular structures are found, but their morphology does not indicate crystallinity. (E) After addition of more cyclodextrin and Lys-C endoprotease, crystalline sheets are observed at 59 h. Scale bar is 200 nm for (A), 400 nm for (B) and (C), 800 nm for (D), and 3 μm for (E). (F) Light scattering profile indicates the reconstitution progress. Samples for micrographs shown in (A)–(E) have been taken at time points indicated by arrows. (G) Cryo electron diffraction of crystals shown in (E). Spots marked by a circle correspond to a resolution of 3.3 Å. Scale bar is 2 nm^{-1}.

2.2.6   Production and 2D crystallization of new AQP8 constructs
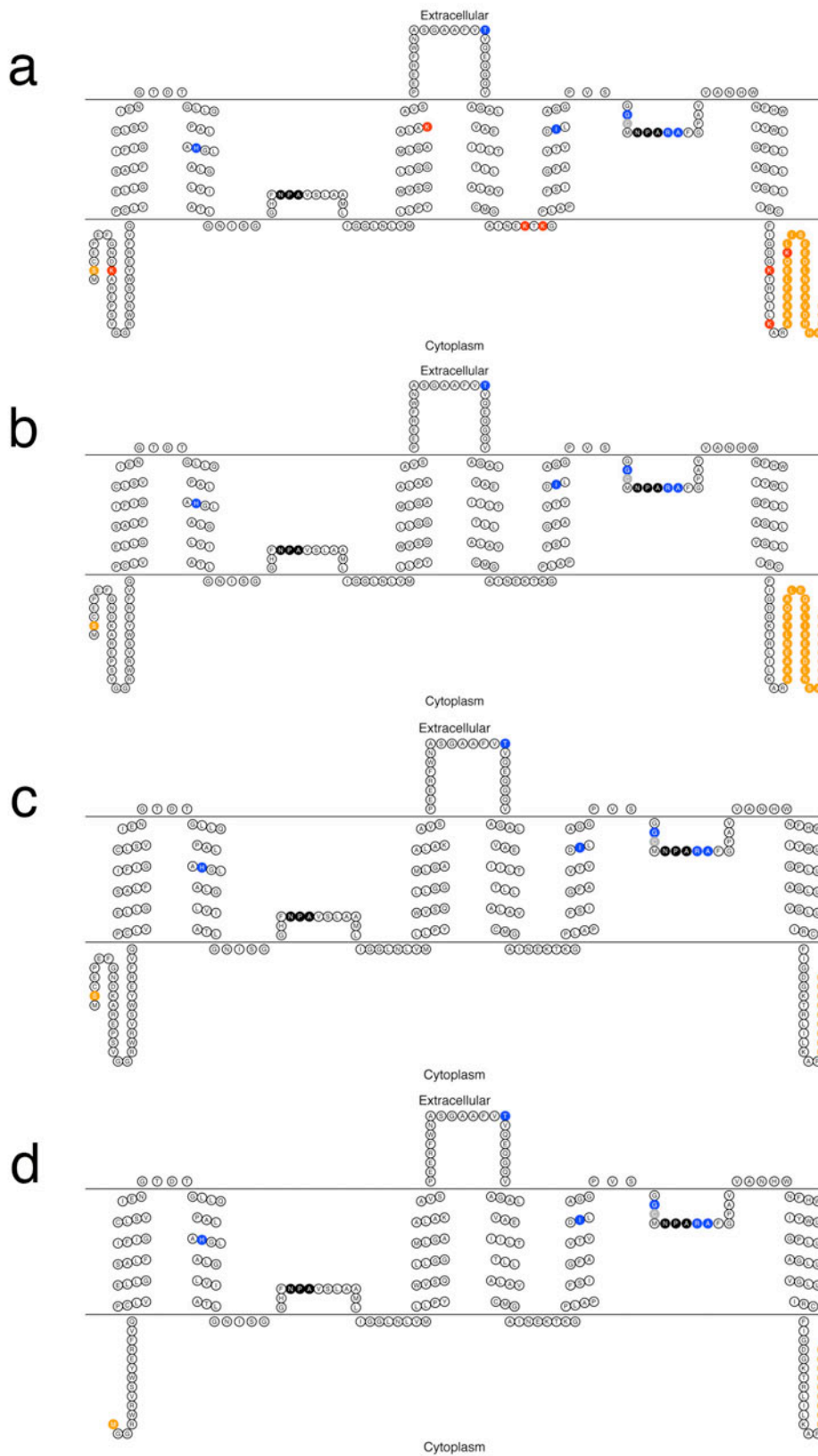
Figure 2.14: Schematic topology of different AQP8 constructs drawn by the TOPO2 software [125]. (a) Construct provided by P.Kjellbom laboratory, Sweden, AQP8-cmyc-His-tag, (b) AQP8-TEV-cmyc-His-tag, (c) AQP8-His-tag, (d) MG$_{18}$-AQP8-His-tag. The colors corresponds to: *orange* − artificial added amino acids (i.e. Ser2 included in *P. pastoris* start codon, TEV, c-myc or 6xHis-Tag), *red* − 7 lysines, *black* − 2 conserved NPA-boxes, *blue* − conserved amino acids in the protein constriction region (from the N-terminal called H2, LC, H5, LE1, LE2 and LE3), *gray* − cysteine responsible for mercury inhibition.

The difficulty in reaching suitable 2D crystals for atomic scale electron crystallography with the initial AQP8 construct was explained by the too long artificial C-terminus of the protein. Thus 3 new shorter AQP8 constructs were designed and introduced into the yeast genome. Until now expression in *P. pastoris* and 2D crystallization trials were performed on two of them (AQP8-TEV-cmyc-His-tag and AQP8-His-tag), and the best results were obtained for AQP8-His-tag.

Cloning, Pichia pastoris electroporation, expression and screening for best expressing clones

The laboratory of Per Kjellbom, Sweden, kindly provided the pPICZB vector with the subcloned construct AQP8-cmyc-His-tag[2] (Figure 2.14 a). Three new AQP8 constructs were designed (Figure 2.14 b-d):

1. AQP8-TEV-cmyc-His-tag (AQP8 gene sequence with the inserted TEV protease cleavage site before the c-myc epitope sequence EQKLISEEDL and the 6xHis-tag on the C-terminus),

---

[2] In the AQP8 construct, provided by Per Kjellbom laboratory (Lund, Sweden) (AQP8-cmyc-His-tag) the first 5 amino acids from N-terminus are missing due to the used cDNA clone, which was the only one accessible on the market. The artificially added Ser2 was inserted for proper initiation of translation in *Pichia pastoris*. Thus, the original construct sequence starts with *MSCEP* while human AQP8 is *MSGEIAMCEP*. The amino acid numbering has been adjusted to the new construct - AQP8-cmyc-His-tag.

2. AQP8-His-tag (AQP8 gene sequence with the inserted 6xHis-tag on the C-terminus),

3. MG$_{18}$-AQP8-His-tag (AQP8 gene sequence starting at Gly18 with the inserted 6xHis-tag on the C-terminus).

Primers were designed to introduce new DNA fragments into the AQP8-cmyc-His-tag construct. The cloning strategy for AQP8-TEV-cmyc-His-tag and AQP8-His-tag constructs was based on the primers annealing, insertion of the modified DNA into the vector and an amplification of AQP8 gene with modified fragments in the pPICZB vector in *E. coli*. To the MG$_{18}$-AQP8-His-tag construct a shorter N-terminus was introduced by PCR using designed primers and the construct AQP8-His-tag as the template. The PCR product was introduced to *E. coli* cells and amplified. Isolated plasmid DNA was linearized and transformed into the *P. pastoris* cells by electroporation and integrated into a yeast genome by homologous recombination.

Since the recombinant *P. pastoris* had the Zeocin resistance gene, the yeast transformants were selected using Zeocin as a selection marker. The Zeocin, which is an antibiotic isolated from *Streptomyces verticillus*, enters the cell removing the copper from its structure, then binds to the DNA and cleaves it causing cell death [126]. The Zeocin resistant protein, a product of the *Streptoalloteichus hindustanus* bleomycin gene, binds stoichiometrically to antibiotic and inhibits the DNA cleavage.

In yeast the gene expression was under the control of the promoter (AQX1) of alcohol oxidase, which is a key enzyme in methanol metabolism. The protein production was induced by change from glycerol (BMGY) to methanol (BMMY) as the carbon source in the medium. Induction was continued for 6 days before cells were harvested and disrupted in the Bead Beater, crude membranes were isolated and stored at -80˚C. The selection of the best expressing *P. pastoris* clones was carried out on YPD agarose plates with increasing Zeocin concentrations. During the homologous recombination multiple insertions into yeast genom are possible, thus it is important to select the clone that is properly growing and expressing protein at the highest possible antibiotic concentration. For each construct a few colonies were picked up from the plates with the highest Zeocin concentration, grown one day in growth medium (BMGY) and 6 days in methanol medium (BMMY) with regular addition of 0.5% methanol. Then cells were harvested by centrifugation, disrupted, membranes collected and analyzed by a western blot. Then the 2 clones of each construct with the most intensive signal were stored in glycerol stocks and

later used for the expression.

Purification of AQP8-TEV-cmyc-His-tag and AQP8-His-tag constructs

Two new AQP8 constructs (AQP8-TEV-cmyc-His-tag and AQP8-His-tag) were overexpressed in *P. pastoris* and purified as described in 2.1.10. In the case of AQP8-His-tag construct, after solubilization in DM, the Ni-NTA affinity chromatography was performed to obtain highly pure protein for 2D crystallization (Figure 2.15 c). For AQP8-TEV-cmyc-His-tag, beside the Ni-NTA agarose purification step, additional TEV protease digestion was performed. Whereas, for construct AQP8-His-tag elution resulted typically in 2 mg of pure protein per 1 l of *P. pastoris* culture, for construct with a TEV protease cleavage site the yield was 3-fold lower. No losses of protein were found during flowthrough or a column wash. In spite of minor impurities detected on the NuPAGE, the single particles EM analysis confirmed that specimens were homogenous, no aggregates occurred and a tetrameric state of protein was observed. Single particles of AQP8-His-tag were more homogenous than particles of AQP8-TEV-cmyc-His-tag.
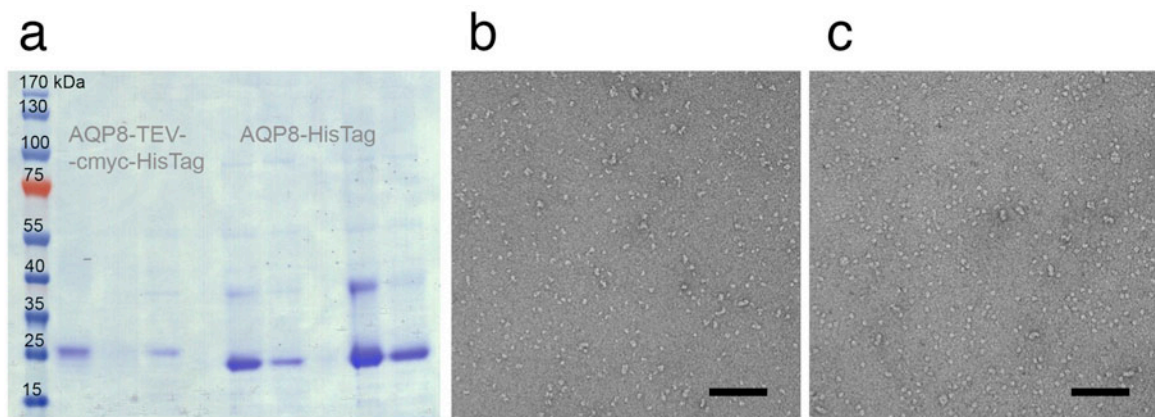
Figure 2.15: Purification of AQP8-TEV-cmyc-His-tag and AQP8-His-tag constructs. (a) NuPAGE. Monomers, dimers, trimers and tetramers of AQP8 are visible on the gel. (b) Single particles of AQP8-TEV-cmyc-His-tag construct. (c) Single particles of AQP8-His-tag construct. Scale bars correspond to 100 nm.

2D crystallization of AQP8-His-tag construct

The protein concentration of eluted His-tagged AQP8 was typically 0.5 – 1.5 mg/ml making it possible to set up 2D crystallization trials at the protein concentration of 0.5 – 1 mg/ml. The conditions that were screened in dialysis buttons were shown in Table 2.4.

Table 2.4: Parameters tested during 2D crystallization of AQP8-His-tag construct

PARAMETER	RANGE TESTED
Lipids	*E. coli* polar, DMPC, DOPC, POPE:POPC (3:7), POPA:POPE:POPC (1:2:7)
LPR	0.1 – 0.5
Protein concentration	0.5 - 1 mg/ml
pH	6.0 - 8.0
Buffer	MES, Hepes, Tris
NaCl / KCl	100 mM – 200 mM
$MgCl_2$	0 - 20 mM
EDTA	0 - 10 mM
DTT	0 - 5 mM
Glycerol	0 - 10%
Temperature profile	20°C→37°C→20°C, 37°C→20°C

In the initial attempt, the 2D crystallization conditions and temperature profile used for the previous construct were tested as described in 2.3.2. The suitable LPR was 0.3, the best crystals were formed at high protein concentrations and glycerol had a positive influence on the crystals formation. In the presence of *E. coli* polar lipids and chelating agent EDTA, at neutral pH, aquaporins formed vesicles that, however, did not exhibit any crystallinity (Figure 2.16 a). Addition of 5 mM $MgCl_2$, instead of EDTA, resulted in formation of crystalline vesicles. Knowing that pI of new construct was shifted from 6.18 to 7.14, the pH was slightly increased to 7.2 and when the reconstitution was done in a mixture of phosphatidyl lipids POPA:POPE:POPC at a ratio of 1:2:7, with 5mM $MgCl_2$, AQP8 formed 2D crystalline sheets (Figure 2.16 b). The unit cell parameters of formed crystals were a = b = 65Å, γ = 90° (Figure 2.17).

Reconstitution in DMPC lipids, which have a phase transition temperature of 23°C, was done directly in 37°C and depending on pH resulted in formation of long tubular crystals of diameter of 1 μm (in pH 6.0, Figure 2.16 c) and 0.5 μm (in pH 6.5, Figure 2.16 d). However, diffraction of these crystals was worse than sheets formed in phosphatidyl lipids. In the presence of 100 mM KCl instead of NaCl, narrow tubes were observed.
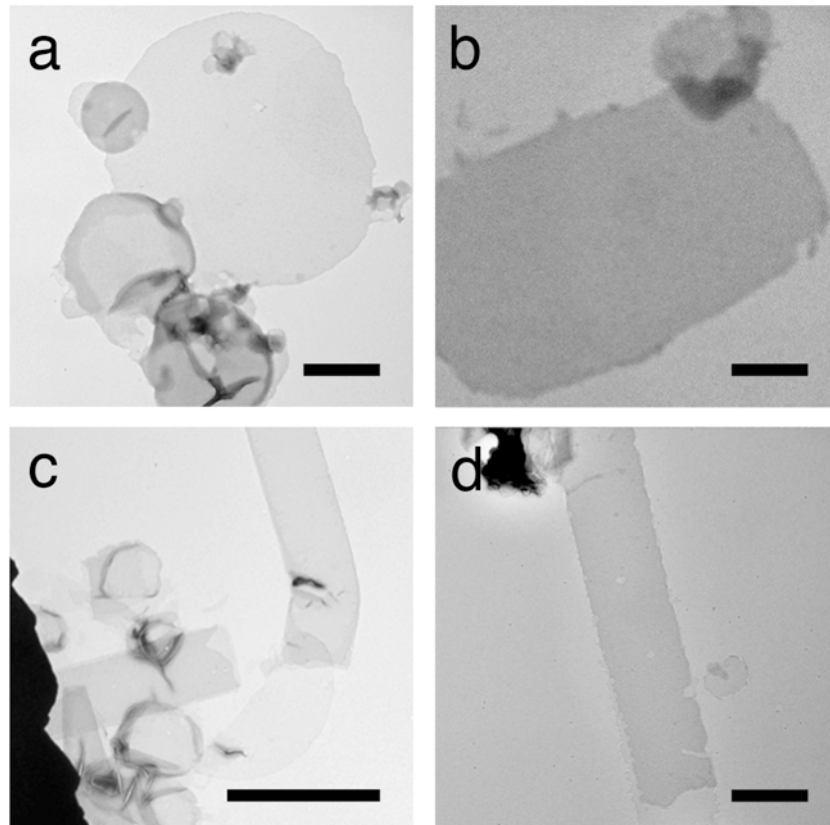
Figure 2.16: 2D crystallization trials of AQP8-His-tag construct. (a) In the presence of *E. coli* polar lipids and chelating agent EDTA aquaporins formed vesicles. (b) In the presence of 5 mM $MgCl_2$ and phosphatidyl lipids large crystalline sheets were observed. In DMPC lipids at pH 6.0 and 6.5 the long tubular crystals of diameter 1 μm (c) and 0.5 μm (d), respectively, were formed. Scale bars correspond to 500 nm in (a) and (d), 200 nm in (b), 2 μm in (c).
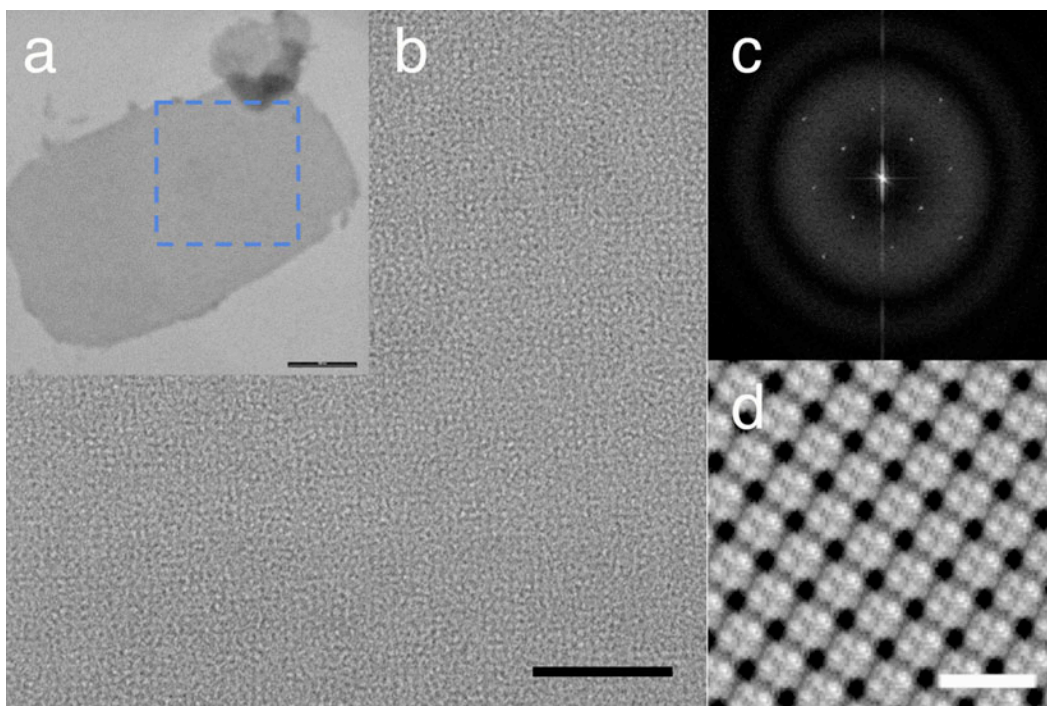
Figure 2.17: Negatively stained AQP8-His-tag crystal imaged in CM-100 EM at magnification 13500x (a) and 50000x (b). (b) is highlighted with blue dashed line square in (a). (c) The Fourier-transformed image. Unit cell parameters are a = b = 65Å, g = 90°. (d) Filtered and backtransformed image of AQP8 tetramers. Scale bars correspond to 200 nm in (a), 100 nm in (b) and 10 nm in (d).

2.3   Discussion

2.3.1   Expression of AQP8 in yeast *Pichia pastoris* and protein purification

Large amounts of the pure and functional protein are needed in order to obtain the high-resolution protein structure. It has been reported that membrane proteins have certain preferences for expression regarding the lipid content of the host [127], which makes the host selection critical. During heterologous expression, in parallel to the host strain, the limiting factors are temperature, culturing media and time of induction. Overexpression of several AQPs using *Pichia pastoris* system has been described in [124]. Accordingly, to obtain the high level of AQP8 expressed in yeast, only the induction time (144h) was optimized.

Treating the membrane fraction with urea and alkali to strip off the peripheral

membrane proteins and thus improve purity of membrane fraction was reported for the native tissue [128] and heterologously expressed AQPs [21, 124]. However, due to the low yield of purified AQP8 after the urea and alkali treatment, we have purified the protein directly from the total membrane fraction. In parallel to activity essays a wide screen of solubilizing detergents was performed (M. Fellert et al., unpublished results). It was found that the low CMC detergent DDM as well as high CMC detergents DM, Cymal-6 and Fos-choline-12 are equally efficient for solubilizing *P. pastoris* crude membranes. In order to use dialysis as 2D crystallization method and taking into account the high-solubilization yield, protein stability and the lowest contamination level, DM was selected. DDM, frequently used detergent with low CMC, which stabilizes the structure of membrane proteins in solution, was used in our 2DX crystallization robot along with methylated β-cyclodextrin to complex the detergent. Extending the solubilization time to 3-4 hours at room temperature significantly increased the AQP8 solubilization yield. Furthermore, a rather high imidazole concentration during the affinity column washing step was necessary to eliminate all contaminants. In summary, a reproducible protocol to obtain well-purified samples of AQP8 adequate to perform 2D crystallization trials was established.

Monodispersity and homogeneity of samples containing human AQP8 purified in DM and DDM was verified by single particle EM. In both detergents, the tetrameric organization of the protein is reflected in the square shape of the particles, specimens did not contain aggregates or contaminants indicating that AQP8 was homogenous and stable.

2.3.2   Two-dimensional crystallization of AQP8

The vast majority of common crystallization conditions that led to successful crystallizations of aquaporins (Table 2.5) were tested for human AQP8. In particular, different pH values and additives were screened, and turned out to be crucial to find the proper crystallization buffer ingredients. When reconstituted in a mixture of neutral phosphatidyl lipids, with optional addition of negatively charged lipids, AQP8-cmyc-His-tag formed ordered 2D crystals. The double-layered crystals appeared as large crystalline protein-lipid arrays about 2 μm wide and with the length of several micrometers. AQP8 sheets were found in the presence of EDTA, at pH 7.0 and LPR 0.25-0.3.

It was reported previously that the addition of Mg^{2+} to the crystallization buffer, either in the form of $MgCl_2$ or as $MgSO_4$, improves the quality of His-tagged protein crystals [129]. Magnesium cations are required for 2D crystallization of the vast majority

of aquaporins (Table 2.5) and other membrane proteins like bovine lens connexin, outer membrane porins and photosystem I reaction center [130-132]. Moreover, it was found that $MgCl_2$ modulates electrostatic interactions by screening charges at the lipid bilayer surface [110] confirming the important role of divalent cations in 2D crystallization of membrane proteins. Here, in case of AQP8 with the long artificial His-tagged terminus, addition of Mg^{2+} to the crystallization buffer resulted in protein aggregation or formation of small vesicles, which often clustered together. In contrast, an addition of a versatile chelating agent (EDTA) resulted in large fractions of well-ordered sheets. Presumably the elimination of divalent and trivalent cations from the crystallization buffer promoted electrostatic interaction of the phosphocholine (PC) and phosphoethanolamine (PE) polar groups [133], which appeared to facilitate AQP8 crystals formation.

Screening for the best crystallization conditions led to the conclusion that the temperature profile had a significant influence on the formation, size, and quality of crystals. A change of temperature, after the appropriate time, from room temperature to 37°C, resulted in formation of crystalline AQP8 arrays. Presumably long tubular crystals of diameter of 2 μm that formed initially at room temperature broke and became open due to the higher rigidity of lipids after the detergent removal. Sheet formation from open tubular crystals would then be subsequently facilitated by the higher flexibility of lipids after the shift to the higher temperature.

In the 2DX robot, using the cyclodextrin-driven method for detergent removal, AQP8 was reconstituted as double-layered crystalline sheets or vesicles of a few micrometers in diameter that were densely packed with proteins. A fully automated crystallization robot used for screening of crystallization conditions, in parallel to the kinetics of the reconstitution processes, controlled the temperature profile and monitored the crystal formation by light scattering. The detergent removal rate has a significant influence on the quality and size of 2D crystals for any membrane protein [134-136]. It was observed that the use of lower cyclodextrin rate during reconstitution resulted in formation of large AQP8 sheets instead of proteoliposomes. The rate should be high to stabilize the protein in the lipid environment, but sufficiently low to allow proteins to pack regularly in the bilayer.

AQUAPORIN	RESOLUTION [Å]	CRYSTALLIZATION CONDITIONS [METHOD, DETERGENT, LIPIDS, LPR, BUFFER]	CRYSTAL SYMMETRY
AQP0	1.9 (3D)	Dialysis buttons DM DMPC 0.25 10mM MES pH 6, 150 mM NaCl, 50 mM MgCl$_2$, 5 mM DTT	P42$_1$2
AQP1	3.8 (3D)	Dialysis machine OG *E. coli* lipids 0.5 20 mM Tris-HCl pH 6, 250 mM NaCl, 1 mM DTT, ± 10-100 mM Mg^{2+} or Ca^{2+}	P42$_1$2
AQP2 6 x His-tag	4.5 (3D)	Dialysis machine OG *E. coli* or Heart polar lipids 0.5 20 mM MES pH 6, 100 mM NaCl, 5 mM MgCl$_2$ or MgSO$_4$, 5 mM Histidine	P22$_1$2
AQP4M23 6 x His-tag	3.2 (3D)	Dialysis buttons OG *E. coli* total lipid extract 1.0 10 mM MES pH 6, 100 mM NaCl, 50 mM MgCl$_2$, 2 mM DTT, 10% glycerol	P42$_1$2
AQP9	7 (2D)	Dialysis buttons OG DMPC 0.4-0.6 20 mM MES pH 6, 150 mM NaCl, 50 mM CaCl$_2$,1 mM DTT	P42$_1$2
AQPZ 10 x His-tag	8 (2D)	Dialysis buttons and *Eppendorf* tube caps OG POPC:DMPC (1:1) 0.3 - 2 20 mM Citrate pH 6, 200 mM NaCl, 100 mM MgCl$_2$, 10% glycerol	P42$_1$2
GlpF 10 x His-tag	3.7 (2D) 6.9 (3D)	Dialysis machine OG *E. coli* total lipid extract 0.6 - 1.4 10 mM Tricine pH 8, 100 mM NaCl, 5 mM MgCl$_2$, 10 mM DTT	P4
SoPIP2;1	5 (3D)	Dialysis machine OG *E. coli* polar lipid extract 0.3-0.4 20 mM Tris-HCl pH 8, 100 mM NaCl, 50 mM MgCl$_2$, 2 mM DTT	P4 rarely P22$_1$2
AtTIP3;1	7.7 (2D)	*Eppendorf* tube caps DHPC *Soybean* phosphatidylcholine 25 mM TEA pH 7.5, 100 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 0.1 mM Butylated hydroxytoluene (antioxidant)	P2

Table 2.5: Two-dimensional crystallization of aquaporins to date. 2D/3D map resolutions, 2D crystallization conditions and crystal symmetries are listed. References: [26, 30, 33, 129, 137-141].

Different 2D crystallization conditions resulted in tubular, vesicle or sheet-like crystals and hence diverse protein packing arrangements in reconstituted lipid bilayers. In the majority of cases, the power spectra showed two sets of diffraction spots indicating two crystalline layers and thus double-layered crystal or collapsed vesicles. The final 2D

projection map of AQP8-cmyc-His-tag at resolution of 8 Å revealed individual monomeric aquaporins in a tetrameric organization with a well visible central depression. AQP8 packs into arrays with p4 symmetry, analogously to GlpF [139] and SoPIP2;1 [140], housing one tetramer per unit cell of size 72 Å.

Atomic force topographs recorded from AQP8 2D crystals adsorbed to the mica showed the crystal thickness to be 12 nm. Taking into account thickness of the single-layered 2D crystals of the bacterial aquaporin AQPZ (5.7 nm) [142] and human AQP1 (5.8 nm) [143, 144] the measured AQP8 crystals consisted of two layers. No single layers were detected during AFM measurements, however cluster areas having a thickness of about 24 and 36 nm made of 2 and 3 double 2D crystal layers, respectively, were frequently observed. In parallel, 2D crystals were gold labeled using Ni-NTA-bound gold beads; the result indicated that C-terminal His-tags were exposed on the surface of the crystal.

Taken together, the EM, AFM and gold labeling results suggest that AQP8 2D crystals expose the cytoplasmic side, N- and C- termini, of AQP8-cmyc-His-tag to the medium, while the extra-cellular side is trapped within the double-layered sandwich, a packing observed for AQP0 [26], AQP4 [33] and SoPIP2;1 [140].

2.3.3   Analysis of digested AQP8

With the purpose of improving the crystallization and the resolution of projection maps, the 2D crystals were exposed to protease Lys-C. In the 283-aa sequence of the human AQP8 gene, seven possible Lys-C cleavage positions were found. However, when the protein is embedded in a lipid bilayer only one Lys on the N terminus (K11) and 3 Lys on the C terminus (K248, K254 and K265) were predicted from the sequence alignment to AQP1 to be exposed to enzymatic treatment. The two Lys (K175, K177) located on the 8-aa long cytosolic D-loop and one Lys located in transmembrane helix 3 (K124) were supposed to be inaccessible for the enzyme. The digestion by Lys-C, when added after 36 h of crystallization, yielded rectangular crystals with sizes of several micrometers in diameter, and diffracting up to 3 Å. Mass spectroscopy and SDS-PAGE analysis revealed that AQP8, ordered in 2D arrays after Lys-C treatment, had lost 11 amino acids of the N-terminal region. To investigate the digestion process at the C-terminal end Ni-NTA-gold labeling and western blot able to detect His-tags were performed. Taken together, the results demonstrated that after protease addition the well diffracting crystal is composed of several populations of protein. Protease removes not only N- and C-terminal moiety but as

well in certain percentage of AQP8 hydrolyzes the Lys peptide bond on the D-loop. Surprisingly, experiments indicated that protein digestion resulted in systematic crystal distortions. This is revealed by the resulting diffraction patterns, which exibit symmetrically distributed stripes. In AQP1 and AQP4 the subunits in the tetramer interact with each other through cytoplasmic loop-D connecting helices 4 and 5 [33], hence it is possible that disruption of AQP8 loop-D resulted in weakening of the tetramer interactions. Although the diffraction peaks were visible even above the resolution 3 Å the regular crystal distortions prevented the exploitation of the high-resolution information by standard crystallographic methods.

In view of the fact that the crystallization conditions were optimized, the removal of the loose termini by proteolytic treatment seems to be crucial for obtaining large, well diffracting rectangular double-layered sheets of AQP8. Thus, proteolytic cleavage of the long C-terminus improved the crystallinity considerably, which might be considered for the 2D crystallization of other recombinant membrane proteins bearing large tags that facilitate their purification.

2.3.4   The reason of crystallization failure and new constructs strategy

Human AQP8 crystallization trials resulted in a large variety of the crystal formation of different sizes and diverse protein packing arrangements in reconstituted lipid bilayers. In optimized crystallization conditions the resolution of electron diffraction of AQP8 crystals improved from 8 to 3 Å after addition of Lys-C protease that removed N- and C-terminal moiety. Unfortunately, in parallel to inducing a better protein arrangement the protease had also hydrolyzed the peptide bond on the D-loop that led to cleavage into two pieces of a certain percentage of AQP8 and possibly to a weakening of crystallographic interactions between the layers. The systematic crystal distortion resulted in diffraction pattern disturbances. Because the crystallization problem was clearly bound with too long, random coil, C-termini of the protein that destabilized the crystal formation, and the addition of Lys-C had other inhibitory effects, the only possible way to obtain better crystals was producing shorter constructs of AQP8.

Three new shorter AQP8 constructs designed with aim to solve the high-resolution channel structure were introduced into the *P. pastoris* genome and expressed. The strategy was to modify the original construct in order to produce a protein having properties as similar as possible to the native human AQP8 and at the same time to allow its successful

crystallization. The first two constructs (AQP8-TEV-cmyc-His-tag and AQP8-His-tag) have modified C-termini, shortened by the introduction of protease TEV cleavage site or the direct removal of the artificial link residues. Only these two constructs have been investigated to date. The expression level of the original construct and AQP8-His-tag were comparable and allowed highly concentrated protein to be purified and to perform 2D crystallization trials. The AQP8-TEV-cmyc-His-tag expression was much lower; moreover the longer purification protocol was associated with some loss of the protein.

Analysis of AQP8-His-tag crystallization results highlighted the role of protein purification tags and that the construct length has an influence on crystal formation. During the first attempts, AQP8-His-tag formed vesicles and crystalline sheets with crystal parameters $a = b = 65$ Å, $\gamma = 90°$. Interestingly, in the AQP8-His-tag construct crystallization addition of 5 mM $MgCl_2$ was crucial to obtain ordered 2D crystals, strikingly different to the results with the original construct. The anticipated correlation between AQP8-His-tag pI of the protein and the optimal crystallization pH is consistent with obtained results since the theoretical pI of AQP-His-tag is 7.14 and optimal crystallization pH was 7.2. The protein had reconstituted easily in several crystallization conditions and in different lipid environments. When after the first screenings in negative stain EM the observed crystals were inspected by cryo EM, this showed that these crystals are not perfect and, unfortunately, are diffracting below expectations.

The structure of AQP4, which is the major water channel in the mammalian brain and, similar to AQP0 [145], plays a key role in membrane adhesion, has been determined by electron microscopy in 2006 [33]. Like the human AQP8 it has an unusually long N-terminus in the comparison to other mammalian aquaporins. It was reported that AQP4 exists in two alternative splicing isoforms: as a full-length protein starting with Met1 (AQP4M1) and an alternative shorter starting with Met23 (AQP4M23). Surprisingly, the large and well-ordered 2D crystals were formed only by the shorter protein version while full-length AQP4 did not crystallize. It was found that the order of AQP4M23 crystals was destroyed after addition of a Met1-Ile22 peptide to the 2D crystal solution, which pointed to the highly conserved Arg9 being responsible for the destabilization of orthogonal arrays. Although human AQP8 does not have any additional Met in the N-terminal sequence suggesting an alternative-splicing version, the unusually long N-terminus and the residue Arg18, which might correspond to the Arg9 in AQP4, could have a serious influence on the 2D crystal formation. After an alignment of AQP8 to the AQP4 sequence it turned out that these water channels both have extraordinary long N-termini (Figure 2.18). The

66

residue AQP4Met23 was aligned with AQP0Met1, AQP1Ala2 and AQP8Gly23. The sequence similarity of the AQP4 and human AQP8 N-termini led to the design and production of the truncated AQP8 construct starting from AQP8Gly23 (called MG$_{18}$-AQP8-His-tag), which, in the near future, may form well diffracting 2D crystals being the major prerequisite for the AQP8 high-resolution structure determination.

Recently, it was reported that when AQP8 was expressed in yeast, in addition to the 28-kDa AQP8 band on the SDS-PAGE, a 26-kDa band was also identified suggesting that protein import into the yeast mitochondria may be accompanied by a proteolysis of AQP8 N-terminus polypeptide [60]. This study is another argument that speaks for the 2D crystallization of the shorter version of AQP8.

```
                                                        α1
AQP4_HUMAN/1-323     1  MSDRPTARRWGKCGPLCTRENIMVAFKGVWTQAFWKAVTAEFLAMLIFVLLSLGSTINWGGTEKP...LP
AQP8_HUMAN/1-261     1  MSGEIAMCEPEFGNDKAREPSVGGRWRVSWYERFVQPCLVELLGSALFIFIGCLSVIENGTDTG......
AQP0_HUMAN/1-263     1  ....................MWELR...SASFWRAIFAEFFATLFYVFFGLGSSL..RWAPG.....P
AQP1_HUMAN/1-269     1  ...................MASEFK...KKLFWRAVVAEFLATTLFVFISIGSALGFKYPVGNNQTAV
                        1    5    10   15   20   25   30   35   40   45   50   55   60   65   70

                                            α2                        α3                        α4
AQP4_HUMAN/1-323    68  VDMVLISLCFGLSIATMVQCFCHISGGHINPAVTVAMVCTRKISIAKSVFYIAACLGAIIGAGILYLVT
AQP8_HUMAN/1-261    65  ..LLQPALAHGLALGLVIATLCNISGGHFNPAVSLAAMLIGGLNLVMLLPYWVSQLLGGMLGAALAKAVS
AQP0_HUMAN/1-263    39  LHVLQVAMAFGLALATLVQSVCHISGAHVNPAVTFAFLVGSQMSLLRAFCYMAAQLLGAVAGAAVLYSVT
AQP1_HUMAN/1-269    47  QDNVKVSLAFGLSIATLAQSVCHISGAHLNPAVTLGLLLSCQISIFRALMYIIAQCVGAIVATAILSGIT
                        75   80   85   90   95   100  105  110  115  120  125  130  135  140

                                η1                    α5                          α6
AQP4_HUMAN/1-323   138  PP...SVVGGLGVTMVHGNLTAGHGLLVELIITFQLVFTIFASCDSKRTDVTGSIALAIGFSVAIGHLFA
AQP8_HUMAN/1-261   133  PEERFWNASGAAFVTVQEQGQVAGALVAEIILTTLLALAVCMGAINEKTK.GPLAPFSIGFAVTVDILAG
AQP0_HUMAN/1-263   109  PP...AVRGNLALNTLHPAVSVGQATTVEIFLTLQFVLCIFATYDERRNGQLGSVALAVGFSLALGHLFG
AQP1_HUMAN/1-269   117  SS...LTGNSLGRNDLADGVNSGQGLGIEIIGTLQLVLCVLATTDRRRDLGGSAPLAIGLSVALGHLLA
                        145  150  155  160  165  170  175  180  185  190  195  200  205  210

                                    α7                      α8
AQP4_HUMAN/1-323   205  INYTCASMNPARSFGPAVIMGNWENHWIYWVGPIIGAVLAGGLYEYVFCPDVEFKRRFKEAFSKAAQQTK
AQP8_HUMAN/1-261   202  GPVSCGCMNPARAFGPAVVANHWNFHWIYWLGPLLAGLLVGLLIRCFIG....................
AQP0_HUMAN/1-263   176  MYYTCAGMNPARSFAPAILTGNFTNHWVYWVGPIIGGGLGSLLYDFLLFPRLK.......SISERLSVLK
AQP1_HUMAN/1-269   184  IDYTGCGINPARSFGSAVITHNFSNHWIFWVGPFIGGALAVLIYDFILAPRSS.......DLTDRVKVWT
                        215  220  225  230  235  240  245  250  255  260  265  270  275  280

AQP4_HUMAN/1-323   275  GSYMEVEDNRSQVETDDLILKPGVVHVIDVDRGEEKKGKDQSGEVLSSV
AQP8_HUMAN/1-261   251  .......DGK.....TRLILKAR..............................
AQP0_HUMAN/1-263   239  GAKPDVSNGQPEVTGEPVELNTQAL............................
AQP1_HUMAN/1-269   247  SGQVEEYDLDADDINSRVEMKPK..............................
                        285  290  295  300  305  310  315  320  325
```

Figure 2.18: The sequence alignment of four human aquaporins (AQP4, AQP8, AQP0, AQP1). High homology regions as well as AQP4 secondary structure are highlighted. Alignment was generated using ClustalW2 [65] (default parameters, matrix Gonnet250, number of iterations 10) and graphically illustrated by ESPript 2.2 [66]. Met23 of AQP4 was marked with black arrowhead.

2.4 Conclusion

The main interest in structure determination of the human AQP8 is connected with its unique position within the aquaporin family. The human AQP8 cDNA has been cloned in 1998 [54] and since then the intensive studies concerning its intravesicular localization and activity have been preformed [60, 146]. Expressed in the inner mitochondrial membrane of several mammalian tissues, found in liver, gastrointestinal tract, testis, airways and kidney cells, the ammonia-permeable AQP8 could be essential for the organism metabolism.

Here, the overexpression in yeast *Pichia pastoris*, purification and 2D crystallization of human AQP8 was presented. Protein purity and homogeneity in different detergents, made it possible to perform wide crystallization screenings. By using different crystallization conditions and methods, ordered crystalline sheets were obtained. AFM, EM and gold labeling experiments revealed the double-layered nature of 2D crystals, tetrameric organization and exposed outwards protein N- and C- termini of AQP8 within 2D crystals. In parallel, an enzymatic removal of N- and C- terminal moiety was the significant approach to obtain the fine diffracting crystals.

The presented results and new designed constructs, AQP8-His-tag and AQP8 truncated version, $MG_{18}AQP8$-His-tag, are a step towards the goal to solve the high-resolution structure of human AQP8.

3   Structure and symmetry determination of the YscC oligomer

Abstract

The translocation of proteins across the biological membrane is an essential part of cellular life. The type III secretion system is a major factor for the virulence or symbiosis of many Gram-negative bacteria that infect plants and animals. Bacterial effector proteins are delivered via T3SS injectisome from the pathogen cytoplasm into the eukaryotic host cells, in which they modulate the host innate immune response.

The outer membrane-localized YscC oligomer belongs to the secretin family and is one of the main components of the injectisome. In this study, the YscC complex was expressed in the avirulent strain of *Yesinia enterocolitica* and purified in order to construct a 3D model from cryo Electron Microscopy single particle images.

Various approaches like rotary metal shadowing of trypsin digested oligomers, mass spectrometry and Scanning Transmission Electron Microscopy were used for oligomer symmetry evaluation. In spite of considerable effort the symmetry could not be precisely assigned and the value was estimated to be either C12 or C13. The generated 3D structure of the YscC complex disclosed the N-terminal flexible domain, which forms the part of a large chamber between the bacterial outer and inner membranes, the conical shape periplasmic domain, and two differently sized ring-shaped domains linked by fine density connectors corresponding to the outer membrane spanning regions. The sample trypsinization revealed the protease-resistant core of the protein.

Comparison of the sequence and structure of YscC to other, close or more distant related secretins made it possible to define homology regions located both on the N- and C-termini of the protein.

3.1   Materials and Methods

3.1.1   YscC expression in bacteria *Yersinia enterocolitica* and the YscC oligomer purification

The pYV-cured *Yersinia enterocolitica* strain carrying plasmids *pMA6* and *pRS6* containing the *yscC* and *yscW* genes, respectively, was grown in 40 ml brain heart infusion (BHI) broth at 37°C overnight. To induce expression of *yscC*, *Y. enterocolitica* strain was inoculated at OD_{600} of 0.1 in 4 liters of BHI broth supplemented with 0.4% glycerol, 20 mM $MgCl_2$, and 20 mM sodium oxalate (BHI-OX). In addition, ampicillin (1 mg/ml), nalidixic acid (25 mg/ml), and tetracycline (10 mg/ml) were used. The culture was grown for 2 h at room temperature, arabinose (final concentration, 0.05%) was added, and the culture was grown for 2 h at 37°C.

The entire YscC purification was performed on ice. Bacterial cells were washed with 0.9% NaCl and resuspended in 100 ml lysis buffer (50 mM Tris-HCl pH 8.5, 1 mM EDTA). After sonication at 4°C, cells were pelleted at 4000 g for 10 min and envelopes were removed by centrifugation at 41000 g for 1 h. *Y. enterocolitica* membranes were solubilized with 6 ml of extraction buffer containing 2% DDM (Anatrace), 50 mM Tris-HCl pH 7.8, 250 mM NaCl, 5 mM EDTA and protease inhibitor (complete protease inhibitor; Roche) for 1 hour at room temperature. Not solubilized debris were removed by centrifugation at 100000 g for 1 h. After the addition of sucrose to the final concentration of 15% (wt/wt), the extracted membrane proteins were layered on top of a 20 to 40% (wt/wt) sucrose gradient in the gradient buffer (0.04% DDM, 50 mM Tris-HCl pH 7.8, 250 mM NaCl, 5 mM EDTA, protease inhibitor) and centrifuged at 38000 rpm in SW41 rotor (Beckman) for 30 h. Sucrose gradient fractions containing the YscC oligomer were combined, dialyzed against chromatography buffer (0.04% DDM, 10 mM Tris-HCl pH 7.8, 100 mM NaCl, 0.1 mM EDTA) and loaded on a MonoQ 5/50 GL (GE Healthcare) ion exchange column. The salt gradient ranged from 100 mM to 1 M NaCl. The YscC oligomer was eluted at 450 to 500 mM NaCl. The pure YscC oligomer was separated from YscC oligomer dimers and other proteins by gel filtration on Superose 6 10/300 GL column (GE Healthcare). Fractions containing the YscC oligomer were controlled by semi-native 3-8% Tris-Acetate NuPAGE and stored in 100 μl aliquots at -20°C for electron microscopy.

3.1.2   Protein reconstitution into the lipids

The purified YscC oligomer was mixed with DDM-solubilized *E. coli* polar, DMPC and phosphatidyl lipids to a final protein concentration of 0.5 mg/ml and a final DDM concentration of 1% in an end volume of 100 μl. The LPRs of 0.5, 1 and 2 were tested. 10 μg of Bio-Beads (Bio-Rad) were added to the ternary mixture that was vigorously mixed overnight at room temperature. Negatively stained samples were checked and images were taken with a Phillips CM-10 electron microscope operated at 80 kV.

3.1.3   Cryo Electron Microscopy

5 μl aliquot of the purified Ysc oligomer sample (0.05 mg/ml) was adsorbed to the thin carbon-coated copper grid, washed with 4 drops of nanopure water and then plunge-frozen in liquid ethane [147] using the Vitrobot (Maastricht Instruments) [148]. The Vitrobot is the fully automated, PC-controlled, rapid cooling device for vitrification of aqueous samples. The grid preparation process takes place in a humidity, and temperature, controlled chamber. The following parameters were controlled: humidity (80%), blotting excess fluid time (3.5 sec) and blotting force (+1). The frozen grids were transferred to a storage box immersed in liquid nitrogen and then to the Philips CM200 FEG electron microscope with Gatan cryoholder. The microscope was operated at an accelerating voltage of 200 kV. Single particle electron micrographs were recorded on Kodak SO163 film at a magnification of 50000x, at low electron doses and with defocus ranged between 2 μm and 5 μm. Negatives were scanned with a Heidelberg Primescan D7100 drum scanner in a resolution of 2 Å /pixel.

Figure 3.1 The vitrification Robot [148].

3.1.4   The YscC oligomer trypsin digestion

100 μg of the purified YscC oligomer (0.1 mg/ml) was incubated with 1 μg/ml of N-p-tosyl-L-phenylalanine chloromethyl ketone-treated trypsin (Sigma) for 2 h at room temperature to cleave YscC N-terminal domain. Pefabloc was then added to concentration of 0.1 mg/ml, and the sample was immediately injected onto a Superose 6 10/300 GL column. A control experiment was performed under the same conditions but the trypsin was omitted. In both cases fractions were collected, aliquots were treated with phenol to dissociate oligomers [149] and analyzed by 4-12% Bis-Tris NuPAGE. Briefly, 0.2 ml of the oligomer sample was treated with 0.5 ml of 88% phenol at 70°C for 10 min, then cooled and centrifuged at 5000 rpm for 10 min at 4°C, and an aqueous phase (without an interface) was removed. The interface and the phenol phase were mixed with nanopure water at 70°C, centrifuged at 5000 rpm for 10 min at 4°C and an aqueous phase was discarded. The phenol phase was treated twice with 0.7 ml acetone and once with 0.7 ml of diethyl ether, the sample was centrifuged and a supernatant discarded after each step. Pellets were dissolved in 100 μl of 1X SDS-PAGE sample buffer, boiled and frozen.

3.1.5   Metal Shadowing

5μl of purified trypsinized sample (0.04 mg/ml) was adsorbed on the carbon-covered copper grid, blotted twice with Whatman filter paper, washed with 5 drops of nanopure water to remove excess detergent and quickly plunged into liquid ethane cooled to -176°C [147]. A frozen grid was transferred to the Balzers BAF400 freezeetching unit and freeze-dried at -90°C for 5 h in a vacuum of ~1.3×10^{-5} Pa. 1-2 nm layer of platinum/carbon was evaporated onto the grid rotating at angle of 45° (Figure 5).

Electron micrographs were recorded on the CCD camera at Philips CM-10 at a nominal magnification of 50000x (at high defocus) and 100000x (at low defocus) as described in 2.2.5. EMAN Boxer [150] was used to collect two sets of particles (1300 particles with box size of 70x70 pixels taken at magnification 50000x; 1000 particles with box size of 140x140 pixels taken at magnification 100000x, respectively) which were normalized, aligned and processed in EMAN software package [151].
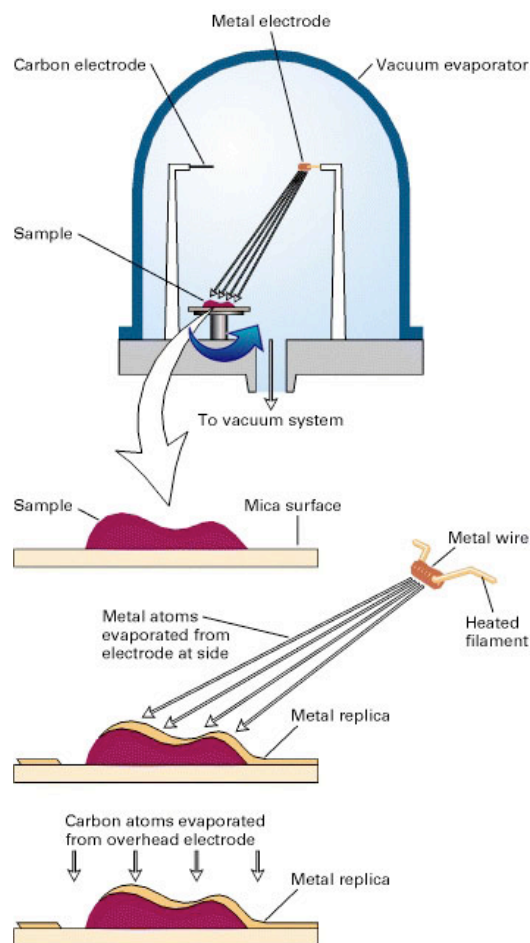
Figure 3.2: The metal shadowing technique. The sample was adsorbed on the grid, frozen in the liquid ethane, transferred to a vacuum evaporator and dried under optimized conditions. When filament of platinum was heated electrically the metal evaporated and felt over the rotating grid in a very thin layer. In order to stabilize the replica, the sample was coated with a carbon film [152].

3.1.6   Scanning Transmission Electron Microscopy

5 µl aliquots of YscC samples were adsorbed for 1 min to the glow-discharged thin carbon-coated STEM grids. The grids were blotted to remove excess of liquid and washed with 8 drops of a quartz double-distilled water to remove detergent frozen by plunging into liquid N_2, and freeze-dried at -80°C and $5x10^{-8}$ Torr overnight in the STEM. Tobacco mosaic virus (TMV) particles adsorbed on the separate grid and washed with 10 mM ammonium acetate were used as the mass standard. The Vacuum Generators (East Grinstead) HB-5 STEM, connected to a modular computer system (Tietz Video and Image Processing Systems) were used [153]. For a mass determination, the series of 512x512

pixels dark-filed images were recorded at the accelerating voltage of 80 kV and a nominal magnification of 200000x. The recording dose was 500 electrons/nm^2. The digital images were evaluated with the IMPSYS program package (Figure 3.3). Particles were selected by circles to calculate the total mass, and then the background was subtracted. The mass of measured TMV was used for scaling the results. The mass values were compared and displayed in histograms and fitted with Gauss curves.
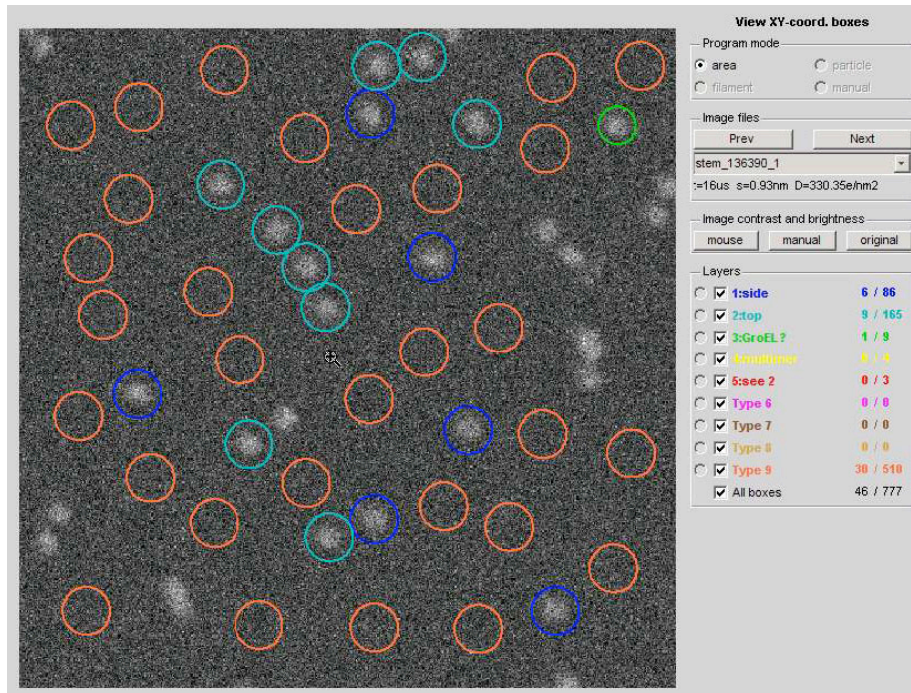
Figure 3.3: STEM calculations were performed in program IMPSYS. Particles were selected by circles to calculate the mass. *Red* circles correspond to the background selection, *dark blue* to side views and *blue* to top views of the YscC oligomer, *green* to a contaminant (possibly GroEL).

3.1.7   Mass spectrometry of the intact and the trypsinized YscC oligomer

Intact and trypsinized YscC oligomers were dissociated after phenol treatment [149] as described in 3.2.4. Single bands on the 4-12% Bis-Tris NuPAGE corresponding to the wild and digested YscC monomers were stained with silver [154], cut out from a gel and send to the Mass Spectrometry Facility, Biozentrum, University of Basel [111]. Proteins isolated from the gel were digested with trypsin and resulting peptides were

74

separated by the capillary reverse phase chromatography coupled to the mass spectrometer. The masses of the peptides were measured and used for the oligomer fragments identification.

3.1.8   Single particle processing using EMAN

Single particles were selected from EM negatives with the EMAN`s Boxer software and processed using the EMAN 1.9 suite (Figure 3.6, [150], [151]). Approximately, 8000 negatively stained particles and 37200 particles from cryo EM images were collected for the processing.

Processing of single particles from cryo EM data

Cryo EM negatives scanned with the resolution of 2 Å/pixel were displayed in the interactive program *Boxer*. Every selected particle was saved in a separate subimage of the size 148x148 pixels. All particle images were saved in a **.img/*.hed* files and normalized (*edgenorm*).

If the reconstructed model is supposed to have a resolution higher than the "first zero" of the Contrast Transfer Function (CTF), a CTF correction, which includes the phase of the images and an amplitude correction, has to be applied. CTF correction (command *ctfit*) was applied by inspecting the Fourier transform of each cryo EM negative. The CTF parameters (Cs – spherical aberration correction fixed by microscope; defocus; B factor /envelope function width – defines the attenuation of amplitudes; amount of amplitude contrast) were determined and the predicted CTF curve was fitted to the displayed profile (Figure 3.5). CTF correction was done for all particle images from one negative.

For the centering and aligning the particles, a command *cenalignint* averaged all of the particles together, and then aligned each particle to this average by shifting it by defined number of pixels. Not centered particles were discarded. The process was iterated until the particles stop moving. The 2D refinement (*refine2d*), which classifies raw particles to produce a set of high-contrast class averages, gave the first good impression on the data quality. An initial model *threed.0a.mrc* was generated with applied cyclic 11-, 12- and 13-fold symmetries (C11, C12 and C13) (*startAny, startcsym*).

To determine orientations of particles and match them to the projection, each particle was compared to a set of reference projections. *Refine* command created the loop

that generated reference projections from starting model (*project3d*), classified images by comparing them to the reference projections (*classesbymra*), averaged classes (*classalignall*) and calculated the reconstruction from class averages (*make3d*). In parallel to the refinements with 11, 12 or 13-fold symmetry supplied as an input parameter (*refine*) the single particle 3D reconstruction based on the multi-refinement protocol (*multirefine*) was performed [155]. In EMAN, the resolution evaluation was performed with the program *eotest*, which splits each of the classes into halves and makes class averaging and 3D reconstruction of two independent sets. From the Fourier Shell Correlation (FSC) curve the 0.5 criterion was used to determine the generated model resolution.

The generated models were visualized in program Chimera [156] for interactive visualization and analysis of molecular structures.
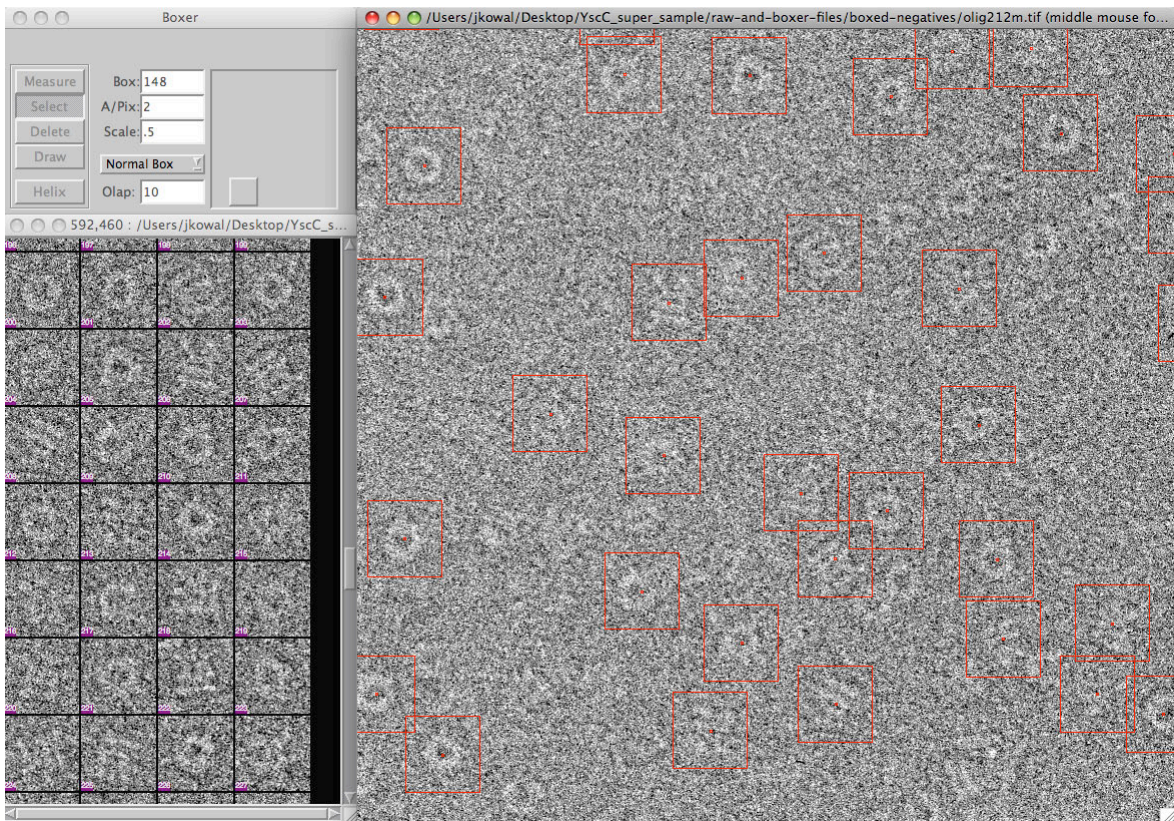
Figure 3.4: The particles were windowed into 148x148 pixels boxes using EMAN's Boxer software [150].
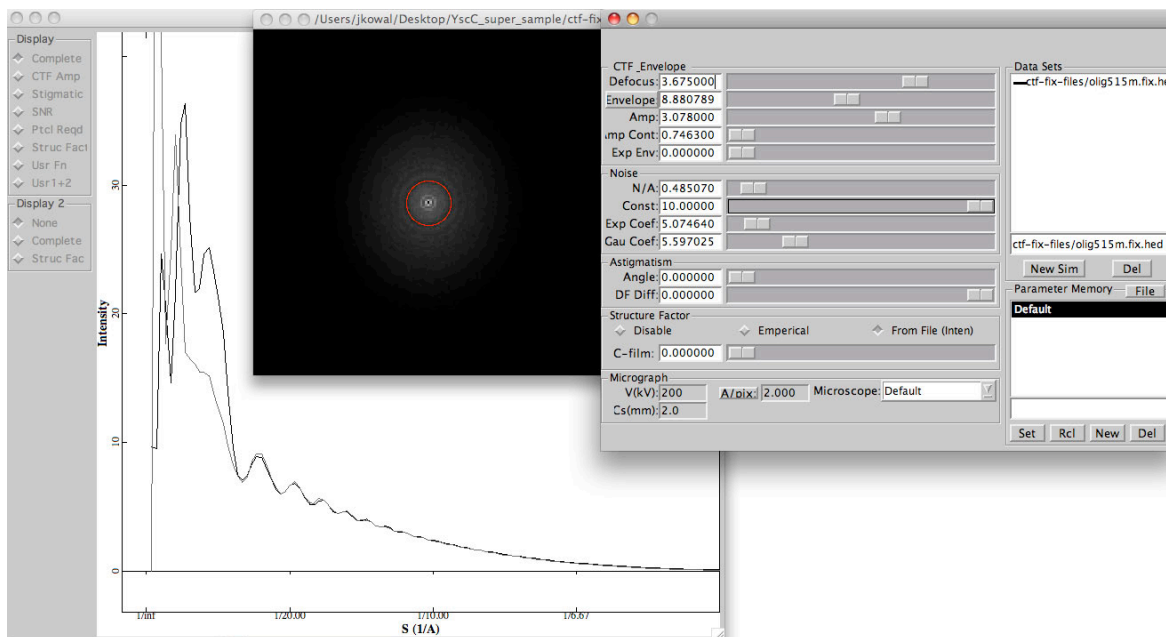
Figure 3.5: Program *ctfit* was run to determine the CTF parameters and apply the phase correction. A plot window, an image window, and the control panel are displayed.
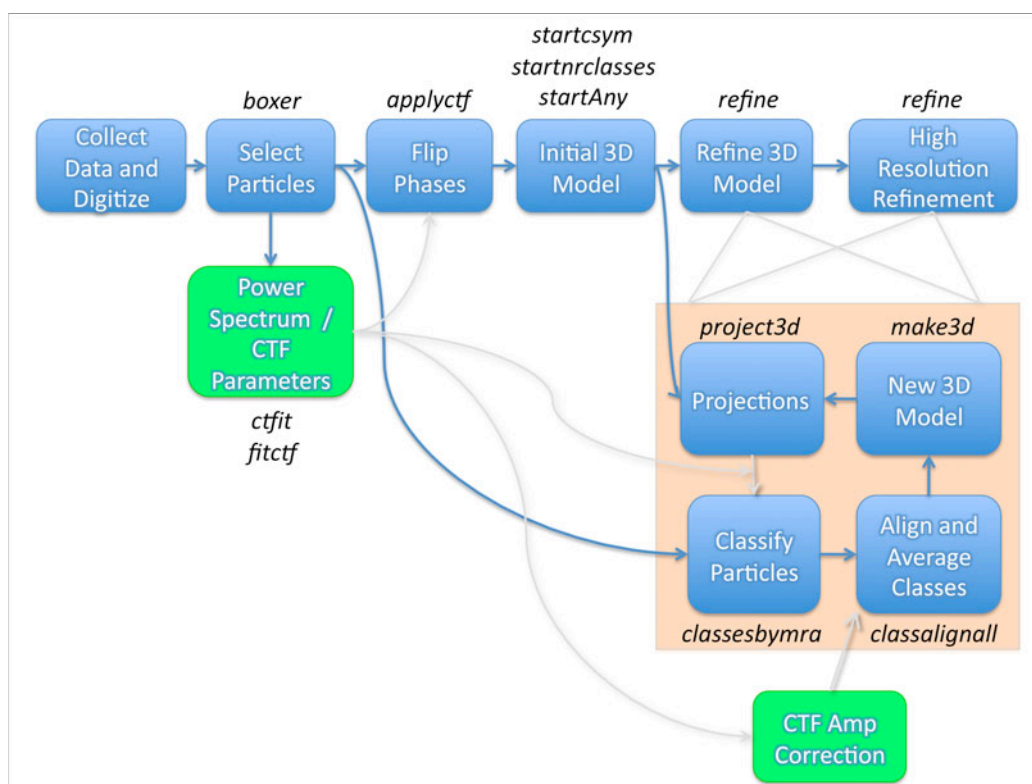
Figure 3.6: The typical EMAN reconstruction strategy pipeline with the commands used during the single particle processing. The strategy includes the following steps: picking of the particles, the CTF correction of images, generating the 3D initial model, finding particle orientations and calculating the 3D reconstruction using *refine* command.

Processing of negatively stained particles

Negatively stained particles were picked from CCD images taken at CM-10 electron microscope at nominal magnification of 50000x, and processed in the similar manner as cryo EM particles except that CTF correction was omitted. Cyclic 12-, 13- and 14-fold symmetries (C12, C13 and C14) were imposed to the 3D initial models. For negatively stained particles not only monomeric but as well dimeric form of the YscC oligomer was generated.

3.2   Results

3.2.1   YscC expression in bacteria *Yersinia enterocolitica* and the YscC oligomer purification

The pYV-cured, avirulent, *Yersinia enterocolitica* strain carrying plasmids *pMA6* and *pRS6* containing the *yscC* and *yscW* genes, respectively, was used to express the YscC oligomer. The YscW, a pilot protein introduced into the strain, was necessary to produce and localize the YscC oligomer in the membranes [96, 99]. After cells sonication and harvesting, the *Y. enterocolitica* membranes were solubilized with 2% DDM, non-ionic detergent which is one of the gentler detergents with the very low CMC of 0.0087%. YscC oligomer was purified using the sucrose gradient, during which outer and inner *Y. enterocolitica* membrane proteins were separated, followed by an ion exchange as described in [95] and size-exclusion chromatography (Figure 3.7). From the MonoQ ion exchange column protein was eluted in two peaks, the first corresponded to the monomeric form and the second to the oligomer. The Superose 6 gel filtration column step was significant to separate the highly pure oligomer from small contaminants <50 kDa, YscC monomeric form (~63 kDa) and dimers made of YscC oligomers (~1.5 MDa). To avoid aggregation of solubilized protein, 100 mM NaCl was added to the all purification buffers. No other proteins were detected in the specimen by semi-native silver stained 3-8% NuPAGE, single particle EM or STEM. The yield of purification was approximately 0.1 mg protein per liter of the *Y. enterocolitica* culture.
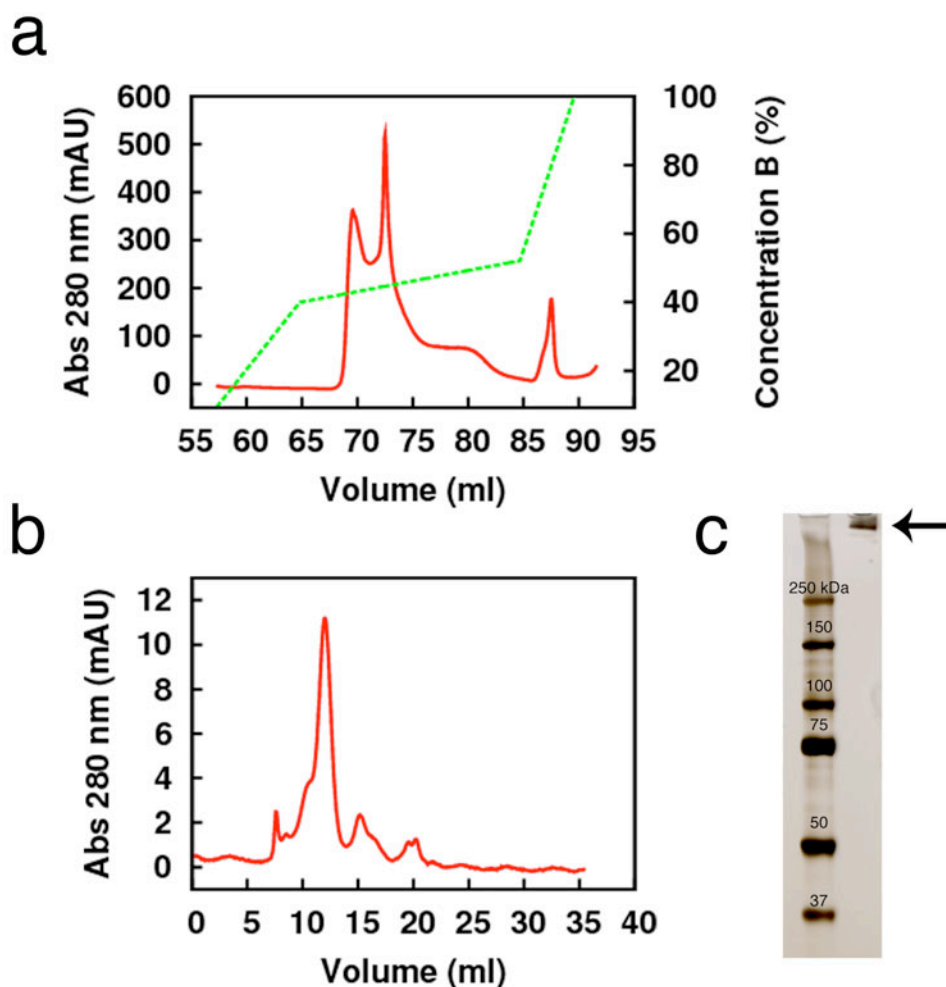
Figure 3.7: Purification of the YscC oligomer. Secretin was purified from *Y. enterocolitica* membranes solubilized with 2 % DDM detergent. Outer and inner membrane proteins were separated using the 20-40% sucrose gradient. (a) Elution profile on MonoQ 5/50 GL ion exchange column. Green dashed line corresponds to the salt gradient profile. Protein was eluted at 450 to 500 mM NaCl. First elution peak corresponds to monomers (64.24 kDa) and second to the oligomer (~800 kDa) YscC. (b) The pure YscC oligomer was separated from YscC oligomer dimers and other proteins by gel filtration on Superose 6 10/300 GL column in 10 mM Tris-HCl pH 7.8, 100 mM NaCl, 0.1 mM EDTA and 0.04% DDM. The highest peak corresponds to oligomer. (c) A silver stained 3-8% NuPAGE with the oligomer marked by black arrow purified on Superose 6 column.

3.2.2 Reconstitution trials of the YscC oligomer into the lipid bilayer

To investigate if the purified YscC oligomer could be reconstituted into the lipids the protein was mixed with DDM-solubilized *E. coli* polar extract, DMPC or phosphatidyl lipids (POPA:POPE:POPC, in ratio 1:2:7) to a final protein concentration of 0.5 mg/ml. The detergent was selectively removed by adsorption onto Bio-Beads. The best results were obtained for *E. coli* polar lipids at LPR 0.5. It was observed that oligomer particles formed small patches made of 10-20 particles in between the lipids (Figure 3.8). Additionally, oligomers were attached to the lipid vesicles of diameter 100-200 nm. Embedded into the lipids with the C-terminal protease-resistant secretin homology regions and keeping the N-terminal flexible domains outwards, oligomers were attached to the surface of vesicles. For DMPC and phosphatidyl lipids aggregates made of several to tens of oligomers were found.
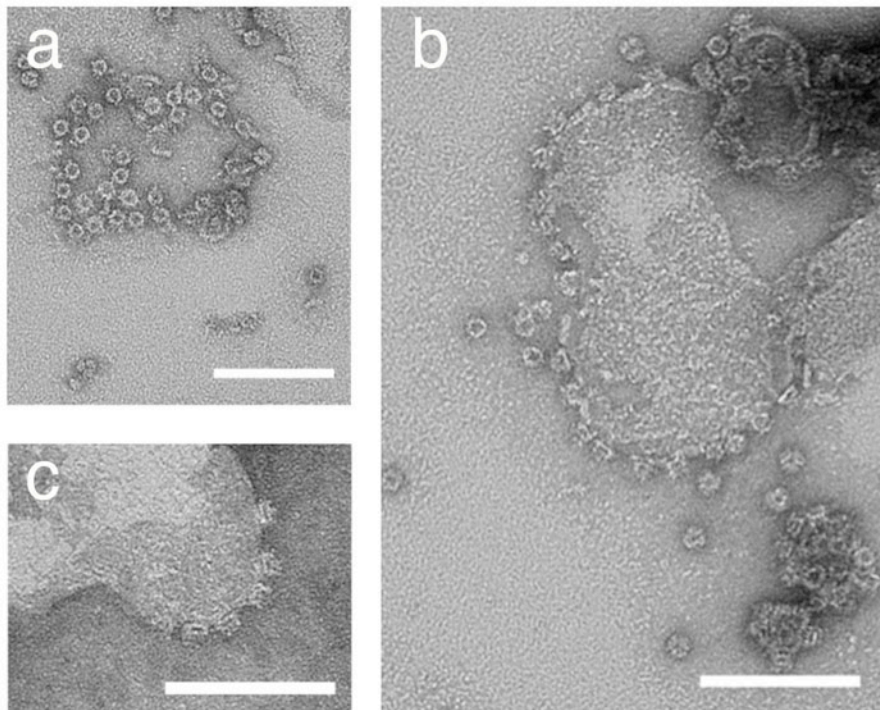
Figure 3.8: Reconstitution trials of the oligomer into the lipids. The purified YscC oligomer was mixed with DDM-solubilized *E. coli* polar lipids extract. The detergent was removed from the ternary mixture by Bio-Beads. (a) Small patches formed by a few particles between the lipids. Top views of particles are well visible. (b) YscC oligomers attached to the surface of the lipid vesicle making it possible to observe oligomers side views. C-terminal part of the protein is embedded into *E. coli* polar lipids and N-terminal

flexible part jut out of the vesicle. In the right bottom corner oligomers formed aggregates made of a number of particles. (c) YscC particles integrated into the lipid vesicle. Scale bars correspond to 100 nm.

3.2.3   Removal of the flexible domain of the YscC oligomer
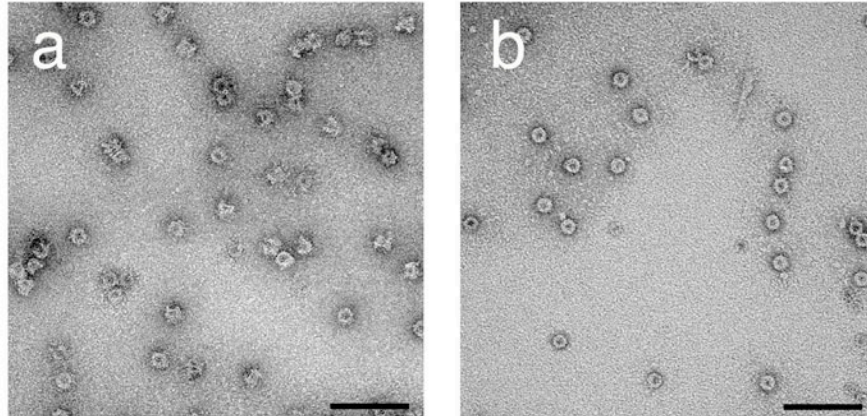
Figure 3.9: Specimens of intact and trypsinized YscC negatively stained oligomers adsorbed to STEM grids visualized by EM. Whereas in (a) the oligomer occurs in side- and top- views orientations, in (b), after removal of N-terminal flexible domain and C-terminal moiety, only top views are present.

The YscC oligomer trypsinization was aimed to reveal the protease-resistant C-terminal core, which is a homology region in all secretins. A removal of flexible domain, which covers part of trypsin-resistant core, could facilitate the discovery of the oligomer symmetry.

The purified trypsinized and intact YscC complexes (Figure 3.9) were subjected to the size exclusion chromatography. The elution profiles are compared on Figure 3.10 a. The intact oligomer eluted at 12 ml and the protease-digested complex at 13 ml. For both samples, the peak at 15.5 ml of elution corresponded to a few monomers of oligomer together. For intact oligomers the peak at 10 ml corresponded to oligomer dimer complexes as confirmed by negative stain EM. STEM analysis (Figure 3.10 c-d) showed that the intact oligomer solubilized in DDM has a mass of 941 kDa (± 74 kDa) and trypsinized of 576 kDa (± 65 kDa).

After the phenol treatment of trypsinized YscC oligomer a single band with an apparent molecular size of ~32 kDa was determined by 4-12% NuPAGE (Figure 3.10 b).
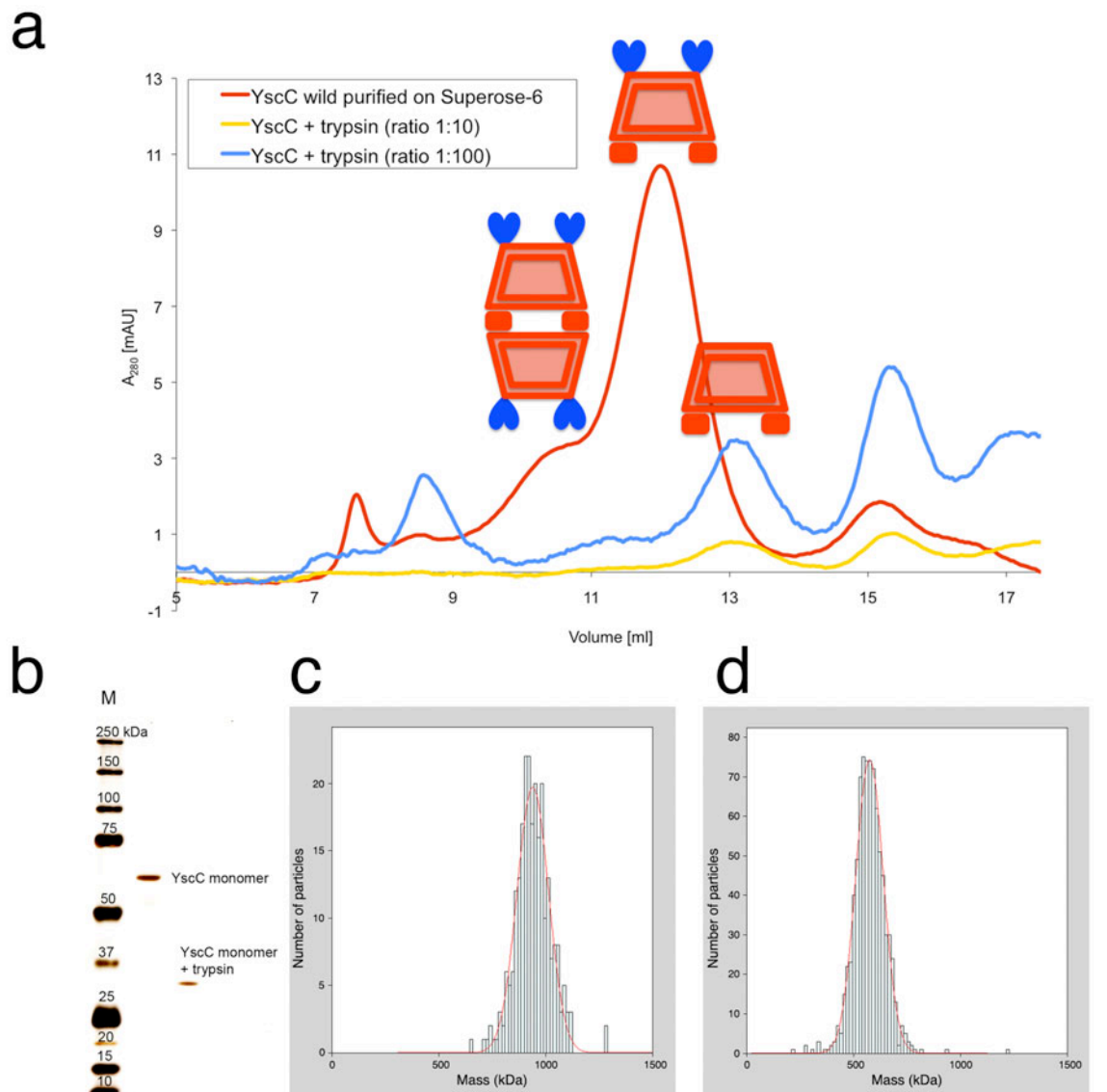
Figure 3.10: Mass measurements of intact and proteolyzed YscC oligomers. Both intact and digested oligomers were purified on Superose 6 10/300 GL column. Graph (a) presents elution profiles of intact and trypsinized (in ratio 1:10 and 1:100) oligomers. Red and blue models of intact monomer and dimer, and trypsinized monomer of YscC oligomer correspond to the elution peaks below. In models the trapezoid shape corresponds to a protease-resistant core region, red squares to the C-terminal part, and blue hearts to the flexible domain located at the N-terminus of the oligomer. b) A silver stained NuPAGE of dissociated by phenol treatment intact (64.24 kDa, mature protein 581 aa) and trypsinized (~ 32 kDa) YscC oligomers. The gel bands were cut off and analyzed by mass

spectroscopy. No additional bands were detected. (c-d) STEM mass analysis. The histograms show the mass distributions for the intact and trypsinized YscC complexes. The Gaussian is at (c) 941 ± 74 kDa (number of measurements, n = 250) for the intact YscC oligomer and at (d) 576 ± 65 kDa (n = 818) after trypsin treatment.

Mass spectroscopy measurements of trypsinized sample indicated that the N-terminal digestion occurred after Arg246 of YscC and the short form of the protein began with sequence $_{247}$VEADPSLNAIIVR. The last identified C-terminal protein fragment was LFIIEPRIID EGIAHHLALG NGQDLR$_{526}$ indicating that the proteolyzed protein finished at Arg526 (Figure 3.11).

Analysis of mass spectroscopy results was based on an alignment to secretin PulD, for which trypsin cleavage sites were reported [97]. It was also taken on the account that the N-terminal trypsin cut could equally well happen at Arg240 or Agr233 and fragments were not identified by mass spectroscopy. Comparing masses of the YscC oligomer cut from N-terminal at 3 possible places it was found that ~32 kDa of the protein mass was retained and 32 kDa lost after the protease treatment (Table 3.1). According to the mass spectroscopy measurements trypsin removed regions from both the N- and C- termini of YscC, leaving the protease-resistant secretin homology core region (C domain) corresponding to residues 247-526. The mass of the trypsin-resistant core (30.6-32.1 kDa) indicates that 48-50% of the mass of the YscC oligomer is removed by proteolysis.

A previous study of YscC oligomer suggested the oligomer symmetry of 13 or 14 [95]. Moreover, here, following the suggestion from the metal shadowing data (chapter 3.3.5) and cryo EM (chapter 3.3.6) the C12 symmetry was also considered. After trypsinization, the oligomer is 365 kDa smaller (941 ± 74 kDa $_{YscC\ intact} - 576 \pm 65$ kDa $_{YscC\ trypsin} = 365 \pm 98$ kDa), which indicates that the average stoichiometry is about 11-12. The detergent contribution was assumed to be the same for intact and trypsinized specimen since the protease-resistant core region includes the membrane integrated part of the protein where most of the DDM detergent is bound. Assuming a micelle mass of 47 kDa (*Anatrace catalogue:* DDM micelle = $40 - 47$ kDa), measured STEM masses of intact (941 kDa) and trypsinized (576 kDa) oligomer, actual monomer masses (intact 64.24 kDa, trypsinized ~32 kDa) and the stoichiometry of 12, the amount of detergent bound to the both, intact and trypsinized, oligomers was calculated to be ~4 DDM micelles per particle. All calculations are combined in Table 3.1.

A	B	C	D	E	F
RETAINED YscC$_{TRYPSIN}$ FRAGMENT [KDA]	LOST YscC FRAGMENT [KDA]	STOICHIOMETRY (941 ± 74 KDA $_{INTACT}$ $- 576 \pm 65$ KDA $_{TRYPSIN}$ $= 365 \pm 98$ KDA) (365 ± 98 KDA $\div$ B)	MASS OF YscC$_{TRYPSIN}$ C12 (A X 12)	DETERGENT CONTRIBUTION IN YscC$_{TRYPSIN}$ (576 ± 65 KDA $- $D)	DDM DETERGENT MICELLE IN YscC$_{TRYPSIN}$ (E $\div$ 47 KDA)
$_{247}$VEAD..DLR$_{526}$ 30.6	33.6	10.9 ± 2.9	403.2	172.8 ± 65	3.7 ± 1.4
$_{241}$ASAQ..DLR$_{526}$ 31.2	33.0	11.0 ± 3.0	396	180 ± 65	3.8 ± 1.4
$_{234}$IPQ..DLR$_{526}$ 32.1	32.1	11.4 ± 3.0	385.2	190.8 ± 65	4.0 ± 1.4

Table 3.1: Analysis of the trypsinized YscC oligomer. According to the mass spectrometry results masses of retained (A) and lost fragments (B) of YscC complex, oligomer stoichiometry (C), mass of truncated oligomer (D), the DDM detergent contribution (E) and number of micelles around trypsinized oligomer (F) were calculated.

```
MAFPLHSFFK RVLTGTLLLL SSYSWAQELD WLPIPYVYVA KGESLRDLLT DFGANYDATV

         70         80         90        100        110        120
VVSDKINDKV SGQFEHDNPQ DFLQHIASLY NLVWYYDGNV LYIFKNSEVA SRLIRLQESE

        130        140        150        160        170        180
AAELKQALQR SGIWEPRFGW RPDASNRLVY VSGPPRYLEL VEQTAAALEQ QTQIRSEKTG

        190        200        210        220        230        240
ALAIEIFPLK YASASDRTIH YRDDEVAAPG VATILQRVLS DATIQQVTVD NQRIPQAATR

        250        260        270        280        290        300
ASAQARVEAD PSLNAIIVRD SPERMPMYQR LIHALDKPSA RIEVALSIVD INADQLTELG

        310        320        330        340        350        360
VDWRVGIRTG NNHQVVIKTT GDQSNIASNG ALGSLVDARG LDYLLARVNL LENEGSAQVV

        370        380        390        400        410        420
SRPTLLTQEN AQAVIDHSET YYVKVTGKEV AELKGITYGT MLRMTPRVLT QGDKSEISLN

        430        440        450        460        470        480
LHIEDGNQKP NSSGIEGIPT ISRTVVDTVA RVGHGQSLII GGIYRDELSV ALSKVPLLGD

        490        500        510        520        530        540
IPYIGALFRR KSELTRRTVR LFIIEPRIID EGIAHHLALG NGQDLRTGIL TVDEISNQST

        550        560        570        580        590        600
TLNKLLGGSQ CQPLNKAQEV QKWLSQNNKS SYLTQCKMDK SLGWRVVEGA CTPAQSWCVS
```

```
APKRGVL
```

Figure 3.11: Amino acid sequence of the YscC protomer. Peptides identified with mass spectrometry are underlined for intact YscC and highlighted with *yellow* for trypsinized YscC protein. Amino acids Arg and Lys at which trypsin cleavage could occur are *orange* and *red*, respectively. The signal peptide 1-26 aa, predicted in SignalP 3.0 Server [157], is highlighted with *gray* color.

3.2.4   Samples homogeneity

Purification of the YscC oligomer was based on the protocol described in [95] which included solubilization of membrane proteins, sucrose gradient and ion exchange chromatography steps. The protocol was modified by the introduction of DDM detergent instead of Elugent detergent and adding the size exclusion chromatography purification step, which allowed to remove impurities, i.e. unbounded YscC monomers, from the

specimen and get the clean background crucial for STEM and cryo EM experiments (Figure 3.12). Exchange of Elugent, a non-ionic detergent composed of a mixture of alkyl glucosides, to DDM detergent resulted in a higher yield, exclusion of aggregates, more homogenous specimen and the clearer background.
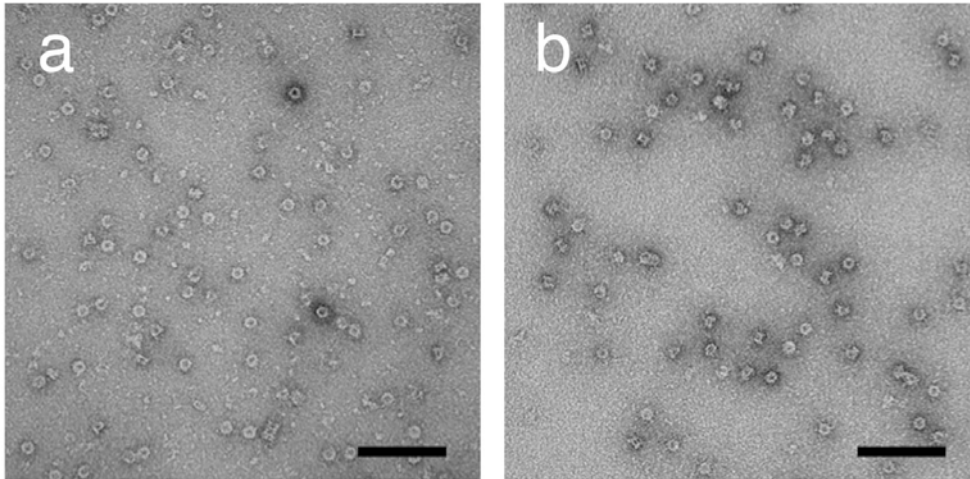
Figure 3.12: Comparison of specimens at different purification steps. (a) Electron micrograph of the YscC oligomer sample after MonoQ 5/50 GL ion-exchange column. (b) Introduction of the size-exclusion Superose 6 10/300 GL column purification step allowed eliminating all small contaminants visible in the specimen (a). Scale bars correspond to 100 nm.
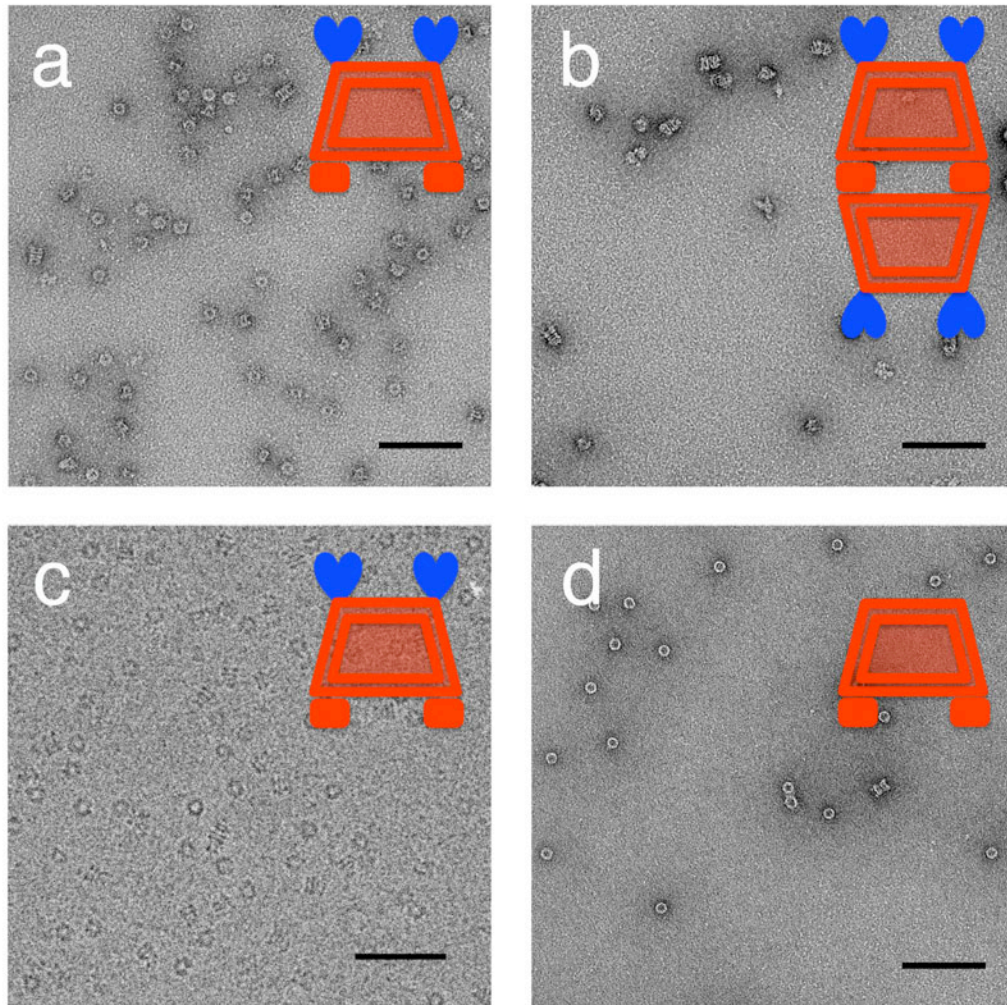
Figure 3.13: Electron microscopic analysis of negatively stained specimens of (a) the purified oligomer YscC, (b) dimer made of YscC oligomers, (d) the trypsinized YscC complex. (c) shows the oligomer YscC embedded in vitreous ice, in cryo EM. In the upper-right corner of each image red models of particles in given samples are shown. The flexible domain could have been observed only in (c). A trapezoid shape corresponds to a protease-resistant core region, red squares to the C-terminal part, and blue hearts to the flexible domain located at the N-terminus of the oligomer. Scale bars correspond to 100 nm.

The purity, homogeneity and monodispersity of the samples were checked using TEM. On the carbon-coated grids the YscC oligomer was present in random orientations. Ladder-like structures (side views) and ring-like structures (top views) were observed. Examination of the negatively stained trypsinized sample revealed mainly ring-like views (trypsin-resistant cores) (Figure 3.13).

Monomeric forms of intact oligomer formed in the buffer solution the dimers due to the C-terminal interactions. It was found that the flexible domain at the N-terminus (represented as blue hearts in Figure 3.14 a) was highly negative charged (Figure 3.11) while in the trypsin-resistant core region and the C-terminus the negative charges are neutralized by positively charged residues in the protein structure.

Interestingly, the physico-chemical properties of the complex were modified by proteolysis. After removal of the N- and C-terminal moiety in the YscC monomers, the C-terminal acquires negative charge, then an electrostatic repulsion excludes C-terminal interaction and the dimers are formed using N-terminal part of the core region (Figure 3.14 b).
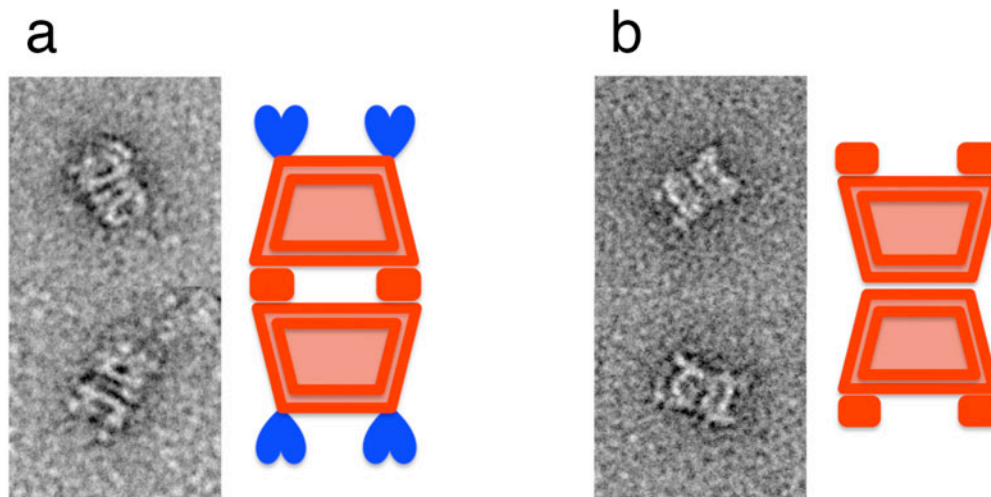
Figure 3.14: Uranyl acetate-stained images of dimers formed by the (a) intact and (b) trypsinized YscC oligomers. Negatively charged parts of the protein are represented in blue. (a) In the intact specimen negatively charged flexible domains of proteins, represented as hearts in the model, are located on the outside. (b) After removal of the flexible domain from N-terminus and C-terminal moiety, the oligomers interact by N-terminal parts of trypsin-resistant cores and trypsinized C-termini protrude outwards.

3.2.5 Metal Shadowing – determination of the oligomer symmetry
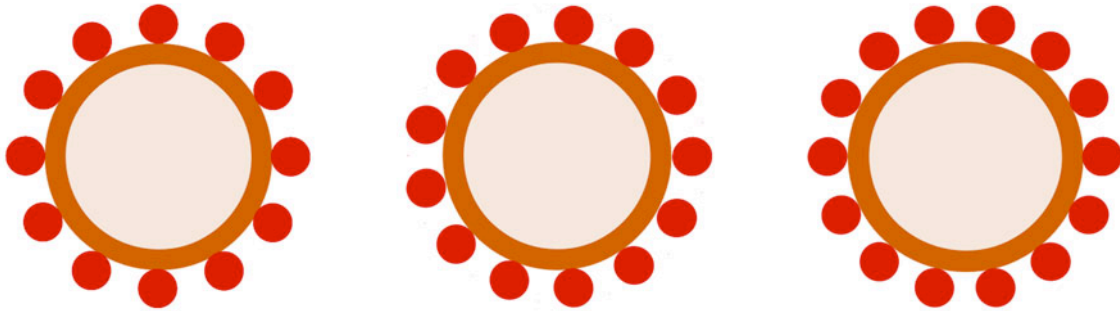
Figure 3.15: From the left: scheme model of a 12-, 13- and 14-mer.

A previous study of YscC oligomer suggested symmetry of C13 or C14 [95]. To verify the stoichiometry, the trypsinized YscC oligomer was freeze-etched, metal-shadowed and analyzed in TEM. The oligomers appeared mainly as ring-like structures (Figure 3.16 a). While for the intact complex the individual YscC monomers were not visible due to the flexible N-termini that masked out the signal, for trypsinized specimen averages were generated with the indicated 12-fold or 13-fold symmetry. Two sets of particles were analyzed. 1300 particles recorded at low defocus (box size of 70x70 pixels) and 1000 particles recorded at high defocus (box size of 140x140 pixels) were collected, normalized, centered and classified (*refine2d*) in EMAN software package [150]. In order to eliminate the possibility of signal loss due to the metal grain size that was similar to the protomer size, the pictures were taken at 2 different focuses and later the results were compared. The classes with indicated symmetry were obtained for both sets of the particles. Clearer averages were produced for particles taken at higher magnification.

Examination of the trypsinized YscC complex in TEM after metal shadowing has revealed the individual YscC subunits, but unfortunately not yet well enough to distinguish between a 12- and 13-mer, the results are summarized on Figure 3.16.
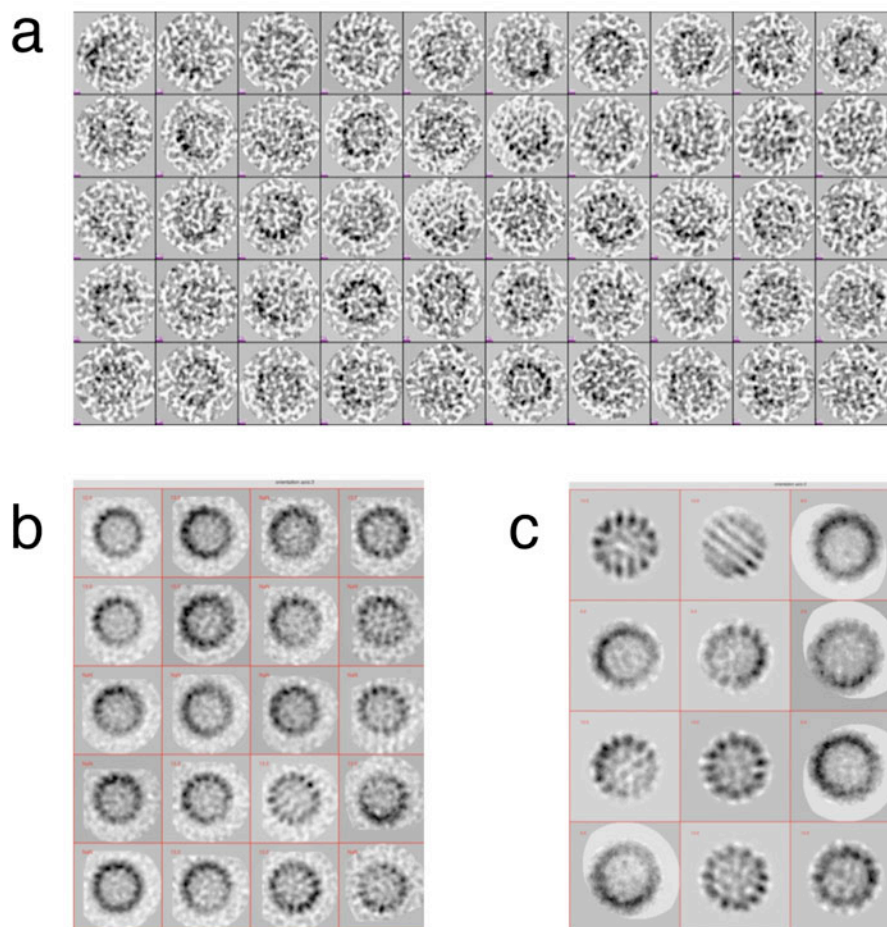
Figure 3.16: Metal shadowing. Single particles analysis of top views from metal-shadowed trypsin digested YscC oligomers. (a) The partial overview of collected in Boxer, the display program associated with the EMAN software package [150], metal-shadowed particle set. Classes generated in EMAN out of (b) 1300 particles with box size of 70x70 pixels taken at magnification 50000x and (c) 1000 particles with box size of 140x140 pixels taken at magnification 100000x indicate the symmetries C12 or C13.

3.2.6   3D structure reconstruction with EMAN

Single particles were selected from EM negatives with the EMAN`s Boxer software and processed using the EMAN 1.9 suite [150]. Approximately 8000 negatively stained particles and 37200 particles from cryo EM images were collected for the processing.

90

TEM

Uranyl-acetate stained particles, both the YscC oligomer monomers and dimers, were boxed and processed in EMAN. Program generated the top view and side view classes (Figure 3.17 a). It was not possible to define the symmetry from the class averages. Additionally, the flexible domain at the N-terminus was not visible for the YscC monomers. In the side view averages of the symmetrized dimer the blurred gray clouds at the ends of particle correspond to the densities of flexible domains (Figure 3.17 b). Thus, 2 different symmetries were imposed to generate initial 3D models (Figure 3.18).
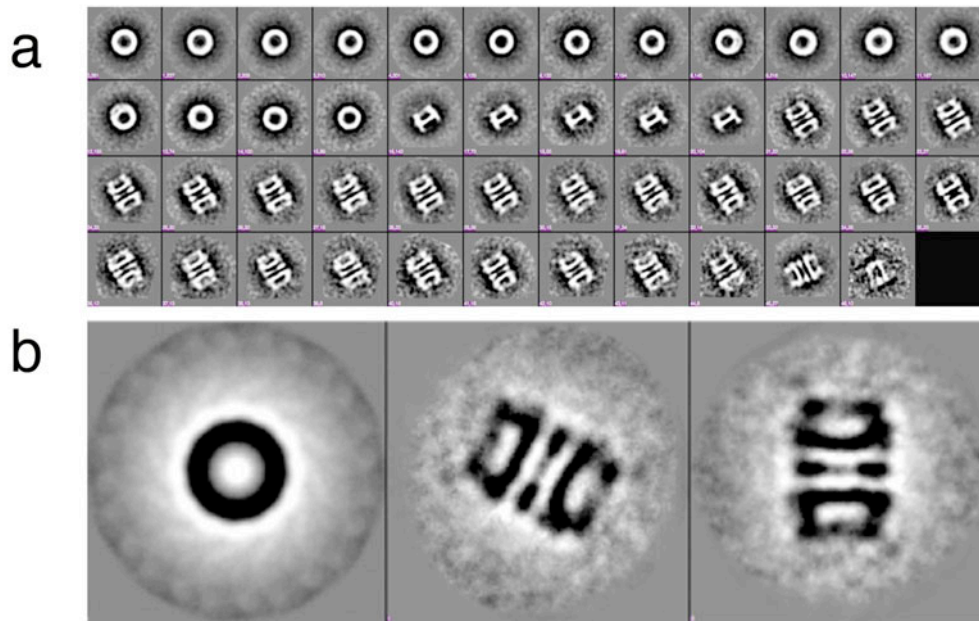
Figure 3.17: (a) Single particle classes generated out of 8000 negatively stained YscC oligomers in EMAN [150]. The program recognized monomers and dimers made of YscC oligomer. (b) Images of the symmetrized dimer top view, the side view, and the side view aligned to the symmetric axis. Cyclic 12-fold symmetry was imposed. Images are displayed as negatives.
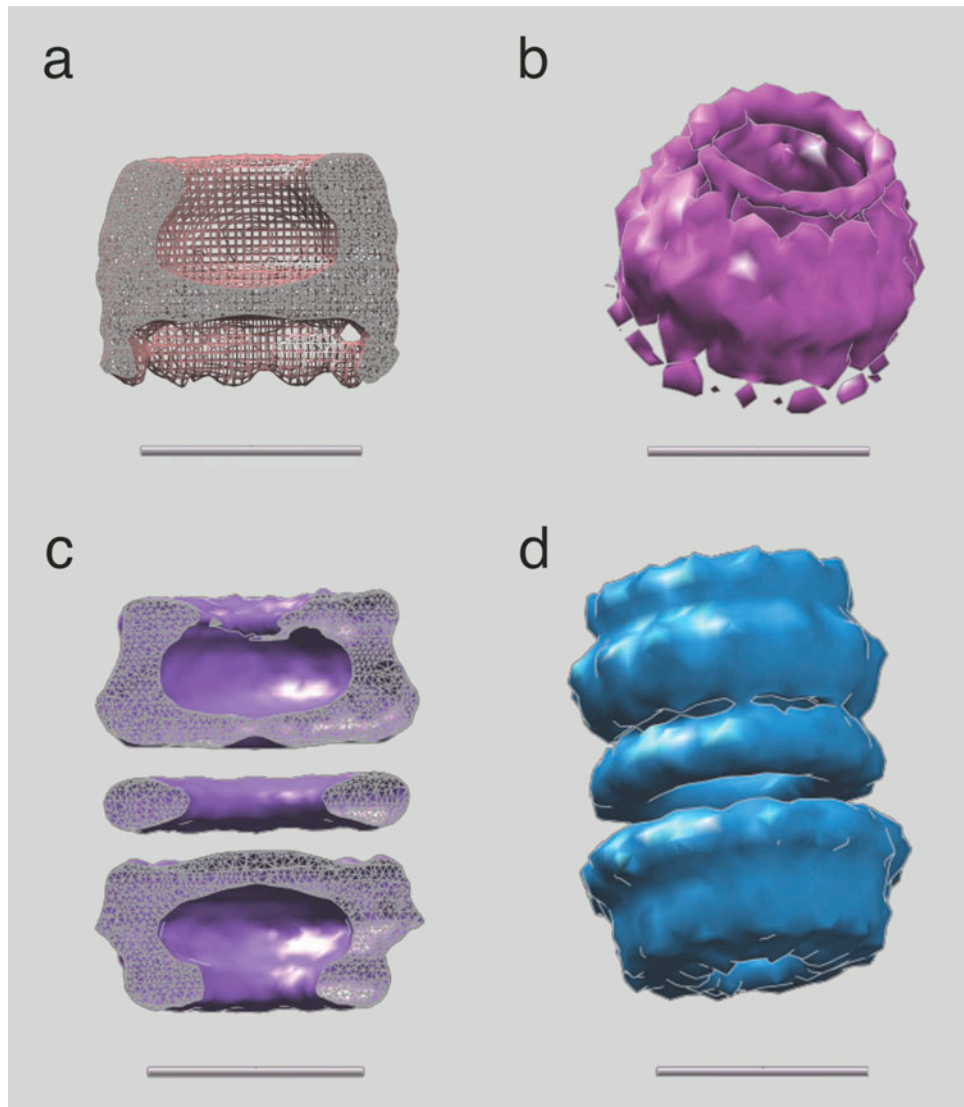
Figure 3.18: For negatively stained particles not only monomer (a-b) but as well dimer (c-d) models of the YscC oligomer were generated. The following symmetries were imposed to the 3D models: 12-fold (a, c), 13-fold (b, d). 3D maps were obtained on contouring at 100% of the mass. Scale bars correspond to 100 Å. (a) and (c) were presented as cross-sections.

Cryo EM

The isolation of the outer membrane ring complex formed by YscC has been optimized for cryo EM.

Intact YscC oligomer samples were adsorbed onto thin carbon films before freezing in liquid ethane using Vitrobot. Cryo EM of unstained, frozen-hydrated protein showed both top-views and side-views of the YscC oligomer (Figure 3.20). Side view of the

complex displayed a ladder-shaped structure with the height of approximately 15 nm and a radius diameter of 6 nm. Occasional side-views of dimeric complexes were also visible (Figure 3.13 c). Cryo EM negatives were digitized, and more than 37000 projections of intact YscC complexes were selected for elucidating their 3D structure.
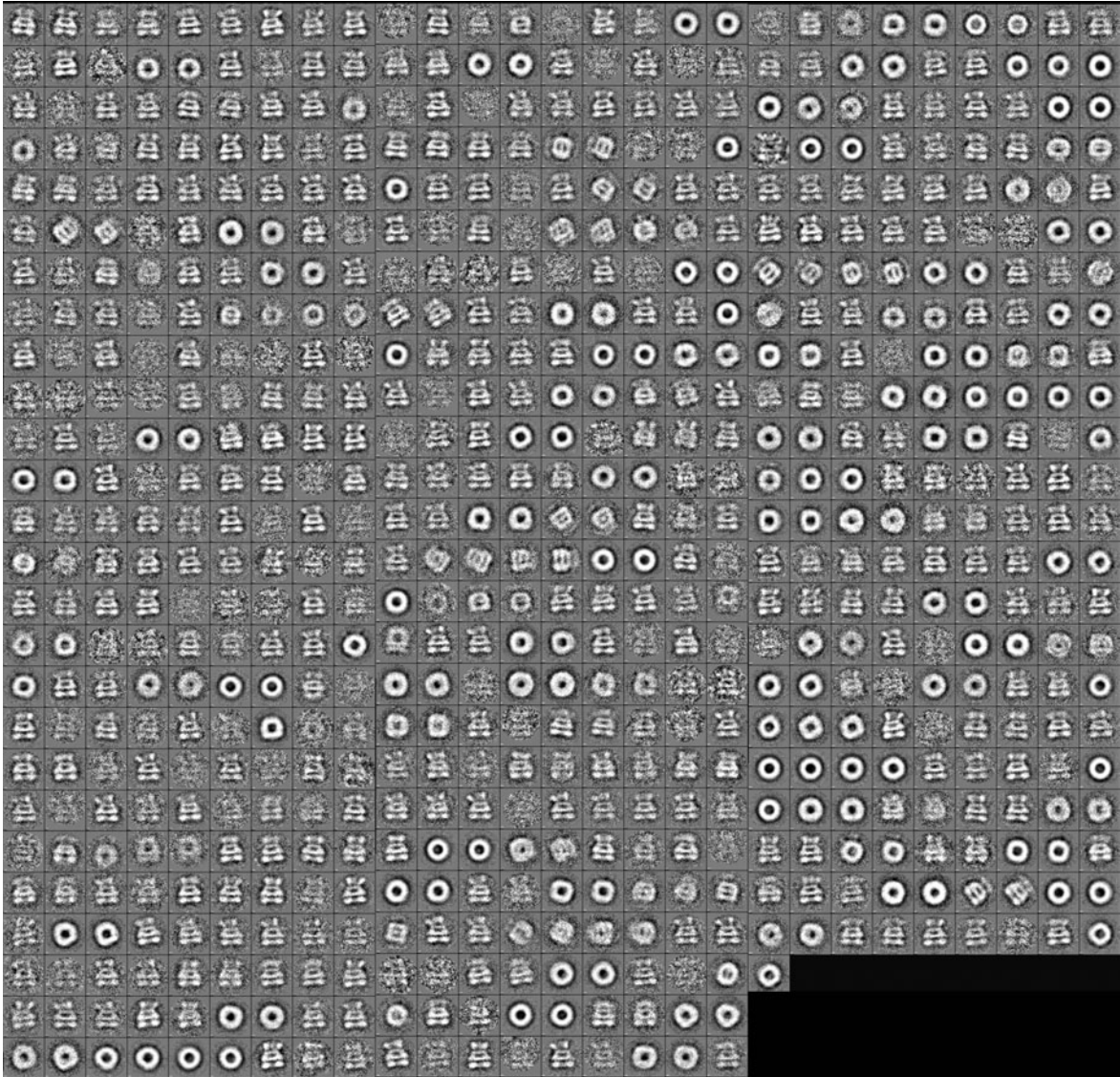
Figure 3.19: 676 class averages of YscC oligomer at different orientations were generated during 2D refinement in EMAN for 37000 single particles.

First, the individual projections of the randomly oriented complexes were classified according to their similarity. Many of the resulting class averages revealed a flexible domain protruding from one side of the complex. Being flexible this region tends to average out during the subsequent processing. Characteristic top-view and side-view class

averages are shown in the Figure 3.19. Top view class averages with 2 different inner radii were observed. The class averages were used for generating the 3D initial model with imposed C12-C13 symmetries (Figure 3.21).
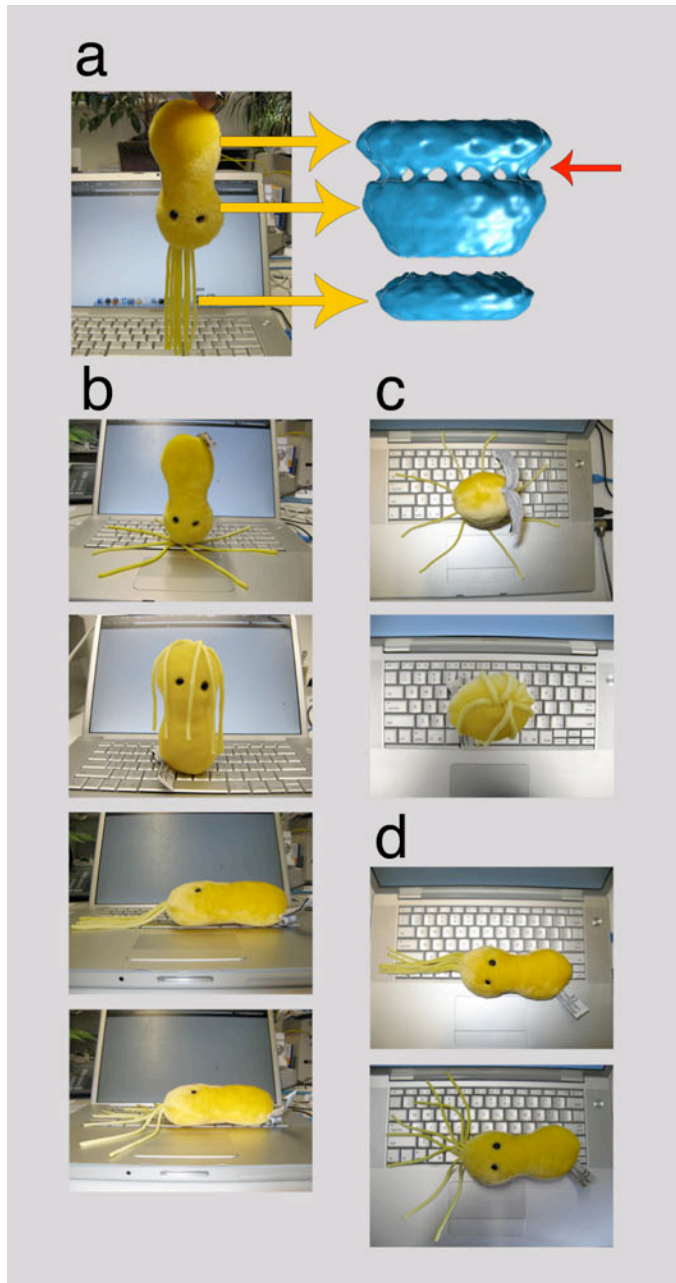
Figure 3.20: *Salmonella typhimurium* (Giantmicrobes) as the schematic representation of the YscC oligomer frozen in the vitreous ice. (a) Oligomer consists of the 3 main parts: flexible domain at the N-terminus (bottom), trypsin-resistant core (partial putative membrane domain, middle) and C-terminal part (top). Probable membrane-spanning region of the YscC initial model (Figure 3.21) is marked with *red* arrow.

In cryo EM the molecules are lying in random orientations in a thin layer of vitreous ice. Column (b) corresponds to possible views from the side. Column (c) shows top views. Column (d) illustrates how flexible domain of the protein artificially raises the number of generated side view classes of the oligomer.
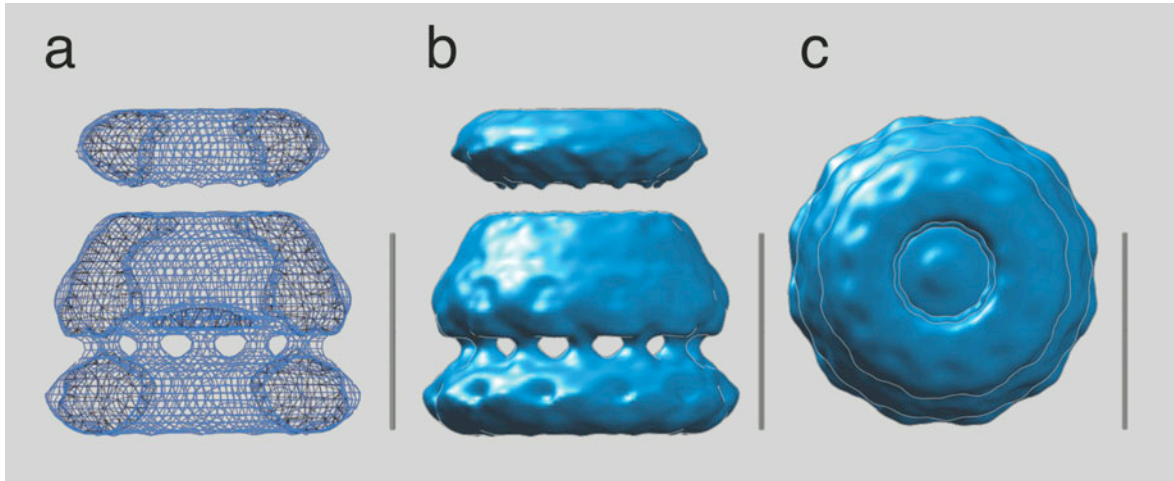
Figure 3.21: The 3D model of the YscC oligomer with applied C12 symmetry after first refinement was used for high-resolution refinement. The images show (a) the cross-section, (b) side view and (c) top view models. 3D map was obtained on contouring at 100% of the mass of a dodecamer. Scale bars correspond to 100 Å.

According to the single particle 3D reconstruction based on the multi-refinement protocol the 13-fold symmetry was preferred. However, when running individual refinements with 12 or 13-fold symmetry, supplied as an input parameter, the resolution obtained in either scenario was the same. Thus, two refined YscC models with symmetries C12 or C13 were generated (Figure 3.22). In both cases, when contouring at 100% of the mass of the complex, the refined model reveals a slightly conical, cylindrical structure of periplasmic domain closed at the N-terminus and two differently sized ring-shaped domains linked by fine densities, which correspond to the membrane spanning regions. Both models have flexible domain at the N-terminal, which tend to average out during the model refinement but could be visualized when the contour level is set to 160 percent of the oligomer mass (Figure 3.22 f). Despite the presence of the flexible domain the generated pore structure is closed from periplasmic side, which suggests that for YscC open conformation other injectisome components are necessary.

Although both structures were solved with the same resolution of 12 Å, greater number of structural details, i.e. more distinct flexible domain or hanging domains in the spanning membrane region, is visible for the YscC complex with the C13 symmetry. These features are blurred or missing in the C12 structure. Interestingly, the putative membrane spanning domains are arranged in clockwise order in the C13 and counter-clockwise in C12 structure (Figure 3.22 d). Additionally, in the 13-fold model are regular clearances,

which in the 12-fold model are jagged. The height of both models without flexible domain, which is 55 Å long, is equal to 110 Å. The inner and outer diameters are slightly different in C12 (inner: 60 Å, outer: 145 Å) than in C13 model (inner: 58 Å, outer: 150 Å).
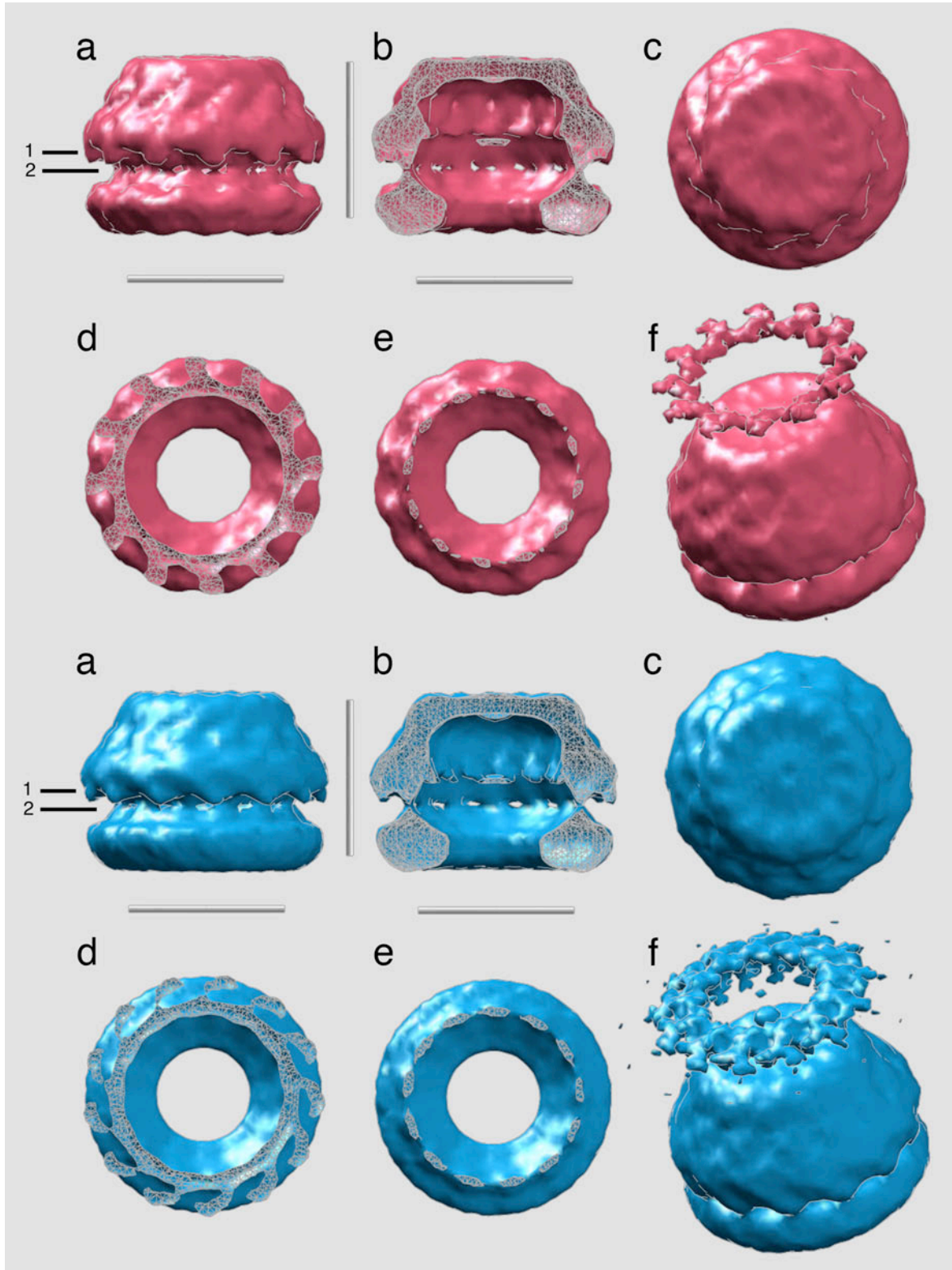
Figure 3.22: 3D structure of the YscC complex with applied symmetries C12 (top, *red*) and C13 (bottom, *blue*) both at resolution of 12 Å. (a-e) 3D map obtained on contouring at 100% of the mass of a C12/C13 complex (not considering the detergent). (a) Side-view. (b) Vertical slice through the centre of the complex. (c) Top-view. (d) Horizontal slice made at position 1 as indicated in (a) and looking down. (e) Horizontal slice made at position 2 as indicated in (a) and looking down. (f) Map obtained on contouring at 160% of the mass of the C12/C13 complex. Scale bars correspond to 10 nm.
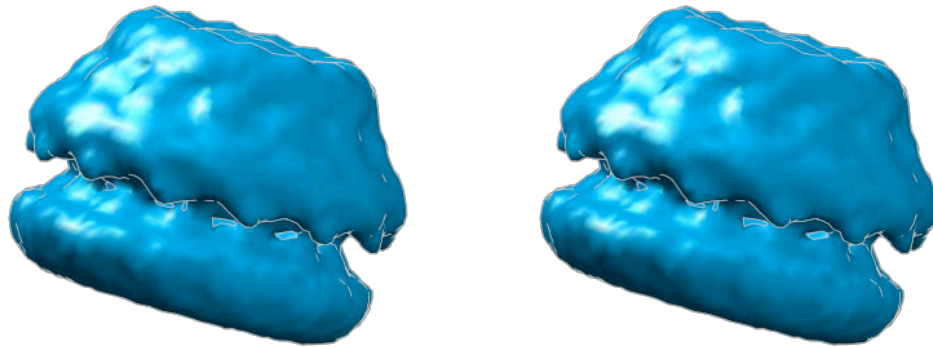
Figure 3.23: Wall-eye stereo view of the YscC oligomer with the C13 symmetry.
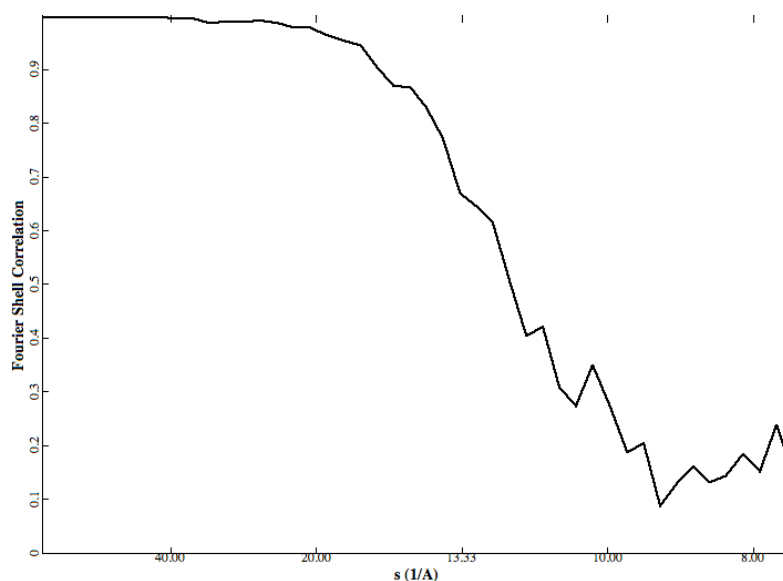
Figure 3.24: Fourier Shell Correlation (FSC) curve. The YscC structure resolution was assessed to be 12 Å with the *FSC=0.5 criterion* from two independent half data sets.

3.3   Discussion

3.3.1   YscC expression in bacteria *Yersinia enterocolitica* and the YscC oligomer purification

The injectisome is part of the *Yersinia enterocolitica* weaponry. The YscC oligomer, which belongs to the family of the secretins, is a major structural injectisome component and forms a stable ring-like multimeric complex in the bacterial outer membrane.

YscC and lipoprotein YscW, a pilot protein necessary to increase the expression yield and properly localize the YscC complex in the membrane [99], were expressed in the avirulent *Y. enterocolitica* strain. The homooligomeric YscC was purified without its chaperone YscW, in contrast to the secretin PulD that copurifies with the pilot protein PulS [158]. Thus, the YscW is not a stable constituent of the formed oligomer. The protein purification protocol established in previous study [95] was modified in two steps, which were fundamental to obtain pure specimens. First, the change of solubilizing non-ionic detergent from Elugent to DDM reduced the amount of aggregates and thus improved the amount and quality of single particles. Second, the size exclusion chromatography purification step, which allowed the pure oligomer to be separated from small contaminants, YscC monomeric form and dimers made of YscC oligomers, and yielded the clean background necessary for STEM and cryo EM experiments. The purity, homogeneity

and monodispersity of the samples were examined by TEM. As result of the required adsorption to carbon film, two predominant oligomer orientations were observed, mostly ring-like structures (top views) and ladder-like structures (side views), which revealed the N-terminal flexible domain of YscC.

3.3.2   Searching for the missing symmetry using various approaches

The rotational symmetry is an essential parameter necessary to construct the correct 3D model of the YscC oligomer from single particle projections. For the intact YscC particles we could not predict the symmetry from the obtained class averages, thus it had to be provided as a parameter from additional experiments. Therefore, we calculated the symmetry using numerous methods: mass spectroscopy of native and trypsinized samples, mass determination by STEM and rotary metal shadowing of freeze-dried trypsinized complexes.

The flexible domain located on the N-terminal part of secretin is visible on the cryo EM negatives, whereas for negatively stained specimens there is only a slight presence of the shadow-like domain manifested by darker pixels. The flexible domain in generated YscC class averages (Figure 3.19) tends to blur the top view projections of particles where each protomer should have a visible distinct position. Therefore, removal of the N-terminal flexible domain was performed to facilitate the visibility of the C-terminal protease-resistant core of the protein in freeze-dried metal shadowed samples.

The YscC complex trypsin cleavage sites, Arg246 and Arg526, were identified by the mass spectrometry. According to the multiple sequence alignment to 4 different secretins (Figure 3.28) it was found that these cleavage sites correspond to the N-terminal trypsin cleavage site of InvG, Lys268 [98], and Lys617, the C-terminal trypsin cleavage residue of PulD [97], respectively. The phylogenetic tree of five secretins presented on Figure 3.26 shows the evolutionary interrelations between these proteins. The primary sequence of YscC is closer related to the InvG and MxiD secretins than to the EscC and PulD, T2SS secretin, which is located on the separate branch of the phylogenetic tree and is the most distant from the *Yersinia* secretin sequence. The large evolutionary distance between YscC and PulD sequences explains different size of N-terminal flexible domains, which in case of PulD consists of multiple additional insertions, which changes the relative positions of the cleavage site.

The attempt to determine the symmetry using the rotary metal shadowing, as it was

performed for PulD [97] (Table 3.2), was based on the assumption that every YscC protomer could be precisely identified and the metal grain size would not collide with the subunit size. Metal shadowing has revealed the individual YscC subunits but regardless of the methods of calculation the symmetry was oscillating between C12 and C13.

STEM analysis gave a mass of 941 kDa for the pure DDM-solubilized native YscC complex, including DDM, thereby strongly suggesting the presence of 12 subunits for which the calculated mass is 959 kDa. However, taking into account more precise calculations including trypsinized sample (for which the bound detergent mass was assumed to be identical as for native protein) and errors, symmetry between C11 and C12 (with the considerable calculation error) was the most likely.

A previous study of YscC oligomer suggested the oligomer symmetry of 13 or 14 [95]. The Elugent-solubilized YscC mass determined by STEM was appreciably higher, 985 kDa, which had to be reduced by mass of the detergent retained by a membrane protein. Elugent, the non-ionic detergent mixture, analogical to the DDM detergent forms 50-kDa detergent micelles. It was concluded that if the 280-kDa photosystem II complex was solubilized with one DDM micelle [159], the YscC oligomer had to be associated with several micelles, but no exact numbers were obtained. With the new knowledge gained from our experiments on non-trypsinized and trypsinized samples, about the number of micelles bound to the YscC complex, the symmetry C12 could be considered also for previous results, presented in [95].

The cryo EM data of the intact YscC oligomer were used to construct 3D model. According to the single particle 3D reconstruction based on the multi-refinement protocol (as implemented in EMAN), the 13-fold symmetry is preferred. However, when running refinements with 12 or 13-fold symmetry, supplied as an input parameter, the resolution obtained in either scenario was the same. A very flexible domain at the N-terminal makes it difficult to precisely resolve the oligomer symmetry. The "floppy" part effectively masks the evenly distributed protomers. Moreover, these floppy parts comprise a significant fraction of the mass (i.e. contrast) and tend to misalign projections. Therefore it would be useful to collect and process the cryo EM data of the trypsinized samples of the YscC, without the N-terminal domain. These experiments are currently carried on in our laboratory.

In spite of the secretins homology, the number of oligomer subunits varies within the secretin family. The presence of 14 subunits has been reported for phage pIV multimer [160] and a dodecameric structure has been proposed for the T2SS secretins PulD [97] and

GspD [165], T3SS secretin MxiD [83] and PilQ that is involved in the assembly of type IV pili [161] (Table 3.2). Additionally, in the recently published study [162], for InvG, it was necessary to use a ring with 14 subunits to fit it into the *Salmonella* cryo EM density map using homology modeling.

In the secretin family, changes in the stoichiometry might arise over the evolution to allow the transport of macromolecules of different diameter like folded extracellular proteins, pseudopilins or bacteriophages. Only small changes in protein packing in the oligomer might be necessary to perform a change from 12 to 14 (or to 13, and vice-versa) secretin subunits. A similar explanation was given to the TatA isolated from the *E. coli,* a component of twin-arginine protein transport system, which forms channel complexes of variable diameter translocating across the inner membrane specific folded proteins of different sizes [163]. The 3D architecture of the TatA complex obtained by single particle cryo EM and the random conical tilt reconstruction suggest 12–35 monomers per complex. However, the observed variation in protein diameter suggests that number of protomers changes to match the size of the channel to the size of the substrate being transported.

BIOLOGICAL SYSTEM	TECHNIQUE	OUTER/INNER DIAMETER [Å]	SUBUNIT DIAMETER [Å]	Mw [kDa]	SYMMETRY
Yersinia enterocolitica T3S system YscC	Cryo EM	145/60 150/58	30 25	771 835	C12 C13
Yersinia enterocolitica T3S system YscC	Rotary metal shadowing	-	-	771/835	C12/C13
Yersinia enterocolitica T3S system YscC	STEM	-	-	707/771	C11/C12
Yersinia enterocolitica T3S system YscC	STEM	155/80	-	835	C13
Shigella flexneri T3S system MxiD	Negative stain TEM	137/81	25	756	C12
Klebsiella oxytoca T2S system PulD Trypsin digested	Rotary metal shadowing	130/86	25	452	C12
Klebsiella oxytoca T2S system PulD-PulS	Cryo EM	137/80	20	825	C12
Escherichia coli filamentus phage pIV multimer	Negative stain TEM	135/87	-	624/1248	C14/D14
Escherichia coli filamentus phage pIV multimer	Cryo EM	137/65	-	624/1248	C14/D14
Neisseria meningitidis type IV pilus secretin PilQ	Negative stain TEM	165/65	-	960	C12
Vibrio cholerae T2S system GspD	Cryo EM	155/55	-	888	C12

Table 3.2: Comparison of the known secretin structures, their outer and inner rings diameter, molecular weights and determined symmetries. Subunit diameters were measured for *S. flexneri* and *K. oxytoca*. Data within *orange* borders corresponds to this study. References: [83, 95, 97, 160, 164, 165].

3.3.3 Structure reconstruction of the oligomer

The secretin is a key component of the T3SS machinery in a number of Gram-negative pathogens. Therefore, it is fundamental to establish the secretin location in the membrane and provide information about its structure.

Previously, the secretins were observed in EM as components of inner membrane vesicles (*K. oxytoca* PulD secretin, T2SS [166]) and as low-diffracting 2D crystals reconstituted in the presence of lipids (*N. meningitidis* PilQ secretin [161]). Top views of the oligomers showed large ring structure with an internal cavity and the side views revealed formation of at least two stacked ring-like regions [167, 168]. The studies on pIV secretin, using the gold labeling technique, revealed the ring-like arrangement of protomers in the oligomer [160].

On the basis of the sequence comparison secretins contain different homologous structural domains, particularly marked within the core region but also in the N-terminal flexible part (Figure 3.28). The limited proteolysis and mutagenesis experiments have shown that C-terminal trypsin-resistant core of secretins participates in the pore formation and protein oligomerization, while the N-terminal flexible domain is possibly involved in specific interactions [169, 170]. The location of the N-domain at the one end of the secretin was proved by gold labeling experiments performed with pIV [160]. Moreover, it was reported that the non-conserved C-terminal of secretin functions as the binding site for so-called pilot proteins [171].

Although no full-length atomic resolution structure of any secretin is known, the general secretin model based on previous studies [95, 97, 98] and the sequence analysis (Figure 3.25) could be built. Assuming that most of the β-structures of the YscC are embedded into the membrane and the α-helices are stabilizing the structure above and below the membrane, the β-barrel structure of the oligomer could be predicted (Figure 3.25).

```
1                                                                                                  100
MAFPLHSFFKRVLTGTLLLLSSYSWAQELDWLPIPYVYVAKGESLRDLLTDFGANYDATVVVSDKINDKVSGQFEHDNPQDFLQHIASLYNLVWYYDGNV

101                                                                                                200
LYIFKNSEVASRLIRLQESEAAELKQALQRSGIWEPRFGWRPDASNRLVYVSGPPRYLELVEQTAAALEQQTQIRSEKTGALAIEIFPLKYASASDRTIH

201                                                                                                300
YRDDEVAAPGVATILQRVLSDATIQQVTVDNQRIPQAATRASAQARVEADPSLNAIIVRDSPERMPMYQRLIHALDKPSARIEVALSIVDINADQLTELG

301                                                                                                400
VDWRVGIRTGNNHQVVIKTTGDQSNIASNGALGSLVDARGLDYLLARVNLLENEGSAQVVSRPTLLTQENAQAVIDHSETYYVKVTGKEVAELKGITYGT

401                                                                                                500
MLRMTPRVLTQGDKSEISLNLHIEDGNQKPNSSGIEGIPTISRTVVDTVARVGHGQSLIIGGIYRDELSVALSKVPLLGDIPYIGALFRRKSELTRRTVR

501                                                                                                600
LFIIEPRIIDEGIAHHLALGNGQDLRTGILTVDEISNQSTTLNKLLGGSQCQPLNKAQEVQKWLSQNNKSSYLTQCKMDKSLGWRVVEGACTPAQSWCVS

601      607
APKRGVL
```

Figure 3.25: Secondary structure prediction of the YscC oligomer generated in program Sable [172]. Red rectangles correspond to α-helices, green arrows to β-strands.
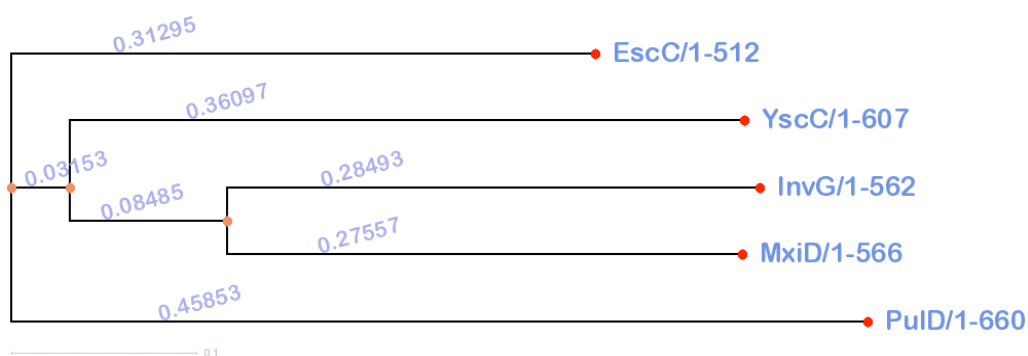
Figure 3.26: Phylogenetic tree as graphical representation of the evolutionary relationship between taxonomic groups. The scale bar represents the evolutionary distance: 0.1 equals 10 substitutions per 100 amino-acid residues. Multiple sequence alignment of secretins was performed in ClustalW2 (Figure 3.28); the guide tree was drawn with PHYLIP [173, 174].

The first step to reach the structural information about the YscC oligomer was its reconstitution into a lipid bilayer. Both, top and side views of *Yersinia* oligomer could be observed as protein has embedded its C-terminal into the lipids and N-terminal domain stuck out of the lipid bilayer. However, no higher ordered 2D crystals were obtained. Subsequent steps to solve the YscC oligomer structure were single particle 3D reconstructions from cryo EM data, based on individual and multi-reference refinements, performed with EMAN.

The single-particle cryo EM model of YscC consists of the conserved C-terminal protease-resistant region and the N-terminal flexible domain not observed in the previously

generated PulD model [97], only barely visible in InvG complexes from *S. typhimurium* [98] and recently observed in GspD (T2SS) from *Vibrio cholerae* [165]. The refined model (Figure 3.22) reveals a slightly conical, cylindrical structure with two differently sized ring-shaped domains linked by fine densities. The latter correspond to the membrane spanning regions, which were suggested in the PulD model. Comparison of the YscC and PulD side views confirmed the structure similarities but also drew attention to the small differences (Figure 3.27). Additionally, taking into account the complex behavior on the carbon-coated grid, it was observed that the YscC particle is less rigid than the PulD oligomer. The PulD model in comparison to the YscC model has more extensive C-terminal structure, while the inner pore diameter remains similar. Furthermore, the YscC middle domain is more conical. Distant relatives, YscC and PulD, share a number of structural features, which suggest YscC to have an even higher similarity to the InvG (*S. typhimurium*) or MxiD (*S. flexneri*) that are much more closely related (Figure 3.26).

The generated YscC model, similarly to the PulD [97] and PilQ [160] complexes, is closed (Figure 3.27). However, the YscC oligomer has the gating unit at the N-domain, while in the PulD and PilQ models plug within the pore is located central in the trypsin-resistant region. The closed aperture in PilQ appeared to be the site to which the type IV pilus binds [175]. It is assumed that the YscC complex undergoes a conformational change to form the closed state as a result of lack of the YscD and YscJ inner membrane complexes.

The structural changes of YscC caused by proteolysis and observed in negatively stained particles allow the flexible domain to be identified, which faces the periplasm. Interestingly, physico-chemical properties of the complex were modified by proteolysis. While trypsinized YscC had capacity to form dimers with its N-terminal part of trypsin-resistant core region, the intact YscC used its C-terminus to form dimers. This phenomenon may be connected with to the exposure of highly negative residues on the C-terminus of the YscC oligomer after removal of the N- and C-terminal moiety.

Results from previous experiments on isolated PulD and InvG secretins suggested that the N-terminal flexible region could not be entirely resolved by EM because the detergent disorders the N-terminal region during purification. In the case of YscC the flexible domain is well visible on the single 2D projections, but tends to average out during the subsequent processing and could be visualized in the YscC 3D model only when the contour level [176] is set at 160 percent of the oligomer mass. Moreover, the densities connecting N-domain with trypsin-resistant core are not visible on the whole length.

104
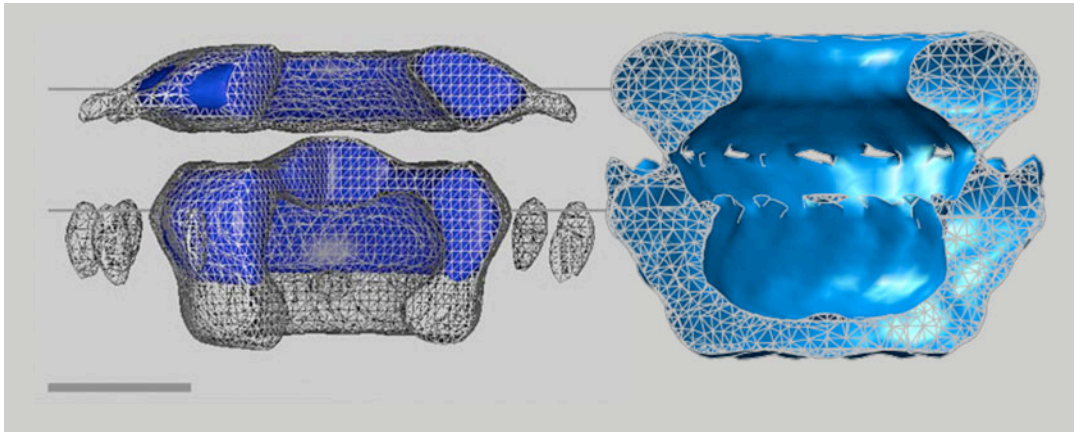
Figure 3.27: 1:1 scale comparison of 3D volumes of two secretins PulD (*K. oxytoca*) [97] and YscC (*Y. enterocolitica*) obtained with the single particle cryo EM technique. On the *left* was presented the intact (*white mesh*) and the digested (*blue*) PulD complex at resolution of 17 Å, on the *right* the 12 Å-resolution YscC complex with symmetry C13. The horizontal lines represent the outer limits of the lipopolysaccharide/phospholipid outer membrane (~4.5 nm wide). The scale bar corresponds to 50 Å.

The high level of conservation between all secretins suggests that sequence-diverse but topologically similar protein folds provide the common structure that allows for assembly of inner membrane rings characteristic for T2SS and T3SS. In the study [98], using EM, molecular modeling, protease digestion and mutagenesis, ring models of flexible domain of EscC and inner membrane rings of PrgH and EscJ were generated and placed into the EM electron density map of the *S. typhimurium* needle resulting in the final ring models within the assembled T3SS. Taking into account the injectisome electron density presented on the Figure 3.29, the YscC structure could cover both the PulD electron density and the crystallized N-terminal domain of EscC, together with additional 20 aa which are in possible contact with inner membrane ring, fitting well in the 3D model of the needle complex.

The observed architecture of the injectisome basal body filled with ring structures and linker regions (both observed in the YscC oligomer) highlights the importance of all components for the function of the assembled complex.

α1 β1 α2 β2

```
EscC/1-512
EscC/1-512    1  ...MKKISFFIFTALFCCSAQAAPSSLEKRLGKSE...YFIITKSSPVRAILNDFAANYSIPVFISSSVN
YscC/1-607    1  .MAFPLHSFFKRVLTGTLLLLSSYSWAQELDWLPIP..YVYVAKGESLRDLLTDFGANYDATVVVSDKIN
InvG/1-562    1  ...MKTHILLARVLACAALVLVTPGYSSEKIPVTG...SGFVAKDDSLRTFFDAMALQLKEPVIVSKMAA
MxiD/1-566    1  ...MKKFNIKSLTLLIVLLPLIVNANNIDSHLLEQNDIAKYVAQSDTVGSFFERFSALLNYPIVVSKQAA
PulD/1-660    1  MIIANVIRSFSLTLLIFAALLFRPAAAEEFS.........ASFKGTDIQEFINTVSKNLNKTVIIDPSVR
```

β3 α3 β4 β5 η1 β6

```
EscC/1-512
EscC/1-512   65  .DDFSGEIKNEKPVK....VLEKLSKLYHLTWYYDENILYIYKT..................NEISRSI
YscC/1-607   68  .DKVSGQFEHDNPQD....FLQHIASLYNLVWYYDGNVLYIFKN..................SEVASRL
InvG/1-562   65  RKKITGNFEFHDPNA....LLEKLSLQLGLIWYFDGQAIYIYDA..................SEMRNAV
MxiD/1-566   68  KKRISGEFDLSNPEE....MLEKLTLLVGLIWYKDGNALYIYDS..................GELISKV
PulD/1-660   62  GTITVRSYDMLNEEQYYQFFLSVLDVYGFAVINMNNGVLKVVRSKDAKTAAVPVASDAAPGIGDEVVTRV
```

α4 β7 β8 β9 α5

```
EscC/1-512
EscC/1-512  111  ITPTYLDIDSLLKYLSDTISVNKNSCNVRKITTFNSIEVRCVPECIKYITSLSESLDKE...........
YscC/1-607  114  IRLQESEAAELKQALQRSGIWEPR.FGWRPDASNRLVYVSGPPRYLELVEQTAAALEQQTQ.........
InvG/1-562  112  VSLRNVSLNEFNNFLKRSGLYNKN.YPLRGDNRKGTFYVSGPPVYVDMVVNAATMMDKQ...........
MxiD/1-566  115  ILLENISLNYLIQYLKDANLYDHR.YPIRGNISDKTFYISGPPALVELVANTATLLDKQ...........
PulD/1-660  132  VPLTNVAARDLAPLLR.QLNDNAGVGSVVHYEPSNVLLMTGRAAVIKRLLTIVERVDNAGDRSVVTVPLS
```

T3S system specific homology region → PulD (F1) ▼

```
EscC/1-512
EscC/1-512  170  .......AQSKAKNKDVVKVFKLNYASATDITYKYRDQNVVVPGVVSILKTMASNGSLPSTGKGAVER..
YscC/1-607  174  .......IRSEKTGALAIEIFPLKYASASDRTIHYRDDEVAAPGVATILQRVLSDATIQQVTVDNQRIP.
InvG/1-562  170  .......NDGIELGRQKIGVMRLNNTFVGDRTYNLRDQKMVIPGIATAIERLLQGEEQPLGNIVSSEPPA
MxiD/1-566  173  .......VSSIGTDKVNFGVIKLKNTFVSDRTYNMRGEDIVIPGVATVVERLLNNGKALSNRQAQNDP..
PulD/1-660  201  WASAADVVKLVTELNKDTSKSALPGSMVANVVADERTNAVLVSGEPNSRQRIIAMIKQLDRQQATQGNTK
```

InvG and
PulD (F2) ▼ YscC ▼

```
EscC/1-512
EscC/1-512  231  .........................SGNLFDNSVTISADPRLNAVVVKDREITMDIYQQLISEL
YscC/1-607  236  .........................QAATRASAQARVEADPSLNAIIVRDSPERMPMYQRLIHAL
InvG/1-562  233  MPAFSANGEKGKAANYAGGM.....SLQEALKQNAAAGNIKIVAYPDTNSLLVKGTAEQVHFIEMLVKAL
MxiD/1-566  234  MPPFNITQKVSEDSNDFS.........FSSVTNSSILEDVSLIAYPETNSILVKGNDQQIQIIRDIITQL
PulD/1-660  271  VIYLKYAKASDLVEVLTGISSTMQSEKQAAKPVAALDKNIIIKAHGQTNALIVTAAPDVMNDLERVIAQL
```

Protease resistant secretin homology region

```
EscC/1-512
EscC/1-512  270  DIEQRQIEISVSIIDVDANDLQQLGVNWS....GTLNAGQG....TIAFNSSTAQANISSSVISNASN..
YscC/1-607  276  DKPSARIEVALSIVDINADQLTELGVDWR....VGIRTGNNHQVVIKTTGDQSNIASNGALGSLVDARGL
InvG/1-562  298  DVAKRHVELSLWIVDLNKSDLERLGTSWS....GSITIGD.......KLGVSLNQSS...ISTLDGSR..
MxiD/1-566  295  DVAKRHIELSLWIIDIDKSELNNLGVNWQ....GTASFGD.......SFGASFNMSSSASISTLDGNK..
PulD/1-660  341  DIRRPQVLVEAIIAEVQDADGLNLGIQWANKNAGMTQFTNSGLPISTAIAGANQYNKDGTVSSSLASALS
```

```
EscC/1-512
EscC/1-512  330  ...........FMIRVNALQQNSKAKILSQPSIITLNNMQAILDKNVTFYTKVSGEKVA......SLE
YscC/1-607  342  DY...........LLARVNLLENEGSAQVVSRPTLLTQENAQAVIDHSETYYVKVTGKEVA......ELK
InvG/1-562  352  ............FIAAVNALEEKKQATVVSRPVLLTQENVPAIFDNNRTFYTKLIGERNV......ALE
MxiD/1-566  352  ............FIASVMALNQKKANVVSRPVILTQENIPAIFDNNRTFYVSLVGERNS......SLE
PulD/1-660  411  SFNGIAAGFYQGNWAMLLTALSSSTKNDILATPSIVTLDNMEATFNVGQEVPVLTGSQTTSGDNIFNTVE
```

```
EscC/1-512
EscC/1-512  381  SITSGTLLRVTPRILDDSSNSLTGKRRERVRLLLDIQD..GNQSTNQSNAQDASSTLPEVQNSEMTTEAT
YscC/1-607  395  GITYGTMLRMTPRVLTQGDKS.......EISLNLHIED......GNQKPNSSGIEGIPTISRTVVDTVAR
InvG/1-562  403  HVTYGTMIRVLPRFSADG.........QIEMSLDIED...GNDKTPQSDTTTSVDALPEVGRTLISTIAR
MxiD/1-566  403  HVTYGTLINVIPRFSSRG.........QIEMSLTIEDGTGNSQSNYNYNNENTSVLPEVGRTKISTIAR
PulD/1-660  481  RKTVGIKLKVKPQINEGD............SVLLEIEQ..EVSSVADAASSTSSDLGATFNTRTVNNAVL
```

```
EscC/1-512
EscC/1-512  449  LSAGESLLLGGFIQDKESSSKDGIPLLSDIPVIGSLFSSTVKQKHSVVRLFLIKATPIKS.......ASS
YscC/1-607  452  VGHGQSLIIGGIYRDELSVALSKVPLLGDIPYIGALFRRKSELTRRTVRLFIIEPRIIDE.......GIA
InvG/1-562  461  VPHGKSLLVGGYTRDANTDTVQSIPFLGKLPLIGSLFRYSSKNKSNVVRVFMIEPKEIVD.......PLT
MxiD/1-566  463  VPQGKSLLIGGYTHETNSNEIISIPFLSSIPVIGNVFKYKTSNISNIVRVFLIQPREIKE.......SSY
PulD/1-660  537  VGSGETVVVGGLLDKSVSDTADKVPLLGDIPVIGALFRSTSKKVSKRNLMLFIRPTVIRDRDEYRQASSG
```

106

```
EscC/1-512                        PulD
                                   ▼▼ YscC
EscC/1-512   512  E...............................................................
YscC/1-607   515  HHLALGNGQDLRTGILTVDEISNQSTTLNKLLGGSQCQPLNKAQEVQKWLSQNNKSSYLTQCKMDKSLGW
InvG/1-562   524  PDASESVNNILKQSGAWSGDDKLQKWVRVYLDRGQEAIK...............................
MxiD/1-566   526  YNTAEYKSLISER..EIQKTTQIIPSETTLLEDEKSLVSYLNY...........................
PulD/1-660   607  QYTAFNDAQSKQRGKENNDAMLNQDLLEIYPRQDTAAFRQVSAAIDAFNLGGNL................


EscC/1-512

EscC/1-512        ........................
YscC/1-607   585  RVVEGACTPAQSWCVSAPKRGVL
InvG/1-562        ........................
MxiD/1-566        ........................
PulD/1-660        ........................
```

Figure 3.28: Multiple sequence alignment of T3S system secretins EscC (*EPEC*), YscC (*Yersinia enterocolitica),* InvG (*Salmonella typhimurium*), MxiD (*Shigella flexneri*) and the T2S system secretin PulD (*Klebsiella oxytoca*) was performed using ClustalW2 [65] (default parameters, matrix Gonnet250, number of iterations 10) and graphically illustrated by ESPript 2.2 [66]. *T3S system specific homology region* (corresponding to YscC (24-224)) and *protease resistant secretin homology* region (corresponding to YscC (245-511), EscC (240-508) [98], PulD (310-597)) were highlighted. Secondary structure of N-terminal flexible domain of EscC (21-174) as described in [98], was presented above the alignment. Trypsin cleavage sites were marked with orange arrowheads for YscC (Val247 and Arg526) and black arrowheads for InvG (Ile269) [98] and PulD (Gln262/Gln298 and Lys617) [97]. PulD trypsin-resistant domain consists of two fragments F1 (262-297) and F2 (298-617), whereas the very conserved domain is CD2 (348-595).
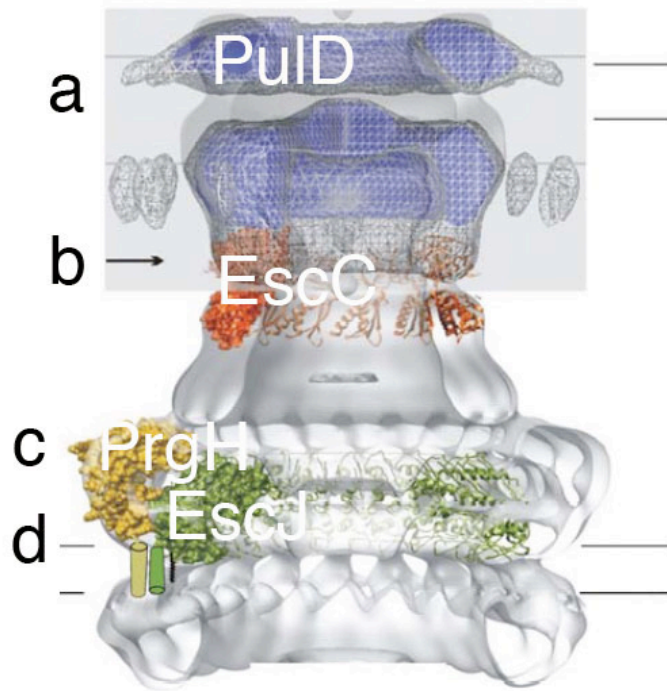
Figure 3.29: Injectisome-chimera based on [98]. Known structures of different host injectisome components were docked into the 3D EM density map of the *S. typhimurium* T3S system basal body [81]. (a) Gray boxed image of full length *(gray mesh)* and trypsin digested *(blue mesh)* cryo EM structure of *K. oxytoca* secretin PulD [97] was imposed on the basal body model. The C-terminal protease resistant secretin homology region was visualized; the N-terminal specific secretin region (marked with black arrow) was nano-gold labeled in the pIV secretin studies [160]. The *gray* radial distributed densities around PulD ring correspond to the last 50 C-terminal residues of secretin and are not conserved among secretins. (b) Flexible domain of *EPEC* EscC (21-174) [98] secretin docked into the model. The generated cryo EM full length YscC secretin model is composed of (a) and (b) parts. (c) *S. typhimurium* PrgH (170-362) [98] corresponds to the YscD which can act as connector between outer and inner rings. (d) *EPEC* EscJ (21-190) [102] corresponds to the YscJ and forms the inner ring.

3.4   Conclusion

Secretins are major components of the type III secretion systems [76]. These systems play the essential role in the interactions between Gram-negative bacteria and eukaryotic hosts. Purified and visualized by EM the T3SS secretins are stable ring-shaped complexes with an internal cavity.

In this study, the YscC oligomer was expressed in *Yersinia enterocolitica*, purified using an improved protocol [95] to obtain highly pure specimens. The 12 Å-resolution 3D structure of the closed YscC complex was calculated from 30000 projections of vitrified YscC. Single particle 3D structure reconstruction revealed the N-terminal flexible domain of YscC, which forms the part of a large chamber between bacterial inner and outer membranes, into which macromolecules might be transported. The core structure consists of the periplasmic conical-shape domain and two differently sized ring-shaped domains linked by fine density connectors corresponding to the membrane spanning regions. The gap between these rings is most likely the location of the OM.

We used various approaches to reveal the oligomer symmetry, however we did not accomplish a precise answer. Although the STEM measurements evidently indicated 4 DDM detergent micelles to be presented around the single secretin particle, due to large statistical errors the symmetry could not be firmly predicted. Both the rotary metal shadowing of trypsinized sample and intact YscC 3D reconstruction in EMAN pointed to C12 or C13 symmetry. Taking into account the sequence analogy to other secretins the symmetry C12 could be assumed, although to have a total confidence it would be necessary to carry out 3D reconstructions of the trypsinized samples in which the trypsin-resistant core is revealed, and not masked by flexible domain of the oligomer. Alternatively, native as well as trypsinized oligomers could be analyzed by mass spectrometry able to resolve complexes in the MDa region.

4   Appendix

The 2DX robot: A membrane protein 2D crystallization Swiss Army knife

The publication "The 2DX robot: A membrane protein 2D crystallization Swiss Army knife" is published in Journal of Structural Biology, Volume 169, Issue 3, March 2010, Pages 370-378.

My contribution to this project consisted in production in *Pichia pastoris* and purification of DDM-solubilized human AQP8, 2D crystallization conditions optimization and protein crystallization in 2DX robot, monitoring the crystallization process as function as methyl-β-cyclodextrin rates during protein reconstitution into the 2D crystals, and the TEM images and electron diffraction data acquisition.

5   Abbreviations

2D	Two Dimensional
3D	Three Dimensional
aa	Amino acid
AFM	Atomic Force Microscope
AOX1	Alcohol Oxidase 1
AQP	Aquaporin
AQP8	Aquaporin 8
AqpZ	Aquaporin Z from *Escherichia coli*
BCIP	5-Bromo-4-chloro-3-indolyl phosphate
BHI broth	Brain Heart Infusion broth
BHI-OX	Sodium oxalate BHI
BMGY	Buffered Glycerol-complex Medium
BMMY	Buffered Methanol-complex Medium
CCD	Charge Coupled Device
CHIP	Channel-forming Integral Protein
CMC	Critical Micelle Concentration
CTF	Contrast Transfer Function
DDM	n-Dodecyl-b-D-maltopyranoside
DM	n-Decyl-b-D-maltopyranoside
DMPC	1,2-Dimyristoyl-sn-Glycero-3-Phosphocholine
DOPC	1,2-Dioleoyl-sn-Glycero-3-Phosphocholine
DOPE	1,2-Dioleoyl-sn-Glycero-3-Phosphatidylethanolamine
DOPG	1-Palmitoyl-2-Oleoyl-sn-Glycero-3-[Phospho-rac-(1-glycerol)]
DOPS	1,2-Dioleoyl-sn-Glycero-3-[Phospho-L-Serine]
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EM	Electron Microscopy
EPEC	Enteropathogenic *Escherichia coli*

FEG	Field Emission Gun
FFT	Fast Fourier Transform
GlpF	Glycerol facilitator from *Escherichia coli*
Glu-C	Endoprotease Glu-C, protease V8 from *Staphylococcus aureus*
IPLT	Image Processing Library and Toolkit
kDa	kilo Dalton
LB	Luria Bertani
LPR	Lipid-to-Protein Ratio
Lys-C	Endoprotease Lys-C, protease from *Lysobacter enzymogenes*
MBCD	Methyl-β-cyclodextrin
MBP	Maltose Binding Protein
MES	2-(N-morpholino)ethanesulfonic acid
MIP	Major Intrinsic Protein
MS	Mass Spectrometry
MW	Molecular Weight
NBT	Nitroblue tetrazolium chloride
Ni-NTA	Ni^{2+}-Nitrilotriacetic acid
NuPAGE	Neutral pH polyacrylamide gel electrophoresis
NPA	Asparagine - Proline - Alanine
OD_{600}	Optical Density at 600 nm
OmpF	Outer Membrane Porin F
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase Chain Reaction
PIP	Plasma Membrane Intrinsic Protein
PMSF	Phenylmethylsulfonylfluoride
POPC	1-Palmitoyl-2-Oleoyl-sn-Glycero-3-Phosphocholine
POPE	1-Palmitoyl-2-oleoyl-sn-Glycero-3-Phosphoethanolamine
POPG	1-Palmitoyl-2-Oleoyl-sn-Glycero-3-[Phospho-rac-(1-glycerol)]
POPS	1-Palmitoyl-2-Oleoyl-sn-Glycero-3-[Phospho-L-Serine]
PVDF	polyvinylidene fluoride
SDS	Sodium-dodecyle sulfate
SNR	Signal-to-Noise ratio
SoPIP2;1	*Spinacia oleracea* Plasma Membrane Intrinsic Protein 2;1

STEM	Scanning Transmission Electron Microscope
T2SS	Type II secretion system
T3SS	Type III secretion system
TaTIP2;1	*Triticum aestivum* Tonoplast Intrinsic Protein 2;1
TEM	Transmission Electron Microscope
TEV	Tobacco Etch Virus
TIP	Tonoplast Intrinsic Protein
Tris	Tis(hydoxymethyl)-aminoethane
Yop	*Yersinia* outer protein
YPD	Yeast Peptone Dextrose
YscC	*Yersinia* secretion component C

6   Acknowledgements

I want to thank the many people who helped me with this thesis. They know who they are.

I especially would like to thank Andreas Engel who was my advisor during the PhD thesis and gave me the opportunity to work in his Laboratory on two very interesting and challenging projects. I am thankful for his great help, support and, particularly, exceptional optimism throughout my PhD.

I am thankful to Wanda Kukulski and Hervé Rémigy, who guided me through the 2D crystallization world, for their help with various wet-lab and cyclodextrin issues, fruitful discussions and friendship.

I would further like to thank Mohamed Chami for patience and sharing with me his huge experience and knowledge about electron microscopy. I would also like to thank Patrick Bosshart for the AFM pictures, Shirley Müller and Philippe Ringler who took the STEM images for the mass measurement and analyzed YscC data and Swetlana Martell who helped me with AQP8 cloning experiments.

During this thesis work I had the pleasure to collaborate with two great research groups leaded by Guy Cornelis and Per Kjellbom. I would like to thank both of them and also the members of their teams.

I am thankful to Henning Stahlberg for his involvement in my projects and to all the members of the C-CINA group for contributing to the good working atmosphere in the lab.

Thanks also go to Andrzej and Szymon on whom you can always rely.

This thesis is dedicated to my parents, Ala and Wojtek, who supported me throughout the ups and downs that come along with a PhD.

7   Bibliography

1.   Gorter, E. and F. Grendel, On Bimolecular Layers of Lipoids on the Chromocytes of the Blood. J Exp Med, 1925. **41**(4): p. 439-43.
2.   Singer, S.J. and G.L. Nicolson, The fluid mosaic model of the structure of cell membranes. Science, 1972. **175**(23): p. 720-31.
3.   Engelman, D.M., Membranes are more mosaic than fluid. Nature, 2005. **438**(7068): p. 578-80.
4.   Simons, K. and E. Ikonen, Functional rafts in cell membranes. Nature, 1997. **387**(6633): p. 569-72.
5.   Takamori, S., et al., Molecular anatomy of a trafficking organelle. Cell, 2006. **127**(4): p. 831-46.
6.   Dreifus, C., Using a Leadership Role to Put a Human Face on Science, in The New York Times. 2009: NY.
7.   Macey, R.I., Transport of water and urea in red blood cells. Am J Physiol, 1984. **246**(3 Pt 1): p. C195-203.
8.   Preston, G.M., et al., Appearance of water channels in Xenopus oocytes expressing red cell CHIP28 protein. Science, 1992. **256**(5055): p. 385-7.
9.   Preston, G.M., et al., The mercury-sensitive residue at cysteine 189 in the CHIP28 water channel. J Biol Chem, 1993. **268**(1): p. 17-20.
10.   Preston, G.M. and P. Agre, Isolation of the cDNA for erythrocyte integral membrane protein of 28 kilodaltons: member of an ancient channel family. Proc Natl Acad Sci U S A, 1991. **88**(24): p. 11110-4.
11.   Borgnia, M., et al., Cellular and molecular biology of the aquaporin water channels. Annu Rev Biochem, 1999. **68**: p. 425-58.
12.   Liu, Z., et al., Arsenite transport by mammalian aquaglyceroporins AQP7 and AQP9. Proc Natl Acad Sci U S A, 2002. **99**(9): p. 6053-8.
13.   Blank, M.E. and H. Ehmke, Aquaporin-1 and HCO3(-)-Cl- transporter-mediated transport of CO2 across the human erythrocyte membrane. J Physiol, 2003. **550**(Pt 2): p. 419-29.
14.   Tornroth-Horsefield, S., et al., Structural insights into eukaryotic aquaporin regulation. FEBS Lett, 2010. **584**(12): p. 2580-8.
15.   Hite, R.K., Z. Li, and T. Walz, Principles of membrane protein interactions with annular lipids deduced from aquaporin-0 2D crystals. EMBO J, 2010. **29**(10): p. 1652-8.
16.   Walz, T., et al., The three-dimensional structure of aquaporin-1. Nature, 1997. **387**(6633): p. 624-7.
17.   Mitsuoka, K., et al., The structure of aquaporin-1 at 4.5-A resolution reveals short alpha-helices in the center of the monomer. J Struct Biol, 1999. **128**(1): p. 34-43.
18.   Jung, J.S., et al., Molecular structure of the water channel through aquaporin CHIP. The hourglass model. J Biol Chem, 1994. **269**(20): p. 14648-54.
19.   de Groot, B.L. and H. Grubmuller, Water permeation across biological membranes: mechanism and dynamics of aquaporin-1 and GlpF. Science, 2001. **294**(5550): p. 2353-7.
20.   de Groot, B.L. and H. Grubmuller, The dynamics and energetics of water permeation and proton exclusion in aquaporins. Curr Opin Struct Biol, 2005. **15**(2): p. 176-83.
21.   Horsefield, R., et al., High-resolution x-ray structure of human aquaporin 5. Proc Natl Acad Sci U S A, 2008. **105**(36): p. 13327-32.
22.   Fushimi, K., S. Sasaki, and F. Marumo, Phosphorylation of serine 256 is required for cAMP-dependent regulatory exocytosis of the aquaporin-2 water channel. J Biol Chem, 1997. **272**(23): p. 14800-4.
23.   Nemeth-Cahalan, K.L., K. Kalman, and J.E. Hall, Molecular basis of pH and Ca2+ regulation of aquaporin water permeability. J Gen Physiol, 2004. **123**(5): p. 573-80.
24.   Hedfalk, K., et al., Aquaporin gating. Curr Opin Struct Biol, 2006. **16**(4): p. 447-56.
25.   Tornroth-Horsefield, S., et al., Structural mechanism of plant aquaporin gating. Nature, 2006. **439**(7077): p. 688-94.
26.   Gonen, T., et al., Aquaporin-0 membrane junctions reveal the structure of a closed water pore. Nature, 2004. **429**(6988): p. 193-7.

27. Tournaire-Roux, C., et al., Cytosolic pH regulates root water transport during anoxic stress through gating of aquaporins. Nature, 2003. **425**(6956): p. 393-7.

28. Nielsen, S., J. Fror, and M.A. Knepper, Renal aquaporins: key roles in water balance and water balance disorders. Curr Opin Nephrol Hypertens, 1998. **7**(5): p. 509-16.

29. Harries, W.E., et al., The channel architecture of aquaporin 0 at a 2.2-A resolution. Proc Natl Acad Sci U S A, 2004. **101**(39): p. 14045-50.

30. Murata, K., et al., Structural determinants of water permeation through aquaporin-1. Nature, 2000. **407**(6804): p. 599-605.

31. Sui, H., et al., Structural basis of water-specific transport through the AQP1 water channel. Nature, 2001. **414**(6866): p. 872-8.

32. Ho, J.D., et al., Crystal structure of human aquaporin 4 at 1.8 A and its mechanism of conductance. Proc Natl Acad Sci U S A, 2009. **106**(18): p. 7437-42.

33. Hiroaki, Y., et al., Implications of the aquaporin-4 structure on array formation and cell adhesion. J Mol Biol, 2006. **355**(4): p. 628-39.

34. Savage, D.F., et al., Architecture and selectivity in aquaporins: 2.5 a X-ray structure of aquaporin Z. PLoS Biol, 2003. **1**(3): p. E72.

35. Fu, D., et al., Structure of a glycerol-conducting channel and the basis for its selectivity. Science, 2000. **290**(5491): p. 481-6.

36. Fischer, G., et al., Crystal structure of a yeast aquaporin at 1.15 angstrom reveals a novel gating mechanism. PLoS Biol, 2009. **7**(6): p. e1000130.

37. Lee, J.K., et al., Structural basis for conductance by the archaeal aquaporin AqpM at 1.68 A. Proc Natl Acad Sci U S A, 2005. **102**(52): p. 18932-7.

38. Newby, Z.E., et al., Crystal structure of the aquaglyceroporin PfAQP from the malarial parasite Plasmodium falciparum. Nat Struct Mol Biol, 2008. **15**(6): p. 619-25.

39. Litman, T., R. Sogaard, and T. Zeuthen, Ammonia and urea permeability of mammalian aquaporins. Handb Exp Pharmacol, 2009(190): p. 327-58.

40. Jahn, T.P., et al., Aquaporin homologues in plants and mammals transport ammonia. FEBS Lett, 2004. **574**(1-3): p. 31-6.

41. Yool, A.J. and A.M. Weinstein, New roles for old holes: ion channel function in aquaporin-1. News Physiol Sci, 2002. **17**: p. 68-72.

42. Ishibashi, K., et al., Molecular cloning and expression of a member of the aquaporin family with permeability to glycerol and urea in addition to water expressed at the basolateral membrane of kidney collecting duct cells. Proc Natl Acad Sci U S A, 1994. **91**(14): p. 6269-73.

43. Yasui, M., et al., Rapid gating and anion permeability of an intracellular aquaporin. Nature, 1999. **402**(6758): p. 184-7.

44. Hazama, A., et al., Ion permeation of AQP6 water channel protein. Single channel recordings after Hg2+ activation. J Biol Chem, 2002. **277**(32): p. 29224-30.

45. Fruhbeck, G., Obesity: aquaporin enters the picture. Nature, 2005. **438**(7067): p. 436-7.

46. Elkjaer, M.L., et al., Immunolocalization of aquaporin-8 in rat kidney, gastrointestinal tract, testis, and airways. Am J Physiol Renal Physiol, 2001. **281**(6): p. F1047-57.

47. Ferri, D., et al., Ontogeny, distribution, and possible functional implications of an unusual aquaporin, AQP8, in mouse liver. Hepatology, 2003. **38**(4): p. 947-57.

48. Carbrey, J.M., et al., Aquaglyceroporin AQP9: solute permeation and metabolic control of expression in liver. Proc Natl Acad Sci U S A, 2003. **100**(5): p. 2945-50.

49. Ishibashi, K., et al., Cloning and identification of a new member of water channel (AQP10) as an aquaglyceroporin. Biochim Biophys Acta, 2002. **1576**(3): p. 335-40.

50. Hatakeyama, S., et al., Cloning of a new aquaporin (AQP10) abundantly expressed in duodenum and jejunum. Biochem Biophys Res Commun, 2001. **287**(4): p. 814-9.

51. Gorelick, D.A., et al., Aquaporin-11: a channel protein lacking apparent transport function expressed in brain. BMC Biochem, 2006. **7**: p. 14.

52. Itoh, T., et al., Identification of a novel aquaporin, AQP12, expressed in pancreatic acinar cells. Biochem Biophys Res Commun, 2005. **330**(3): p. 832-8.

53. King, L.S., D. Kozono, and P. Agre, From structure to disease: the evolving tale of aquaporin biology. Nat Rev Mol Cell Biol, 2004. **5**(9): p. 687-98.

54. Koyama, N., et al., Cloning and functional expression of human aquaporin8 cDNA and analysis of its gene. Genomics, 1998. **54**(1): p. 169-72.

55. Ma, T., B. Yang, and A.S. Verkman, Cloning of a novel water and urea-permeable aquaporin from mouse expressed strongly in colon, placenta, liver, and heart. Biochem Biophys Res Commun, 1997. **240**(2): p. 324-8.

116

56. Koyama, Y., et al., Molecular cloning of a new aquaporin from rat pancreas and liver. J Biol Chem, 1997. **272**(48): p. 30329-33.

57. Liu, K., et al., Purification and functional characterization of aquaporin-8. Biol Cell, 2006. **98**(3): p. 153-61.

58. Gradilone, S.A., et al., Glucagon induces the plasma membrane insertion of functional aquaporin-8 water channels in isolated rat hepatocytes. Hepatology, 2003. **37**(6): p. 1435-41.

59. Calamita, G., et al., Expression and immunolocalization of the aquaporin-8 water channel in rat gastrointestinal tract. Eur J Cell Biol, 2001. **80**(11): p. 711-9.

60. Soria, L.R., et al., Aquaporin-8-facilitated mitochondrial ammonia transport. Biochem Biophys Res Commun, 2010. **393**(2): p. 217-21.

61. Bienert, G.P., et al., Specific aquaporins facilitate the diffusion of hydrogen peroxide across membranes. J Biol Chem, 2007. **282**(2): p. 1183-92.

62. Holm, L.M., et al., NH3 and NH4+ permeability in aquaporin-expressing Xenopus oocytes. Pflugers Arch, 2005. **450**(6): p. 415-28.

63. Saparov, S.M., et al., Fast and selective ammonia transport by aquaporin-8. J Biol Chem, 2007. **282**(8): p. 5296-301.

64. Yang, B., D. Zhao, and A.S. Verkman, Evidence against functionally significant aquaporin expression in mitochondria. J Biol Chem, 2006. **281**(24): p. 16202-6.

65. Larkin, M.A., et al., Clustal W and Clustal X version 2.0. Bioinformatics, 2007. **23**(21): p. 2947-8.

66. Gouet, P., et al., ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics, 1999. **15**(4): p. 305-8.

67. Coburn, B., I. Sekirov, and B.B. Finlay, Type III secretion systems and disease. Clin Microbiol Rev, 2007. **20**(4): p. 535-49.

68. Deakin, W.J. and W.J. Broughton, Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. Nat Rev Microbiol, 2009. **7**(4): p. 312-20.

69. Cornelis, G.R. and H. Wolf-Watz, The Yersinia Yop virulon: a bacterial system for subverting eukaryotic cells. Mol Microbiol, 1997. **23**(5): p. 861-7.

70. Galan, J.E. and A. Collmer, Type III secretion machines: bacterial devices for protein delivery into host cells. Science, 1999. **284**(5418): p. 1322-8.

71. Cornelis, G.R. and F. Van Gijsegem, Assembly and function of type III secretory systems. Annu Rev Microbiol, 2000. **54**: p. 735-74.

72. Mota, L.J., I. Sorg, and G.R. Cornelis, Type III secretion: the bacteria-eukaryotic cell express. FEMS Microbiol Lett, 2005. **252**(1): p. 1-10.

73. Cornelis, G.R., The type III secretion injectisome, a complex nanomachine for intracellular 'toxin' delivery. Biol Chem, 2010. **391**(7): p. 745-51.

74. Cornelis, G.R., Molecular and cell biology aspects of plague. Proc Natl Acad Sci U S A, 2000. **97**(16): p. 8778-83.

75. Imagery From the History of Medicine. Available from: http://www.fransmensonides.nl/pest.htm.

76. Kubori, T., et al., Supramolecular structure of the Salmonella typhimurium type III protein secretion system. Science, 1998. **280**(5363): p. 602-5.

77. Blocker, A., et al., Structure and composition of the Shigella flexneri "needle complex", a part of its type III secreton. Mol Microbiol, 2001. **39**(3): p. 652-63.

78. Tamano, K., et al., Supramolecular structure of the Shigella type III secretion machinery: the needle part is changeable in length and essential for delivery of effectors. EMBO J, 2000. **19**(15): p. 3876-87.

79. Kimbrough, T.G. and S.I. Miller, Contribution of Salmonella typhimurium type III secretion components to needle complex formation. Proc Natl Acad Sci U S A, 2000. **97**(20): p. 11008-13.

80. Marlovits, T.C., et al., Structural insights into the assembly of the type III secretion needle complex. Science, 2004. **306**(5698): p. 1040-2.

81. Marlovits, T.C., et al., Assembly of the inner rod determines needle length in the type III secretion injectisome. Nature, 2006. **441**(7093): p. 637-40.

82. Bayan, N., I. Guilvout, and A.P. Pugsley, Secretins take shape. Mol Microbiol, 2006. **60**(1): p. 1-4.

83. Hodgkinson, J.L., et al., Three-dimensional reconstruction of the Shigella T3SS transmembrane regions reveals 12-fold symmetry and novel features throughout. Nat Struct Mol Biol, 2009. **16**(5): p. 477-85.

84. Kubori, T., et al., Molecular characterization and assembly of the needle complex of the Salmonella typhimurium type III protein secretion system. Proc Natl Acad Sci U S A, 2000. **97**(18): p. 10225-30.

85. Hoiczyk, E. and G. Blobel, Polymerization of a single protein of the pathogen Yersinia enterocolitica into needles punctures eukaryotic cells. Proc Natl Acad Sci U S A, 2001. **98**(8): p. 4669-74.

86. Mueller, C.A., et al., The V-antigen of Yersinia forms a distinct structure at the tip of injectisome needles. Science, 2005. **310**(5748): p. 674-6.

87. Broz, P., et al., Function and molecular architecture of the Yersinia injectisome tip complex. Mol Microbiol, 2007. **65**(5): p. 1311-20.

88. Neyt, C. and G.R. Cornelis, Role of SycD, the chaperone of the Yersinia Yop translocators YopB and YopD. Mol Microbiol, 1999. **31**(1): p. 143-56.

89. Goure, J., et al., The V antigen of Pseudomonas aeruginosa is required for assembly of the functional PopB/PopD translocation pore in host cell membranes. Infect Immun, 2004. **72**(8): p. 4741-50.

90. Deane, J.E., et al., Molecular model of a type III secretion system needle: Implications for host-cell sensing. Proc Natl Acad Sci U S A, 2006. **103**(33): p. 12529-33.

91. Katsura, I., Determination of bacteriophage lambda tail length by a protein ruler. Nature, 1987. **327**(6117): p. 73-5.

92. Journet, L., et al., The needle length of bacterial injectisomes is determined by a molecular ruler. Science, 2003. **302**(5651): p. 1757-60.

93. Moriya, N., et al., The type III flagellar export specificity switch is dependent on FliK ruler and a molecular clock. J Mol Biol, 2006. **359**(2): p. 466-77.

94. Wagner, S., et al., The helical content of the YscP molecular ruler determines the length of the Yersinia injectisome. Mol Microbiol, 2009. **71**(3): p. 692-701.

95. Burghout, P., et al., Structure and electrophysiological properties of the YscC secretin from the type III secretion system of Yersinia enterocolitica. J Bacteriol, 2004. **186**(14): p. 4645-54.

96. Koster, M., et al., The outer membrane component, YscC, of the Yop secretion machinery of Yersinia enterocolitica forms a ring-shaped multimeric complex. Mol Microbiol, 1997. **26**(4): p. 789-97.

97. Chami, M., et al., Structural insights into the secretin PulD and its trypsin-resistant core. J Biol Chem, 2005. **280**(45): p. 37732-41.

98. Spreter, T., et al., A conserved structural motif mediates formation of the periplasmic rings in the type III secretion system. Nat Struct Mol Biol, 2009. **16**(5): p. 468-76.

99. Burghout, P., et al., Role of the pilot protein YscW in the biogenesis of the YscC secretin in Yersinia enterocolitica. J Bacteriol, 2004. **186**(16): p. 5366-75.

100. Lario, P.I., et al., Structure and biochemical analysis of a secretin pilot protein. EMBO J, 2005. **24**(6): p. 1111-21.

101. Silva-Herzog, E., et al., Membrane localization and topology of the Yersinia pestis YscJ lipoprotein. Microbiology, 2008. **154**(Pt 2): p. 593-607.

102. Yip, C.K., et al., Structural characterization of the molecular platform for type III secretion system assembly. Nature, 2005. **435**(7042): p. 702-7.

103. Spaeth, K.E., Y.S. Chen, and R.H. Valdivia, The Chlamydia type III secretion system C-ring engages a chaperone-effector protein complex. PLoS Pathog, 2009. **5**(9): p. e1000579.

104. Imada, K., et al., Structural similarity between the flagellar type III ATPase FliI and F1-ATPase subunits. Proc Natl Acad Sci U S A, 2007. **104**(2): p. 485-90.

105. Cornelis, G.R., The Yersinia Ysc-Yop 'type III' weaponry. Nat Rev Mol Cell Biol, 2002. **3**(10): p. 742-52.

106. Diepold, A., Model of the Yersinia Type III Secretion Injectisome. 2009: Weissenstein. p. Sinergia Meeting.

107. Kimbrough, T.G. and S.I. Miller, Assembly of the type III secretion needle complex of Salmonella typhimurium. Microbes Infect, 2002. **4**(1): p. 75-82.

108. Diepold, A., et al., Deciphering the assembly of the Yersinia type III secretion injectisome. EMBO J, 2010. **29**(11): p. 1928-40.

109. Gill, S.C. and P.H. von Hippel, Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem, 1989. **182**(2): p. 319-26.

110. Jap, B.K., et al., 2D crystallization: from art to science. Ultramicroscopy, 1992. **46**(1-4): p. 45-84.

111. Mass Spectrometry Facility. Available from: http://www.biozentrum.unibas.ch/mass_spec/.

112. Nanoprobes. Available from: http://nanoprobes.com/instructions/Inf2080.html.

113. Hirai, T., et al., Trehalose embedding technique for high-resolution electron crystallography: application to structural study on bacteriorhodopsin. J Electron Microsc (Tokyo), 1999. **48**(5): p. 653-8.

118

114. Philippsen, A., et al., Collaborative EM image processing with the IPLT image processing library and toolbox. J Struct Biol, 2007. **157**(1): p. 28-37.
115. IPLT. Available from: http://www.iplt.org/wiki.
116. Schabert, F.A. and A. Engel, Reproducible acquisition of Escherichia coli porin surface topographs by atomic force microscopy. Biophys J, 1994. **67**(6): p. 2394-403.
117. Kaufmann, T.C., A. Engel, and H.W. Remigy, A novel method for detergent concentration determination. Biophys J, 2006. **90**(1): p. 310-7.
118. Iacovache, I., et al., The 2DX robot: a membrane protein 2D crystallization Swiss Army knife. J Struct Biol, 2010. **169**(3): p. 370-8.
119. Expasy. ProtParam. Available from: http://www.expasy.ch/tools/protparam.html.
120. Sigma-Aldrich. Protocol for Annealing Oligonucleotides. Available from: http://www.sigmaaldrich.com/life-science/custom-oligos/custom-dna/learning-center/annealing-oligos.html.
121. Phan, J., et al., Structural basis for the substrate specificity of tobacco etch virus protease. J Biol Chem, 2002. **277**(52): p. 50564-72.
122. Kapust, R.B., Tözsér, J., Copeland, T. D., and Waugh, D. S., The P1' specificity of tobacco etch virus protease. Biochem. Biophys. Res. Commun., 2002. **294**: p. 949-955.
123. Kapust, R.B. and D.S. Waugh, Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. Protein Science, 1999(8): p. 1668-1674.
124. Karlsson, M., et al., Reconstitution of water channel function of an aquaporin overexpressed and purified from Pichia pastoris. FEBS Lett, 2003. **537**(1-3): p. 68-72.
125. Johns, S.J. Transmembrane protein display software TOPO2. Available from: http://www.sacs.ucsf.edu/TOPO2/.
126. Berdy, J., Bleomycin-Type Antibiotics. Amino Acid and Peptide Antibiotics, 1980: p. 459-497.
127. Opekarova, M. and W. Tanner, Specific lipid requirements of membrane proteins--a putative bottleneck in heterologous expression. Biochim Biophys Acta, 2003. **1610**(1): p. 11-22.
128. Fotiadis, D., et al., Structural characterization of two aquaporins isolated from native spinach leaf plasma membranes. J Biol Chem, 2001. **276**(3): p. 1707-14.
129. Schenk, A.D., et al., The 4.5 A structure of human AQP2. J Mol Biol, 2005. **350**(2): p. 278-89.
130. Zhuang, J., et al., The Campylobacter jejuni porin trimers pack into different lattice types when reconstituted in the presence of lipid. Eur J Biochem, 1997. **244**(2): p. 575-9.
131. Ford, R.C., A. Hefti, and A. Engel, Ordered arrays of the photosystem I reaction centre after reconstitution: projections and surface reliefs of the complex at 2 nm resolution. EMBO J, 1990. **9**(10): p. 3067-75.
132. Lampe, P.D., et al., In vitro assembly of gap junctions. J Struct Biol, 1991. **107**(3): p. 281-90.
133. Seelig, J., P.M. Macdonald, and P.G. Scherer, Phospholipid head groups as sensors of electric charge in membranes. Biochemistry, 1987. **26**(24): p. 7535-41.
134. Dolder, M., A. Engel, and M. Zulauf, The micelle to vesicle transition of lipids and detergents in the presence of a membrane protein: towards a rationale for 2D crystallization. FEBS Lett, 1996. **382**(1-2): p. 203-8.
135. Remigy, H.W., et al., Membrane protein reconstitution and crystallization by controlled dilution. FEBS Lett, 2003. **555**(1): p. 160-9.
136. Signorell, G.A., et al., Controlled 2D crystallization of membrane proteins using methyl-beta-cyclodextrin. J Struct Biol, 2007. **157**(2): p. 321-8.
137. Viadiu, H., T. Gonen, and T. Walz, Projection map of aquaporin-9 at 7 A resolution. J Mol Biol, 2007. **367**(1): p. 80-8.
138. Ringler, P., et al., Structure of the water channel AqpZ from Escherichia coli revealed by electron crystallography. J Mol Biol, 1999. **291**(5): p. 1181-90.
139. Braun, T., et al., The 3.7 A projection map of the glycerol facilitator GlpF: a variant of the aquaporin tetramer. EMBO Rep, 2000. **1**(2): p. 183-9.
140. Kukulski, W., et al., The 5A structure of heterologously expressed plant aquaporin SoPIP2;1. J Mol Biol, 2005. **350**(4): p. 611-6.
141. Daniels, M.J., M.J. Chrispeels, and M. Yeager, Projection structure of a plant vacuole membrane aquaporin by electron cryo-crystallography. J Mol Biol, 1999. **294**(5): p. 1337-49.
142. Scheuring, S., et al., High resolution AFM topographs of the Escherichia coli water channel aquaporin Z. EMBO J, 1999. **18**(18): p. 4981-7.
143. Walz, T., et al., Surface topographies at subnanometer-resolution reveal asymmetry and sidedness of aquaporin-1. J Mol Biol, 1996. **264**(5): p. 907-18.

144.   Muller, D.J. and A. Engel, The height of biomolecules measured with the atomic force microscope depends on electrostatic interactions. Biophys J, 1997. **73**(3): p. 1633-44.

145.   Gonen, T., et al., Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. Nature, 2005. **438**(7068): p. 633-8.

146.   Calamita, G., et al., The inner mitochondrial membrane has aquaporin-8 water channels and is highly permeable to water. J Biol Chem, 2005. **280**(17): p. 17149-53.

147.   Dubochet, J., et al., Cryo-electron microscopy of vitrified specimens. Q Rev Biophys, 1988. **21**(2): p. 129-228.

148.   FEI. Vitrobot. Available from: http://www.fei.com/products/sample-prep/vitrobot.aspx.

149.   Hardie, K.R., S. Lory, and A.P. Pugsley, Insertion of an outer membrane protein in Escherichia coli requires a chaperone-like protein. EMBO J, 1996. **15**(5): p. 978-88.

150.   Ludtke, S.J., P.R. Baldwin, and W. Chiu, EMAN: semiautomated software for high-resolution single-particle reconstructions. J Struct Biol, 1999. **128**(1): p. 82-97.

151.   EMAN. Available from: http://blake.bcm.tmc.edu/eman/eman1.

152.   Harvey Lodish, A.B., S Lawrence Zipursky, Paul Matsudaira, David Baltimore, and James Darnell, Molecular Cell Biology, ed. t. edition. 2000, NY.

153.   Müller, S.A., K. N. Goldie, R. Bürki, R. Häring, and A. Engel, Factors influencing the precision of quantitative scanning transmission electron microscopy. Ultramicroscopy, 1992. **46**: p. 317-334.

154.   Merril CR, G.D., Sedman SA, Ebert MH., Ultrasensitive stain for proteins in polyacrylamide gels shows regional variation in cerebrospinal fluid proteins. Science, 1981. **211**(4489): p. 1437-8.

155.   3D reconstruction in EMAN, part 3. 2008; Available from: www.microbio.uab.edu/mic774/lectures/dokland9.12.08.pdf.

156.   UCSF Chimera. Available from: http://www.cgl.ucsf.edu/chimera/

157.   Emanuelsson, O., Brunak, S., von Heijne, G., Nielsen, H., Locating proteins in the cell using TargetP, SignalP, and related tools. Nature Protocols, 2007. **2**: p. 953-971.

158.   Nouwen, N., et al., Secretin PulD: association with pilot PulS, structure, and ion-conducting channel formation. Proc Natl Acad Sci U S A, 1999. **96**(14): p. 8173-7.

159.   Hasler, L., et al., Structural Analysis of Photosystem II: Comparative Study of Cyanobacterial and Higher Plant Photosystem II Complexes. J Struct Biol, 1997. **119**(3): p. 273-83.

160.   Opalka, N., et al., Structure of the filamentous phage pIV multimer by cryo-electron microscopy. J Mol Biol, 2003. **325**(3): p. 461-70.

161.   Collins, R.F., et al., Structure of the Neisseria meningitidis outer membrane PilQ secretin complex at 12 A resolution. J Biol Chem, 2004. **279**(38): p. 39750-6.

162.   Sanowar, S., et al., Interactions of the Transmembrane Polymeric Rings of the Salmonella enterica Serovar Typhimurium Type III Secretion System. MBio, 2010. **1**(3).

163.   Gohlke, U., et al., The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. Proc Natl Acad Sci U S A, 2005. **102**(30): p. 10482-6.

164.   Collins, R.F., et al., Analysis of the PilQ secretin from Neisseria meningitidis by transmission electron microscopy reveals a dodecameric quaternary structure. J Bacteriol, 2001. **183**(13): p. 3825-32.

165.   Reichow, S.L., et al., Structure of the cholera toxin secretion channel in its closed state. Nat Struct Mol Biol, 2010. **17**(10): p. 1226-32.

166.   Guilvout, I., et al., Bacterial outer membrane secretin PulD assembles and inserts into the inner membrane in the absence of its pilotin. EMBO J, 2006. **25**(22): p. 5241-9.

167.   Nouwen, N., et al., Domain structure of secretin PulD revealed by limited proteolysis and electron microscopy. EMBO J, 2000. **19**(10): p. 2229-36.

168.   Linderoth, N.A., M.N. Simon, and M. Russel, The filamentous phage pIV multimer visualized by scanning transmission electron microscopy. Science, 1997. **278**(5343): p. 1635-8.

169.   Shevchik, V.E., J. Robert-Baudouy, and G. Condemine, Specific interaction between OutD, an Erwinia chrysanthemi outer membrane protein of the general secretory pathway, and secreted proteins. EMBO J, 1997. **16**(11): p. 3007-16.

170.   Daefler, S., M. Russel, and P. Model, Module swaps between related translocator proteins pIV(f1), pIV(IKe) and PulD: identification of a specificity domain. J Mol Biol, 1997. **266**(5): p. 978-92.

171.   Daefler, S. and M. Russel, The Salmonella typhimurium InvH protein is an outer membrane lipoprotein required for the proper localization of InvG. Mol Microbiol, 1998. **28**(6): p. 1367-80.

172.   SABLE. Accurate sequence-based prediction of relative Solvent AccessiBiLitiEs, secondary structures and transmembrane domains for proteins of unknown structure. Available from: http://sable.cchmc.org/.

173.   Felsenstein, J., PHYLIP - Phylogeny Inference Package (Version 3.2). Cladistics, 1989. **5**: p. 164-166.

174.   Saitou, N. and M. Nei, The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol, 1987. **4**(4): p. 406-25.

175.   Collins, R.F., et al., Interaction with type IV pili induces structural changes in the bacterial outer membrane secretin PilQ. J Biol Chem, 2005. **280**(19): p. 18923-30.

176.   Harpaz, Y., M. Gerstein, and C. Chothia, Volume changes on protein folding. Structure, 1994. **2**(7): p. 641-9.

8   Curriculum Vitae

Julia Kowal

Personal details

Date of birth: 06.05.1982

Place of birth: Kraków, Poland

Nationality: Polish

Education

2006 – 2010:   PhD studies in the group of Prof. Andreas Engel, Center for Cellular Imaging and Nano Analytics (C-CINA), Biozentrum, University of Basel, Switzerland, within the Marie Curie Research Training Network *Aqua(glycero)porins*

2006:   Master's Degree in Biochemistry

2001 – 2006:   Undergraduate studies, Faculty of Biochemistry, Biophysics and Biotechnology, Jagiellonian University, Kraków, Poland

1997 – 2001:   Secondary school; Jan III Sobieski 2nd General College, Kraków, Poland

1988 – 1997:   Primary school; 113th Elementary School, Kraków, Poland

Practical experience

2006 – 2010:   PhD studies in the Structural Biology group of Prof. Andreas Engel, Center for Cellular Imaging and Nano Analytics (C-CINA), Biozentrum, University of Basel, Switzerland

Thesis: Structure determination of membrane-located complexes: Aquaporin 8 and YscC secretin

2005:   6 month research project within the EU Student Exchange Program (Erasmus Programme) in the NMR group of Dr. Tad Holak, Max-

Planck-Institut für Biochemie, Department of Structural Research, Martinsried, Germany

2004 – 2006:	Master`s project in the Biochemistry group of Prof. Adam Dubin, Department of Analytical Biochemistry, Jagiellonian University, Kraków, Poland

Thesis: Biochemical properties of the two types of aureolysin, metalloprotease from _Staphylococcus aureus_

Publications

I Iacovache, Biasini M, **Kowal J**, Kukulski W, Chami M, van der Goot FG, Engel A, and Remigy HW. 2010. The 2DX robot: a membrane protein 2D crystallization Swiss Army knife. J Struct Biol. 169:370-8.

AJ Sabat, Wladyka B, Kosowska-Shick K, Grundmann H, van Dijl JM, **Kowal J**, Appelbaum PC, Dubin A, and Hryniewicz W. 2008. Polymorphism, genetic exchange and intragenic recombination of the aureolysin gene among Staphylococcus aureus strains. BMC Microbiol. 8:129.

MA Fellert, **Kowal J**, Kukulski W, Norden K, Gustavsson N, Johanson U, Engel A, and Kjellbom P. Reconstitution of water channel function and crystallization of human aquaporin 8. Manuscript submitted to Proc. Natl. Acad. Sci. USA.

Presentations

2010:	Sinergia Meeting, Basel: The YscC oligomer

2009:	Aquaglyceroporins Progress Report Meeting, Göttingen, Germany: "Aquaporin 8 … the last battle"

Sinergia Meeting, Weissenstein, Switzerland: YscC EM imaging: "The ups, downs and dangers in the field"

2008:	Aquaglyceroporins Progress Report Meeting, Engelberg/Basel: "2D crystallization of AQP8. "

Per Kjellbom laboratory, Department of Biochemistry, Center for Molecular Protein Science, Lund, Sweden: "Cryo Electron Microscopy and

analysis of AQP8 projection map"

Aquaglyceroporins Progress Report Meeting, Aarhus, Denmark: "Aquaporin 8 Studies"

2007: Aquaglyceroporins Progress Report Meeting, Manchester Interdisciplinary Biocentre (MIB), University of Manchester, Manchester, UK: "AQP8: purification and 2D crystallization. Part 1"

EMBO Practical Course on Image Processing for Cryo EM, Birkbeck College, London, UK: "Electron microscopy of human Aquaporin 8. Purification and 2D crystallization"

Aquaglyceroporins Progress Report Meeting, Institute of Marine Sciences, CSIC, Barcelona, Spain: "AQP8: purification and 2D crystallization. Part 2"

2006: 33th International Winter School of Biotechnology in Krynica, Poland: "Biochemical properties of the two types of aureolysin."

Teaching

2009: Instructor in the *Blokkurs in Biophysik und Strukturbiologie* in the image processing and electron crystallography exercises. Biozentrum, University of Basel

2008: Instructor in the *Blokkurs in Biophysik und Strukturbiologie* in the image processing and electron crystallography exercises. Biozentrum, University of Basel

Tutor at the *7th NCCR Practical Course and EMBN Summer School Membrane Protein 2D Crystallization and Observation*

Tutor at the *Course in Protein Crystallography-EM (2D)* within the Marie Curie Research Training Network, University of Basel, Engelberg/Basel

2007: Instructor in the *Blokkurs in Biophysik und Strukturbiologie* in the image processing and electron crystallography exercises. Biozentrum, University of Basel

Selected Posters

2D Crystallization and Electron Crystallography of two Aquaporins: AQP8 and SoPIP2;1

J Kowal, Kukulski W, Kjellbom P and Engel A

Biozentrum Annual Symposium, Jan 2008, Basel, Switzerland

Two dimensional crystallization of human Aquaporin 8

J Kowal, Fellert M, Kukulski W, Johanson U, Kjellbom P, and Engel A

VIII European Symposium of The Protein Society, Jun 2009, Zurich, Switzerland