Structural investigation into recombinant eye lens aquaporin (AQP0) and the effector proteins (BEPS) from *Bartonella henselae*

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von
Dinesh Vellore Palanivelu
aus
Indien

Basel, 2005

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Tilman Schirmer
Prof. Dr. Christoph Dehio
Prof. Dr. Andreas Engel

Basel, den 5.7.2005

Prof. Dr. Hans-Jakob Wirz
Dekan

Declaration

I declare that I wrote this thesis, **Structural investigation into recombinant eye lens aquaporin (AQP0) and the effector proteins (BEPS) from *Bartonella henselae***, with the help indicated and only handed it into the Faculty of Science of the University of Basel and to no other faculty and no other university**.**

Acknowledgements

I would like to thank, My “Doktorvater”, Prof. Dr. Tilman Schirmer, for his guidance and encouragement, which he provided throughout my PhD studies. I thank him for his interesting discussions, suggestions, advice throughout my student life and special thanks for giving me the opportunity to work in his group. I am grateful to Prof. Dr. Andreas Engel and Prof. Dr. Christoph Dehio for their enthusiastic motivation with constant support which helped me to work in their projects with full determination.

I would like to extend my special thanks to Kitaru Suda who trained me to work in the field of membrane protein. Without his knowledge it could be a milestone to achieve success in this field. It was a great pleasure in working with my group members, Dr. George Orriss, Dr. Zora Housely, Dr. Caroline Peneff, Dr. Carmen Chan, Sivaraman, Dr. Arnaud Basle and Dietrich Samoray. Their support and motivation scientifically and personally made a good relationship which will remain in my memories forever. Dr. Dimitri Fotiadis, the greek guy, who loves Indian food so much is one of my best friend in Basel. His tireless efforts and great patience for long discussions has helped me to gain knowledge in the field of aquaporins. Special thanks to Jun who is my scientific advisor, shown great aspects of showing science in a unique dimension. Thanks to Mrs. Ute Grutter for her great help with the administration work and to find a stay during the first year of my life in Basel. She helped me a lot in translating all the German letters. I thank Roland Buerki and Margrit Jenny for their support in computer installations.

Many thanks to Senthil, Carmen, Siva and Satish for supporting me throughout my thesis writing. Without their help it could have been difficult to complete my thesis within the time frame. My stay in Basel was made so lively by many people whom I met and made good friends. My Basel friends - Senthil, Indu, Bans, Satish, Siva, Sachin, Arundhathi, Niranjan, Joe, Anurag and Shyam for all the memorable moments which we shared during these four years. My special regards to Pilar Garcia, my very first friend that I

discovered in Basel, for all her support and help during my stay in Basel. Thanks to Brian and Zora Housley for many great evenings and funfilled chatting which made Switzerland looked homely. Thanks for Tilman's family for all their support for my family life.

My humble thanks to Dr. Suguna for giving me the opportunity as a research assistant and introducing me into the field of crystallography. Thanks to Prof. Dr.Venugopal P. Menon, my Masters supervisor, helped in my research career and supported for my PhD position.

I am very thankful to all my colleagues in Zurich- Anand, Koushick, Rajesh, Bala, Aravind, Mathi, Ram, Bhoopathi and co., for giving me lot of support and fun. They made a huge difference for my stay in Switzerland.

My life started blossoming when Kavitha (my sweet wife) entered inside into my life. Our sweet son, Charan's arrival in this world made our life meaningful. My wife extended all her support during my final year of PhD life. She gave all the motivation and encouragement for helping me to focus on research. My special thanks to my parents and my sweet brother. I got all the support from them for finding my PhD career. They deserve all the credit for my work.

Finally, I thank the University of Basel and the Swiss National Foundation for their financial support.

… to my parents

Abbreviations

T4SS	Type IV secretion system
Bep	Bartonella-translocated effector protein
BID	Bep Intracellular Delivery
ABC	ATP binding cassette
CRAFT	Cre recombinase assay for translocation
YTHS	Yeast two hybrid studies
HUVAC	Human umbilical vein endothelial cells
GFP	Green fluorescent protein
IMAC	Ion metal affinity chromatography
AU	Analytical ultracentrifugation
EM	Electron microscopy
SEM	Scanning electron microscopy
AFM	Atomic force microscopy
OM	Outer membrane
IM	Inner membrane
His-tag	Hexahistidine tag
FACS	Fluorescence activated cell sorter
HPLC	High pressure liquid chromatography
MAD	Multiwavelength anomalous diffraction
MW	Molecular weight
IPTG	Isopropyl-β-D-thiogalactopyranoside
PEG	Polyethylene glycol
LB	Luria-Bertani medium
MALDI-TOF	Matrix assisted laser desorption ionization-Time of flight
LC-MS	Liquid chromatography-Mass spectrometry
SDS-PAGE	Sodium dodecylsulphate-polyacrylamide gel electrophoresis
BN-PAGE	Blue native- polyacrylamide gel electrophoresis
R-factor	Crystallographic residual for working set of reflections

R_{free}	Crystallographic residual for test set of reflections
rpm	Rotation per minute
OD	Optical density
kDa	Kilo Dalton
RT	Room temperature

PREAMBLE

Structural biology is a fundamental aspect of modern biology which aims to determine the three-dimensional atomic structure of biological molecules, especially proteins, and hence give insight into the function of these molecules in the living cell. In particular, knowledge of structures of key human and pathogen proteins is vital for speeding up the difficult task of discovering new drugs or antibiotics. My project is comprised of performing structural investigation on two different proteins, an integral transmembrane protein and effector proteins from a Gram-negative bacteria, respectively. I present my thesis therefore in two chapters.

The first part of Chapter I describes in detail the expression of bovine eye lens aquaporin AQP0 and its crystal structure. The second part contains the manuscript of this work which is to be submitted.

In Chapter II, I describe in detail the structural studies on the various effector proteins (Beps) of the VirB-Type IV secretion system in *Bartonella henselae*.

Table of Contents

Chapter IA

Co-axial association of recombinant eye lens aquaporin AQP0

ABSTRACT

Aquaporin-0 (AQP0) is the major membrane protein present in the vertebrate eye lenses. It has been proposed that AQP0 tetramers mediate contact between membranes of adjacent lens fiber cells, which would be consistent with the extraordinarily narrow inter-cellular space. Indeed, octamer formation with purified AQP0 in solution was observed by native gel electrophoresis and analytical ultracentrifugation methods. We obtained 3D crystals of AQP0 that diffract to 7.0 Å resolution and molecular replacement was performed using the recently determined 3D structure of AQP0 from native source. The result shows that, within the cubic lattice, tetramers (point symmetry 42) are associated head-to-head. There are no direct octamer-octamer contacts and the crystal integrity is most probably maintained by detergent belts surrounding the membrane protein. Within the octamer, extracellular loops A and C interdigitate at the center and the perimeter of the octamer, respectively. The octamer formation has been compared with the AQP0 structure derived from 2D crystals using electron diffraction. Intriguingly, the mutual orientation of tetramers within the octamer is significantly different to that previously reported for 2D crystals. Clearly, the low resolution of the X-ray data permits only a comparison of the oligomeric arrangement. The interactions observed in the loosely-packed 3D crystals presented here possibly represent the *in vivo* association mode between AQP0 tetramers from juxtaposed membranes in the eye lens.

1.0 Introduction

1.1 Aquaporins: The water channel proteins

Aquaporins (AQPs) are a family of water and small molecule channels found in organisms ranging from bacteria to animals (Figure 1). AQPs are integral transmembrane proteins that play critical roles in controlling the water content of cells depending on their osmotic gradient. AQP gene superfamily encodes over 150 isoforms in nature [1]. The family of AQPs has so far been divided into two subgroups: (a) pure water pores (AQP cluster) and (b) pores through which small uncharged solutes and gases pass (GLP cluster). AQPs play a central role in water homeostasis in both plants and animals and is exquisitely selective for the transport of water and repulsive to hydronium ion (H_3O^+) [2]. Some members of the AQP superfamily are permeated by a variety of small neutral solutes such as glycerol and urea and hence are termed as aquaglyceroporins (Figure 1).

1.2 Types and functions of AQPs in *Homo sapiens*

The diversity of the aquaporin is embodied by the human proteome, where at least ten different AQPs are expressed in tissues and cells such as brain, kidneys and erythrocytes (Figure 2). The sites of expression of these AQPs predict the clinical phenotypes in humans as shown in Table 1. AQP0-AQP9 can be subdivided into two major groups. AQP0-AQP2, AQP4-AQP6 and AQP8 are permeable to water and AQP3, AQP7 and AQP9 are permeable to glycerol or urea in addition to water. They play a fundamental role in osmoregulation and mutations of these proteins are responsible for several human diseases. AQP0 is the founding member of the AQP superfamily which is present in the vertebrate eye lenses. It plays a vital role in maintaining the transparency and the structural integrity of the eye lens. Mutations in AQP0 (Glu134Gly or Thr138Arg) lead to congenital cataract formation [3]. AQP1, the major water channel of the human system is found expressed in multiple tissues such as kidney, red blood cells etc. Individuals lacking Colton blood group antigens have mutations in the AQP1 gene. When deprived of water, AQP1-null individuals exhibit a defect in urine concentration and a marked

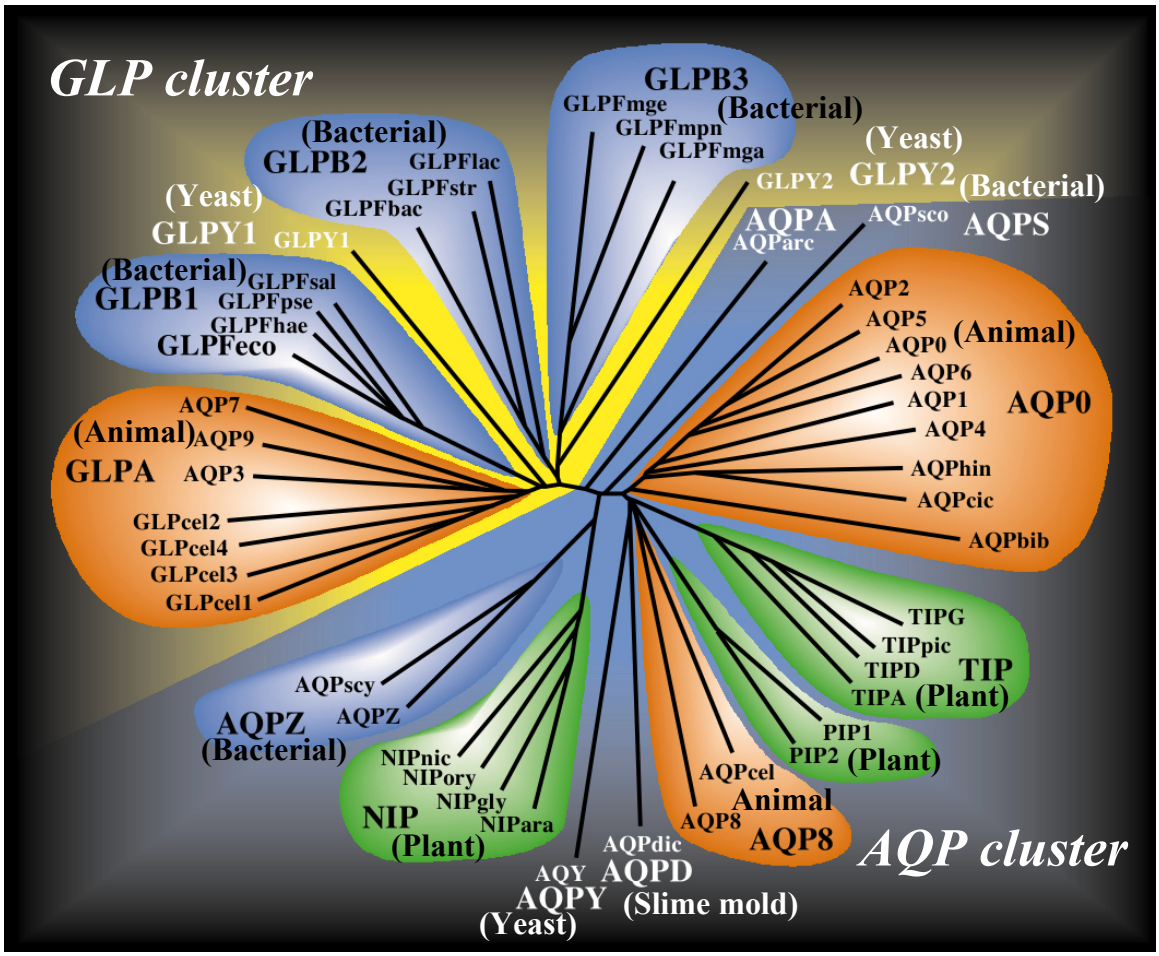


Figure 1: The aquaporin superfamily is classified by phylogenetic analysis in two subgroups, the AQP and the GLP cluster, 16 subfamilies and 46 types. The figure adapted from Heymann et al., 1999.

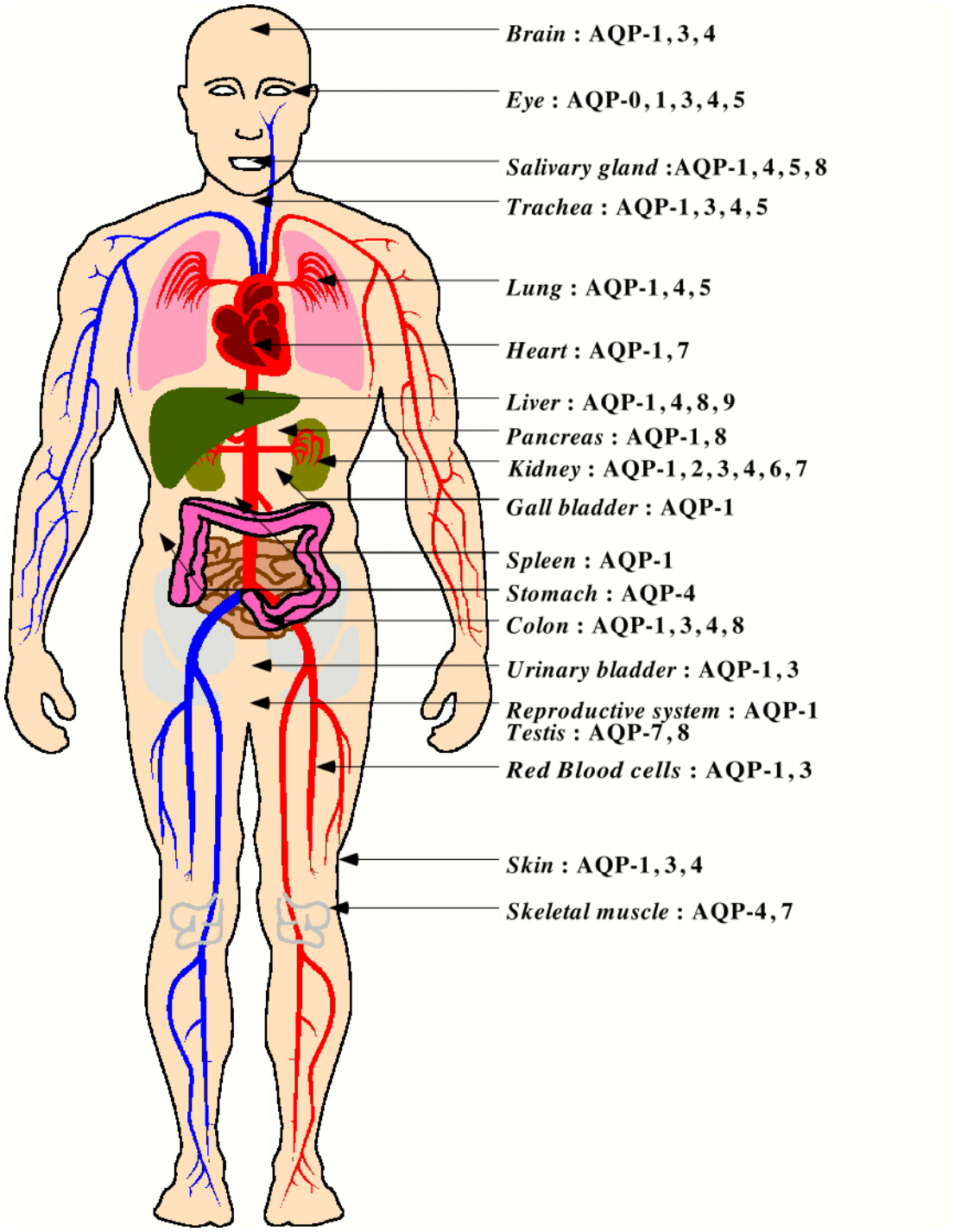


Figure 2: The localization of different types of AQPs in the human system. (adapted from Federique Tacnet et al., 2000).

Aquaporins	Major sites of expression	Functions
Aquaporin-0	Eye: lens fiber cells	Fluid balance within the lens
	Red blood cells	Osmotic protection
	Kidney: proximal tubule	Concentration of urine
Aquaporin-1	Eye: ciliary epithelium	Production of aqueous humor
	Brain: choriod plexus	Production of cerebrospinal fluid
	Lung: alveolar epithelial cells	Alveolar hydration state
Aquaporin-2	Kidney: collecting ducts	Mediates antidiuretic hormone activity
	Kidney: collecting ducts	Reabsorption of water into blood
Aquaporin-3*	Trachea: epithelial cells	Secretion of water into trachea
	Kidney: collecting ducts	Reabsorption of water
	Brain: ependymal cells	CSF fluid balance
Aquaporin-4	Brain: hypothalamus	Osmosensing function?
	Lung: bronchial epithelium	Bronchial fluid secretion
	Salivary glands	Production of saliva
Aquaporin-5	Lacrimal glands	Production of tears
Aquaporin-6	Kidney	Very low water permeability
Aquaporin-7*	Testis and sperm	
Aquaporin-8	Testis, pancreas, liver, others	
Aquaporin-9*	Leukocytes	

(*) corresponds to the aquaglyceroporins.

Table 1: A summary of different types of aquaporins and their corresponding localization and functions in the mammalian system.

reduction in fluid exchange between capillary and interstitium in lungs. AQP2 is responsible for the vasopression-regulated water channel of renal collecting ducts [4]. Mutations in this protein cause nephrogenic diabetes insipidus (NDI) disease. AQP3, 7 and 9 are distributed in multiple organs such as kidney, skin, eye, brain, liver etc. They play a physiological role as a glycerol transporter which may be important during fasting or starvation [5]. AQP4 is found in blood brain barrier, stomach, lungs, skeletal muscle etc. Its deficiency causes duchenne muscular dystrophy [6]. AQP5 is present in the secretory glands (salivary glands, lacrimal glands and submucosal glands), lung and eye. Malfunctioning of this protein leads to Sjogren's syndrome, lymphocyte infiltration, dry eye and mouth [7]. AQP6 functions as a gated ion channel of renal collecting ducts and plays a pivotal role in the control of urine acidification [8]. In summary, AQPs play a vital role in human physiology and are believed to be potential targets for therapeutic intervention for drug targeting.

1.3 Eye lens aquaporin AQP0

The AQP0, formerly known as major intrinsic protein (MIP26) is a transmembrane water channel protein present in vertebrate eye lenses. It is the most abundant protein in the plasma membrane of lens fiber cells constituting almost 50 % of the total protein and is considered to be a lens differentiation marker [9]. AQP0 has been identified in abundance in lens fiber cells for its various physiological functions [9, 10]. It was proposed to perform a dual function: a water channel to maintain lens transparency and a cell-to-cell adhesion to perform a structural role. The optically dense biconvex body of the lens is a major element in the eye physiology. Lens consists of symmetric shells of fiber cells which meet end-to-end along well defined planes forming the lens sutures (Figure 3). Two principle properties of the lens are transparency and refractive power. Lens represents unique features in structure, shape, physiology and biochemistry to achieve the transmittance of light. In order to maintain the lens transparency, the content of the water has to be limited in the cells as well as in the extracellular space[11]. The architecture of the eye lens shows several distinct features:

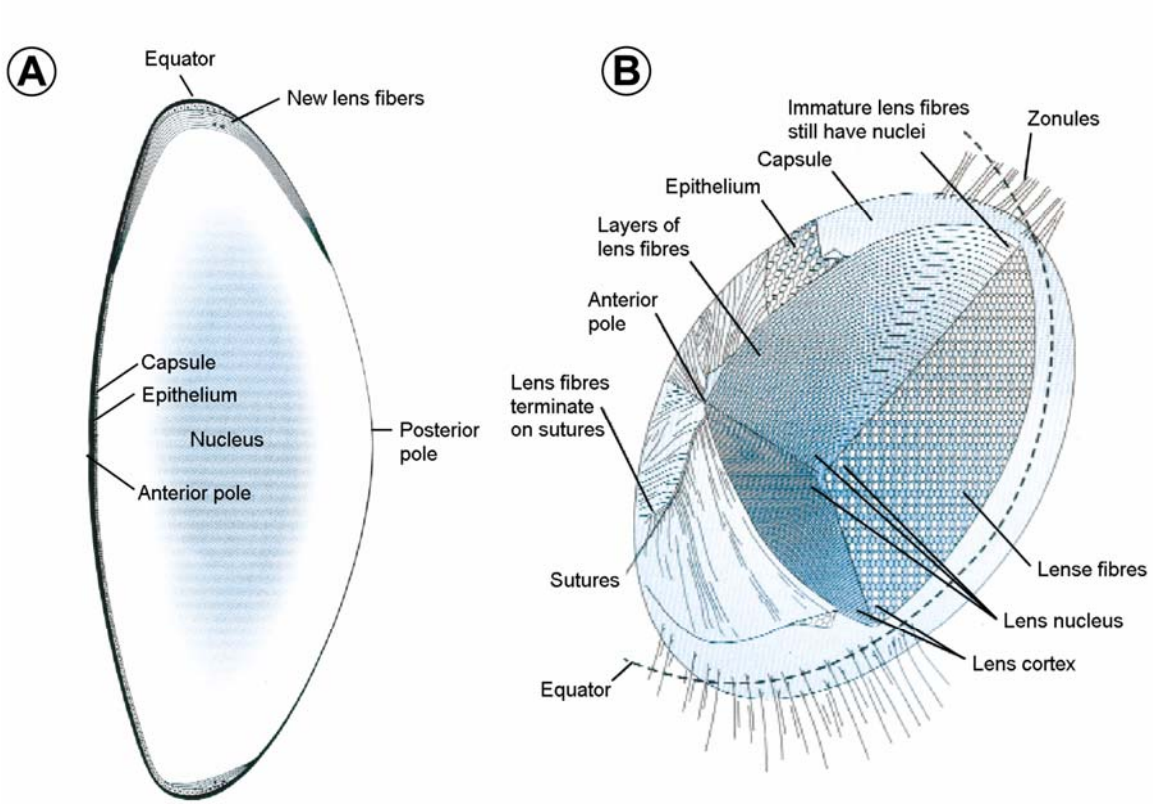


Figure 3: The assembly of the vertebrate eye lens. (A) Vertical cross section through the main focusing element of the crystalline lens. (B) Three dimensional diagram of the lens, which reveals the internal construction of lens fiber cells. The figure adapted from Stafford, M., 2001[11].

- the maintenance of a highly ordered molecular structure of the crystalline proteins;
- terminally differentiated fiber cells containing very few organelles; and
- intracellular and intercellular spaces being kept smaller than the wavelength of the ambient light[12].

To perform these unique functions, lens core contains high amounts of AQP0 and other gap junction proteins (Example: connexins). The genetic sequence of AQP0 identified it as a member of the AQP, rather than the connexion family[13]. The function of AQP0 besides the water channel activity is believed to provide cell adhesive property. This putative role in mediating cell-cell interaction appears unique amongst aquaporins. The structural and functional studies for AQP0 are discussed in the following chapters in detail.

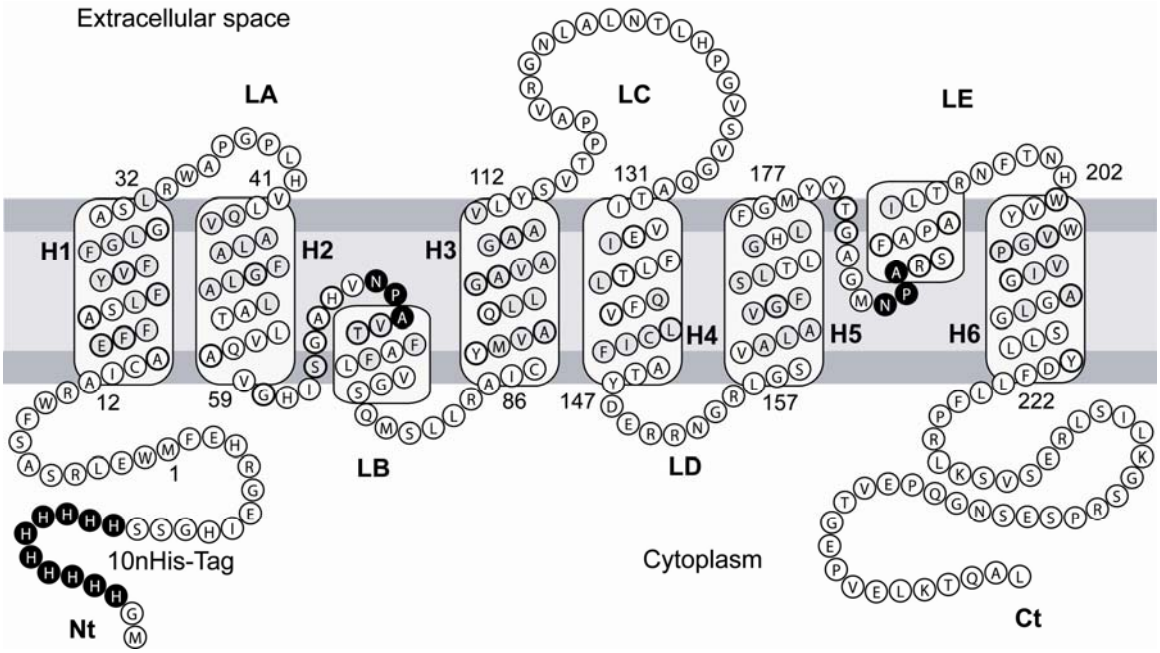


Figure 4. Topological model of the amino acid sequence of recombinant bovine AQP0 showing six membrane-spanning helices and the two short functional loops B and E with the NPA motive partially embedded in the membrane with an internal symmetry. This model was derived from hydropathy analysis by TmPred K. Hofmann & W. Stoffel (1993).

AQP0 (monomer mass 26 kDa) are homotetramers containing four independent hour glass-shaped channels [14-16]. AQP0, like other aquaporins form a common membrane topology having six transmembrane segments forming helices H1-H3 and the repeated analogues H4-H6 (Figure 4) and five connecting loops obeying an internal two-fold symmetry. Both loops B and E contain the consensus motif Asn-Pro-Ala and dip into the membrane from opposite sides contributing to form the water channel. The structural unit of AQP0 is a tetramer.

1.3.1 Water channel function

AQP1 has been shown to allow facilitated passive diffusion of water across the erythrocyte membrane [3]. Experimental evidence to show that AQP0 also acts as a water channel protein has been obtained from oocyte swelling assays. The water channel function of AQP0 was 40 times less efficient compared to AQP1 [17] but AQP0 water conductance doubles under mildly acidic conditions. In addition to the water channel property, exogenous expression studies demonstrated a glycerol transport function for AQP0 [18]. Sequence alignment shows AQP0 is closely related to the water channel AQP1 (44 % identiy, 62.6% structural similarity) (Figure 5). The presence of His 172 in AQP0, a residue conserved only in aquaporins but substituted by glycines in aquaglyceroporins also suggests that AQP0 is a water channel. AQP0 water channels are considered essential for the lens microcirculation system, proposed to supply deeper-lying fiber cells with nutrients and to clear waste products [19]. AQP0 forms functional water pores with a maximal conductance around pH 6.5 [20] . The water channel function has been investigated from the two high resolution structures of AQP0 obtained from electron microscopy (EM) and X-ray respectively. Interestingly, the AQP0 EM structure at 3.0 Å resolution (1SOR) [21] of AQP0 showed the closed pore conformation at pH 6.0. In contrast, the AQP0 X-ray structure solved at 2.2 Å resolution (1YMG) showed the presence of eight water molecules in the channel [22]. The X-ray and EM structures are remarkably identical to an α-carbon rms deviation of 1.12 Å. In the EM structure, the water pore was found to be more constricted and it was suggested to be due to the packing of AQP0 molecules in the 2D crystals.

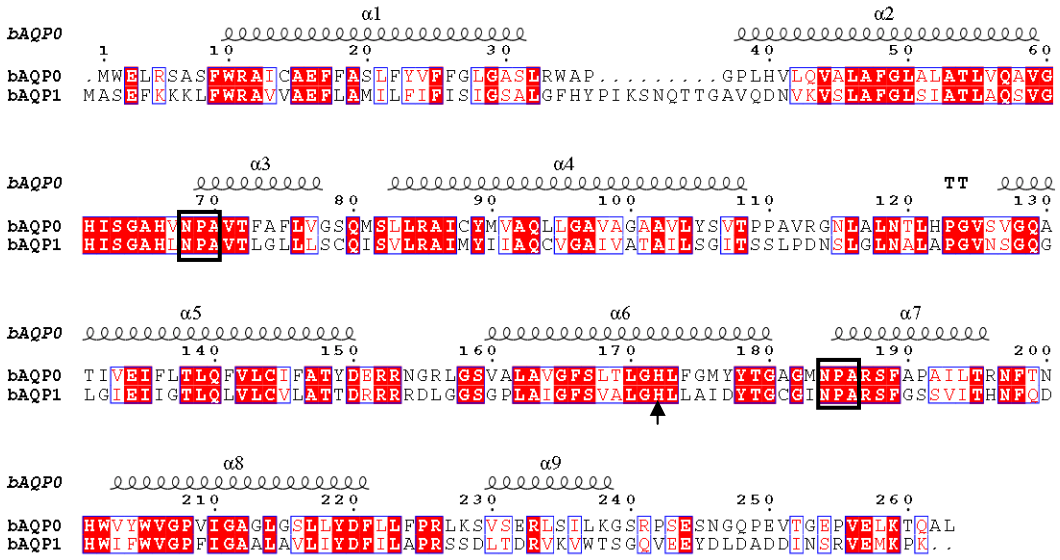


Figure 5: Sequence alignment of bAQP0-bAQP1. The secondary structure of bAQP0 is shown on top. The alignment shows 44 % identity and 62 % structural similarity. The consensus NPA motifs are shown in boxes. The conserved His 172 in aquaporins is marked with an arrow.

1.3.2 Cell-to-cell adhesion function

The adhesive function of AQP0 was proposed since they are arranged in the eye lens as crystalline square arrays (Figure 6). Zamphigi and co workers have investigated the structural organization and protein composition of lens fiber junctions isolated from adult bovine and calf lenses using electron microscopy, immunolocalization from freeze-fracture and label-fracture methods (Figure 7) [10]. It was observed that there are three types of lens junctions (thick, thin or straight and wavy). They form the pentalamellar structure which consists of two plasma membranes apposed through their external surfaces. AQP0 was found only in tightly abutting membranes (thin and wavy junctions) but not in thick junctions (Figure 7A). The wavy junctions are asymmetric structures, where AQP0 was found only in one membrane and flips into another membrane forming a wave like appearance (Figure 7B and 7C).

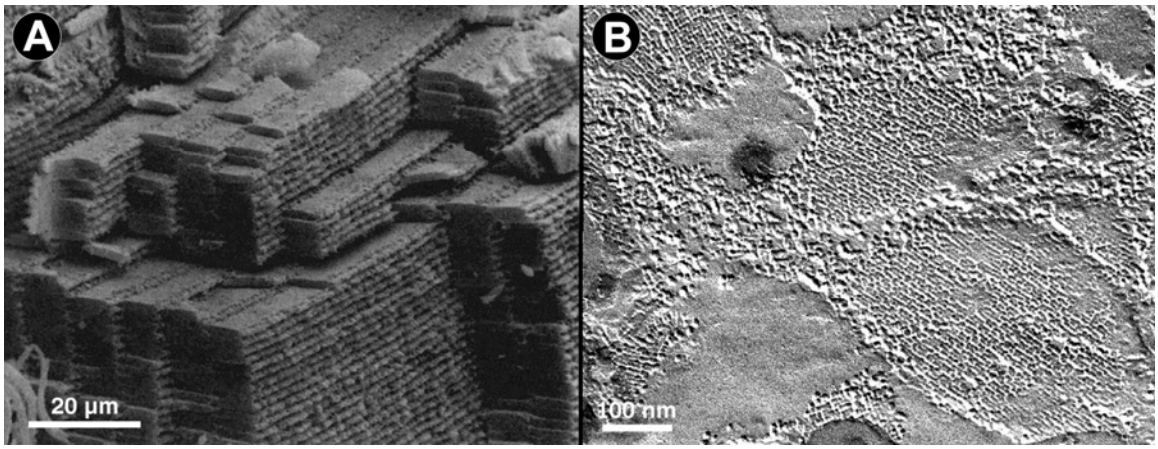


Figure 6: AQP0, the main component of lens fiber membranes is highly enriched in junctional domains connecting fibers. (A) Adhesive proprieties of AQP0 lead to stacking of fiber cells in the lens nucleus and cortex. (B) Arrays of AQP0 in lens fiber cell membranes are visualized by the freeze fracture technique. The figure (A) and (B) adapted from Molecular Biology of the cell and Kistler et al., 1980 [23] respectively.

In thin or straight junctions AQP0 was found localized in both the membranes forming a symmetric arrangement. These junctions have an overall thickness of 11 to 13 nm containing tetragonal crystalline arrays of AQP0. This arrangement is completely unique compared to the arrangement on the thick junctions or "gap junction proteins" (16-17 nm), where the arrays of channels on the plasma membrane of the adjacent cells join across the intercellular space to provide direct cell-cell communication. Interactions between AQP0 containing proteoliposomes have been investigated by fluorescent resonance energy transfer (FRET) and turbidity measurements [24]. No evidence was obtained for direct AQP0-AQP0 interactions, whereas aggregation was observed upon mixing the proteoliposomes with phosphatidylserine liposomes. This indicates that *in vivo* AQP0 might interact with the lipids of the juxtaposed cell membrane which is consistent with the extraordinarily narrow extracellular spaces. From the above results, direct contact between AQP0 molecules could not be proven. The elimination of extracellular spaces between fibers is considered an essential part of transforming the lens into a series of co-axial refractive surfaces. Thus, AQP0 has been speculated to function as electrostatic glue minimising extracellular space in the lens and facilitates the formation of specialised membrane

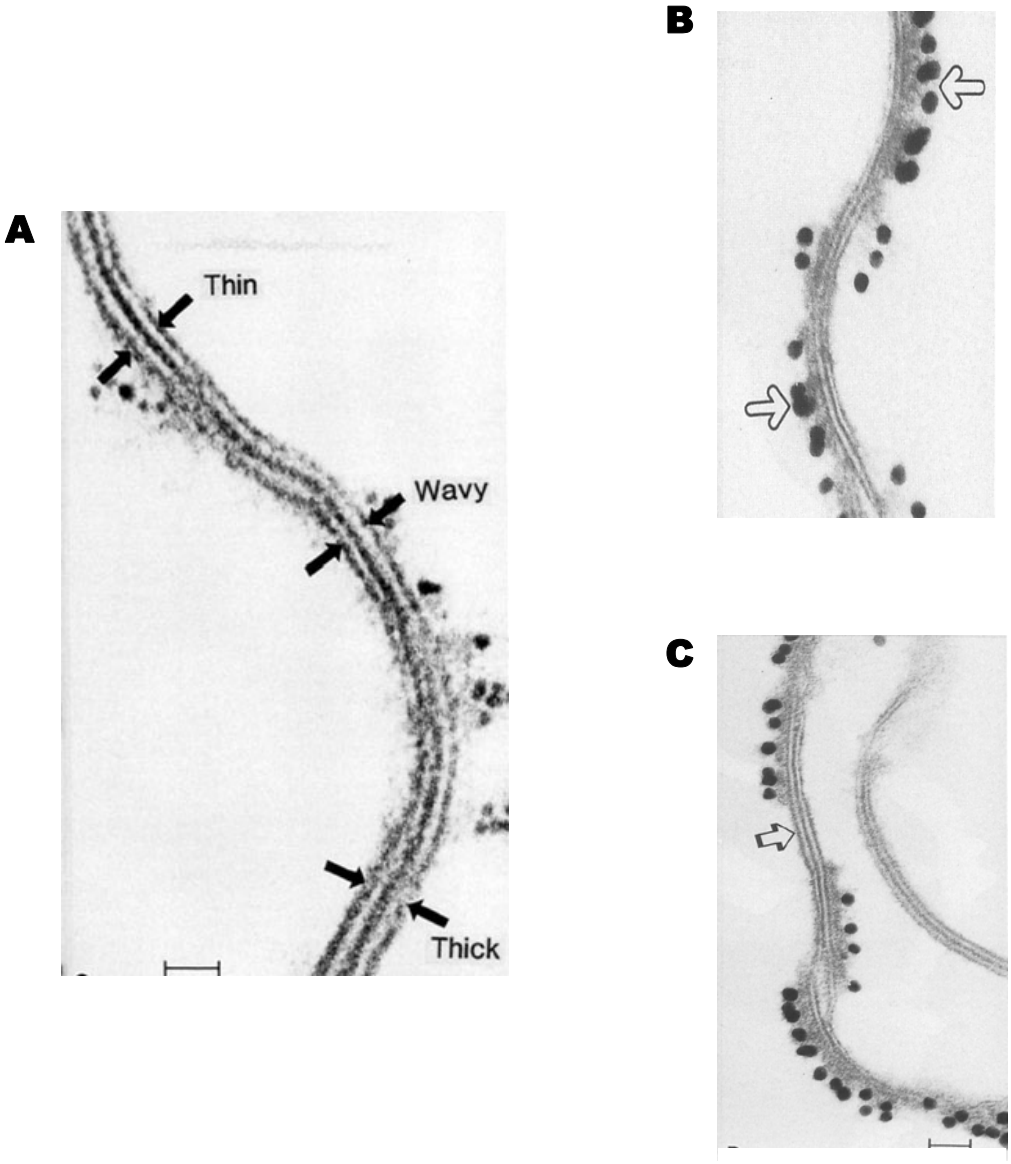


Figure 7: Higher magnification views of three types of lens junctions (A) Shows the presence of wavy, thin and thick-coexisting in the same profile. The concave membrane of wavy junctions appeared substantially thinner than the convex membrane. Scale bar, 20 nm. (B) and (C) Junctions isolated from calf lenses labeled with the polyclonal anti-AQP0 and visualized with protein A-gold (15-nm particle diameter). These views demonstrate specific labeling on the cytoplasm surface of wavy junctions, the lack of staining of thick junctions, and the labeling of single membranes. When the curvature of the wave reverses, the location of the gold particles also reverses to become associated with the convex membrane of the opposing membrane (B, arrows). At these regions where the curvature reverses, short junctional segment display gold particles on both junctional membranes. Scale bars: (B) 0.1 μm; and (C) 50 nm (adapted from Zampighi et al., 1989) [10].

junctions between the lens cells. This interaction most probably represents the *in vivo* aggregation mode between AQP0 tetramers from juxtaposed membranes in the eye lens to mediate cell-cell interaction (Figure 8).

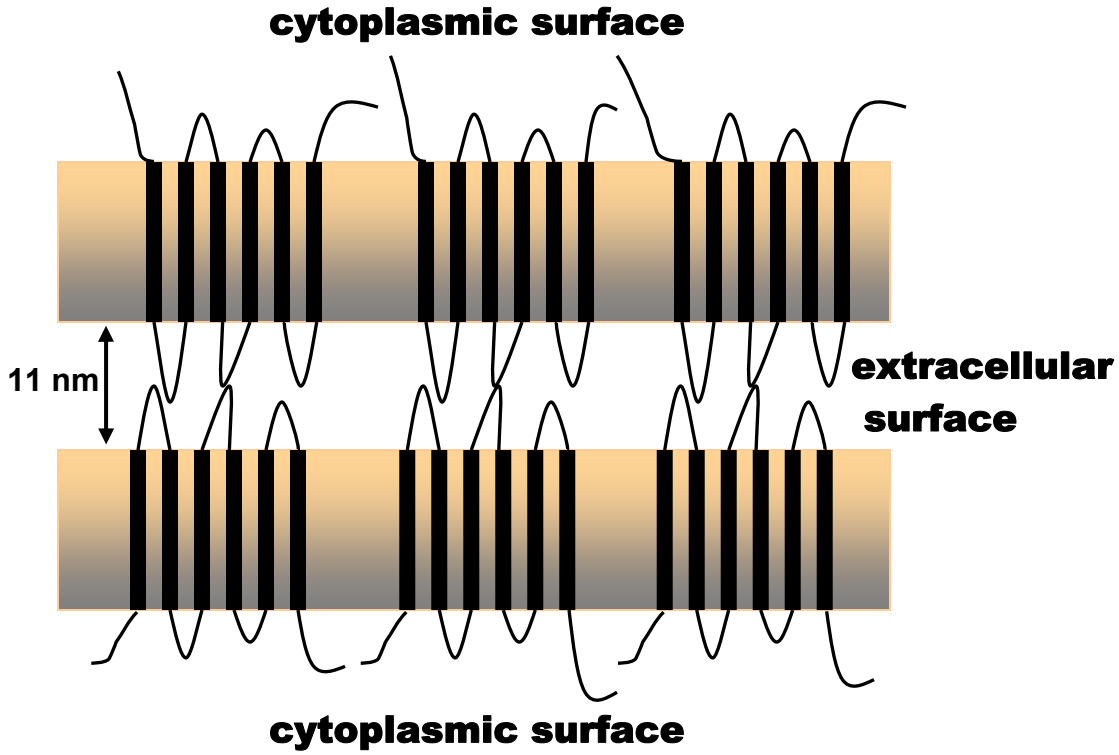


Figure 8: The schematic view of AQP0 tetramers localized in the membrane junctions of lens fiber cells with the narrow extracellular space of 11 nm. *In vivo* aggregation mode of AQP0 tetramer is shown mediating the putative cell-to-cell adhesion mechanism.

Knock-out studies have been conducted for this protein in Cat^{Fr} (cataract Fraser) mice which cause cataract formation (Figure 9) [25]. Optical dysfunction of the crystalline lens in AQP0-deficient mice has shown that, the heterolzygous loss of AQP0 is sufficient to trigger cataractogenesis [26]. Mutations in AQP0 that prevent proper folding or targeting cause lens opacification, hence supporting a crucial role of AQP0 in the development of the transparent lens [27]. Thus, AQP0-deficiency results in an increased extracellular space between fiber cells, leading to a loss of proper fiber cell shape, size, and order necessary for the optical quality of the lens (Figure 9).

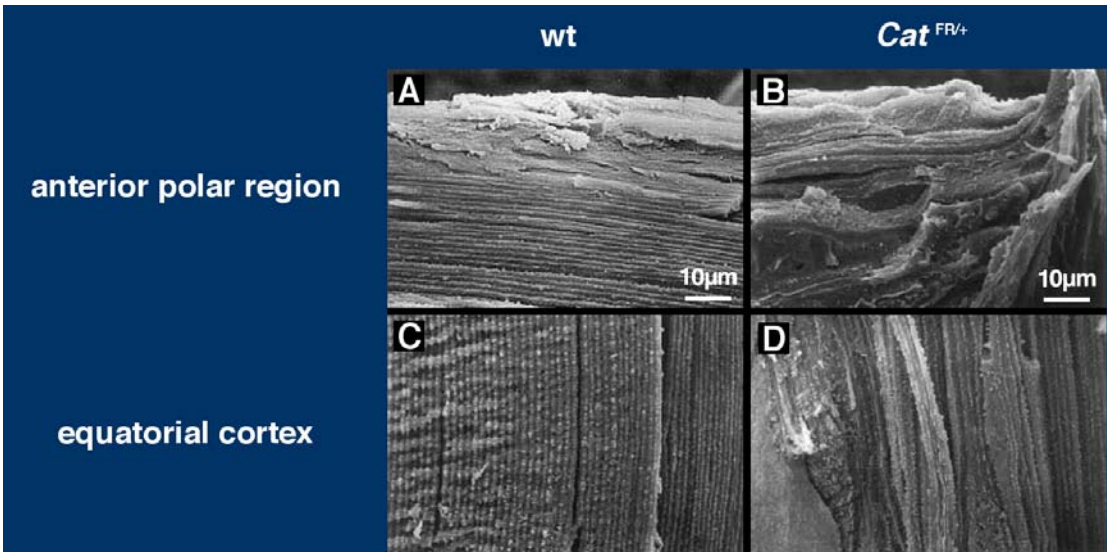


Figure 9: Scanning electron microscopy of fiber cell pathology in the heterozygous Cat^{Fr} lens at three weeks of age. (A) Anterior polar region of wild-type lens and (B) $Cat^{Fr/+}$ lens showing loss of uniform fiber cell stratification in the mutant. (C) Equatorial cortex of wild-type lens and (D) $Cat^{Fr/+}$ lens showing irregular swollen fibers in the mutant. The figure is adapted from Shiels et al., 2000 [25].

These molecular details have shown the importance of AQP0 for maintaining the structural integrity within the lens fiber cells. The first structural study for AQP0 was carried out using atomic force microscopy (AFM) and cryo-electron crystallography which unveiled distinct peptide domain structures at subnanometer resolution. Engel and co-workers studied the AFM topographs of AQP0 obtained from bilayered 2D crystal sheets. The orientation of the AQP0 molecule was detected by specific proteolysis and antibody detection. The model demonstrates that the AQP0 tetramers from two layers were found to be present in precise register facing to each other. They interact with each other extracellularly by forming a “tongue-and-groove” contact (Figure 10) [28]. This proposed model demonstrates that AQP0 tetramers meet each other in the juxtaposed membrane surface by means of loop interactions. Later, the high resolution structures of

AQP0 from bovine (X-ray) and sheep eye lenses (EM) were solved. Both the structures are similar to each other, but the latter shows the tetramer association.

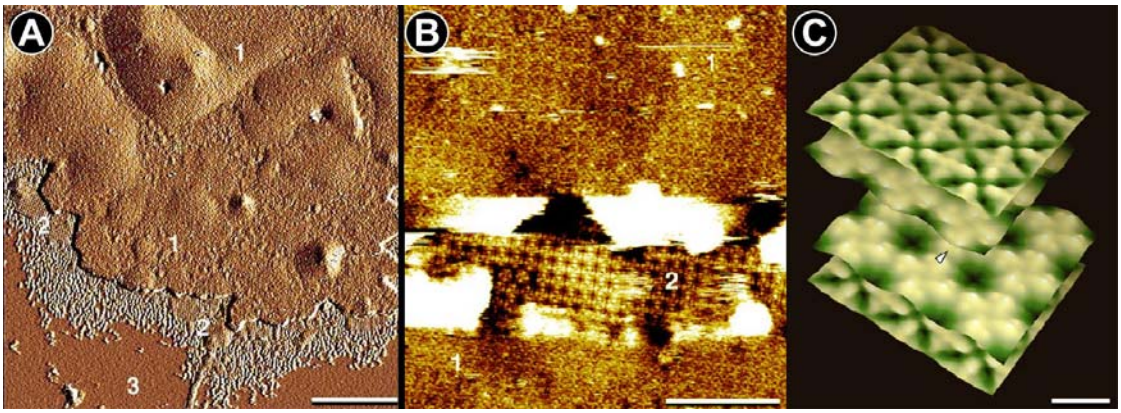


Figure 10: Surface tongue-and-groove contours on lens AQP0 facilitate cell-to-cell adhesion visualized by AFM. (A) Reconstituted double-layered sheets of AQP0 showing three different regions: (1) a "smooth" crystalline, (2) a "rough" crystalline and (3) a surface composed of lipids only. Scale bar 250 nm. (B) By partial removal of the upper layer using the AFM tip both layers could be visualized and compared. Scale bar 50 nm. (C) Computer reconstructed model of the double-layered AQP0 sheet showing the "tongue-and-groove" interactions between extracellular surfaces. Scale bar 5 nm. The figure adapted from Fotiadis et al., 2000 [28].

The EM structure has shown that AQP0-AQP0 can mediate contacts determined from the double-layered 2D crystal which are formed by tight head-to-head association of tetramers coming from the two layers. These specific interactions in the extracellular space have been proposed to resemble the situation in membranes of lens fiber cells. Here, we report a similar, though not identical, association of the tetramers obtained from recombinant bovine AQP0 in a loosely packed 3D crystal. Overexpression and purification of large amounts of pure and homogenous AQP0 was achieved for obtaining 3D crystals and structure analysis was performed up to a resolution of 7 Å. Insights into the structural features of AQP0 is discussed in comparison to the known structures of AQP0 from the EM and X-ray studies. A direct comparison of the arrangement of AQP0 tetramers in the 3D crystal lattice with that in the 2D crystal lattice yielded several

distinct features of the oligomeric state of the protein to resemble that in the native eye lenses.

2.0 Materials and Methods

2.1 Clones

Recombinant Bovine AQP0 was cloned in pYES2.0 yeast vector with Gal6 promotor and transformed in *Sacchromyces cereveciae* strain Sc334 with fusion at the N-terminus hexahistidine affinity tag and a Factor-Xa cleavage site (Kozono D., Johns Hopkins University, U.S.A.). The C-terminal construct was designed to attach the hexahistidine tag and the Factor-Xa cleavage site at the c-terminus region of the protein sequence. All the constructs were obtained from Prof. Peter Agre's lab (Johns Hopkins University, U.S.A.) for the structural studies.

2.2 Overexpression

The yeast strain Sc334 (a GAL4 GAL80 ure3-52 lue2-3 112 reg1-501 gal1 pep4-3) (Hovalnd et al., 1989) harboring the plasmids MTL4C and pXU607 was grown in SD – URA media (0.67 % Yeast Nitrogen Base, 2 % Glucose, 0.5 % CASA amino acid and 40 μg/mL (Adenine, histidine and tryptophan)) (6 x 100 mL) lacking uracil and leucine for 3 days at 30 ^{o}C. When the cell density reaches 1.6 O.D. at OD_{600} nm, the induction was performed with YPGal media (1 % Yeast extract, 2 % Bacto peptone and 2 % Galactose) (6 x 1.25 L) for an overnight expression at 30 ^{o}C. The cells were then harvested and washed in 1/ 25 culture volume of breaking buffer (100 mM K-phosphate buffer pH 8.0, 200 mM NaCl, 10 % glycerol, 5 mM β-mercaptoethanol, 5mM EDTA and 1mM phenyl methyl-sulphonyl fluoride (PMSF)). This yielded 70 g of cells.

2.3 Purification

Cells were resuspended in the breaking buffer and they were broken by a Dyanomil using glass beads for 3 X 2 minutes at 4 ^{o}C. The cell lysate was clarified by centrifugation at 9000 rpm (Sorval GSA) and the cell debris and the supernatant were separated. For membrane protein extraction, the supernatant was further subjected to ultracentrifugation step at 40,000 rpm (TFT45.94) at 4 ^{o}C for 2 hours (Kontron Instruments, Switzerland).

The membrane fractions were resuspended and homogenized in solubilisation buffer without detergent (SoliB) (100 mM K-phosphate buffer pH 8.0, 300 mM NaCl, 10 % glycerol, 5 mM β-mercaptoethanol, 5mM EDTA and 1mM PMSF) and stored at –70 ^{o}C. AQP0 was extracted from membrane fractions by homogenizing in presence of 2% n-decyl-β-D-maltoside (DM) in SoliB buffer at 4 ^{o}C for 1 hour. After solubilization, they were subjected to Ni-NTA affinity column which was preequilibrated with SoliB buffer containing 2 % DM and 1 mM histidine. Addition of histidine was to prevent the non-specific binding of the contaminants to the nickel column. The binding was carried out for 1 hour at 4 ^{o}C. The Ni-NTA beeds were collected in wizard Midi Columns (Catalys AG, Switzerland) for batch method purification. The column was washed with the washing buffer (WB) (20 mM Tris buffer pH 8.0, 300 mM NaCl, 10 % glycerol, 7 mM β-mercaptoethanol, 2 mM histidine and 0.4 % DM) and the protein was eluted at 100 mM histidine. The purified fractions were then concentrated to around 10 mg/mL using an Amicon Ultra device with a cut-off of 100 kDa (Millipore AG, Volketswil, Switzerland). During this step, the detergent was getting concentrated (from 0.4% to about 6–7 % for DM) along with the protein as measured by a home made device aimed to measure detergent concentration (Kaufmann et al., manuscript in preparation). This was overcome by resuspending the solution (starting volume 4 mL) 4 times every 3 minutes with 1 mL buffer solution resulting in a final detergent concentration below 1% DM.

2.4 Polyacrylamide gel electrophoresis, Silver staining and Western blot

Electrophoresis was performed using 15 % SDS-PAGE gels, according to the method of Laemmli. Protein bands were visualized with Coomassie Brilliant Blue staining and the molecular weight determination was done by referring to LMW-SDS markers as standard (Prestained SDS-PAGE Standards, Bio-Rad Laboratories, U.S.A.).

Silver stain analysis was performed using the procedure from Morrisey et al (Morrisey et al., 1981, Anal. Biochem). The SDS-gel was placed in a fixer solution (50 % methanol + 0.1 mL 37 % formaldehyde per 100 mL) for 1 hour at room temprature. The obtained gel was washed with distilled water containing 1M DTT for 10 minutes. Prepared fresh silver nitrate solution was added to the gel for 10 minutes on a shaker. A quick rinse with water and then added the developer solution (3 % Na_2CO_3 + 100 µL of 37 % formaldehyde per

100 mL). The bands appeared within 2 to 5 minutes and the reaction was stopped by adding a quenching solution (1 % acetic acid.).

For western blot analysis, protein samples were electroblotted onto a nitro-cellulose membranes (BA85, Schleicher & Schuell, Dassel, Germany) using the Bio-Rad Mini-PROTEAN 2 Electrophoresis/Mini Trans-Blot Module. Transfer was conducted in presence of blotting buffer (20 mm Tris, 150 mM glycine and 20 % methanol) at 100 V (expected current: 350 mA) for 1 hour at 4 ^{o}C. The blot was subjected to immunoanalysis which was carried out according to the manufacturer's protocol of ECL western blotting analysis system (Amersham Biosciences, Freiburg, Germany). Briefly, non-specific binding sites on the blotting membrane were blocked with Tris-buffered saline (TTBS) (20 mM Tris-HCl pH 7.6, 150 mM NaCl, 0.1 % Tween) containing 5 % skimmed milk powder. The blot was washed for three times in TTBS buffer, each time for 10 minutes and then incubated in TTBS containing the primary antibody, monoclonal anti-polyhistidine antibody raised in mouse (Sigma Chemie, Buchs, Switzerland), at 1:3000 dilution for 1 hour at room temperature. The blot was washed as described before in TTBS buffer and then incubated in TTBS containing the secondary antibody, anti-mouse IgG (Fc specific) antibody-peroxidase conjugate (Sigma), at 1:5000 dilution. The blot was washed for three times with TTBS buffer and incubated in ECL western blotting detection reagents one and two (Amersham Biosciences, Freiburg, Germany) mixed at the ratio of 1:1 for one minute. The blot was subsequently exposed on KodakX-OMAT XAR-5 radiography film for 15 seconds to 15 minutes until the signal appeared.

2.5 Absorption spectroscopy

Protein concentrations were determined by absorption spectroscopy (Hewlett Packard 8453, HP845X UV-Visible system, Switzerland) at 280 nm. The theoretical extinction coefficient ε at 280 nm was calculated using the programme. ProtParam from the ExPASy server. For AQP0, ε was found to be 38780.

2.6 Single particle analysis

The protein solution was diluted to 10 µg/mL with 10 mM Tris/HCl pH 7.5 with 0.04% DM to the grids. The carbon-coated grids are glow-discharged for 10 seconds and loaded

with 3 µl protein sample. After an incubation of 10-20 seconds the grid is blotted with a drying paper. Then the grid was washed three times and stained on top with 6 µl of negative stain (0.75 % uranyl formiate) for 10 seconds. Micrographs were scanned (20 µm/pixel) using a Joyce-Lobel microdensitometer. The recorded micrographs are digitized and analysed by a single particle tools such as Spider (System for processing image data in electron microscopy and related fields) in combination with Web (graphical user interface for displaying and interacting spider images). For a typical two-dimensional analysis the picked particles were aligned by shifting and rotating algorithms that can be combined with references and low-high pass filters. Multivariate statistical analysis (MSA) and subsequent automated hierarchical classification schemes were applied for further analysis on the single particle measurements.

2.7 Analytical ultracentrifugation

Molecular masses were measured for the protein at 20 ^{o}C using a Beckman Optima XL-A analytical ultracentrifuge. The protein was tested at different concentration of 0.3, 0.4 , 0.5 and 2.0 mg/mL. Velocity data were collected at 54,000 rpm for checking the monodispersity. Subsequently, sedimentation equilibrium experiments were carried out at different angular velocities (13000, 17000, and 24000 rpm; protein concentration 1.8 mg/mL) to investigate the self-association process for AQP0. The relative extinction coefficient (at 277 nm) was recorded during the run along the radius of the cell (6.65 - 7.05 cm). Density of the solvent was adjusted with sucrose solution (final concentration; 28 %) to match the density of the hydrated micelle [29]. Density of the solvent was measured using DMA02C digital density-meer (A. Paar AG, Austria), and was 1.128 g/mL and the partial specific volume of the solute was 0.757 cm^3/g. Fitting was performed with non-linear regression using Levenberg-Marquart algorithm in the program- PROFIT (version 6.0). Three different models were used for analyzing the data from different speeds.

Ctotal = c1*Exp(b) + e --> 1-statemodel

Ctotal = c1*Exp(b) + c1^8*ka*Exp(8*b) + e --> 2-state model

Ctotal = c1*Exp(b) + c1^4*ka1*Exp(4*b) + c1^8*ka2*Exp(8*b) + e --> 3-state model

Ctotal - concentration at radius r [cm]

c1 - conc. of the monomer at the reference radius r0 [M]

ka, ka1and ka2 - dissociation constant ($[M^{-7}]$, $[M^{-1}]$ and $[M^{-6}]$ respectively)

b - (2*3.14*omega^2/60)^2*M*(1-ῦρ)*((r2/100)^2-6.75e-2^2)/(2*R*T)

M - monomer molecular weight [kDa]

ω - angular velocity $[S^{-1}]$

ῦ - partial specific volume of the solute $[cm^3/g]$

ρ - density of the solvent [g/mL]

R - gas constant [8.314 J/mol.K]

T - Temperature in [K]

e - base line

All analytical ultracentrifugation experiments were performed by Ariel Lustig, Biozentrum, Univeristy of Basel.

2.8 Blue-Native gel electrophoresis (BN-PAGE)

The technique of blue native gel electrophoresis was developed by Schagger H. and von Jagow G. in 1991 [30], and is a powerful application for the separation of native membrane protein complexes. The separation was performed on a gradient gel (5 – 13 %), in the presence of blue cathode buffer (15mM Bis-Tris / HCl pH 7.0, 50 mM Tricine and 0.005 % Commassie Brilliant Blue G-250) and anode buffer (50 mM Bis-Tris / HCl pH 7.0). Special standard markers (thyroglobulin 660 kDa, ferritin 440 kDa, catalase 320 kDa, lactate dehydrogenase 150 kDa and albumin with 66 kDa) were used for identification. The experiment was conducted overnight at 100 V.

2.9 Glycoprotein detection system

Electrophoresis was performed for AQP0 samples. AQP1 (glycosylated and deglycosylated samples) and transferrin were used as a positive control. Then, the gel was transferred to a PVDF membrane by western blot method (See section 2.4). The blot was then incubated for 10 minutes in phosphate buffered saline (PBS). Then it was subjected to oxidation step using 10 mM sodium metaperiodate dissolved in 10mM

acetate buffer pH 5.5 in dark for 20 minutes. The blot was washed for five times in PBS buffer, each time for 10 minutes and then incubated for one hour in 100 mM acetate buffer pH 5.5 containing 0.125 mM biotin hydrazide. The blot was washed as described before in PBS buffer and then incubated for one hour in PBS containing 5 % membrane blocking agent (milk). Washing step was performed as described before and then incubated for 30 minutes in PBS containing the streptavidin (1:6000). The blot was washed for three times with PBS buffer and incubated in ECL western blotting detection reagents one and two (Amersham Biosciences, Freiburg, Germany) mixed at the ratio of 1:1 for one minute. The blot was subsequently exposed on KodakX-OMAT XAR-5 radiography film for 15 seconds to 15 minutes until the signal appeared.

2.10 Bio-beads treatment

Bio-beads SM2 (Bio-Rad) are neutral hydrophobic macroporous polystyrene beads with a high surface area for adsorbing organics, including detergents from aqueous solutions. These beads are capable to bind huge amount of detergents relatively independent of the CMC-value. Biobeeds were mixed with the protein sample (1:10) and incubated overnight at 4 ^{o}C. The biobeeds were separated by filtering and the detergent concentration of the sample was measured with a home made device aimed to measure detergent concentration.

2.11 His-tag cleavage

Purified AQP0 from Ni-NTA column was subjected to buffer exchange by performing dialysis with TAGZyme Buffer (20 mM bis-Tris pH 7.0, 300 mM NaCl, 10 % glycerol, 5 mM β-mercaptoethanol, 0.4 % DM) using Spectra-Por 7 dialysis membrane of cut-off 12–14 kDa at 4 ^{o}C overnight. Digestion was performed with TAGZyme (DAPase : diamino peptidase) (Qiagen AG, Switzerland) (100 mU, twice the concentration mentioned by the manufacture) for 1 hour at 4 ^{o}C with gentle agitation. Separation of the cleaved AQP0 was carried out using 2nd Ni-NTA column. TAGZyme was efficiently removed during the 2nd nickel column, since it possess his-tag at the C-terminal end. Cleaved AQP0 was concentrated to 10 mg/mL as descried before.

2.12 Crystallization

Crystals were obtained at room temperature by vapour diffusion using the sitting drop method. Crystals were grown by sitting drop method at 20 ºC under similar conditions as found earlier for native AQP0. Equal volumes (3 + 3 μL) of protein solution (8.5 mg/ml; assuming an ε_{280} of 38780 M^{-1} cm^{-1}) in Tris 10 mM pH 8.0, 300 mM NaCl, 10% (v/v) glycerol, 5 mM β-mercaptoethanol, 0.4 % DM) and reservoir solution (100 mM bicine NaOH pH 8.9, 350 mM NaCl, 34% (v/v) PEG400) were mixed.

2.13 Data collection

Data collection was performed without any cryoprotectant (mother liquor contains 34 % PEG400). The crystals were flash frozen in liquid nitrogen. Diffraction data were collected from a single native crystal (about 150 μm in diameter). All the measurements were performed at the beamline X06SA (PX) at the Swiss Light Source (SLS), Villigen, Switzerland.

2.14 Data processing and structural elucidation

Diffraction data were processed with **IPMOSFLM (version 6.2.4) / SCALA (version 5.0)** [31]. **Self rotation** functions were calculated using programme **POLARRFN (version 5.0)** to validate the spacegroup. Molecular replacement was performed using **AMoRe (version 5.0)**. The scheme for molecular replacement is summarized in Figure 22. The model was built using interactive graphics in the programme '**O**' (version 8.0) and refined by **REFMAC (version 5.0)**. The figures were prepared using the programme **DINO** (*http://www.dino3d.org*).

2.15 Bioinformatic analysis

Pairwise sequence alignment for AQP0 was performed using ClustalW, a multiple sequence alignment programme [32]. The secondary structures on top of the aligned sequence were displayed using ESPript [33].

3.0 RESULTS

3.1 AQP0 constructs

Bovine AQP0 was heterologously expressed in *Saccharomyces cerevisiae* strain. We worked on two different constructs of AQP0 (N-His_{10}-AQP0 (N-AQP0) and AQP0-His_6-C (AQP0-C)) for the structural analysis. We investigated expression properties of both the constructs individually. For the removal of the His-tag, Factor-Xa protease cleavage site was inserted between the His-tag and the protein sequence (Figure 11). All the constructs were provided by Prof. Peter Agre's laboratory (Kozono D., Johns Hopkins University).

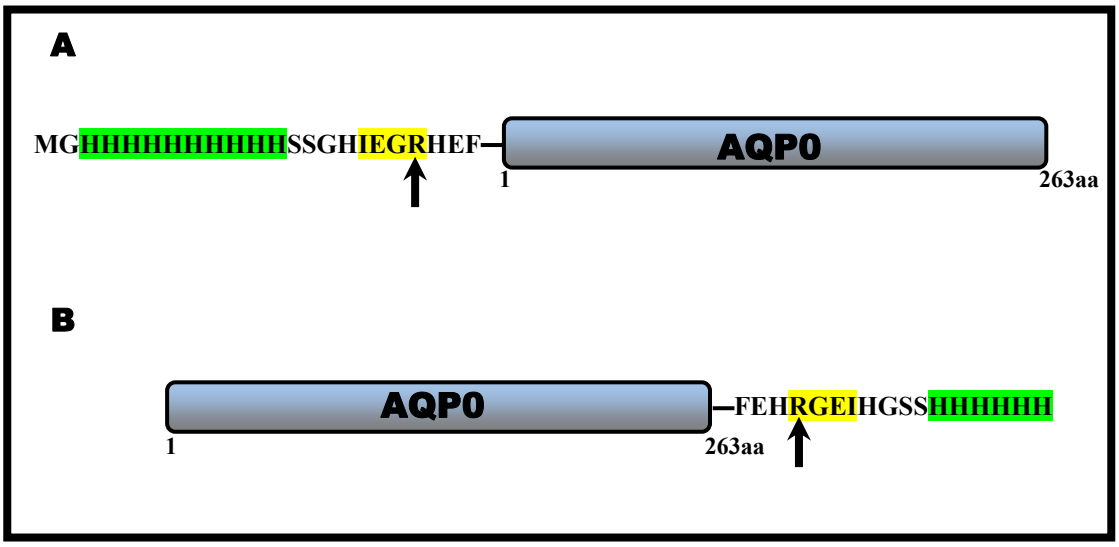


Figure 11: Design of N-His_{10}-AQP0 (N-AQP0) (A) and AQP0-His_6-C (C-AQP0) (B) constructs; the hexahistidine tag is marked in green and the Factor Xa cleavage site (arrow) with the recognition sequence is marked in yellow. The constructs were designed by Prof. Peter Agre's laboratory (Johns Hopkins University, U.S.A.).

3.2 Expression studies on N-AQP0 and AQP0-C constructs

A small scale culture was tested for the expression levels of N-AQP0 and AQP0-C constructs. The isolated membrane fractions were subjected to the Western blotting analysis for estimating the yield of the protein. The result shows that N-AQP0 was highly expressed in comparison to the AQP0-C as observed from the blot (Figure 12). Since the yield from N-AQP0 was high, we worked extensively on this construct for structural analysis. We performed a large scale expression for N-AQP0 construct with 9 L culture size. The overall yield was estimated to be 20 mg of protein. Since all our experiments were carried out with N-AQP0, now it is referred to as AQP0.

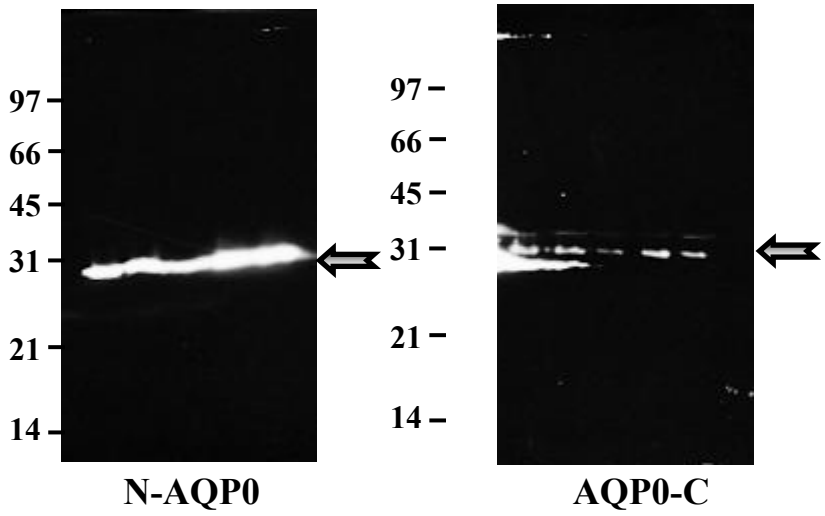


Figure 12: The Western blot analysis of N-AQP0 and C-AQP0 expression in *Sacchromyces cereveciae*. Expression levels of both the constructs are indicated (arrows).

3.3 Purification of AQP0

AQP0 was purified in a single chromatography step using Ni-NTA (IMAC) affinity chromatography column. Initially, AQP0 was extracted from membrane fractions using 10% n-octyl-β-D-glucoside (OG). Subsequently, several detergents were tested to yield monodisperse protein preparation as judged by analytical ultracentrifugation, blue native gel electrophoresis and single particle analysis. The results obtained from these methods are shown separately.

3.3.1 Detergent studies

After the extraction of AQP0 from the membrane fraction, detergent exchange step was performed on the Ni-NTA columns against n-decyl-β-D-maltoside (DM: 0.3 %; 3 times the CMC), n-dodecyl-β-D-maltoside (DDM: 0.15 %; 15 times the CMC), n-nonyl-β-D-glucoside (NG: 1 %; 5 times the CMC), n-nonyl-β-D-maltoside (NM : 1 %: 15 times the CMC) or n-decyl-β-D-glucoside (DG: 1 %: 10 times the CMC). AQP0 in DM (0.3 %) or DDM (0.15 %) was found only partially aggregated whereas severe aggregation was observed with the other tested detergents (Table 2). Further improvement, i.e. monodispersity, was achieved by employing DM or DDM for extraction (at a concentration of 2 %) as well as purification (0.4 % or 0.15 % respectively). SDS-PAGE and silver stain analysis was performed for estimating the purity of the protein. The Ni-NTA fractions showed 95 % purity as observed from SDS-gel (coomassie stain) and silver stained gel (Figure 13A and B). Upon concentrating the protein, the detergent also was getting concentrated in the concentrating unit (from 0.4 % to about 6 – 7 % for DM), although a large cut-off membrane (100 kDa cut-off) was used. In order to remove the excess detergent molecules, we performed bio-beeds treatment. These beads are capable to bind huge amount of detergents relatively independent of the CMC-value, because they do not only absorb free monomers but also micellar detergent structures. Bio-beeds removed the excess of detergent from the concentrated sample by ten folds to the starting concentration. Finally, we solved the detergent concentration problem by resuspending the protein solution 4 times every three minutes with one mL buffer solution during the concentration step.

AQP0 with different detergents	Not Aggregated	Partially Aggregated	Aggregated
Extraction with10 % O.G. and exchanged to 0.3% DM			
Before concentration	-	yes	-
After concentration	-	-	yes
Extraction with10 % O.G. and exchanged to 0.15% DDM			
Before concentration	-	yes	-
After concentration	-	yes	-
Extraction with10 % O.G. and exchanged to 1% NG			
Before concentration	-	-	yes
After concentration	-	-	yes
Extraction with 2% DDM and exchanged to 0.15% DDM			
Before concentration	yes	-	-
After concentration	yes	-	-
Extraction with 2 % DM and exchanged to 0.3% DM			
Before concentration	yes	-	-
After concentration	yes	-	-
Extraction with 2% DDM and exchanged to 0.3% DM			
Before concentration	yes	-	-
After concentration	yes	-	-

‘-‘ : not applicable

Table 2: The effect of different detergents on solubilization studies of AQP0. The aggregation state of the protein before and after concentration step is shown.

This resulted in a final detergent concentration below 1 % DM (10 times the CMC). AQP0 was concentrated up to 8.5 mg / mL for several crystallization trials.

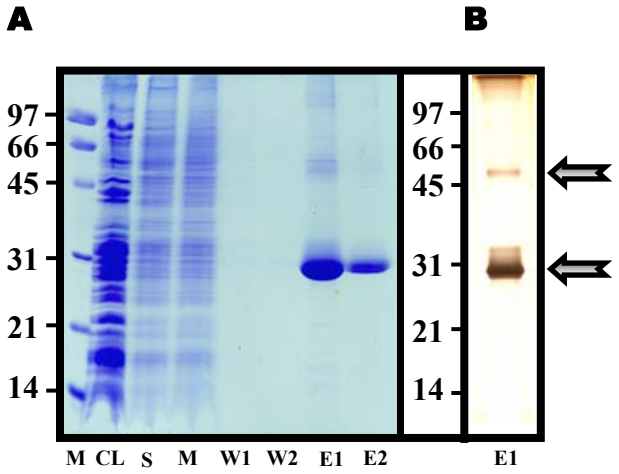


Figure 13: SDS-PAGE and silver stain analysis of solubilization and purification of AQP0 from yeast cells. (A) Different steps involved in purifying AQP0 using Ni-NTA affinity chromatography column. (B) Ni-NTA fractions were subjected to silver stain procedure for checking the purity of the protein.

3.4 Biophysical characterization

In order to obtain homogenous AQP0 for crystallization experiments, we performed several biophysical characterization experiments.

3.4.1 Single particle analysis

Electron micrographs obtained from single particle analysis were analyzed for monodispersity of AQP0 in DM (Figure 14). The images contained small particles and large particles exhibiting irregular outlines that are, most probably, formed by aggregation of the smaller particles. Some of the smaller particles appeared as squares, a shape suggesting that AQP0 is organized as tetramers. The results indicate the presence of monodispersed tetramers with the possible occurrence of coaxial octamers in solution (arrows in Figure 14).

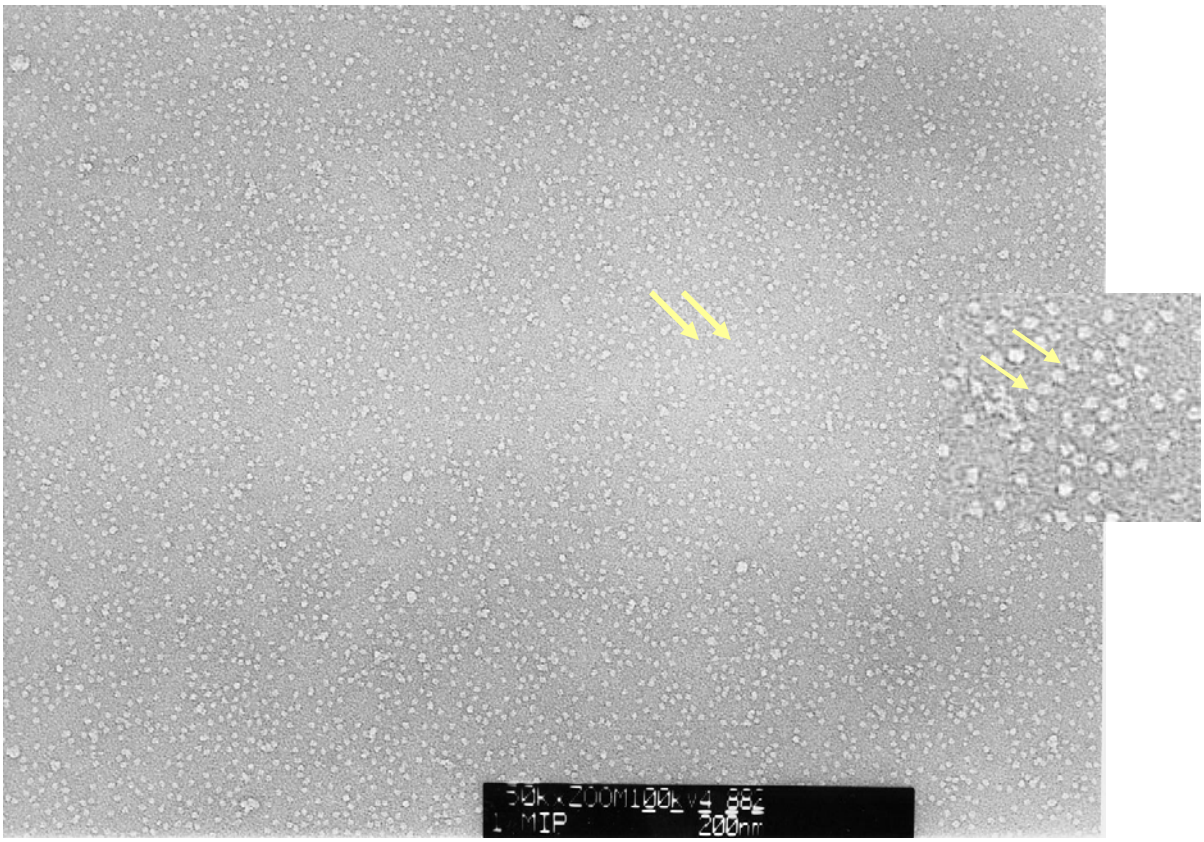


Figure 14: Electron micrographs of purified AQP0. AQP0 in negative stain preparation appears as a rather uniform population of particles. Arrows (large) indicate a particle with a square appearance indicating the presence of tetramers. The zoomed version is shown as an inset showing the possible appearance of coaxial octamers (small arrows).

3.4.2 Blue Native-PAGE electrophoresis

Purified AQP0 was tested to study the oligomeric state using blue native gel electrophoresis (BN-PAGE). BN-PAGE revealed a major band with an apparent mass between 160 and 180 kDa (protein and detergent mass) which probably corresponds to the tetramer form of AQP0. In addition, a strong band around 400 kDa (protein and detergent mass) was observed which could correspond to the octamer form of AQP0 (Figure 15). The result showed that AQP0 exists as two different species in solution.

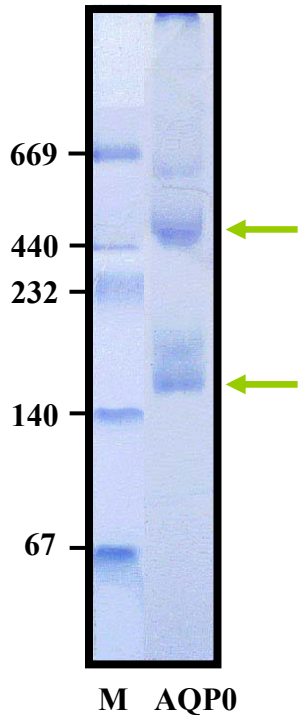


Figure 15: Blue Native-PAGE (5 – 13 %) analysis of purified AQP0 samples. Separation of two forms of AQP0 is indicated by arrows.

3.4.3 Analytical ultracentrifugation

Analytical ultracentifugation was used to assess the homogeneity and oligomeric state of solubilized AQP0. To obtain monodisperse AQP0, several detergents (DM, DDM, NM, NG and DG) were tested for solubilization of the protein. Sedimentation velocity (SV)

method was used to judge the solubilization property for each of this detergent upon detergent exchange. The results showed that solubilization with OG followed by detergent exchange to DM or DDM yielded partially aggregated protein (Figure 16B and 16C). Incase of detergent exchange with NM, NG and DG yielded polydisperse samples has shown in the Figure 16A. Finally, monodisperse AQP0 was obtained using the same detergent (DM and DDM) for extraction and purification steps (Figure 16D and 16E). SV measurements for monodisperse AQP0 was performed with protein concentrations of 0.3 to 2 mg/mL. The result showed the presence of one main species with an observed sedimentation coefficient between 4.3 and 4.6 S in detergent/buffer solution. This most probably corresponds to the tetramer, since from these values an upper mass boundary of approximately 110 kDa can be calculated for the protein-detergent complex assuming spherical shape and water viscosity [34]. A marginal contribution of a smaller species (probably monomers) was apparent at the lower protein concentrations. Sedimentation equilibrium measurements were carried out in the presence of 28 % sucrose to match the density of the bound detergent (1.128 g/mL) such that the buoyancy of the mixed micelles was due to the protein component alone [29]. Data from three runs (13000, 17000, and 24000 rpm; protein concentration 1.8 mg/mL) were fitted simultaneously. The low apparent mean mass of m = 52 kDa (chi^2 = 0.075) was consistent with the presence of monomers (M = 28 kDa) in addition to oligomers. Therefore, several self-association models were tested. While the fit to a 1-state model was unsatisfactory (m = 62 kDa, chi^2=0.075), resorting to a tetramer-octamer model did not improve the fit either (m = 50 kDa, chi^2=0.071). Unexpectedly, in the 3-state model (monomer-tetramer-octamer) the association rate for tetramer formation refined to negligible values. Whereas, the 2-state model (monomer-tetramer and monomer-octamer) gave equally good agreements (m = 38 kDa, chi^2 = 0.059 and m = 37 kDa, chi^2 = 0.052, respectively). Therefore, in summary, monomer-tetramer and monomer-octamer self-association models fitted the data equally well, while a 3-state model did not improve the fit. A monomer - tetramer equilibrium has been demonstrated previously by Konig et al. [35]. It appears that the different oligomeric forms of solubilized AQP0, up to octamers, are in an equilibrium that is largely dependent on the experimental conditions.

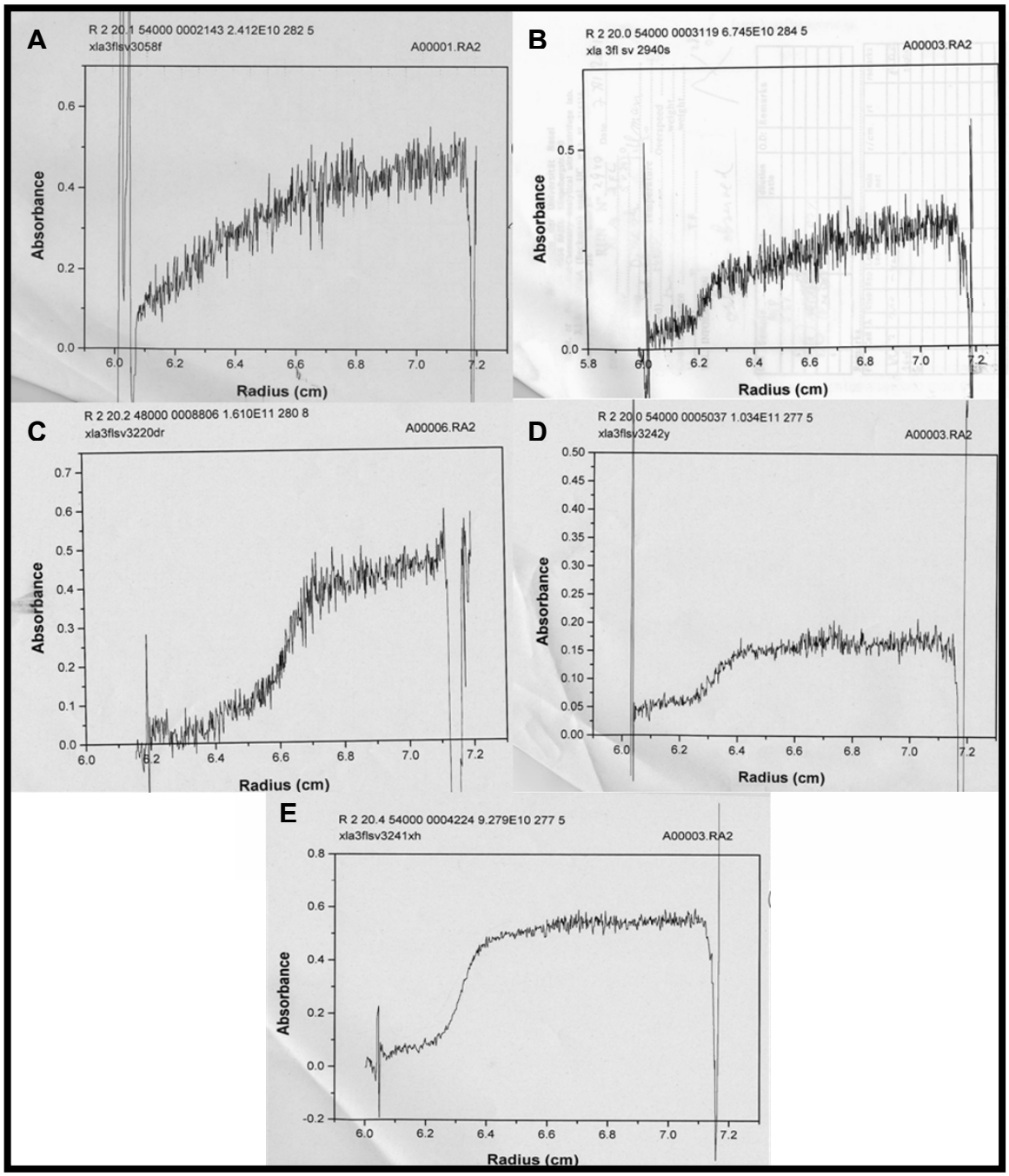


Figure 16: Graphs showing the boundary formation for AQP0 solubilized in different detergents for sedimentation velocity measurements. (A) Aggregation curve formed for AQP0 solubilized in 10 % OG and exchanged with 1 % NG. The same pattern was observed for DG (1 %) and NM (1 %) respectively. (B) and (C) Shows 50 % aggregation of AQP0 solubilized from 10 % OG and exchanged to DDM (0.15 %) and DM (0.3 %) respectively. (D) and (E) No aggregation when DDM and DM was used for solubilization (at a concentration of 2 %) and purification (0.15 % and 0.4% respectively).

3.5 Glycosylation detection

Since glycosylated protein could be a hindrance for crystal formation, we did investigations on the glycosylation of AQP0, since proteins of the aquaporin families are known to get glycosylated. From our observation for AQP0 in the SDS-gel, we didn't find a protein smear which is a characteristic feature for a glycosylated protein. This observation favored that recombinantly expressed AQP0 is not glycosylated. In order to confirm that AQP0 is not glycosylated, we used an ECL glycoprotein detection system. We ran SDS-gel for AQP0 samples along with AQP1 as a control (glycosylated and deglycosylated) and transferrin (control given by the manufacture). The blot image showed a strong signal for transferrin and AQP1 glycosylated sample (positive control). One of the low molecular weight marker, ovaalbumin also showed a positive signal for glycosylation. No signal was shown for AQP0 samples and AQP1 deglycosylated sample (negative control) (Figure 17). This confirms that AQP0 is not glycosylated.

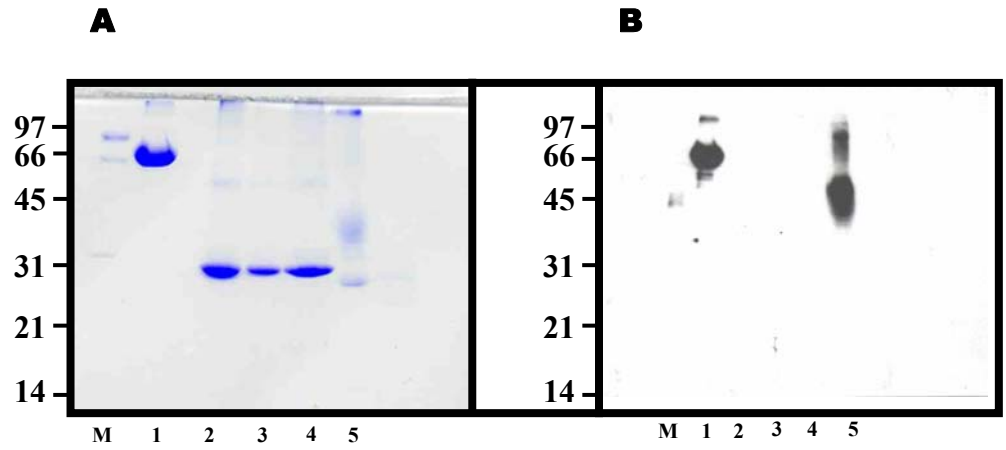


Figure 17: SDS-PAGE analysis of AQP0 for glycoprotein detection. (A) SDS-gel showing the samples ((M): Marker, (1): transferrin (positive control), (2) and (3): AQP0, (4): deglycosylated AQP1 (negative control) and (5) glycosylated AQP1 (positive control)) used for the glycoprotein detection. (B) The corresponding western blot picture shows the signal for the positive controls and no signal for the negative control and AQP0.

3.6 His-tag cleavage

The His-tag cleavage for AQP0 was carried out to improve the chances for crystallization, since presence of the His-tag could be an obstacle for crystal formation. His-tag cleavage of AQP0 was performed using TAGZyme (DAPase). This enzyme is a dipeptidase that cuts every two amino acids from the N-terminus and stops its action at a specific amino acid (DAPase stop point: either K, R, P or Q). This enzyme has a C-terminal His-tag which can be exploited to remove the enzyme during the second Ni-NTA step. In the AQP0 construct, there is an intrinsic stop site (R 20) found in the AQP0 sequence which was identified as the termination region (Figure 11). A time course study was conducted for TAGZyme digestion from 1 to 60 minutes at 4 ^{o}C. SDS-PAGE analysis of the digested samples showed a significant shift of 2 kDa when compared to the control (Figure 18A). Second Ni-NTA column was used to separate the cleaved from the uncleaved and the TAGZyme. Cleaved AQP0 was found in the flow through. The cleavage was confirmed using the Western blotting which showed negative signal for the cleaved material (Figure 18B). The cleaved AQP0 was then concentrated and several crystallization trials were made.

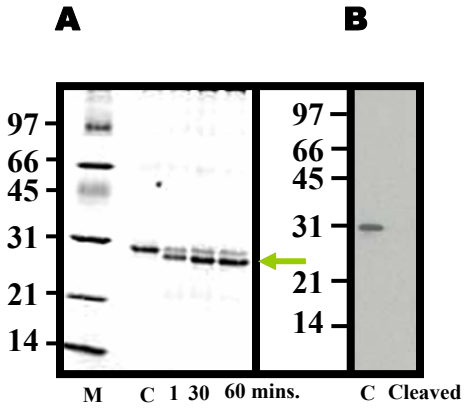


Figure 18: SDS-PAGE (coomassie stain) and silver stain analysis of the TAGZyme digestion for AQP0. (A) A time course study (1 – 60 minutes) was carried out 4 ^{o}C. The gel shift of 2 kDa is indicated by an arrow. (B) Western blot analysis of the cleaved AQP0 obtained from 2nd Ni-NTA column.

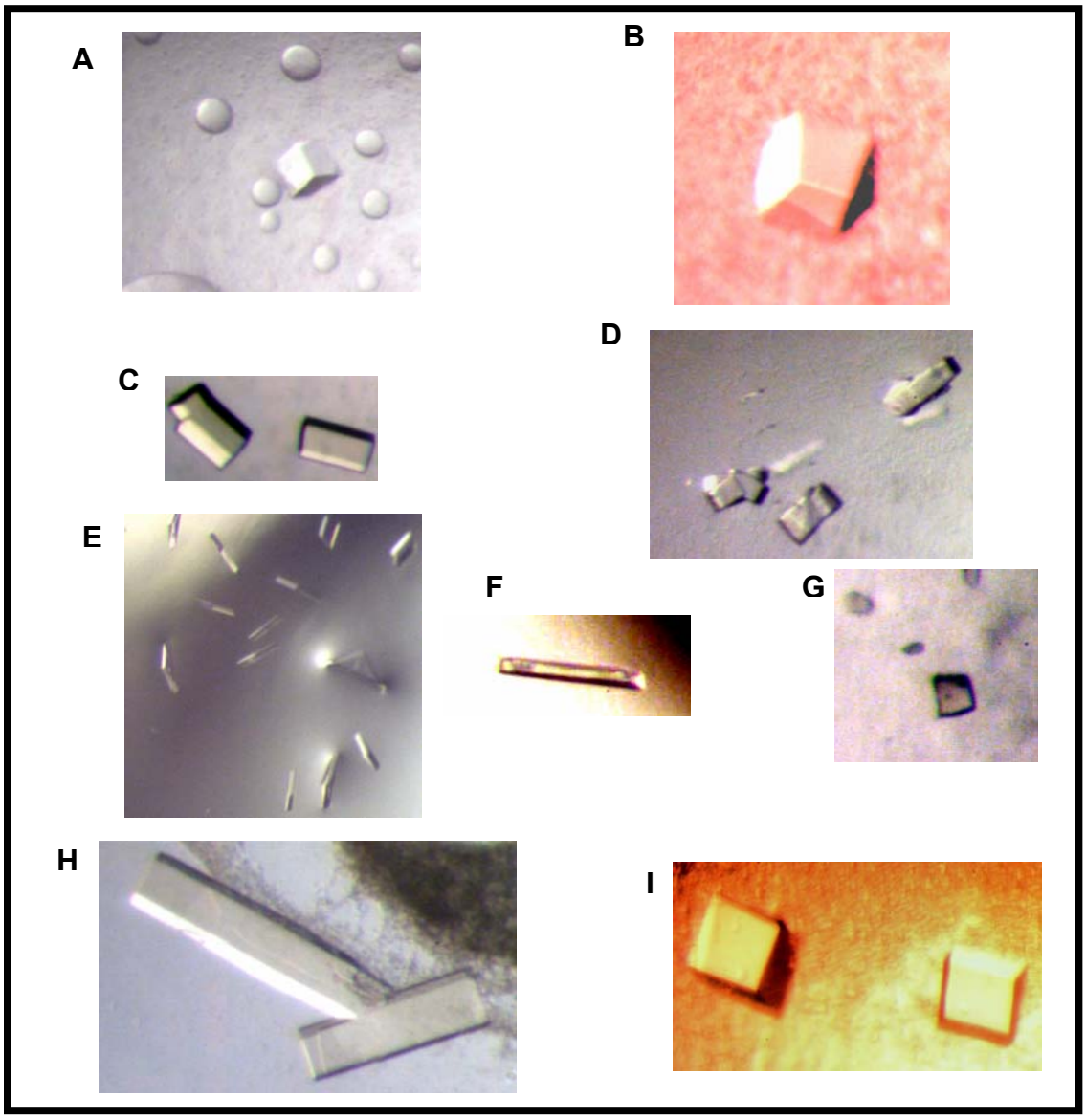


Figure 19: Crystals of AQP0 obtained from different crystallization conditions using sitting drop method at 20 ^{o}C. 3 µL of protein solution was mixed with 3 µL of reservoir solution. The reservoir contained 500 µL. (A) Cubic crystal obtained from 28 % PEG550MME, 550 mM NaCl 100 mM Tris pH 8.5. (B) Cubic crystal obtained from 34 % PEG400, 350 mM NaCl, 100 mM bicine NaOH pH 8.9. (C) Rod shaped crystals from 30% PEG2K, 50 mM Tris pH 8.6, 50 mM Li_2SO_4. (D) Rod or horse shoe shaped morphologies from 30% PEG2K, 50 mM bicine NaOH pH 9.5, 50 mM Na_2SO_4. (E) Arrow shaped crystals from 30% PEG2K, 50 mM bicine NaOH pH 9.5, 100 mM K Na tartarate. (F) Rod or needle shaped crystal obtained from 30% PEG2K, 50 mM Tris pH 8.6, 1 M NaCl. (G) Small cubic crystal of the His-tag cleaved AQP0 obtained from 34 % PEG400, 350 mM NaCl, 100 mM Tris pH 8.3. (H) Bigger crystals where obtained upon optimizing condition (E). (I) Cubic crystals obtained using NG from 38 % PEG400, 350 mM NaCl, 100 mM Tris pH 8.6. All the pictures were taken at the same magnification. Diffraction properties of each crystal form are shown in Table 3.

3.7 Structural analysis of AQP0 by X-ray crystallography

3.7.1 AQP0 crystals

Crystals of AQP0 were grown from sitting drops at 20 ºC under similar conditions as found earlier for native AQP0 [36] and had similar shape. The diameter of the crystals was about 150 μm. They were grown in a period of 3 months to 1 year and had cubic and rod morphologies. They were grown at different detergent conditions such as 0.4 % DM, 0.15 % DDM and 1 % NG. Fine screening of PEG 400 (34 – 46 %) yielded cubic crystals of different size (Figure 19). The crucial parameter for optimization of crystal growth was found to be the drop size and pH (from (1 + 1 μL) to (3 + 3 μL) and pH (8.3 – 9.2) respectively). We also managed to find new conditions for AQP0, by performing a systematic screening on various PEGs/pH/salts. We obtained new crystallization conditions from the above screen which showed different crystal morphologies such as arrow or flakes (Figure 19). Crystals of the His-tagged cleaved AQP0 were also obtained under the above conditions.

3.7.2 Data collection

AQP0 crystals were tested at the synchrotron facility at Swiss Light Source (SLS) in Villigen, Switzerland. The overall diffraction property for each crystal shape is summarized in the Table 3. Among all the crystal form which we obtained, only cubic crystals showed isotropic diffraction. All the other crystal shapes showed very weak diffraction at SLS. A complete dataset was collected at SLS for structural investigation.

3.7.3 Data processing

AQP0 cubic crystals showed strong diffraction to 7 Å resolution (Figure 20). The modest resolution seems to be inherent to the crystal packing, since, despite extensive trials on crystallization conditions, no improvement was achieved. Autoindexing showed that it belongs to the cubic space group *I432* (Table 4). Self rotation was executed to confirm the crystallography 422 symmetry for the authentication of the space group. The cubic arrangement of molecules in the self-rotation projections of point group symmetry 42 at sections kappa 90^{o} and 180^{o} is shown in the Figure 21. Assuming 1 monomer / asymmetric unit (a.u.), the Matthews coefficient V_m was determined to be 5.32 $Å^3$/Da using the following equation

No.	Crystallization condition	Detergent	Crystal shapes	Diffraction limit
1	28 % PEG550MME, 550 mM NaCl	DDM (0.15%)	cubic	17 Å
2	27 % PEG400, 350 mM NaCl 100mM Tris pH 7.7	DM (0.3%)	cubic	Very few spots
3	34 % PEG400, 350 mM NaCl 100mM Tris pH 8.3	DM (0.4%)	cubic	11-12 Å
4	37% PEG400, 350 mM NaCl 100mM Tris pH 8.6	DM (0.4%)	cubic	7.2 Å
5	34 % PEG400, 350 mM NaCl 100 mM Bicine NaOH (pH 8.9 , 9.2)	DM (0.4%)	cubic	7 Å
6	30 % PEG2000, Tris pH 8.6 & Bicine NaOH pH-9.5 - 50mM, 200 mm NH_4SO_4	DM (0.4%)	cubic	12 Å
7	30 % PEG2000, Tris pH 8.6 & Bicine NaOH pH 9.5 - 50mM, 100 mm K Na tartarate	DM (0.4%)	thin flakes (like arrows)	-
8	30 % PEG2000, Tris pH 8.6 & Bicine NaOH pH-9.5 - 50mM, 50 mM Li_2SO_4	DM (0.4%)	rods	No diffraction
9	30 % PEG2000, Tris pH 8.6 & Bicine NaOH pH-9.5 - 50mM, 50 mM Na_2SO_4	DM (0.4%)	rods	16 – 17 Å
10	30 % PEG2000, Tris pH 8.6 & Bicine NaOH pH 9.5 - 50mM 1 M Ammonium formate	DM (0.4%)	cubic	-
11	30 % PEG2000, Tris pH 8.6 & Bicine NaoH pH 9.5 - 50mM, 1M NaCl	DM (0.4%)	Needles	-
12	PEG2000 – 30%, Tris pH 8.6 & Bicine NaoH pH 9.5 - 50mM, 500 mM KCl	DM (0.4%)	Needles	-
13	38% PEG400, NaCl - 350mM 100mM Tris pH 8.6	NG (1%)	cubic	7.5 Å

'-': crystals were not tested in the X-ray beam

Table 3: An overall summary of the different crystallization conditions for AQP0 crystals and their diffraction limits. Data obtained for structure determination were highlighted in green.

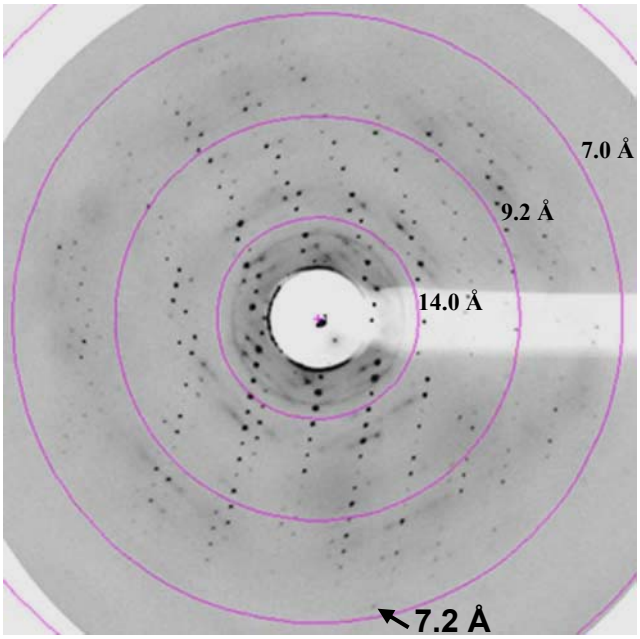


Figure 20: Diffraction pattern of a AQP0 crystal. Data were collected at the Swiss Light Source (SLS), synchrotron source, Villigen, Switzerland. The images were recorded on a MAR-CCD image plate detector (Crystal-detector distance 500 mm, rotation angle 1.0°). The crystal showed isotrophic diffraction, with spots diffracting to 7 Å resolution.

$$\mathbf{V_m} = \frac{\textbf{Volume of unit cell } (\text{Å}^3)}{\textbf{Total weight of protein in the unit cell (Da)}}$$

$$= \frac{\boldsymbol{abc}}{\mathbf{mnZ}}$$

where $\boldsymbol{a} = \boldsymbol{b} = \boldsymbol{c} = 188$ Å; **m** = 26000 Da; the molecular weight of AQP0; **n** is the number of molecules in the asymmetric unit (assuming 1), **Z** is the number of asymmetric units in the unit cell, 48 for *I432*.

The $\mathbf{V_m}$ value is very close to the upper border of the observed range for soluble protein (1.66 – 5.0 Å^3 / Da) [37].

The solvent content was derived by the following equation:

$$\textbf{\% solvent content} = 1 - \frac{1.23}{\mathbf{V_m}}$$

$$= \textbf{76.8 \%}$$

Table 4. Data collection and refinement statistics

Data Collection	
X-ray source; λ	**SLS, Villigen; 0.97970 Å**
Detector	**MAR345**
Temperature	**100 °K**
No. of crystals	**1**
Space group	***I*432**
a* = *b* = *c	**188 Å**
Content of a.u.	**1 monomer**
Resolution range (Å)	**15.0 – 7.0 (7.4 - 7.0)**
Number of unique reflections	**891 (137)**
R_{sym} (%) [a]	**9.4 % (36.3 %)**
$\langle I \rangle / \langle \sigma(I) \rangle$	**5.2 (2.1)**
Completeness (%)	**98.3 % (100 %)**
Multiplicity	**6.7 (7.0)**
Molecular replacement	
Search Model	**AQP0 (1YMG)**
Correlation of F	**0.61 (0.28)** [b]
R-factor	**43.8 % (59.0 %)** [b]
Rigid Body Refinement	
R , R_{free} [c]	**39.8 % , 40.0 %**

Values in parentheses refer to the highest resolution shell.

a $R_{sym} = \sum_{hkl} \sum_i (| I(hkl) - \langle I(hkl) \rangle |) / \sum_{hkl} \sum_i \langle I(hkl) \rangle$.

b 2nd highest peak.

c R_{factor} calculated with 5 % of the data that were not used for refinement.

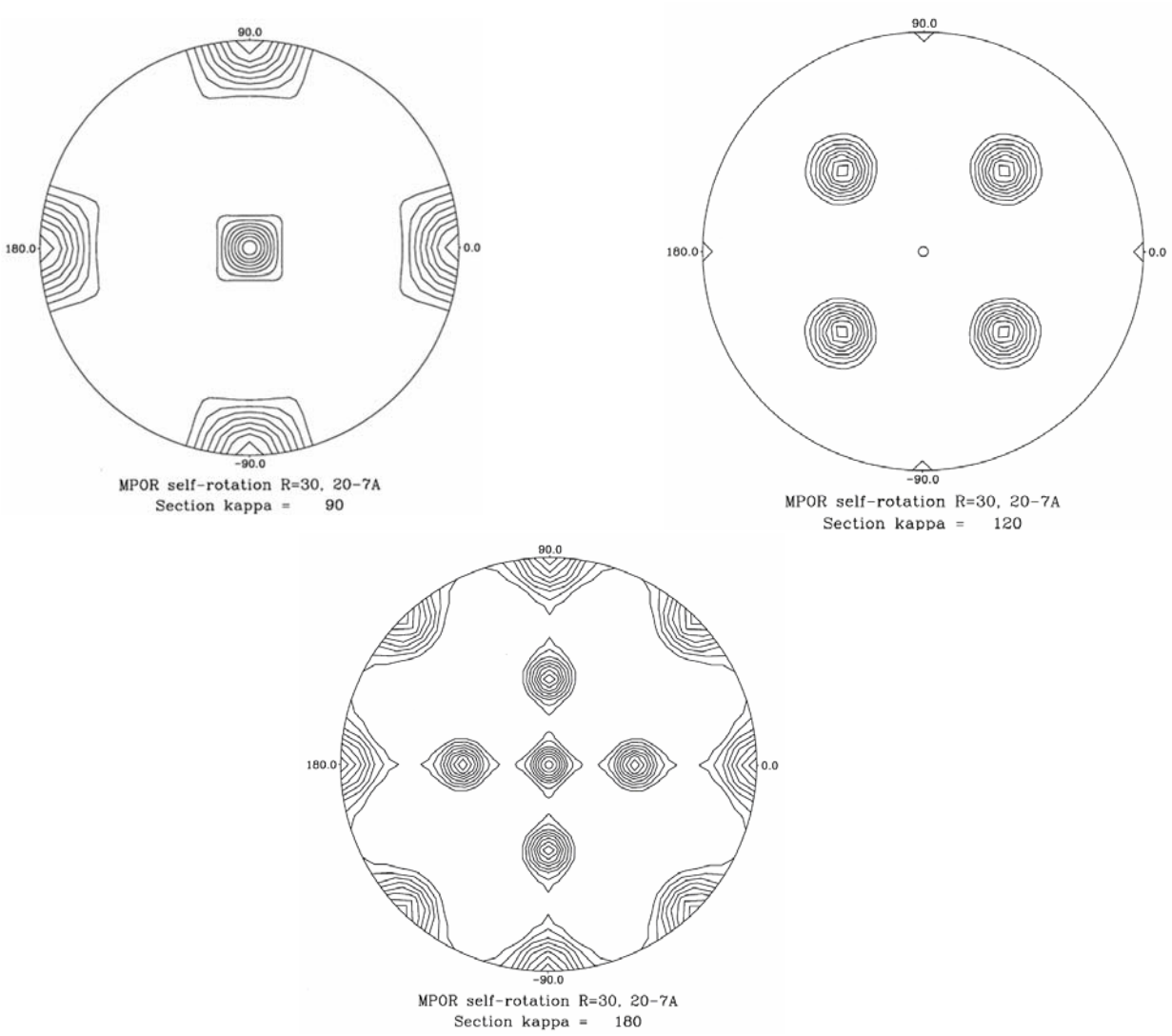


Figure 21: Self rotation diagram at sections kappa 90^{o}, 120^{o} and 180^{o}. The projections at kappa 90^{o} and 180^{o} shows the point group 422 and the presence of three-fold peaks are shown at kappa 120^{o}.

3.7.4 Molecular replacement

The structure was solved by molecular replacement (MR) method using the high resolution structure of native bovine AQP0 (1YMG) as a search model [22]. We generated a tetramer model as an input for MR search. We performed the cross rotation function at RESOLUTION 15 TO 8.0 Å and SPHERE radius of 35 Å. The obtained cross-rotation list was given as an input for performing translation search. The top solutions were used for fit function. We performed the same procedure using monomer as a search model. The results did not give satisfactory results. Finally, to obtain the crystal structure of AQP0, the fit function was rerun with the monomer as the search model using the fitting values obtained from the tetramer solution. By performing this step, we obtained the best fit values. The resulting MR model was subjected to rigid body refinement. The refinement gave an overall Rfactor of 39.8 % and R_{free} of 40 %. The final AQP0 model obtained was used for analyzing the packing in the crystal lattice. The overall scheme in determining the AQP0 crystal structure by MR is represented in the flowchart (Figure 22). The overall statistics obtained from data collection, MR and refinement are summarized in Table 4.

3.7.5 Crystal structure

The crystal structure obtained for AQP0 is illustrated in the Figure 23. The quarternary structure of the tetramer is found virtually unchanged when the model was superimposed with the search model (rmsd = 0.1 Å for all Cα positions). In the crystal, tetramers are tightly associated head-to-head forming octamers of point group 42 symmetry (Figure 24). In contrast and rather intriguingly, there are no direct protein - protein contacts between octamers (minimal distance 19 Å). Rather, their interactions seem to be mediated by the detergent belts that are known to cover the hydrophobic surface of membrane proteins. Indeed, rough modeling suggests that detergent molecules covering adjacent octamers can easily meet (Figure 24). Detergent-mediated crystal contacts have been shown in tetragonal OmpF porin by neutron diffraction [38].

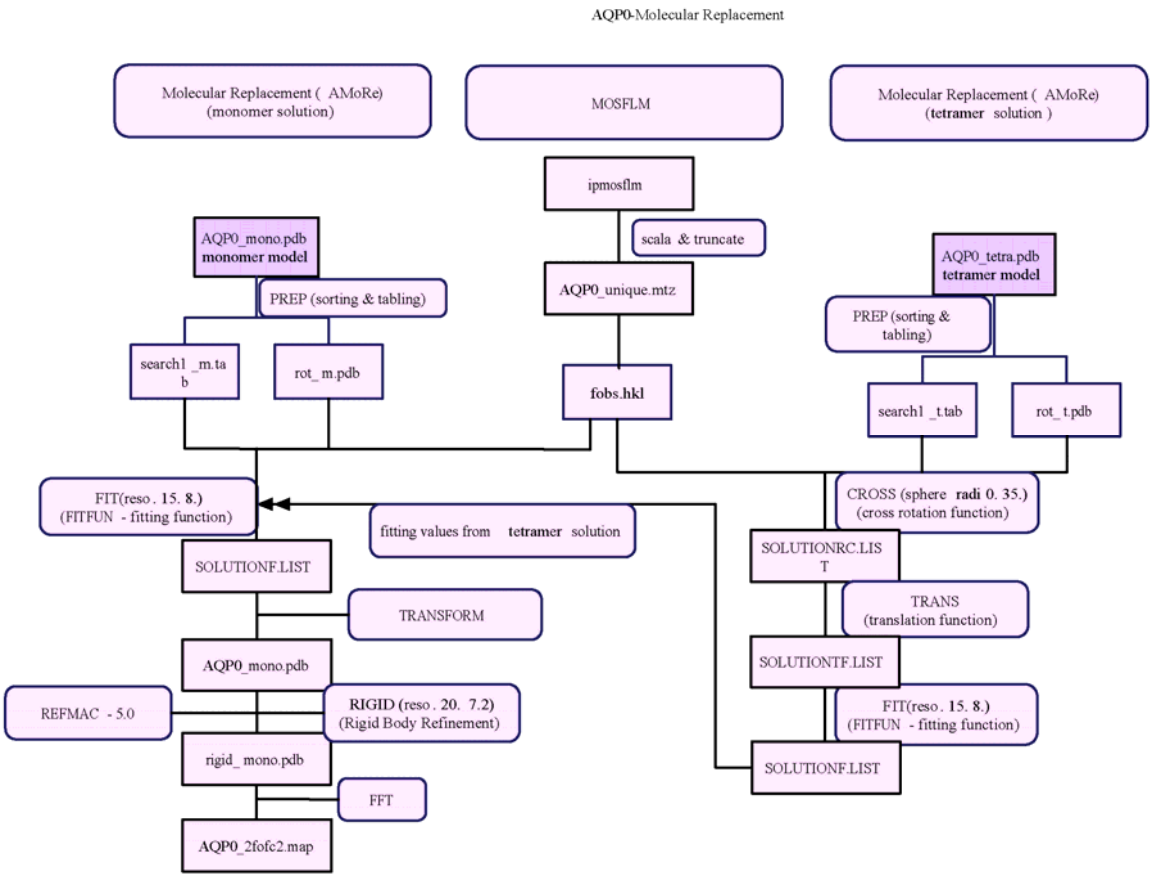


Figure 22: Scheme detailing the procedures in AMoRe (Molecular Replacement (MR)) for obtaining the crystal structure of AQP0.

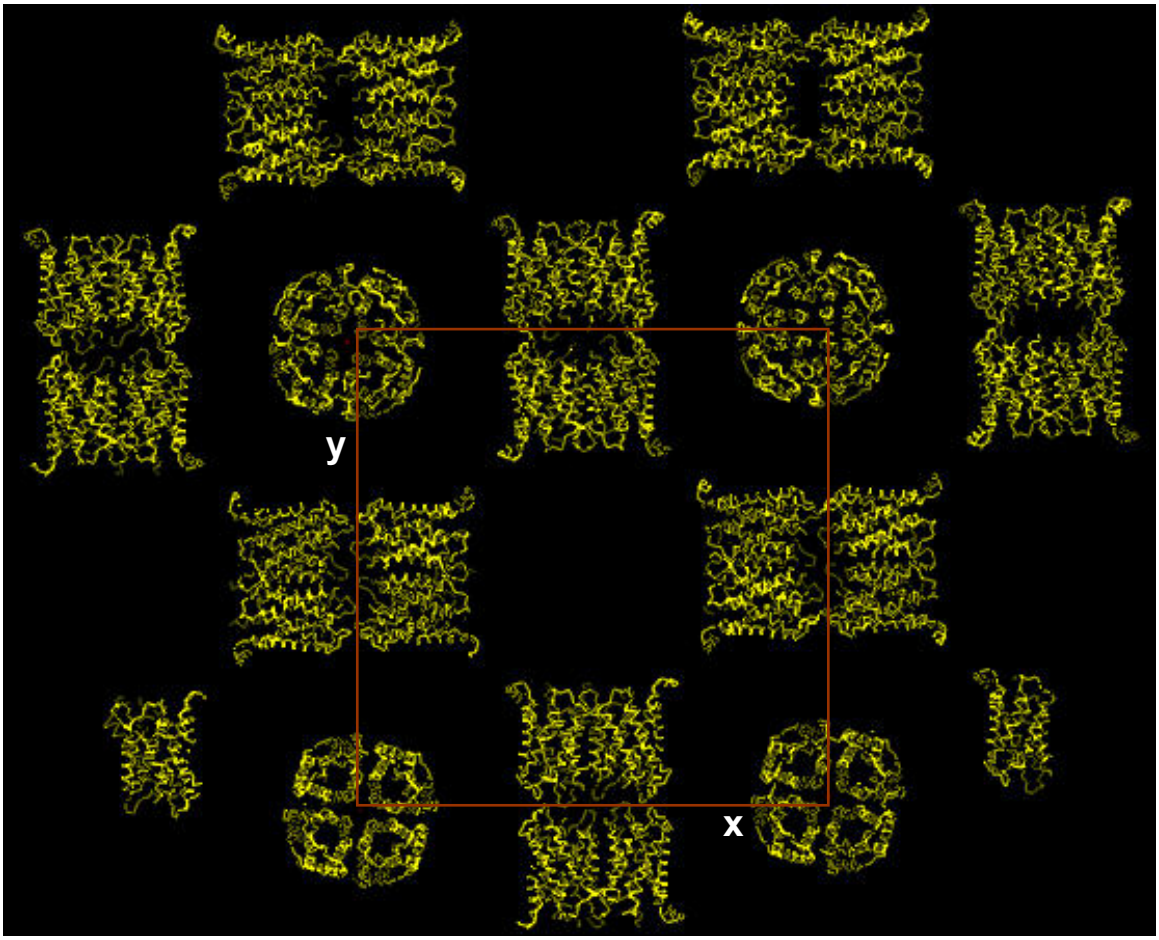


Figure 23: Crystal structure of AQP0. Shown are the Cα traces of the search model (AQP0 from bovine eye lenses, PDB code: 1YMG) positioned and oriented according to the molecular replacement solution**.**

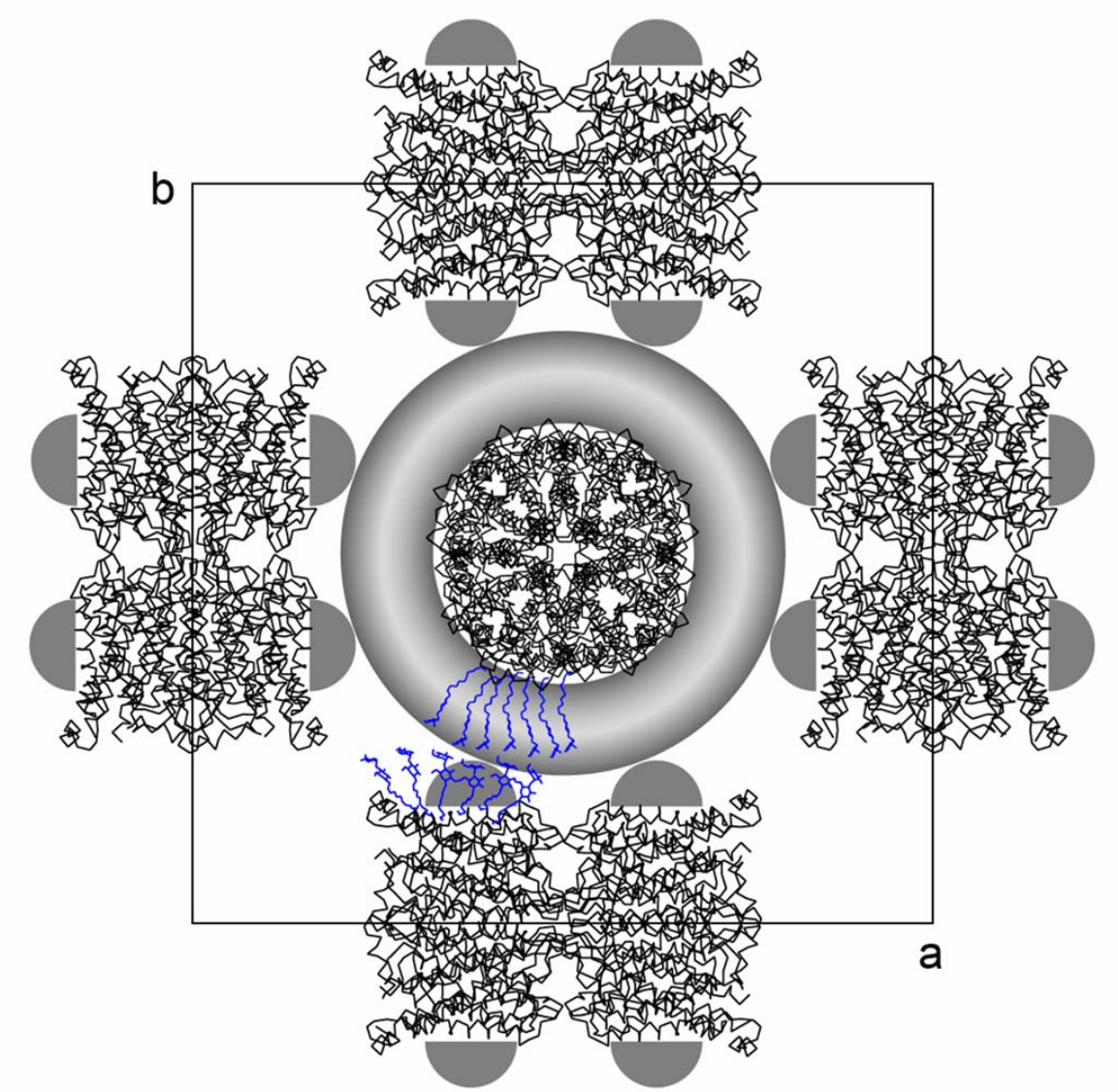


Figure 24: Crystal packing viewed along the c-axis with unit cell indicated. The slab shown extends from – c/2 to c/2. All five octamers shown are centered at a height of Z = 0. The detergent belts covering the transmembrane region of the protein are shown schematically.

Figure 25 shows that AQP0 tetramers are associated *via* their extracellular surfaces. At the center of the octamer, the loop A of one tetramer interact tightly through a close contact (2.6 Å) between Trp34 residues and their symmetry mates of the other tetramer. Part of the loop A is shown as a broken line, since it was found disordered in the search

model. At the perimeter, lobe C2 is inserted neatly into a pocket formed by C1 and C2 of the neighboring tetramer (loop 'C' is represented as lobe C1 and C2). In detail, as shown

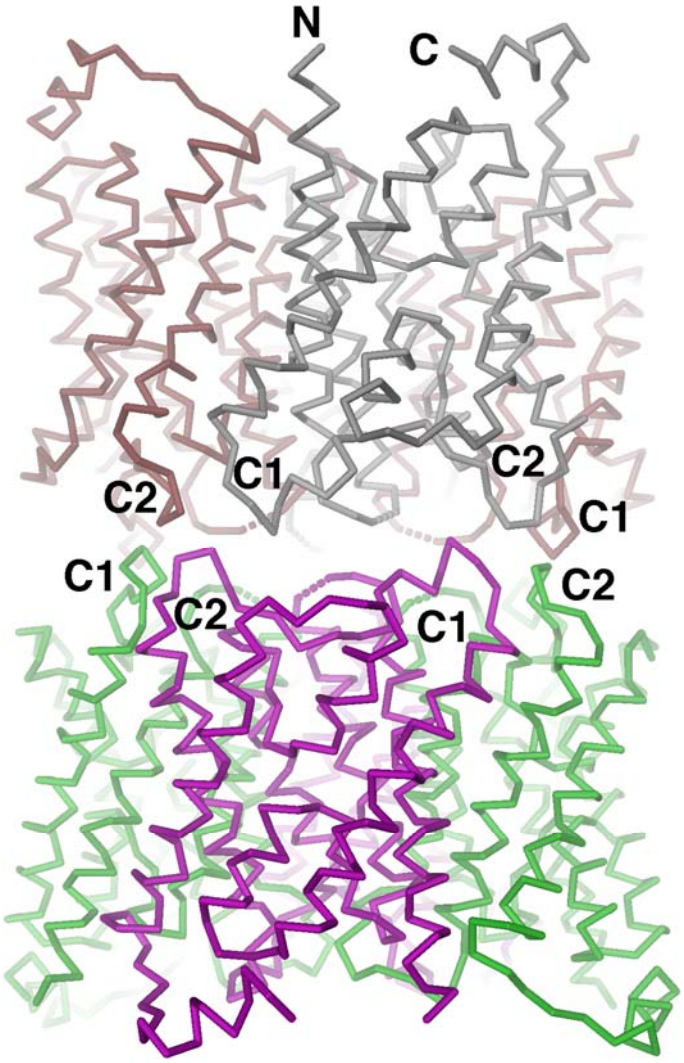


Figure 25: Side-view of the AQP0 octamer (point-group symmetry 42) with neighboring subunits distinguished by different colors. Two tetramers are associated head-to-head *via* their extracellular surfaces. For clarity, two monomers have been clipped off at the rear. At the center of the octamer, loops A interact. Part of loop A is disordered in the search model (broken line). At the perimeter, two-fold related interdigitating contacts (lobes C1 (residues 107 - 116) and C2 (121 - 127)) of loop C are formed.

in Figure 26, P123 and G124 are tucked into the pocket formed by residues P109, A111, V112 (C1) of one monomer and P123, G124 (C2) of the adjacent monomer. Due to the symmetry of the octamer, this interaction occurs eight times.

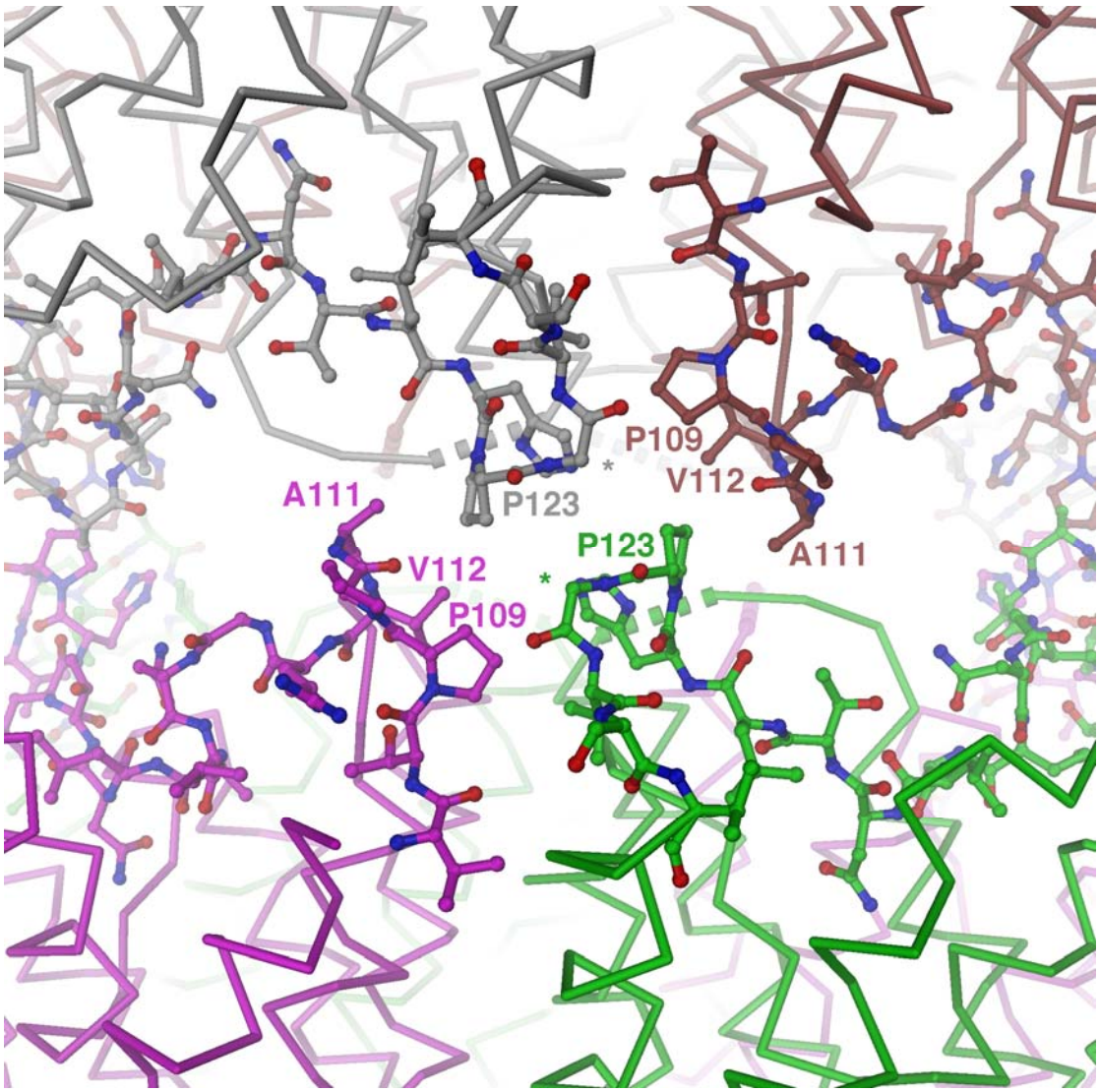


Figure 26: Close-up view of the interdigitating C loops. The view is along a molecular two-fold symmetry axes and is rotated by 45 degrees about the vertical 4-fold axes with respect to panel (b). The asterisk denotes residue G124.

Tetramers are associated such that the monomers and their intrinsic pores come to lie on top of each other (Figure 27). Such interactions might indicate that this tetramer-tetramer arrangement may facilitate the water movement and maintain the osmotic equilibrium between lens fiber cells, which is critical for maintaining transparency of the optic lens.

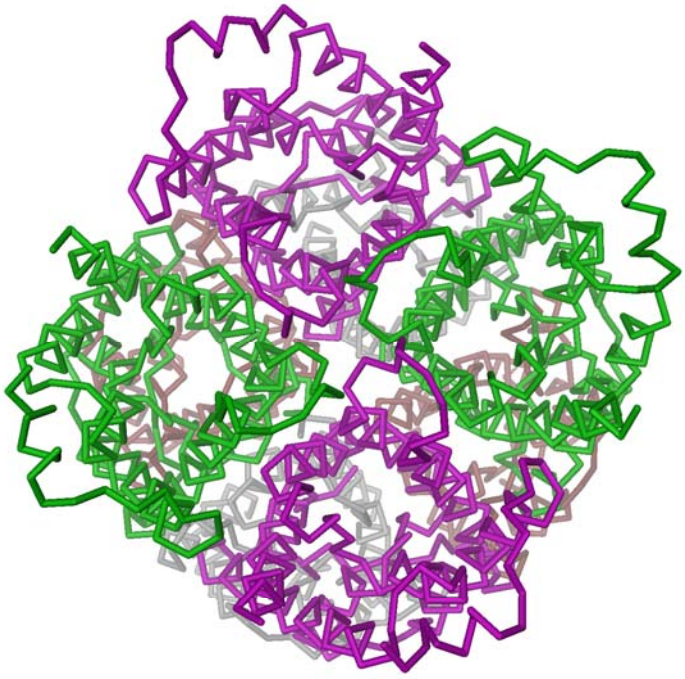


Figure 27: Same as (Figure 25), but viewed along the 4-fold axis. Tetramers are associated such that the monomers and their intrinsic pores come to lie on top of each other.

4.0 Discussion

Studies on AQP0 have revealed that it has two important characteristic functions in eye lenses. It plays an important role as a water channel protein in maintaining the lens transparency and as a cell-to-cell adhesion function to maintain the interior architecture of the lens. The latter function became evident from knock-out studies, which showed that the deficiency of this protein causes disarray of the membrane leading to cataract formation [25]. Several groups visualized the architecture of this protein in eye lens using different methodologies. All the structural investigations of AQP0 so far have been performed using native source. In our study, we have recombinantly expressed Bovine AQP0 for our structural investigations.

4.1 Evaluvation of oligomeric state of AQP0

Biophysical results have shown that purified AQP0 may exist in solution as two species, tetramers and octamers, respectively. BN-PAGE revealed the presence of two components for AQP0. This gave us a first hint that AQP0 might form higher order oligomer. Electron micrographs obtained from single particle analysis indicated monodispersed tetramers with the occasional occurrence of coaxial octamers. The dilution of the sample for single particle analysis might shift the equilibrium towards tetramer formation. Investigation from sedimentation velocity measurements showed the presence of one main species. However, from sedimentation equilibrium (SE) experiments we obtained the monomer-tetramer and monomer-octamer self-association models which suggests that different oligomeric forms of solubilized AQP0, up to octamers, are in equilibrium. From this analysis we could not precisely determine the equilibrium between the two states, since it is largely dependent on the experimental conditions. From all the biophysical observation, we may suggest that octamer formation for AQP0 could be a concentration dependent process.

4.2 Evaluation of AQP0 crystal formation

Recombinant AQP0 was the first mammalian membrane protein expressed in a eukaryotic system and crystallized. AQP0 crystals of different morphologies showed very weak diffraction at SLS. Intriguingly, the cubic crystals diffracted up to 7 Å resolution after a period of one year. This is probably being due to the process of dehydration of the crystal which improved the diffraction quality [39]. Other crystal shapes did not show any diffraction pattern.

The crystal structure of AQP0 shows head-to-head association of tetramers forming octamers of crystallographic 422 symmetry. From the crystal packing, we could not observe protein mediated contacts, since the distance between the two octamers were found to be around 19 Å. Our rough modeling suggests that it is mediated by the detergent belts. Intriguingly, AQP0 crystals previously obtained using 1 % NG also showed diffraction up to 7 Å resolution (Ref. : Table 3) and showed the same packing.

4.3 Comparison of octamer model from X-ray and EM

Head-to-head association of tetramers is not observed in the high-resolution structure of AQP0, which has been determined from a different 3D crystal form ($P42_12$) [22]. Whereas the AQP0 structure obtained from EM studies have shown the association of tetramers to form octamers (*P422*) in 2D crystals [21]. How does, however, our X-ray model compare with the recently obtained EM structure? Unfortunately, the low resolution of the X-ray data permits only a comparison of the oligomeric arrangement in 3D crystals with respect to the 2D crystals obtained from EM studies. The arrangement of the monomers in the tetramer is virtually identical, and in both cases the tetramers are associated to form tight octamers. However, the mutual arrangement of the tetramers is significantly different, due to different isologous interactions between the C loops (Figure 28). Into the same pocket (outlined by P109, A111, and P123), a different residue is inserted, which is P110 in the EM, and P123 in our X-ray structure. This results in a change of 24^o (about the 4-fold axis) of the relative tetramer - tetramer orientation, equivalent to a relative shift of 9 Å at the perimeter. The difference in packing is clearly

significant, although derived from low resolution diffraction data. In fact, while the EM tetramer yielded the same molecular replacement solution as the high-resolution X-ray model, no solution was found, when using the EM octamer model.

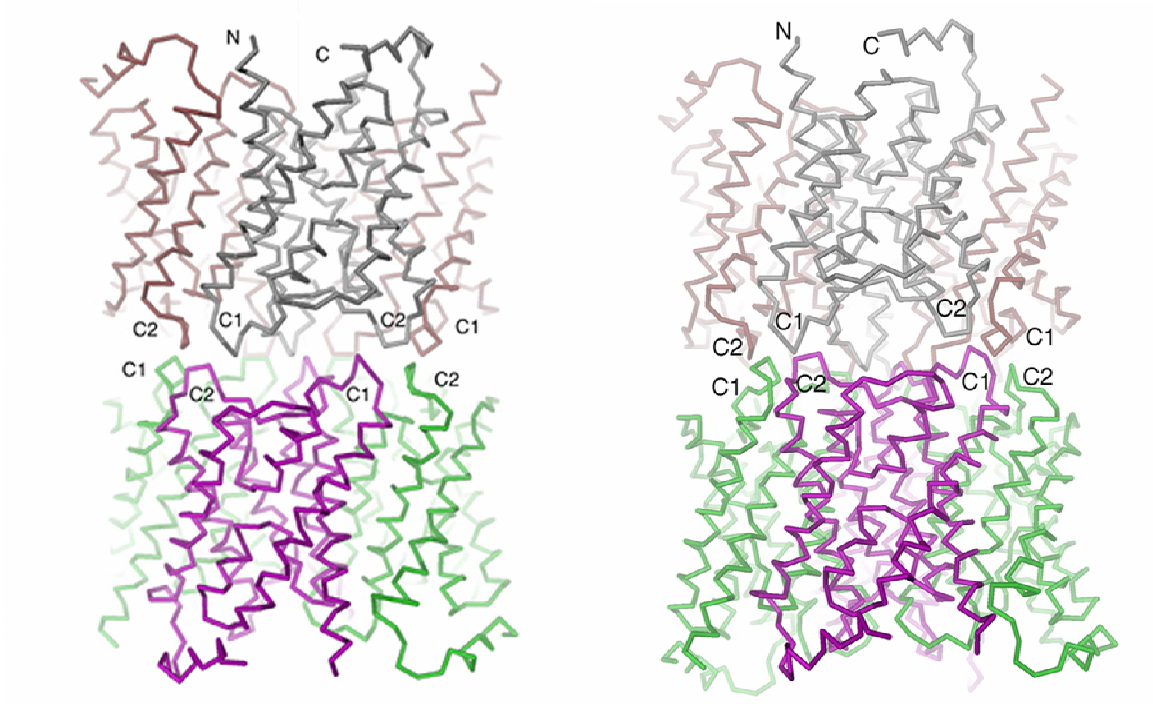


Figure 28: Comparison of X-ray and EM octamer models. Between the two models, there is a relative shift of tetramer – tetramer orientation of 24^{o} (about the 4-fold axis), equivalent to a relative shift of 9 Å at the perimeter.

4.4 Does our AQP0 octamer model resemble its function in eye lenses?

From our study, it appears that the extracellular surfaces of AQP0 tetramers can snap in with two distinct azimuthal orientations. Which one of these corresponds to the situation in solution and, possibly, in the native membrane junction? It has to be considered that the tight lateral packing in the two-layered 2D crystals may severely restrain possible tetramer - tetramer associations. In fact, these crystals have probably been formed by head-to-head association of the two layers and not by incorporation of preformed octamers. This can be inferred from the occurrence of single layers at the edge of 2D crystals. Moreover, AQP0 obtained from the core of sheep lenses for EM studies contain a proportion of the truncated protein (proteolytic cleavage at the C-terminus) which mediated the double layered crystals. Whereas the full length AQP0 isolated from lens cortex yielded only single layer crystals [40]. In our X-ray model, the lack of any lateral contacts between the octamers in the 3D crystals of full length AQP0, makes it unlikely that the octamers represent a crystallization artifact. Therefore, if there is a direct interaction of AQP0 molecules of adjacent membranes in the lens nucleus, it appears most likely that it would occur in the same way as seen in the 3D crystals. Hence, in nature, this unique arrangement of head-to-head association of AQP0 aligns the aqueous pores of the channel from adjacent membranes which may facilitate water movement and maintain osmotic equilibrium between lens fiber cells. This may be one of the reasons why water channel was “chosen” to mediate cell-cell interactions which may in turn have a physiologic role in maintaining optic transparency and preventing cataracts.

5.0 Conclusions

Recombinant bovine AQP0 was crystallized and showed diffraction up to 7 Å resolution. Our crystal structure of the recombinant bovine AQP0 has shown a new insight for the proposed cell-cell adhesion property in eye lenses apart from its water channel function. We obtained only few experimental evidences from biophysical studies which suggested the presence of octamers in solution. In our crystal structure, AQP0 forms head-to-head association of the tetramers (forming octamers) at the extracellular surface in the loosely packed 3D crystals. The contacts at the extracellular region were mediated by tight loop "C" interactions at the perimeter of the octamer. Comparing our X-ray model with the EM model, we found a difference in loop "C" interactions which mediates the AQP0-AQP0 contact. In order to envisage the residues involved in the tetramer-tetramer association, mutant studies have to be performed. We hope that the outcome of the mutant analysis will shed the light on these two proposed models to inspect and to elucidate its interactions in the extracellular surface to resemble its arrangement in the native eye lenses.

Bibliography

1. Heymann, J.B. and A. Engel, Structural clues in the sequences of the aquaporins. J Mol Biol, 2000. **295**(4): p. 1039-53.
2. Murata, K., et al., Structural determinants of water permeation through aquaporin-1. Nature, 2000. **407**(6804): p. 599-605.
3. Agre, P. and D. Kozono, Aquaporin water channels: molecular mechanisms for human diseases. FEBS Lett, 2003. **555**(1): p. 72-8.
4. Deen, P.M., et al., Requirement of human renal water channel aquaporin-2 for vasopressin-dependent concentration of urine. Science, 1994. **264**(5155): p. 92-5.
5. Echevarria, M., et al., Cloning and expression of AQP3, a water channel from the medullary collecting duct of rat kidney. Proc Natl Acad Sci U S A, 1994. **91**(23): p. 10997-1001.
6. Furman, C.S., et al., Aquaporin-4 square array assembly: opposing actions of M1 and M23 isoforms. Proc Natl Acad Sci U S A, 2003. **100**(23): p. 13609-14.
7. Nielsen, S., et al., Aquaporins in complex tissues. II. Subcellular distribution in respiratory and glandular tissues of rat. Am J Physiol, 1997. **273**(5 Pt 1): p. C1549-61.
8. Ikeda, M., et al., Characterization of aquaporin-6 as a nitrate channel in mammalian cells. Requirement of pore-lining residue threonine 63. J Biol Chem, 2002. **277**(42): p. 39873-9.
9. Gorin, M.B., et al., The major intrinsic protein (MIP) of the bovine lens fiber membrane: characterization and structure based on cDNA cloning. Cell, 1984. **39**(1): p. 49-59.
10. Zampighi, G.A., et al., The structural organization and protein composition of lens fiber junctions. J Cell Biol, 1989. **108**(6): p. 2255-75.
11. Stafford, M., The histology and biology of the lens, Association of Optometrists. 2001.

12. Kinoshita, J.H., Kador, P. & Catiles, M, Aldose reductase in diabetic cataract. J. Am. Med. Assoc, 1981. **246**(3): p. 257–261.
13. Park, J.H. and M.H. Saier, Jr., Phylogenetic characterization of the MIP family of transmembrane channel proteins. J Membr Biol, 1996. **153**(3): p. 171-80.
14. Walz, T., et al., Projection map of aquaporin-1 determined by electron crystallography. Nat Struct Biol, 1995. **2**(9): p. 730-2.
15. Walz, T., et al., The three-dimensional structure of aquaporin-1. Nature, 1997. **387**(6633): p. 624-7.
16. Preston, G.M., et al., Membrane topology of aquaporin CHIP. Analysis of functional epitope-scanning mutants by vectorial proteolysis. J Biol Chem, 1994. **269**(3): p. 1668-73.
17. Chandy, G., et al., Comparison of the water transporting properties of MIP and AQP1. J Membr Biol, 1997. **159**(1): p. 29-39.
18. Varadaraj, K., et al., The role of MIP in lens fiber cell membrane transport. J Membr Biol, 1999. **170**(3): p. 191-203.
19. Donaldson, P., J. Kistler, and R.T. Mathias, Molecular solutions to mammalian lens transparency. News Physiol Sci, 2001. 16: p. 118-23.
20. Nemeth-Cahalan, K.L. and J.E. Hall, pH and calcium regulate the water permeability of aquaporin 0. J Biol Chem, 2000. 275(10): p. 6777-82.
21. Gonen, T., et al., Aquaporin-0 membrane junctions reveal the structure of a closed water pore. Nature, 2004. **429**(6988): p. 193-7.
22. 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.
23. Kistler, J. and S. Bullivant, Lens gap junctions and orthogonal arrays are unrelated. FEBS Lett, 1980. **111**(1): p. 73-8.
24. Michea, L.F., M. de la Fuente, and N. Lagos, Lens major intrinsic protein (MIP) promotes adhesion when reconstituted into large unilamellar liposomes. Biochemistry, 1994. **33**(24): p. 7663-9.
25. Shiels, A., et al., Disruption of lens fiber cell architecture in mice expressing a chimeric AQP0-LTR protein. Faseb J, 2000. **14**(14): p. 2207-12.

26. Shiels, A., et al., Optical dysfunction of the crystalline lens in aquaporin-0-deficient mice. Physiol Genomics, 2001. **7**(2): p. 179-86.
27. Shiels, A. and S. Bassnett, Mutations in the founder of the MIP gene family underlie cataract development in the mouse. Nat Genet, 1996. **12**(2): p. 212-5.
28. Fotiadis, D., et al., Surface tongue-and-groove contours on lens MIP facilitate cell-to-cell adherence. J Mol Biol, 2000. **300**(4): p. 779-89.
29. Lustig, A., et al., Molecular weight determination of membrane proteins by sedimentation equilibrium at the sucrose or nycodenz-adjusted density of the hydrated detergent micelle. Biochim Biophys Acta, 2000. **1464**(2): p. 199-206.
30. Schagger, H. and G. von Jagow, Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. Anal Biochem, 1991. **199**(2): p. 223-31.
31. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr, 1994. **50**(Pt 5): p. 760-3.
32. Thompson, J.D., D.G. Higgins, and T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res, 1994. **22**(22): p. 4673-80.
33. Gouet, P., et al., ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics, 1999. **15**(4): p. 305-8.
34. Lebowitz, J., M.S. Lewis, and P. Schuck, Modern analytical ultracentrifugation in protein science: a tutorial review. Protein Sci, 2002. 11(9): p. 2067-79.
35. Konig, N., G.A. Zampighi, and P.J. Butler, Characterisation of the major intrinsic protein (MIP) from bovine lens fibre membranes by electron microscopy and hydrodynamics. J Mol Biol, 1997. **265**(5): p. 590-602.
36. Dutzler, R., PhD thesis, in Department of Structural biology. 1998, University of Basel.
37. Matthews, B., Solvent content of protein crystals. J Mol Biol, 1968. **33**(2): p. 491-7.
38. Pebay-Peyroula, E., et al., Detergent structure in tetragonal crystals of OmpF porin. Structure, 1995. **3**(10): p. 1051-9.

39. Heras, B., et al., Dehydration converts DsbG crystal diffraction from low to high resolution. Structure (Camb), 2003. **11**(2): p. 139-45.
40. Gonen, T., et al., Aquaporin-0 membrane junctions form upon proteolytic cleavage. J Mol Biol, 2004. **342**(4): p. 1337-45.

Chapter 1B

(Manuscript to be submitted)

Co-axial association of recombinant eye lens aquaporin AQP0

Dinesh. V. Palanivelu[1], David E. Kozono[2], Andreas Engel[1], Kitaru Suda[1], Ariel Lustig[3], Peter Agre[2], Tilman Schirmer[1,*]

Keywords: eye lens, aquaporin, X-ray crystallography, membrane protein, protein - protein interaction

[1] Division of Structural Biology, Biozentrum, University of Basel, Klingelbergstr. 70, CH - 4056 Basel, Switzerland

[2] Department of Biological Chemistry and Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

[3] Division of Biophysical Chemistry, Biozentrum, University of Basel, Klingelbergstr. 70, CH - 4056 Basel, Switzerland

[*] Corresponding author, tilman.schirmer@unibas.ch.

Abstract

Aquaporin-0 (AQP0) is the major membrane protein in vertebrate eye lenses. It has been proposed that AQP0 tetramers mediate contact between membranes of adjacent lens fiber cells, which would be consistent with the extraordinarily narrow inter-cellular space. Indeed, octamer formation with purified AQP0 in solution was observed by native gel electrophoresis and analytical ultracentrifugation methods. We obtained 3D crystals of AQP0 that diffract to 7.0 Å resolution have been obtained and the molecular replacement was performed using the recently determined 3D structure of AQP0 from native source. The result shows that, within the cubic lattice, tetramers (point symmetry 42) are associated head-to-head. There are no direct octamer-octamer contacts and the crystal integrity is most probably maintained by detergent belts surrounding the membrane protein. Within the octamer, extracellular loops A and C interdigitate at the center and the perimeter of the octamer, respectively. The octamer formation has been compared with the AQP0 structure derived from 2D crystals using electron diffraction. Intriguingly, the mutual orientation of tetramers within the octamer is significantly different to that previously reported for 2D crystals. Clearly, the low resolution of the X-ray data permits only a comparison of the oligomeric arrangement. The interactions observed in the loosely-packed 3D crystals presented here possibly represent the *in vivo* association mode between AQP0 tetramers from juxtaposed membranes in the eye lens.

The major intrinsic protein (MIP26, AQP0), an integral membrane protein present in vertebrate eye lenses[1] is the founding member of the aquaporin superfamily[2; 3]. AQP0 is the most abundant protein in the plasma membrane of lens fiber cells constituting almost 50% of the total protein[4]. There is extensive intercellular coupling in this tissue, which renders AQP0 a possible candidate to mediate cell-cell interactions[5; 6]. Knock-out of this protein in Cat^{Fr} mice[7] causes cataract formation. Similarly, mutations in AQP0 that prevent proper folding or targeting cause lens opacification supporting a crucial role of AQP0 in the development of the transparent lens[8]. Water channel function of AQP0 (though about 40 times less efficient compared to AQP1[9]) and facilitation of glycerol translocation[10] have been demonstrated for AQP0, but its putative role in mediating cell-cell interactions appears unique amongst aquaporins.

The arrangement of AQP0 in native membranes has been examined in freeze- and label-fractured preparations by electron microscopy using immunolabeling. AQP0 was found only in tightly-abutting membranes (thin and wavy junctions) but not in thick junctions[11]. However, a direct contact between AQP0 molecules could not be proven in this case. Interactions between AQP0 containing proteoliposomes have been investigated by fluorescent resonance energy transfer and turbidity measurements[12]. No evidence was obtained for direct AQP0-AQP0 interaction, whereas aggregation was observed upon mixing the proteoliposomes with phosphatidylserine liposomes. This might indicate that *in vivo* AQP0 interacts directly with the lipids of the juxtaposed cell membrane. Later, studies from 2D-Crystalline sheets of ovine AQP0[15] analyzed by atomic force microscopy (AFM) have shown a precise interaction of AQP0 in the extracellular surfaces. Analysis of AFM togographs made them to propose a tongue-and-groove fit model for AQP0 (Dimi

JMB). AQP0 (monomer mass 26 kDa) is a homolog of AQP1 from bovine red blood cells (44 % sequence identity). The structure of AQP0 has been determined by X-ray crystallography[13] and is very similar to AQP1[14]. It is composed of six transmembrane helices and five connecting loops obeying an internal two-fold symmetry. Both loops B and E contain the consensus motif Asn-Pro-Ala and dip into the membrane from opposite sides contributing to form the water channel[13]. The structural unit of AQP0 is a tetramer with four independent hour glass-shaped channels. Recently, structure of AQP0 from sheep eye lens was solved by electron microscopic (EM) structure at 3.0 Å resolution[17] determined from the same double-layered 2D crystal form which showed tight head-to-head association of tetramers from the two layers. This has been proposed to resemble the situation in membrane junctions. Here, we report a similar, though not identical, association of AQP0 tetramers in loosely-packed 3D crystals.

Biophysical characterization

The His-tagged bovine AQP0 was heterologously expressed in *Saccharomyces cerevisiae* and purified by nickel affinity chromatography. This had several advantages including straight forward expression and purification of milligram quantities of protein. Also, heterologous expression lessened the likelihood that protein interactions are mediated by the co-factors found in the native environment, rather than driven by intrinsic structural properties. To assess for its homogeneity and oligomeric state, the solubilized AQP0 was subjected to several tests. Electron micrographs indicated monodispersed tetramers with the occasional occurrence of coaxial octamers (data not shown). Native gel electrophoresis[18] revealed, a major (tetramer) band with an apparent mass between 160

and 180 kDa (protein and detergent mass) and in addition, a strong band of about 400 kDa (protein and detergent mass) (data not shown).

Analytical ultracentrifugation performed with protein concentrations of 0.3 to 2 mg / mL indicated one main species with an observed sedimentation coefficient between 4.3 and 4.6 S in detergent / buffer solution. This most probably corresponds to the tetramer, since from these values an upper mass boundary of approximately 110 kDa can be calculated for the protein-detergent complex assuming spherical shape and water viscosity. A marginal contribution of a smaller species (probably monomers) was apparent at the lower protein concentrations.

Sedimentation equilibrium measurements were carried out in the presence of 28 % sucrose to match the density of the bound detergent (1.128 g / mL) such that the buoyancy of the mixed micelles was due to the protein component alone[19]. Data from three runs (13000, 17000, and 24000 rpm; protein concentration 1.8 mg / mL) were fitted simultaneously. The low apparent mean mass of m = 52 kDa (chi^2 = 0.075) was consistent with the presence of monomers (M = 28 kDa) in addition to oligomers. Therefore, several self-association models were tested. Monomer-tetramer and monomer-octamer self-association models fitted the data equally well (m = 38 kDa, chi^2 = 0.059 and m = 37 kDa, chi^2 = 0.052, respectively), while a 3-state model did not improve the fit. A monomer - tetramer equilibrium has been demonstrated previously by Konig et al.[20]. In summary, it appears that the different oligomeric forms of solubilized AQP0, up to octamers, are in an equilibrium that is largely dependent on the experimental conditions.

Crystal structure

The recombinant protein readily formed cubic crystals (Figure 1) under similar conditions as reported for the wild-type protein from sheep eye lenses[21] and showed strong diffraction to 7 Å resolution. The modest resolution seems to be inherent to the crystal packing (see below), since, despite extensive trials on crystallization conditions, no improvement was achieved. A complete dataset was collected to 7 Å resolution and the statistics are given in Table 1. The structure was solved by molecular replacement with the high resolution structure of native bovine AQP0[13]. The quarternary structure of the tetramer is found virtually unchanged when compared with the search model (rmsd = 0.1 Å for all Cα positions). In the crystal, tetramers are tightly associated head-to-head forming octamers of crystallographic 422 symmetry (Figure 2a). In contrast and rather intriguingly, there is no direct protein - protein contacts between octamers (minimal distance 19 Å). Rather, their interactions seem to be mediated by the detergent belts that are known to cover the hydrophobic surface of membrane proteins. Indeed, rough modeling suggests that detergent molecules covering adjacent octamers can easily meet (Figure 2a). Detergent-mediated crystal contacts have been shown in tetragonal OmpF porin[22] by neutron diffraction[23].

Figure 2b shows that AQP0 tetramers are associated *via* their extracellular surfaces. At the center of the octamer, the loop A of one tetramer interact tightly through a close contact (2.6 Å) between Trp34 residues and their symmetry mates of the other tetramer. At the perimeter, lobe C2 is inserted neatly into a pocket formed by C1 and C2 of the neighboring tetramer (loop 'C' is represented as lobe C1 and C2). In detail, as shown in Figure 2c, P123 and G124 are tucked into the pocket formed by residues P109, A111,

V112 (C1) of one monomer and P123, G124 (C2) of the adjacent monomer. Due to the symmetry of the octamer, this interaction occurs eight times. Tetramers are associated such that the monomers and their intrinsic pores come to lie on top of each other (see Figure 2d). Such interactions might indicate that this tetramer-tetramer arrangement may facilitate the water movement and maintain the osmotic equilibrium between lens fiber cells, which is critical for maintaining transparency of the optic lens.

Head-to-head association of tetramers is not observed in the high-resolution structure of AQP0, which has been determined from a different 3D crystal form ($P42_12$)[13]. How does, however, our model compare with the recently obtained EM structure[17]? Unfortunately, the low resolution of the X-ray data permits only a comparison of the oligomeric arrangement. The arrangement of the monomers in the tetramer is virtually identical, and in both cases the tetramers are associated to form tight octamers. However, the mutual arrangement of the tetramers is significantly different, due to different isologous interactions between the C loops (compare Figure 2d of ref.[17] and Figure 2c). Into the same pocket (outlined by P109, A111, and P123), a different residue is inserted, which is P110 in the EM, and P123 in the X-ray structure. This results in a change of 24^o (about the 4-fold axis) of the relative tetramer - tetramer orientation, equivalent to a relative shift of 9 Å at the perimeter. The difference in packing is clearly significant, although derived from low resolution diffraction data. In fact, while the EM tetramer yielded the same molecular replacement solution as the high-resolution X-ray model, no solution was found, when using the EM octamer model.

Conclusions

It appears that the extracellular surfaces of AQP0 tetramers can snap in with two distinct azimuthal orientations. Which one of these corresponds to the situation in solution and, possibly, in the native membrane junction? It has to be considered that the tight lateral packing in the two-layered 2D crystals may severely restrain possible tetramer - tetramer associations. In fact, these crystals have probably been formed by head-to-head association of the two layers and not by incorporation of preformed octamers. This can be inferred from the occurrence of single layers at the edge of 2D crystals[17]. In contrast, the lack of any lateral contacts in the 3D crystals makes the crystal packing to be an artifact unlikely. Therefore, if there is a direct interaction of AQP0 molecules of adjacent membranes in the lens nucleus, it appears most likely that it would occur in the same way as seen in the 3D crystals. Hence, in nature, this unique arrangement of head-to-head association of AQP0 aligns the aqueous pores of the channel from adjacent membranes, which may facilitate water movement and osmotic equilibrium between lens fiber cells. This may be one of the reasons why a water channel was "chosen" to mediate cell-cell interactions which may in turn have a physiologic role in maintaining optic transparency and preventing cataracts.

Acknowledgements

This work was funded by the Swiss National Science Foundation grant 31-53727.98 to T.S. We thank Zora Markovic-Housley and Carmen Chan for the critical reading of the manuscript, Dimitrios Fotiadis, Herve Remigy and George Orriss for their advice and support. Special thanks to Caroline Peneff-Verheyden and the staff of beamline X06SA of the SLS synchrotron (Villigen, Switzerland) for assistance with data acquisition.

References

1. Gorin, M. B., Yancey, S. B., Cline, J., Revel, J. P. & Horwitz, J. (1984). The major intrinsic protein (MIP) of the bovine lens fiber membrane: characterization and structure based on cDNA cloning. *Cell* 39, 49-59.

2. Park, J. H. & Saier, M. H., Jr. (1996). Phylogenetic characterization of the MIP family of transmembrane channel proteins. *J Membr Biol* 153, 171-80.

3. Zardoya, R. & Villalba, S. (2001). A phylogenetic framework for the aquaporin family in eukaryotes. *J Mol Evol* 52, 391-404.

4. Broekhuyse, R. M. & Kuhlmann, E. D. (1978). Lens membranes. IV. Preparative isolation and characterization of membranes and various membrane proteins from calf lens. *Exp Eye Res* 26, 305-20.

5. Sas, D. F., Sas, M. J., Johnson, K. R., Menko, A. S. & Johnson, R. G. (1985). Junctions between lens fiber cells are labeled with a monoclonal antibody shown to be specific for MIP26. *J Cell Biol* 100, 216-25.

6. Bok, D., Dockstader, J. & Horwitz, J. (1982). Immunocytochemical localization of the lens main intrinsic polypeptide (MIP26) in communicating junctions. *J Cell Biol* 92, 213-20.

7. Shiels, A., Mackay, D., Bassnett, S., Al-Ghoul, K. & Kuszak, J. (2000). Disruption of lens fiber cell architecture in mice expressing a chimeric AQP0-LTR protein. *Faseb J* 14, 2207-12.

8. Shiels, A. & Bassnett, S. (1996). Mutations in the founder of the MIP gene family underlie cataract development in the mouse. *Nat Genet* 12, 212-5.

9. Chandy, G., Zampighi, G. A., Kreman, M. & Hall, J. E. (1997). Comparison of the water transporting properties of MIP and AQP1. *J Membr Biol* 159, 29-39.

10. Varadaraj, K., Kushmerick, C., Baldo, G. J., Bassnett, S., Shiels, A. & Mathias, R. T. (1999). The role of MIP in lens fiber cell membrane transport. *J Membr Biol* 170, 191-203.

11. Zampighi, G. A., Hall, J. E., Ehring, G. R. & Simon, S. A. (1989). The structural organization and protein composition of lens fiber junctions. *J Cell Biol* 108, 2255-75.

12. Michea, L. F., de la Fuente, M. & Lagos, N. (1994). Lens major intrinsic protein (MIP) promotes adhesion when reconstituted into large unilamellar liposomes. *Biochemistry* 33, 7663-9.

13. Harries, W. E., Akhavan, D., Miercke, L. J., Khademi, S. & Stroud, R. M. (2004). The channel architecture of aquaporin 0 at a 2.2-A resolution. *Proc Natl Acad Sci U S A* 101, 14045-50.

14. Sui, H., Han, B. G., Lee, J. K., Walian, P. & Jap, B. K. (2001). Structural basis of water-specific transport through the AQP1 water channel. *Nature* 414, 872-8.

15. Hasler, L., Walz, T., Tittmann, P., Gross, H., Kistler, J. & Engel, A. (1998). Purified lens major intrinsic protein (MIP) forms highly ordered tetragonal two-dimensional arrays by reconstitution. *J Mol Biol* 279, 855-64.

16. Fotiadis, D., Hasler, L., Muller, D. J., Stahlberg, H., Kistler, J. & Engel, A. (2000). Surface tongue-and-groove contours on lens MIP facilitate cell-to-cell adherence. *J Mol Biol* 300, 779-89.

17. Gonen, T., Sliz, P., Kistler, J., Cheng, Y. & Walz, T. (2004). Aquaporin-0 membrane junctions reveal the structure of a closed water pore. *Nature* 429, 193-7.

18. Schagger, H. & von Jagow, G. (1991). Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal Biochem* 199, 223-31.

19. Lustig, A., Engel, A., Tsiotis, G., Landau, E. M. & Baschong, W. (2000). Molecular weight determination of membrane proteins by sedimentation equilibrium at the sucrose or nycodenz-adjusted density of the hydrated detergent micelle. *Biochim Biophys Acta* 1464, 199-206.

20. Konig, N., Zampighi, G. A. & Butler, P. J. (1997). Characterisation of the major intrinsic protein (MIP) from bovine lens fibre membranes by electron microscopy and hydrodynamics. *J Mol Biol* 265, 590-602.

21. Dutzler, R. (1998). Ph.D thesis, University of Basel.

22. Cowan, S. W., Garavito, R. M., Jansonius, J. N., Jenkins, J. A., Karlsson, R., Konig, N., Pai, E. F., Pauptit, R. A., Rizkallah, P. J., Rosenbusch, J. P. & et al. (1995). The structure of OmpF porin in a tetragonal crystal form. *Structure* 3, 1041-50.

23. Pebay-Peyroula, E., Garavito, R. M., Rosenbusch, J. P., Zulauf, M. & Timmins, P. A. (1995). Detergent structure in tetragonal crystals of OmpF porin. *Structure* 3, 1051-9.

24. Ralston, G. (1993). Introduction to Analytical Ultracentrifugation. *Beckman Instruments Inc,*.

25. CCP4. (1994). Collaborative computational project, number 4. The CCP4 suite: programs for protein crystallography. *Acta. Cryst.* D50, 760-763.

Legends for the Figures

Figure 1. Crystal of recombinant bovine AQP0. The diameter of the cubic crystal is about 150 μm. Bovine AQP0 (native sequence with a N-terminal $(His)_{10}$-tag and a Factor Xa cleavage site) was cloned in yeast vector pYES2.0 with a GAL6 promoter and expressed in yeast strain Sc334. Yeast cells were grown in SD –URA media (6 x 100 mL) and expressed in YPGal media (6 x 1.25 L). This yielded 70 g of cells that were broken by a Dyanomil using glass beads. Cell debris was removed by centrifugation at 9000 rpm (Sorval GSA) and the membrane fraction was obtained by ultracentrifugation at 40,000 rpm (TFT45.94). Initially, AQP0 was extracted from membrane fractions using 10% n-octyl-β-D-glucoside (OG). Subsequently several detergents were tested to yield monodisperse protein preparation as judged by ultracentrifugation. Detergent exchange was performed on Ni-NTA Wizard Midi Columns (Catalys AG, Switzerland) against n-decyl-β-D-maltoside (DM; 0.3 %), n-dodecyl-β-D-maltoside (DDM; 0.15 %), n-nonyl-β-D-glucoside (NG; 1 %), n-nonyl-β-D-maltoside (1 %) or n-decyl-β-D-glucoside (1 %). AQP0 in DM (0.4 %) or DDM (0.15 %) was found only partially aggregated whereas severe precipitation was obtained with the other tested detergents. Further improvement, i.e. monodispersity, was achieved by employing DM or DDM for extraction (at a concentration of 2 %) as well as purification (0.4 % or 0.15 % respectively). Protein concentration was accompanied by detergent concentration (from 0.4 % to about 6 – 7 % for DM) as measured by a home made device aimed to measure detergent concentration (Kaufmann et al., manuscript in preparation), although a large cut-off membrane (100 kDa cut-off, Ultrafree or Amicon Ultra centricon from Millipore) was used. This was overcome by resuspending the solution (starting volume 3 mL) 4 times every 3 minutes with 1 mL

buffer solution resulting in a final detergent concentration below 1 % DM. Sedimentation velocity was measured using a Beckman Optima XL-A analytical ultracentrifuge at 54,000 rpm for checking sample monodispersity. Sedimentation equilibrium experiments were carried out at different angular velocities to investigate the self-association behavior of the sample. Fitting to various self-association models[24] was performed by non-linear regression using the program PROFIT (version 6.0).

Crystals were grown from sitting drops at 20 ºC under similar conditions as found earlier for native AQP0 and had similar shape[21]. Equal volumes (3 + 3 μL) of protein solution (8.5 mg / mL in Tris 10 mM pH 8.0, 300 mM NaCl, 10 % (v/v) glycerol, 5 mM β-mercaptoethanol, 0.4 % DM) and reservoir solution (100 mM bicine NaOH pH 8.9, 350 mM NaCl, 34 % (v/v) PEG400) were mixed. Using 1% NG, crystals of similar quality were obtained.

Figure 2. Crystal structure of AQP0. Shown are the Cα traces of the search model (AQP0 from bovine eye lenses[17], PDB code : 1YMG) positioned and oriented according to the molecular replacement solution. (a) Crystal packing viewed along the c-axis with unit cell indicated. The slab shown extends from – C/2 to C/2. All five octamers are centered at Z = 0. The detergent belts covering the transmembrane region of the protein are shown schematically. (b) Side-view of the AQP0 octamer (point-group symmetry 42) with neighboring subunits distinguished by different colors. Two tetramers are associated head-to-head *via* their extracellular surfaces. For clarity, two monomers have been clipped off at the rear. At the center of the octamer, loops A interact. Part of loop A is disordered in the search model (broken line). At the perimeter, two-fold related interdigitating contacts

(lobes C1 (residues 107 - 116) and C2 (121 - 127)) of loop C are formed. (c) Close-up view of the interdigitating C loops. The view is along a molecular two-fold symmetry axis and is rotated by 45 degrees about the vertical 4-fold axis with respect to panel (b). The asterisk denotes residue G124. (d) Same as (b), but viewed along the 4-fold axis. The figures were prepared with program DINO (*http://www.dino3d.org)*.

Table 1. Data collection and refinement stastics

Data Collection	
X-ray source; λ	SLS, Villigen 0.97970 Å
Detector	MAR345
Temperature	100 K
No. of crystals	1
Space group	*I*432
$a = b = c$	188 Å
Content of a.u.	1 monomer
Resolution range (Å)	15.0 – 7.0 (7.4 - 7.0)
Number of unique reflections	891 (137)
R_{sym} (%) [a]	9.4 % (36.3 %)
$\langle I \rangle / \langle \sigma(I) \rangle$	5.2 (2.1)
Completeness (%)	98.3 % (100 %)
Multiplicity	6.7 (7.0)
Molecular replacement	
Search Model	AQP0 (1YMG)
Correlation of F	0.61 (0.28) [b]
R-factor	43.8% (59.0%) [b]
Rigid Body Refinement	
R , R_{free} [c]	39.8% , 40.0%

Values in parentheses refer to the highest resolution shell.

[a] $R_{sym} = \sum_{hkl} \sum_i (| I(hkl) - \langle I(hkl) \rangle |) / \sum_{hkl} \sum_i \langle I(hkl) \rangle$. [b] 2nd highest peak. [c] R_{factor} calculated with 5 % of the data that were not used for refinement.

Data were processed with MOSFLM 5.0/SCALA[25]. The structure was determined by molecular replacement using program AMORE[25] using all data between 15 and 7 Å resolution. The tetrameric model of AQP0 (PDB code 1YMG) was used as the initial search model. From the resulting solution, a monomer was used for subsequent refinement employing the FIT option of AMORE. Final rigid body refinement of the monomer was performed with REFMAC5[25].

Abbrevations

AQP - Aquaporin

CMC - critical micellar concentration

OG - n-Octyl-β-D-Glucoside

DM - n-Decyl-β-D-Maltoside

DDM - n-Dodecyl-β-D-Maltoside

NG - n-nonyl-β-D-glucoside

EM - electron microscope

Figure 1

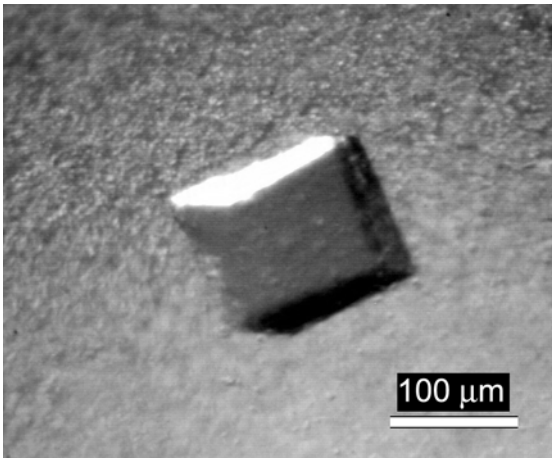


Figure 2a

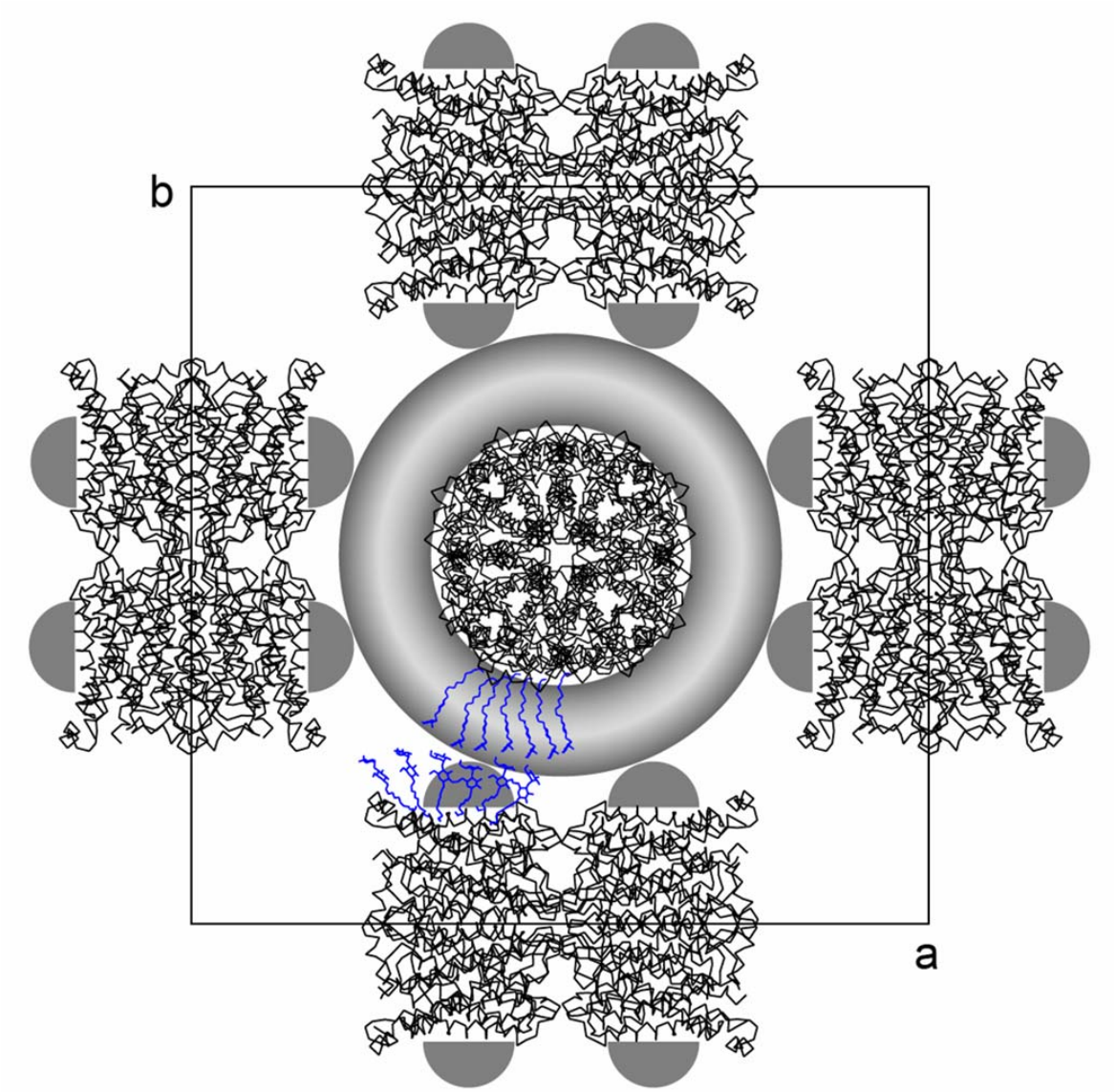


Figure 2b

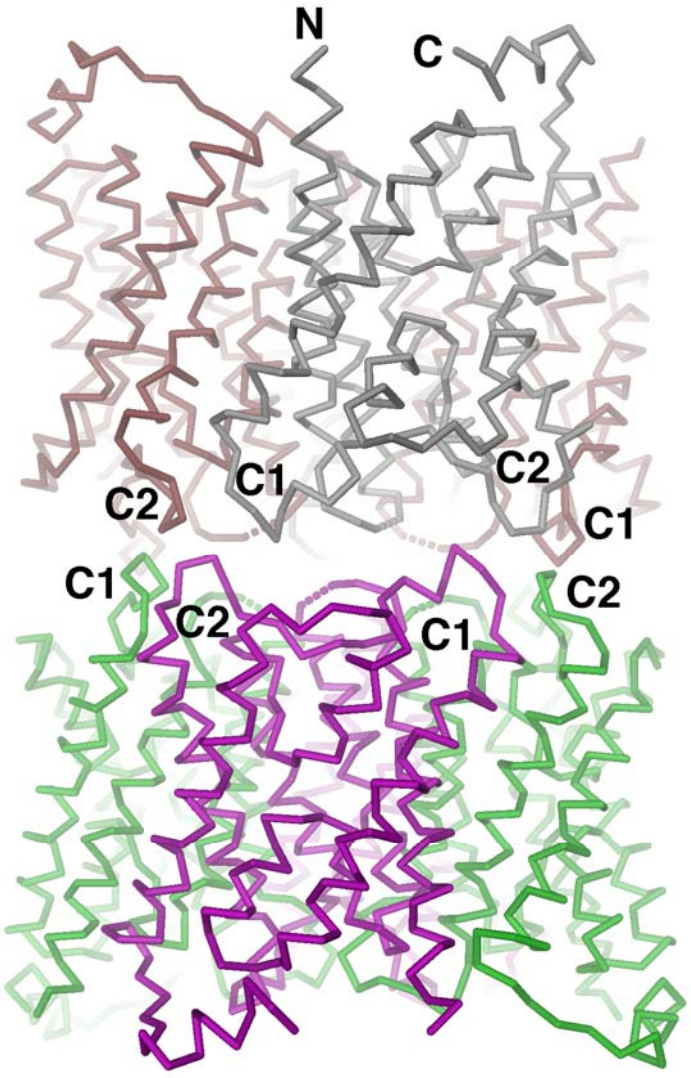


Figure 2c

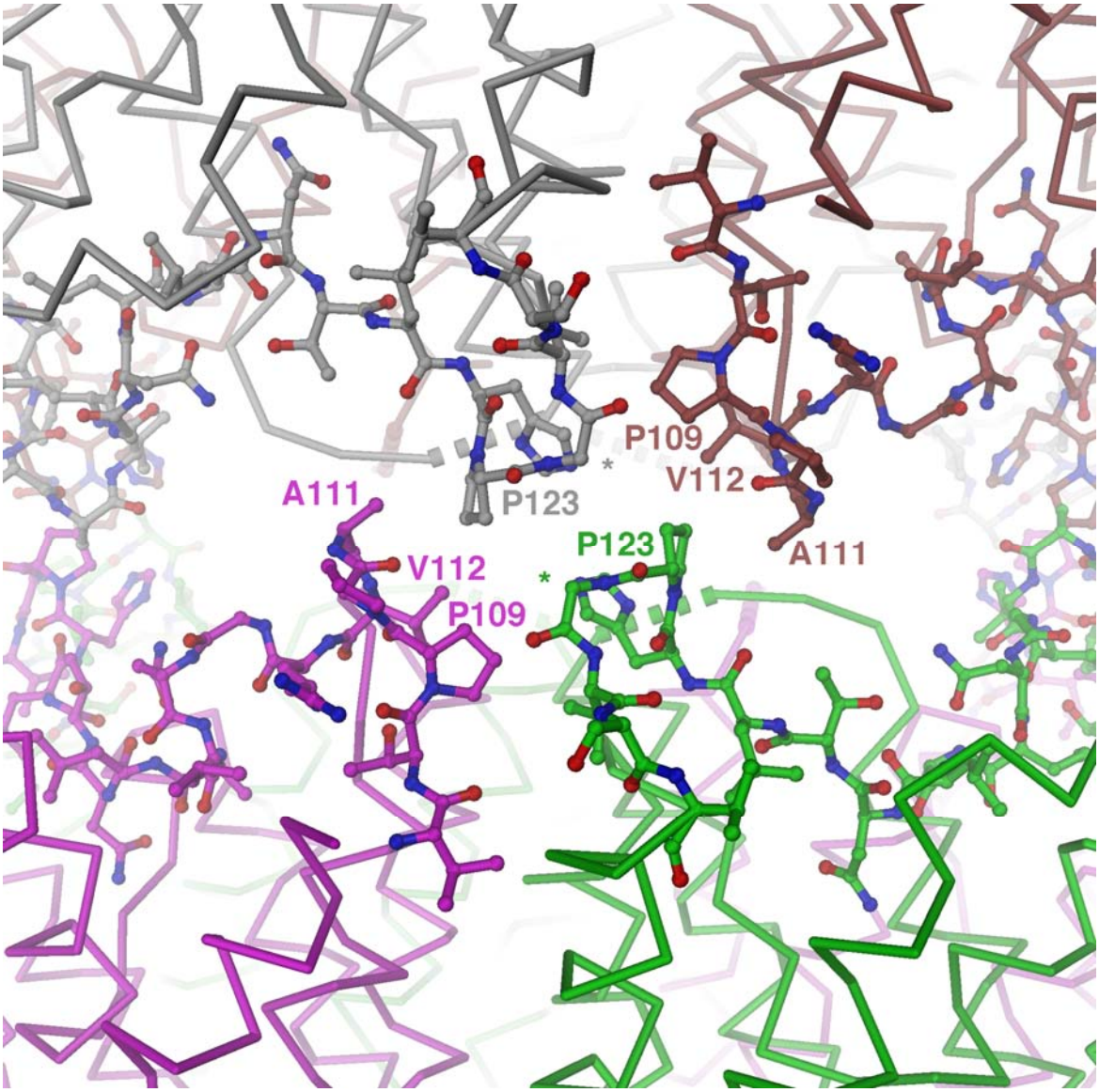
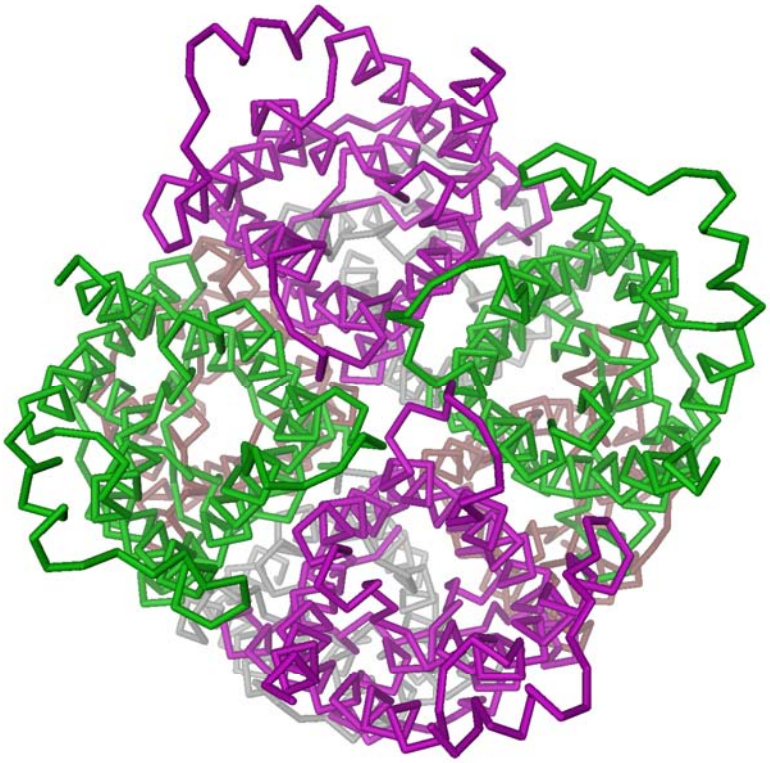


Figure 2d



Chapter II

Structural studies on the effector proteins (Beps) of the VirB-Type IV secretion system in Bartonella henselae

Abstract

Type IV transporters are one of the five major families of transporters that are capable of exporting virulence factors across the membranes of pathogenic bacteria. Members of the Type IV secretion system (T4SS) share structural homologies but they display diversity in the nature of the substrates that they transport. T4SS are multicomponent transporters of Gram-negative bacteria with functions as delivery of effector proteins into eukaryotic target cells in pathogenesis or DNA transfer in bacterial conjugation through the long pilus structure (VirB apparatus). Type IV transporters are produced by several bacterial pathogens. *Bartonella henselae* is one among the T4SSs which mediates the delivery of effector proteins into the eukaryotic cell *via* the process of translocation.

Bartonella henselae is a Gram-negative, arthopod-borne zoonotic pathogen causing a broad range of clinical manifestations in incidentally infected humans. Transmission to humans occurs by cat scratch or bite of an infected cat flea causing vasoproliferative lesions, which result in the formation of tumours in the skin and/or inner organs. *In vitro* studies showed that the VirB T4SS of *B. henselae* mediates most of the virulence attributes associated with the interaction of *B. henselae* with human endothelial cells and they interfere with the signalling network of the host. We have characterized few of the Bartonella effector proteins (Beps) BepA, B, C and D and their domains of *B. henselae* for structural analysis. Crystals were obtained for the truncated version of BepA and it diffracted to 3.3 Å resolution. Currently, we are in the process of investigating the data for structural elucidation of the BepA protein. Investigations on other proteins are under progress.

In this study, we explain the preliminary results obtained from all the constructs and in detail for the BepA protein. Ultimately, obtaining structural information of these effector proteins will help us to investigate its functional importance in the eukaryotic host.

1.0 Introduction

1.1 Bacterial Secretion Systems

Bacteria capable of resisting the inborn immune response of the host typically cause an acute infection. In general, the pathogenic bacteria employ many devices to subvert the defenses of host cells by means of secretion machineries, mediating the transport and injection of effector molecules into the target cells. They have evolved sophisticated strategies to infect their eukaryotic hosts. As bacterial colonization of a given host often depends on extracellular proteins, the investigation of bacterial protein secretion is a major focus of microbiological research. Secretion is generally defined as active protein transport across the bacterial membrane(s), whereas translocation refers to the transport of proteins across the host cell membrane. So far, six major protein-secretion pathways have been identified [1] which are described as follows.

Autotransporters: Autotransporters (also known as type V secretion systems), need the Sec mechanism for mediating the transport of a passenger domain across the outer membrane (OM). The C-terminus of the protein forms a pore by building a β-barrel in the OM, through which the rest of the protein (passenger domain) is translocated. The N-terminus remains anchored to the C-terminal β-barrel in the membrane or are released by proteolytic cleavage. IgA1 from *Neisseria gonorrhoea* is the prototypical member of this family.

Chaperone / usher pathway: Secretion of the chaperone / usher pathway requires a chaperone and OM protein, named as usher. They are dedicated to the transport of pilus subunits, which are assembled at the bacterial surface. The prototypical organelles assembled by this pathway are P and type I pili of the uropathogenic *Escherichia coli*.

Type I Secretion Systems: Type I or ATP-binding cassette (ABC) protein exporter is a Sec-independent pathway for mediating substrate(s) directly from the bacterial cytoplasm to the outside of the cell [2]. The prototypical substrate of type I exporter is α-hemolysin of *Escherichia coli*.

Type II Secretion Systems: The more complex type II system is a Sec-dependent,

consisting of 12 to 16 accessory proteins which are found associated with the inner membrane (IM). Four IM proteins are proposed to form a pilus-like structure that could act as a piston for mediating the transport of extracellular enzymes and toxins through the OM pore. The prototypical example is the secretion of pullulunase (pul) by *Klebsiella oxytoca* [2].

Type III Secretion Systems: Type III system is a Sec-independent secretion system which shows virulence mechanism for various Gram-negative bacteria. They are multicomponent system which injects protein into the cytosol of eukaryotic cells, where the translocated proteins facilitate bacterial pathogenesis. The secretion of Yersinia outer proteins (Yops) by *Yersinia spp*. represents the prototypical pathway [3].

Type IV Secretion Systems: Bacterial Type IV Secretion Systems (T4SS) are multi-component systems referred to as prokaryotic nanomachines which mediate the transfer of macromolecular substrates into various target cells, e.g., the transfer of DNA into bacteria or the transfer of effector proteins into eukaryotic target cells. Type IV transporters are one of the best studied among the families of transporters, which are ancestrally related to conjugation systems of Gram-negative bacteria [4]. In bacterial conjugation systems exchange of genetic material occurs between two mating bacteria. This interesting phenomenon leads to the spread of antibiotic resistance genes among pathogenic bacteria. Bacterial T4SS are present in a wide range of bacterial species causing pathogenicity by interacting with many different hosts; they translocate their effector molecules into virtually every cell type – prokaryotic, yeast, filamentous fungi, plant or mammalian system. The architecture of Type IV secretion pathway and their clinical implications in the host cells are discussed in detail in the following paragraphs.

1.2 Molecular architecture of Type IV tranporters

Bacterial T4SS are multiprotein complexes formed by eleven proteins, for example VirB1 to VirB11 in *Agrobacterium tumefaciens*, the bacterium that harbors the prototype of the family of type IV transporters. Two types of factors, namely the coupling proteins (CPs) and chaperones are found necessary for translocating DNA and protein substrates into the target cells. The CPs are ubiquitous components of T4SSs machines and play a crucial role in translocation of effector molecules. Nester and colleagues made the

fascinating discovery that at least a subset of VirB proteins forms a long pilus structure [4] (Figure 1).

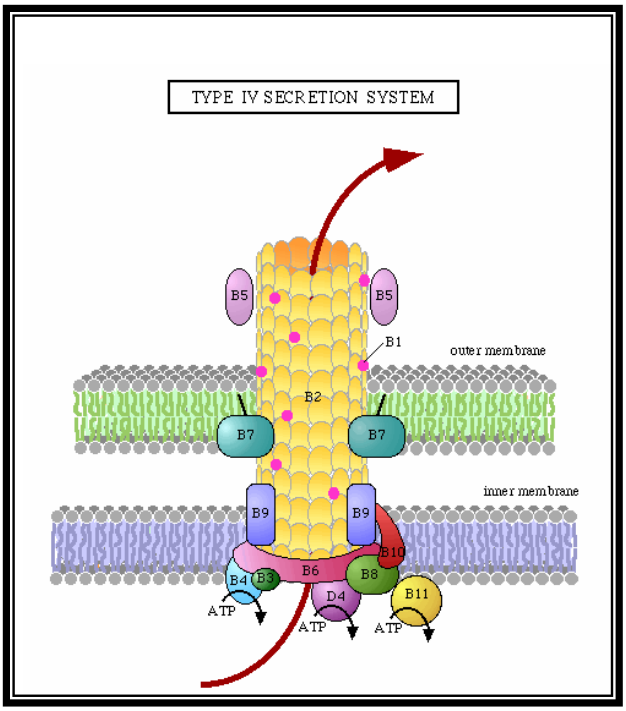


Figure 1: Schematic representation of the putative structure of the type IV secretion pathway depicting the different components of the VirB system. The scheme is drawn with the conjugative pilus of *E. coli* and *A. tumefaciens* as a prototype. The putative pathway is denoted with a red arrow. The figure is adopted from the webpage: http:// www.genome.jp/kegg/pathway/map/map03080.html

Ten out of the eleven VirB proteins (VirB2 - VirB11) may associate to form the pilus structure and associated transport apparatus that spans both the inner and outer membranes, to the outside of the cell. The basic components of T4SS are the core complexes (VirB1 and VirB6 to VirB10), pilus subunits (VirB2 and VirB5), the cytoplasmic ATPases (VirB4 and VirB11) and the coupling protein (VirD4) [5]. These basic components associate to form a channel across the bacterial envelope for macromolecule secretion. The pili or fimbriae are surface filaments of varying diameters and lengths. Specific homogenous protein subunits called as the pilin proteins (pilin interlock) of the T4SS form the rigid, helical, tube like pilus. Pili are composed of a single pilin subunit that forms a helical filament of ~8 – 16 nm in diameter. They are

present outside of bacteria and are responsible for attaching bacteria to specific tissues of the host organism that they infect. Typically, these proteins are synthesized as proproteins with an unusually long signal peptide of 25 - 50 residues [6]. Processing to mature pilin requires the action of cellular factors and often one or more Tra proteins [7]. Mature pilin is thought to localize initially at the inner membrane, where, upon the action of an unknown signal, it is recruited for pilus assembly. Yeast two-hybrid studies (YTHS) and complementary biochemical assays have shown the protein-protein interactions between different VirB proteins forming the VirB complex structure which ultimately forms the putative pathway for assembly of the transporter [8]. From the above studies, the process of substrate translocation in *A. tumefaciens* VirB/VirD4 system is presented as two models. They are the "channel" model in which the pilus acts as a channel for passage of the substrate across the cell envelope, and the "piston" model in which the pilus acts as a piston motor, pushing the substrates to the medium or into the eukaryotic cell [5] (Figure 2).

1.3 Types and functions of different T4SS

T4SS are multicomponent transporters of Gram-negative bacteria. This versatile translocation family can be classified into three subfamilies [5] (Figure 3). The largest subfamily, 'the conjugation systems', is found in most species of Gram-negative and Gram-positive bacteria. These systems mediate DNA transfer both within and between phylogenetically diverse species. The second subfamily is the 'DNA uptake and release systems' functions independently of contact with a target cell. The third subfamily is designated as 'effector translocator systems' which are resemblance of the Type III secretion systems, where they deliver the substrates directly in to the eukaryotic target cell. Type IV transporters are produced by several pathogens such as *Agrobacterium tumefaciens*, *Bordetella pertusis*, *Brucella spp.*, *Bartonella spp.*, *Helicobacter pylori* and *Legionella pneumophilia*. These transporters are critical for their pathogenic process. The T4SS produced by these pathogens are evolutionary and structurally related but their characteristic effector molecules are highly diverse. Intracellular delivery of solely protein substrates (effector proteins) subverting host cell function is considered as the primary function of T4SS in human pathogenic bacteria.

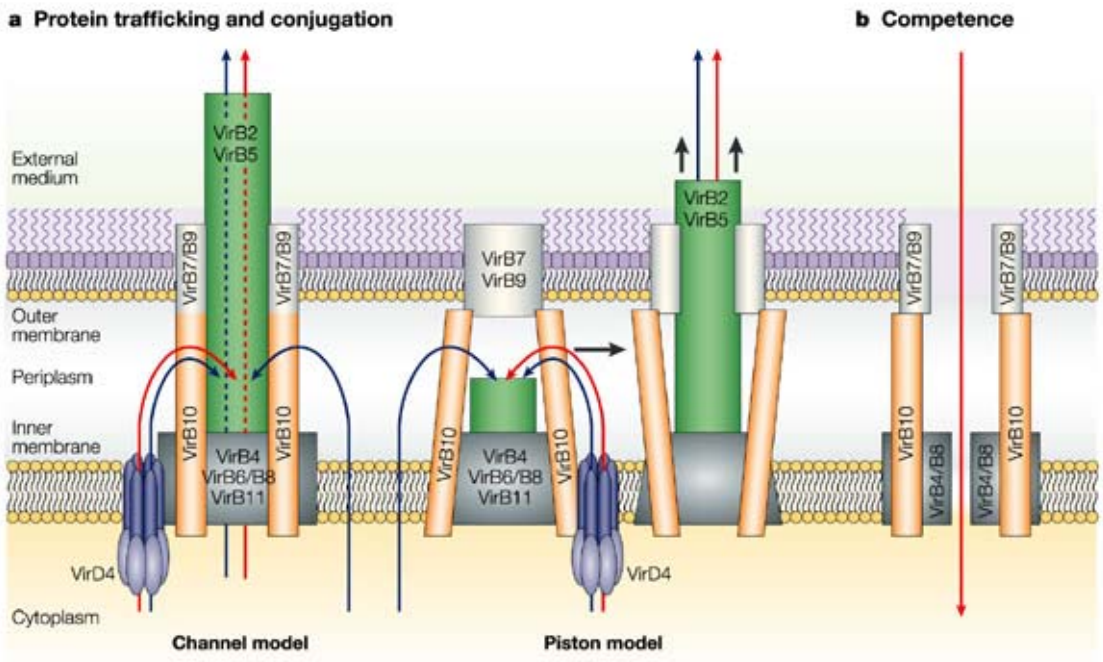


Figure 2: The models of Type IV secretion systems (T4SSs) mediated substrate translocation. The *A. tumefaciens* VirB/VirD4 system is presented as a model. (a), Two working models are depicted – the "channel" model in which the pilus acts as a channel for passage of the substrate across the cell envelope, and the "piston" model in which the pilus acts as a piston motor, pushing the substrates to the medium or into the eukaryotic cell. The blue arrows represent protein substrate translocation and red arrows represent DNA substrate translocation. DNA and protein substrates might be translocated through the same or different pathways, using the coupling protein (VirD4), the mating-pore-formation (Mpf) complex. (b) The competence model is shown as a comparison (adapted from Cascales and Christie, 2003) [5].

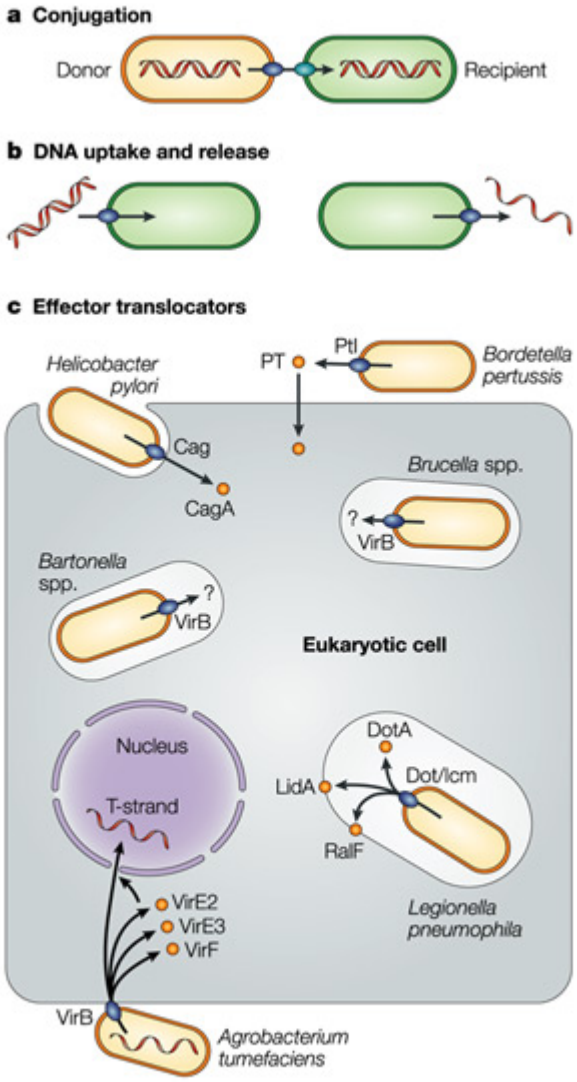


Figure 3: Schematic representation of the different type-IV-dependent mechanisms. The three subfamilies of type IV secretion systems (T4SSs) are shown. Conjucation machines donate DNA to recipient bacteria and other cell types by cell-to-cell contact. DNA uptake and release systems exchange DNA with the extracellular milieu independently of contact with target cells. Effector translocators deliver DNA or protein substrates into the eukaryotic cells during infection. The effector molecules for different pathogenic bacteria are shown in the figure. The figure was taken from the review article (adapted from Cascales and Christie, 2003) [5].

Here are some of the examples of bacterial pathogens which possess the T4SS is discussed below.

- *Agrobacterium tumefaciens*: The best studied member of this transporter family is the VirB/VirD4 system served as a prototype for T4SSs. They are responsible for the transfer of oncogenic T-DNA and several effector proteins (VirD2, VirE2 and VirF) to the nuclei of plant cells, leading to disruption in the levels of plant growth hormones. This often incites crown gall disease in the plant which can destroy agriculturally important crop species. The known effector proteins, VirD2, VirE2 and VirF contribute virulence in various ways. VirD2 is a relaxase required in the bacterium for nicking the T-DNA border repeat sequences and remains covalently associated with a single-stranded T-DNA (T-strand) destined for transfer across the bacterial envelope [9]. VirE2 is a single-stranded DNA-binding protein (SSB) transferred to plant cells independently and there they bind co-operatively to the VirD2 T-strand particle (T-complex) and delivers to the plant nucleus. The function of VirF is still unknown. The VirB system of *A. tumefaciens* shows divergence of T4SS transporters, a switch from systems involved in transfer of DNA strictly from bacteria to bacteria (conjugation) to systems involved in interkingdon transport of DNA [10].
- *Bordetella pertussis*: The T4SS of *Bordetella pertussis* is the causative agent of the childhood respiratory disease whooping cough. It secretes a pertussis toxin (PT), a multicomponent protein complex composed of five subunits. The subunits of PT are secreted *via* the general sec-mechanism [2] into the periplasm, where they are assembled into the holotoxin. Assembly of holotoxin in the periplasm is required for recognition and export across the outer membrane by the Pertussis toxin liberation (Ptl) system. The Ptl system is unique among the known effector translocators where delivery of the toxin is mediated by cell-contact-independent mechanism [11, 12].
- *Helicobacter pylori*: The pathogenesis of *Helicobacter pylori*-associated disorders is strongly dependent on a specialized T4SS encoded by Cag pathogenicity island (PAI). Cytotoxi-associated gene A (CagA: 145 kDa) is the only known

Helicobacter pylori protein translocated into the cytoplasm of mammalian cells which causes gastric ulcers. Secretion and translocation of CagA into gastric epithelial cells is dependent on the PAI and the VirD4-like coupling protein HP0524. Upon transfer, CagA gets tyrosine phosphorylated through host protein kinases and induces changes in host cell signal transduction and cytoskeletal rearrangements [13]. It infects more than 50 % of the human population causing gastritis, ulcer and gastric carcinoma.

- *Legionella pneumophila* (LP): *Legionella pneumophila* is the causative agent of legionnaire's disease, is a facultative intracellular pathogen. In the lung, they survive within host macrophages where it induces pore formation and enters a vacuole. The infection cycle involves host-cell entry by phagocytosis by creation of a specialized vacuole for replication, macrophage lysis and infection of neighboring cells [14]. The genes required for intracellular growth have been named Dot (defective organelle trafficking) and Icm (Intracellular multiplication). The Dot / Icm system functions as a conjugation system. The Dot / Icm system is implicated in controlling phagosome trafficking by creating a niche for intracellular survival. So far, few Dot / Icm substrates (LidA, RalF) have been identified which has been proposed to function in vesicle recruitment during biogenesis of the replication vacuole[14].
- *Brucella spp.*: Brucella spp. causes infections in several animal species as well as in humans. It is considered to be an intracellular pathogen and is required for the establishment of an intracellular replication niche in macrophages. Although the substrate(s) for this transporter have not been identified, it has unique mechanism of survival and multiplication of the bacteria in both macrophages and epithelial cells [15].
- *Bartonella spp.*: More recently T4SS genes were found in several species of the genus *Bartonella* which is similar to the VirB system of *A. tumefaciens*. The genus *Bartonella* comprises of an increasing number of human and animal pathogens. They cause long-lasting intraerythrocytic infections in their mammalian reservoir host and incidentally infect mammalian hosts. The *Bartonella spp.* which infects humans are *Bartonella bacilliformis* (Carrion's

disease), *B. quintana* (trench fever, bacillary angiomatosis, endocarditis) and *B. henselae* (cat-scratch-disese) [16, 17]. Common features of *Bartonella* include transmission by blood-sucking arthropods and the specific interaction with endothelial cells and erythrocytes of their mammalian hosts. *Bartonella*-endothelial cell interactions has been best studied for *B. henselae* which uses the VirB/VirD4 T4SS for delivering the effector molecules (Bartonella effector proteins, Beps) into the eukaryotic target cells which mediate subversion of endothelial cell functions [18].

In summary, the family of related T4SS has the ability to export macromolecules such as nucleoprotein complexes, multicomponent protein complexes or monomeric proteins across the Gram-negative bacterial cell envelope into bacterial, plant or mammalian cells. Recent work in the laboratory of Prof. C. Dehio (University of Basel) on animal and *in vitro* models has shown the molecular and cellular basis of *Bartonella* pathogenesis. They have identified two distinct T4SSs, VirB /VirD4 and Trw for causing pathogenicity in the eukaryotic hosts [19].

Among the various *Bartonella spp.*, *Bartonella henselae* was the best studied member of VirB/VirD4 system, which showed novel insights into the molecular and cellular basis of the Beps in the eukaryotic system where they interfere with the signaling pathways. The VirB/VirD4 mediated delivery system of *B. henselae* represents an essential virulence determinant for vascular endothelial cell infection both in mammalian reservoir hosts (i.e. for establishing intraerythrocytic infection) and in incidentally infected humans (i.e. for establishing the prominent vascular pathologies). Detailed *in vitro* studies have shown novel mechanisms of the trafficking of Beps into the host cell for various functions which are discussed in the upcoming paragraphs.

1.4 *Bartonella henselae*

Bartonella are Gram-negative oxidase-negative, fastidious bacteria belonging to the alpha2-subclass of *Proteobacteria*.

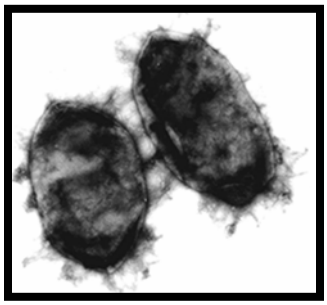


Figure 4: Image of *Bartonella henselae*

Bartonella henselae is a ubiquitous flea-borne feline pathogen with a seroprevalence of 3-6 % in blood donors in the USA and Australia [20, 21] (Figure 4). It is capable of infecting other animals and causing serious illness in humans. It causes an asymptomatic bacteremia in the feline reservoir, whereas it's an important pathogen for incidental human host. Pathological implications caused by this organism are discussed in the next paragraph.

1.5 Pathological disorders

Bartonella henselae is the causative agent of Cat-Scratch Disease (CSD) which is usually present as a benign lymphadenopathy in the proximal region of a cat's bite or scratch [22]. They are unique among pathogenic bacteria because of its painless nature of the infection and the poor response of the bacteria to antibiotic therapy. *B. henselae* escapes from the immune system by residing in the erythrocytes and is transmitted from cat to cat *via* the bloodsucking flea *Ctenocephalides felis* (Figure 5). Humans are infected incidentally, either directly by a bite / scratch of a cat or indirectly by a bite of the cat fleas. In immunocompetent hosts, the disease is self-limited to CSD. In immunocompromised hosts, the bacteria are often present in the blood and are involved in angioproliferative disorders such as human bacillary angiomatosis (vascular proliferation and angioproliferative tumour formation) [23, 24], hepatic peliosis (blood-

filled cysts in the liver and/or spleen), inflammation of lymph nodes and trench fever. Bacillary angiomatosis-peliosis is characterized by the formation of vasoproliferative tumors, which result from bacterial colonization and activation of human endothelial cells. The diagnosis of CSD is mainly obtained clinically (serological and histopathological examinations) and epidemiologically.

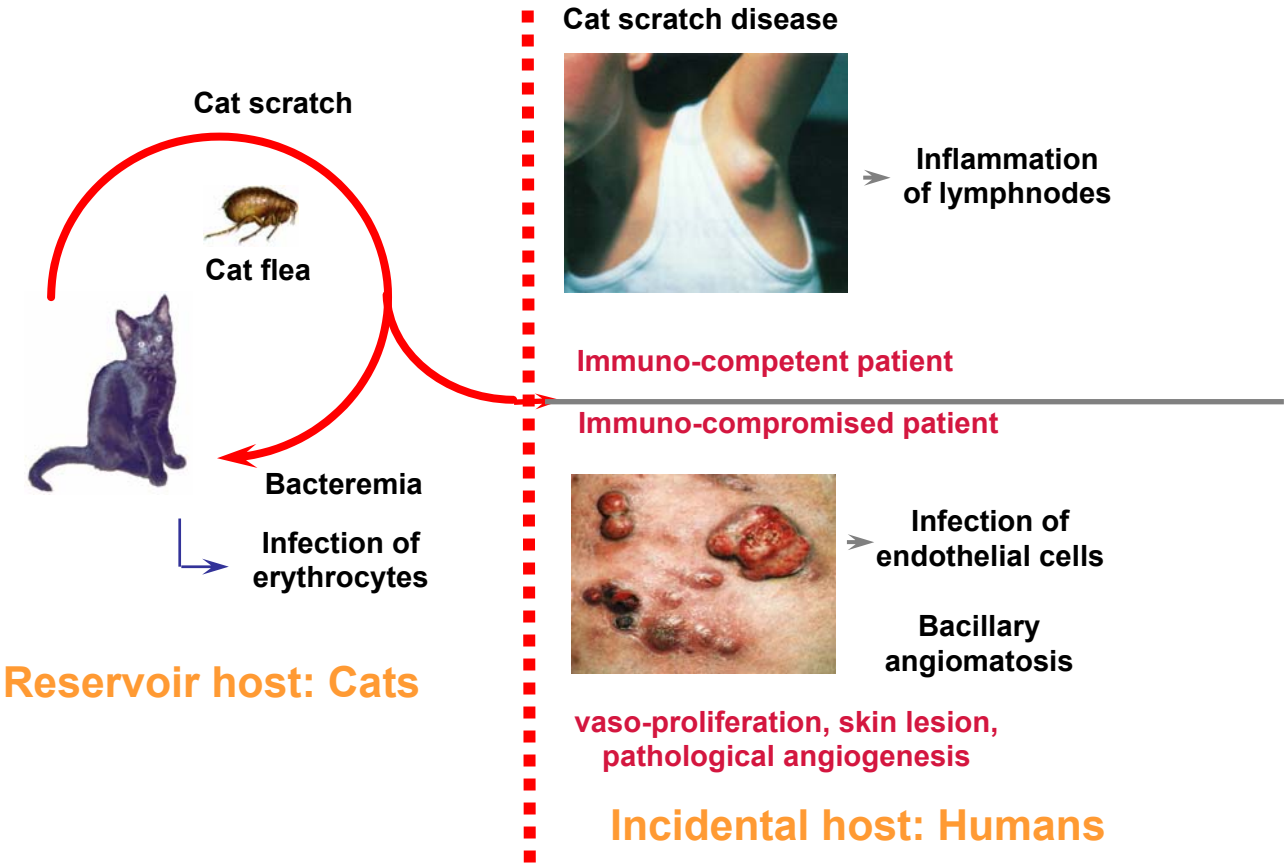


Figure 5: Infections caused by *Bartonella henselae:* On the left, the pictures shown are the routes of infections (infections of erythrocytes) caused in cats. On the right, shown are the diseases caused in immuno-competent (inflammation of lymph nodes) and immuno-compromised patients (Bacillary angiomatosis and pelliosis).

1.6 Role of Beps in VirB/VirD4 T4SS of *B. henselae* in the eukaryotic cell

The VirB/VirD4 T4SS apparatus of the *B. henselae* causes subversion of human endothelial cell function. This mechanism of action is mediated by means of Beps, the effector proteins which are responsible for their pathogenicity. The infection of vascular endothelial cells by *B. henselae* has been studied using Primary Human Umbilical Vein Endothelial Cells (HUVEC) as an *in vitro* infection model [16, 25]. The virulence attributes associated with *B. henselae* show different types of interactions to the host cells (Figure 6). These includes:

(i) Massive cytoskeletal rearrangements resulting in cell-surface aggregation and uptake of large bacterial aggregates by a defined structure termed the invasome (Figure 7A) [26]. Internalization *via* the VirB-dependent invasome represents a novel example for the invasion of bacteria into host cells and may serve as a cellular colonization mechanism for endothelial cells associated with the formation of vasoproliferative lesions [25].

(ii) Inhibition of cell apoptosis (anti-apoptosis) causes enhanced endothelial cell survival which leads to the formation of vasoproliferative lesions *in vivo* [27, 28]. This anti-apoptotic protection of HUVEC cells are shown by the suppression of the caspase activation and DNA fragmentation [26] (Figure 7B). The caspase-3 / 7 activity was measured of the infected HUVEC cells followed by induction of apoptosis by actinomycin D.

(iii) Induction of a proinflammatory phenotype by activation of nuclear factor-κB resulting in surface expression of the cell adhesion molecules (i.e. ICAM1 and E-selectin) [23] resulting in increased neutrophil rolling and adherence [26] (Figure 7C).

(iv) The VirB system of *B. henselae* mediates cytostatic and cytotoxic effects at high bacterial titres, which interferes with a potent VirB-independent mitogenic activity. This function is required for targeting multiple endothelial cells resulting in endothelial cell proliferation [25, 27] (Dehio et al., 2003) (Figure 7D).

These results have shown that the VirB/VirD4 T4SS represent an essential virulence determinant of *B. henselae* mediating subversion of endothelial cell functions. The

discovery of the VirB/VirD4/Bep PAI (Bartonella-specific pathogenicity) in *B. henselae* has led to the identification of seven VirB/VirD4-translocated effector proteins (BepA-BepG). Similar to VirB/VirD4, BepA-BepG is required for invasion, proinflammatory activation and antiapoptotic protection of HUVEC. Dehio and Co-workers have shown the VirB/VirD4/Beps dependent cellular phenotypes of HUVEC by performing deletion mutants of ΔvirD4 and ΔBepA-BepG. The results showed that the deletion constructs failed to trigger all the above mentioned phenotypes in HUVEC [26] These results have been characterized by performing a reporter (CRAFT: Cre Recombinase Assay For Translocation) assay which was used to monitor the translocation of NLS-Cre_Bep fusion proteins from *Bartonella* into the HUVEC cells. In summary, VirB-mediated phenotypes of HUVEC require the coupling protein VirD4 and the putative effector proteins BepA–BepG.

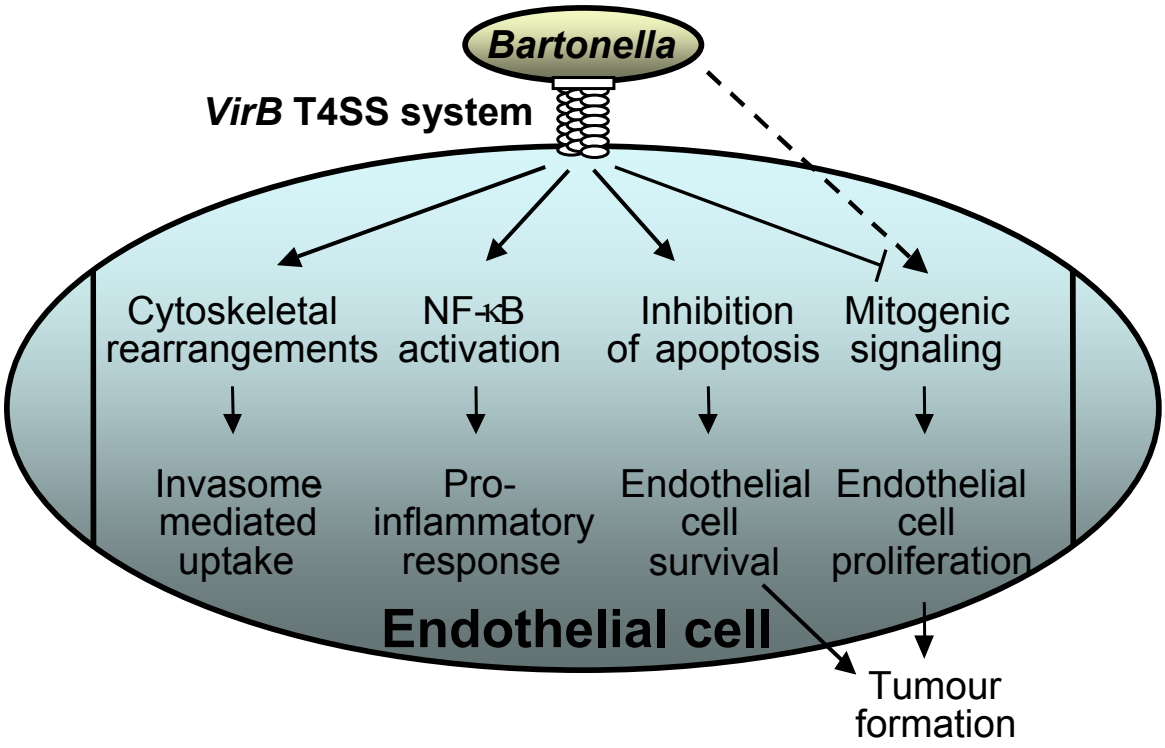


Figure 6: Scheme of the role of the *VirB*-T4SS in mediating *Bartonella henselae*-vascular endothelial cell interactions (adapted from Schmid et al., Mol Micro, 2004) [27].

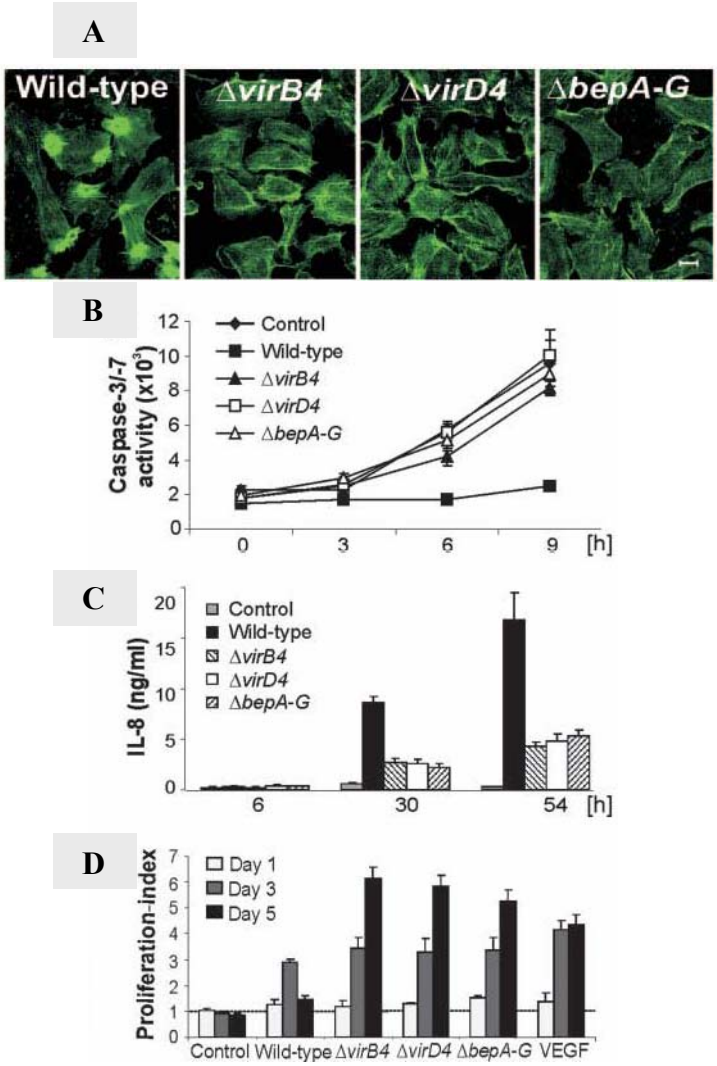


Figure 7: The Bartonella effector proteins (Beps) mediate VirB/VirD4-dependent invasion, antiapoptotic protection, proinflammatory activation and control of proliferation of HEC. (A) VirB4/VirD4/Beps are required for mediating actin rearrangements resulting in uptake of *Bartonella* aggregates by the process of invasomes. HUVEC infected with the indicated *Bartonella* strains were stained for F-actin (Scale bar, 10 µm). (B) VirB4/VirD4/Beps are required for antiapoptosis. Caspase-3/7 activity of HUVEC was measured after infection with the indicated *Bartonella* strains for 24 hours, followed by induction of apoptosis by actinomycin D for the indicated times. (C) VirB/VirD4/Beps are required for NF-κB-dependent proinflammatory activation. HUVEC were infected for the indicated time with the indicated *Bartonella* strains, followed by quantification of IL-8 in the culture medium. (D) VirB/VirD4/Beps are required for controlling *Bartonella*-stimulated HUVEC proliferation. HUVEC infected with the indicated *Bartonella* strains were counted at the indicated time points and proliferation indices were calculated (adapted from Schulein at al., 2005) [26].

1.7 Modular architecture of the effector proteins (Beps) in *B. henselae*

The Bep family of proteins consists of seven putative Bartonella effector proteins (BepA - BepG) which are encoded together with the VirB system and the coupling protein VirD4 on a Bartonella-specific PAI [26] (Figure 8). The Beps display a modular architecture, suggesting an evolution by extensive domain duplication and reshuffling. Each of them consists of different modular domains, the Fic domain (Filamentation induced by cAMP), pTyr domain (tyrosine phosphorylation) and an approximately 140 aa largely conserved domain termed BID (Bep Intracellular Delivery) which is present as one or several copies of the domain that is conserved among the Beps (Figure 9). Later, the BID domain was identified as one part of a translocation signal. All Beps harbor atleast one BID domain which is localized closely to the C-terminus with a short tail sequence. The sequence of this short C-terminal tail is highly positive charged and they are not conserved among the Bep family. Intriguingly, the C-terminus tail holds the other part of the transfer signal for mediating translocation of substrates *via* VirB/VirD4-dependent intracellular delivery. They were used for construction of reporter fusion proteins of Beps (B, C, D and F) which were shown to hold a transfer signal for the reporter assay (CRAFT) [24]. The whole translocation process is related to the BID like domains in the relaxases of rhizobial conjucation systems. The phylogenetic analysis of BID domains of the Beps and the relaxases suggests that *Bartonella* has adopted this putative secretion signal of relaxases for translocation process [19].

These interesting finding of the VirB/VirD4/Beps interactions of *B. henselae* in HUVEC system have initiated in-house collaboration with Prof. C. Dehio's group for structural analysis on these effector proteins (Beps). We started our work on the BepA-BepD constructs and the BID domains of BepA and BepB, since the functional characterization for each of these proteins was under investigation. Moreover, our interest was mainly on BepA, B and C, since they are the closest homologues having common domain architecture (Figure 9) and sharing a strong sequence homology of around 43–57 %. They contain an N-terminal Fic domain, BID domain and a partly conserved C-terminal tail. The putative role of Fic domain is set to be involved in the process of cell division [29] and it is remarkably conserved in many bacterial organisms.

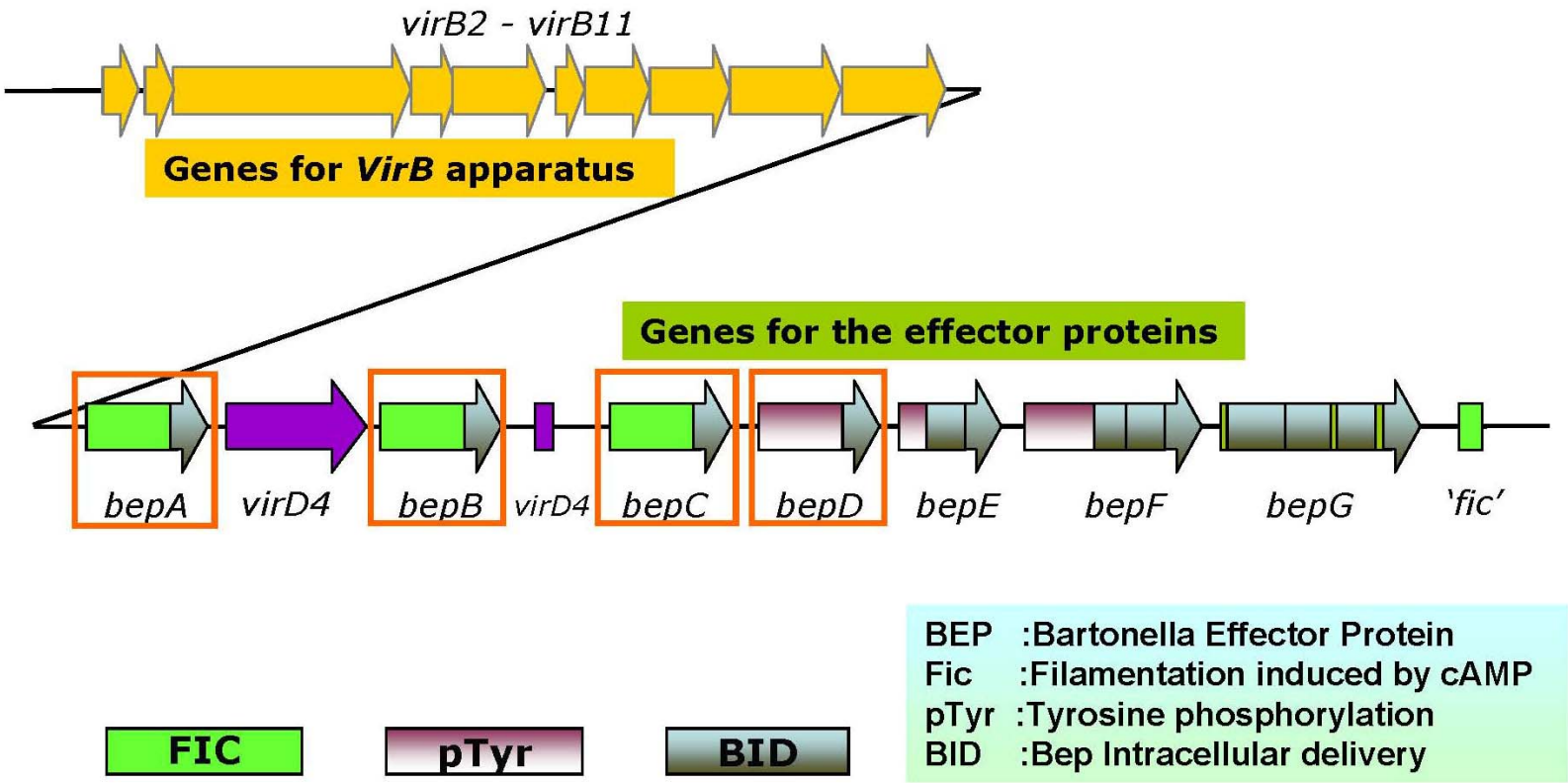


Figure 8: Schematic representation of the genetic organisation of the T4SS apparatus in *Bartonella henselae*, showing the genes involved in VirB (VirB2 – VirB11) apparatus and the downstream genes for the various Bartonella effector proteins (BepA-BepG). Among the Beps, BepA-BepD (shown in boxes) were the target proteins of our study.

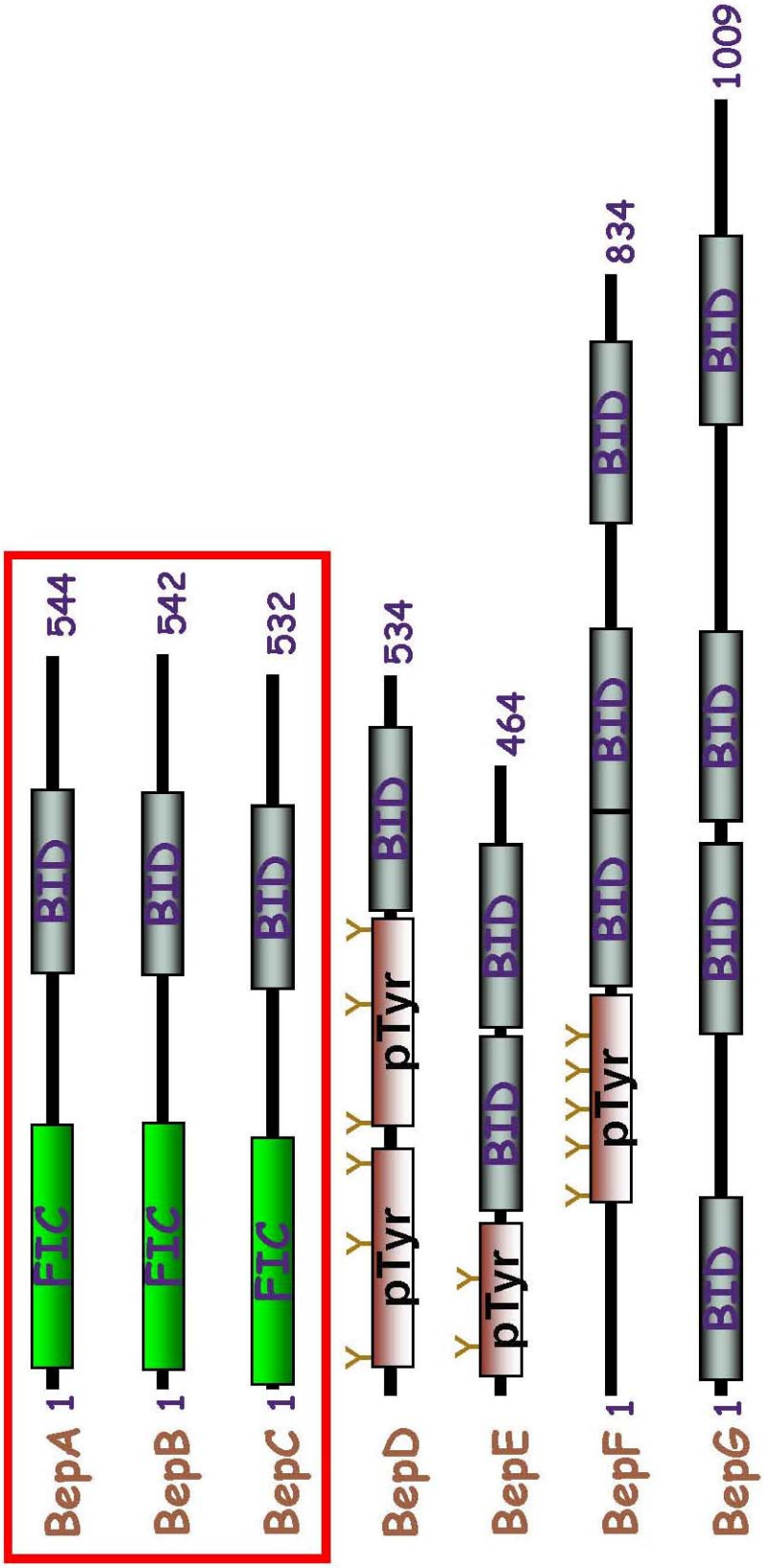


Figure 9: Schematic representation of the modular architecture of Beps which are composed of the Fic (green), BID (grey) and pTyr (red) domains. Among the seven Bep proteins shown, Bep A, B and C (box) exhibit an overall sequence identity between 43 – 57 %.

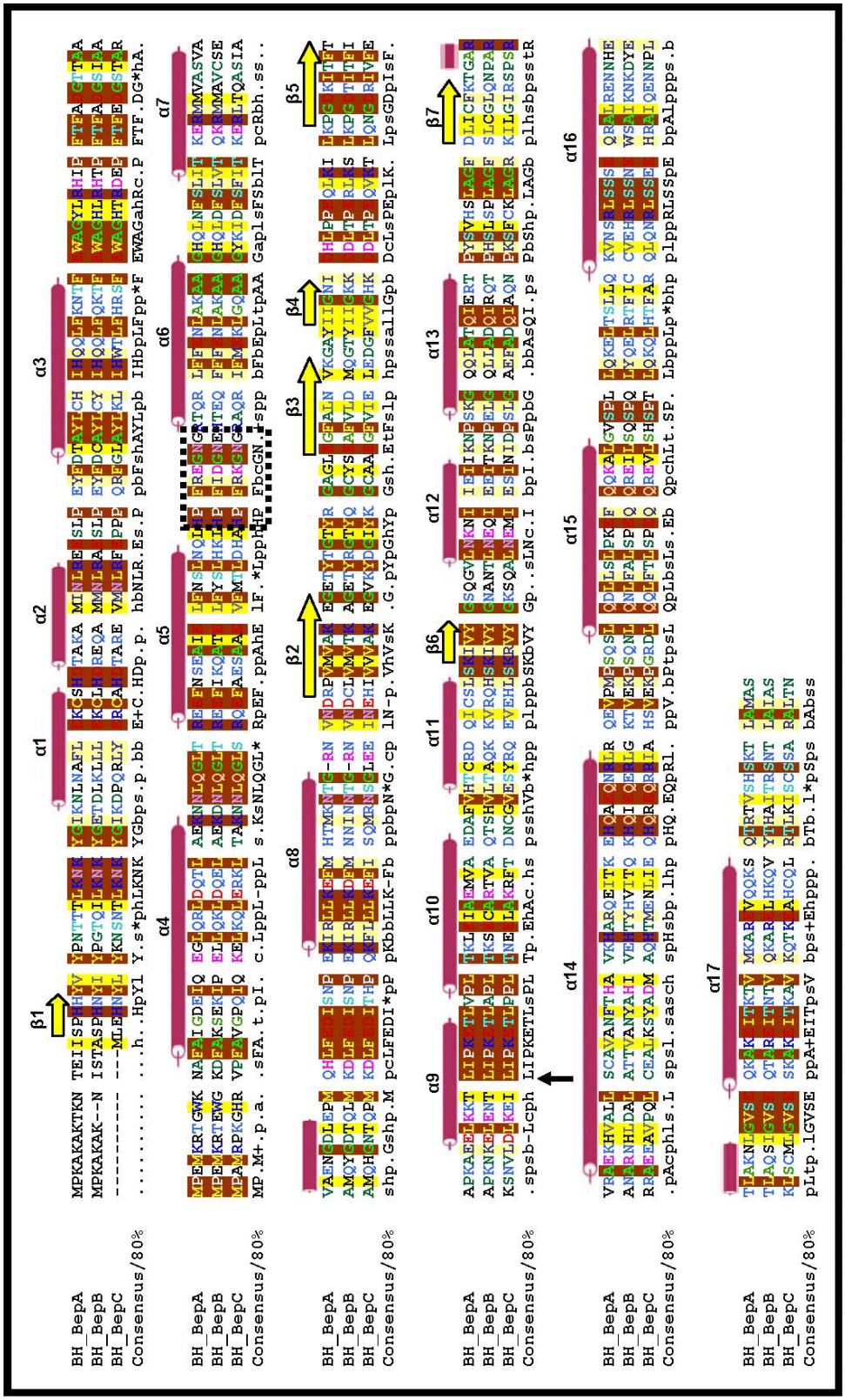


Figure 10: Multiple sequence alignment and secondary structure prediction of BepA, B and C from *Bartonella henselae*. The secondary structures predicted are shown at the top of the alignment which is obtained from the *JPRED* program [33]. Prediction was performed based on the sequence conservation of the three proteins. The strictly conserved residues are in red background. The Fic motif HPFXXGNG is well conserved in these proteins which are shown in box. The probable BID starting region is marked with a black arrow.

Secondary structure predictions for BepA, B and C based on sequence conservation showed a strong prediction of α-helices and few β-sheets (Figure 10). BepD is another candidate which we are interested in, since it has a different modular domain (pTyr) containing tyrosine-phosphorylation sites for the human tyrosine kinase c-src. We started working on these four proteins extensively to understand their characteristic behavior for structural analysis. Interestingly, recent results have revealed the functional phenotypes for BepA and BepD protein for their role in the eukaryotic system. BepA has been proven to perform anti-apoptotic function; in particular, the BID domain of BepA was responsible for performing the anti-apoptotic function apart from its translocation activity [30]. This was proved by performing various deletion constructs of Beps and delineation of BepA. The function of FIC domain is yet to be characterized.

The functional studies on BepD have been shown to get tyrosine-phosphorylated by host cellular kinases upon translocation into HUVEC cells [26]. These interesting observations for BepA and BepD gave us valuable information into this project. In this study, we report the structural insights into these Beps extensively, and, in detail, the modular structure and function of BepA.

2.0 Materials and Methods

2.1 Clones

BepA, B, C & D (full length) and individual BID domains of BepA & B were cloned in pET-15b vector and transformed in *Escherichia coli* strain BL21 (DE3) with fusion at the N-terminus with a hexahistidine affinity tag and a thrombin cleavage site (Schmid M. and Guye P., Prof. C. Dehio's laboratory, University of Basel). All the constructs were obtained from Prof. C. Dehio's lab for the structural studies (Table 1).

2.2 Overexpression of Bep proteins

Cell cultures were grown in LB medium with 0.1 mg/mL of ampicillin at 37 ^{o}C until they reached an OD_{600} of 0.6. Inductions was performed with IPTG and the concentrations used for each construct is described in the Table 2. The cells were then harvested at 5000 rpm (Sorval GSA) and washed in 1/25 culture volume of washing buffer (Tris pH 8.0 – 50 mM, NaCl – 150 mM) and subsequently pelleted and frozen for overnight storage at – 20 ^{o}C.

2.3 Purification scheme

Cells were thawed and resuspended with the lysis buffer (20 mM Tris HCl, pH 8.0, 500 mM NaCl, 5 mM β-mercaptoethanol, 10 mM imidazole, 0.5 mg/mL AEBSF, 'Complete' EDTA-free protease inhibitor cocktail tablets used according to the manufacture's instruction (Roche Diagnostics AG, Rotkreuz, Switzerland)) and lysed using French Press (FRENCH® PRESSURE CELL – 40,000 PSI, Thermo Electron Corporation, Switzerland). The cell lysate was clarified by centrifugation in a TFT 70.38 rotor at 45,000 rpm at 4 ^{o}C (Kontron Instruments, Switzerland). The supernatant was then loaded onto a 4 mL charged Ni-NTA Superflow™ resin (Qiagen AG, Switzerland) pre-equilibrated with the protein buffer (20 mM Tris pH 8.0, 500 mM NaCl, 5 mM β-mercaptoethanol, 10 mM imidazole).

Group	Construct	Residue range	No. of aa.	MW (Da)	pI	ε
Full-length	BepA	1 – 544	567	63878	9.11	25280
	BepB	1 – 542	566	64487	6.98	40710
	BepC	1 – 532	556	63616	8.95	25940
	BepD	1 – 534	557	61738	7.11	37100
BID	BID_A	305 – 544	268	29948	9.57	2800
	BID_B	303 – 542	268	30110	7.83	13610

Table 1: A summary of the physical properties of the various Beps and BID domain constructs of BepA & BepB. All the constructs were made with a hexahistidine fusion at the N-terminus. MW: theoretical molecular weight; pI: theoretical isoelectric point and ε: Theoretical extinction coefficients at 280 nm were calculated using the programme ProtParam from the ExPASy server [37]. The above constructs were obtained from Prof. C. Dehio's laboratory (Schmid M. and Guye P., University of Basel).

The protein was eluted on a Bio-Rad chromatography systems (Model 2110 Fraction Collector, Model 1327 Chart Recorder and Model EM-1 Econo UV Monitor) using a step-wise gradient of 20 – 250 mM imidazole solution. The eluted fractions were pooled and dialyzed against the gel filteration buffer (20 mM Tris pH 8.0, 500 mM NaCl, 1 mM DTT) using Spectra-Por 7 dialysis membrane of cut-off 12 – 14 kDa at 4 ^{o}C overnight. The protein solution was then concentrated to around 30 mg/mL using an Amicon Ultra device with a cut-off of 30 kDa (Millipore AG, Volketswil, Switzerland). The concentrated protein was spun at 14,000 rpm (Vaudaux-eppendorf, Centrifuge 5804 R, Switzerland) at 4 ^{o}C for 10 minutes, the sample was loaded on a Hiload^{TM} 26/60 Superdex^{TM} 75 (prep grade) gel filtration column (Amersham Biosciences Europe, Freiburg, Germany) pre-equilibrated with the gel filtration buffer. The gel filtration step was very crucial to obtain pure protein devoid of aggregates and degraded material. The protein was then concentrated in the same way as before for crystallization experiments.

2.4 Polyacrylamide gel electrophoresis, Silver staining and Western blot

Electrophoresis was performed using 12 % SDS-PAGE gels, according to the method of Laemmli. Protein bands were visualized with Coomassie Brilliant Blue staining and the molecular weight determination was done by referring to LMW-SDS markers as standard (Prestained SDS-PAGE Standards, Bio-Rad Laboratories, U.S.A. and BlueRanger® Prestained Protein Molecular Weight Marker Mix, PIERCE, U.S.A.).

Silver stain analysis was performed using the procedure from Morrisey et al. The SDS-gel was placed in a fixer solution (50 % methanol + 0.1 mL 37 % formaldehyde per 100 mL) for 1 hour at room temprature. The obtained gel was washed with distilled water containing 1M DTT for 10 minutes. Prepared fresh silver nitrate solution was added to the gel for 10 minutes on a shaker. A quick rinse with water and then added the developer solution (3 % Na_2CO_3 + 100 µL of 37 % formaldehyde per 100 mL). The bands appeared within 2 to 5 minutes and the reaction was stopped by adding a quenching solution (1 % acetic acid.).

For western blot analysis, protein samples were electroblotted onto a nitro-cellulose membranes (BA85, Schleicher & Schuell, Dassel, Germany) using the Bio-Rad Mini-PROTEAN 2 Electrophoresis/Mini Trans-Blot Module. Transfer was conducted in

presence of blotting buffer (20 mm Tris, 150 mM glycine and 20 % methanol) at 100 V (expected current: 350 mA) for 1 hour at 4 ^{o}C. The blot was subjected to immunoanalysis which was carried out according to the manufacturer's protocol of ECL western blotting analysis system (Amersham Biosciences, Freiburg, Germany) and Alkaline Phosphatase method (SIGMA®, U.S.A.)

ECL method: Briefly, non-specific binding sites on the blotting membrane were blocked with Tris-buffered saline (TTBS) (20 mM Tris-HCl pH 7.6, 150 mM NaCl, 0.1 % Tween) containing 5 % skimmed milk powder. The blot was washed for three times in TTBS buffer, each time for 10 minutes and then incubated in TTBS containing the primary antibody, monoclonal anti-polyhistidine antibody raised in mouse (Sigma Chemie, Buchs, Switzerland), at 1:3000 dilution for 1 hour at room temperature. The blot was washed as described before in TTBS buffer and then incubated in TTBS containing the secondary antibody, anti-mouse IgG (Fc specific) antibody-peroxidase conjugate (Sigma), at 1:5000 dilution. The blot was washed for three times with TTBS buffer and incubated in ECL western blotting detection reagents one and two (Amersham Biosciences, Freiburg, Germany) mixed at the ratio of 1:1 for one minute. The blot was subsequently exposed on KodakX-OMAT XAR-5 radiography film for 15 seconds to 15 minutes until the signal appeared.

Alkaline phosphatase method: The procedure was very similar to the above method except it has a different secondary antibody, Anti-Mouse IgG (Fc specific) antibody-alkaline phosphatase conjugate (Sigma®, U.S.A.) with a working dilution of 1:5000 with an incubation period of 1 hour at room tempearture. Three washes: 3, 5 and 10 minutes with TTBS buffer containing 10 % skimmed milk powder was performed before staining. The enzymatic substrate solution was prepared by dissolving one tablet of FAST (Sigma *FAST*^{TM} – BCIP/NBT) in 10 mL of water. The alkaline phosphatase enzymatic reaction was performed by adding the substrate solution to the blot. The bands appeared within a period of 1 to 10 minutes.

2.5 Limited proteolysis

Limited proteolysis treatment was performed only for purified BepA_tr samples for checking their domain solidity. BepA_tr protein at 1 mg/mL in protein buffer was

proteolysed by trypsin (Sequencing Grade Modified Trypsin, Promega, U.S.A.) at a ratio of 5000: 1 on ice. Aliquots of 9 µL were removed from the reaction mixture after 5, 10, 30, 60, 90 and 120 minutes for analysis. The reaction was stopped with 1 µL of 10 % AEBSF (final concentration 1 %). The aliquots were then subjected to SDS-PAGE and western blot analysis against his-tag.

2.6 Mass spectral analysis

Liquid chromatography (LC) / Mass spectrometric (MS) analysis of the Bep protein and its digest were carried out on 100 mm i.d. capillary columns with C_{18} material (5 mm particle size, MONITOR, Column Engineering, Ontario, USA). Bound peptides were eluted with a linear 30 min. gradient 0.05 % TFA to 60 % acetonitrile containing 0.05 % TFA at a flow rate of 1.5 mL/min. The microspray needles used for electrospray ionization were pulled from 100 µm i.d. × 280 µm o.d. fused silica capillaries (LC Packings, Amsterdam, Netherlands) on a Model P-2000 quartz micropipette puller (Sutter Instrument Co., Novato, CA). For LC/MS analysis, the needles were placed into an XYZ micropositioner, and the voltage was applied directly to the sample stream through the capillary union. Spray voltages were usually between 1100 and 1400 V. Mass determinations were carried out on a TSQ7000 triple quadrupole mass spectrometer (Thermo Finnigan, San José, CA). All measurements were carried out in the positive ion mode. Precursor ion scanning was between 200 and 2000 Da m/z for 3 s at unit resolution. For operation in the MS / MS mode, the resolution of Q1 was set to transmit a mass window of 4 Da and the resolution of Q3 was adjusted to 1.5 Da. Scanning was between 50 and 2250 Da in 3.5 s. Argon was used as a collision gas at a pressure of 3.5 mTorr.

MALDI-TOF spectral analysis was performed using a Bruker REFLEX III mass spectrometer (Bruker Dalatonik Gmbh, Bremen, Germany). 2 µL of BepA_tr sample was mixed with 1 µL of matrix solution (10 mg/mL of α-cyano-4-hydroxycinnamic acid (Aldrich) in 80 % (v/v) acetonitrile, 0.1 % (v/v) trifluoroacetic acid) and placed on the sample plate to dry. Calibration of the instrument was done with bovine serum albumin (BSA, Sigma Chemicals) by using the molecular masses of the singly, doubly, and triply charged form of BSA. Mass measurements of BepA_tr were carried out in the linear

mode. All mass spectrometric measurements were performed by the group of Dr. Paul Jenoe, Biozentrum, Basel.

2.7 Mass fingerprinting

Protein bands of the SDS-gel corresponding to the molecular weight of the target protein were cut and were subjected to trypsin (Sequencing Grade Modified Trypsin, Promega, U.S.A.) digestion. Trypsined fragments belonging to each intermediate were subsequently subjected to matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometric analysis. MALDI-TOF mass spectra were acquired on a Bruker Reflex III instrument (Bruker Daltonik, Bremen, Germany). Peptides were analysed either in linear or in reflector mode by using a-cyano-4-hydroxycinnamic acid (1 mg/mL in 80% acetonitrile / 0.1 % TFA) matrix. Samples were prepared by mixing 1 mL peptide solution with 1 mL matrix solution and 300 nL were deposited onto anchor spots of a Scout 400 mm / 36 sample support (Bruker Daltonic, Bremen Germany). The droplet was left to dry at room temperature. The instrument was calibrated with angiotensin II, substrate P, bombesin, and ACTH-18-39. For each proteolysis intermediate, the molecular weights of the ejected species were measured and searched in the sequence library of the theoretically trypsinised Bep proteins using the programme MASCOT. These fragments were used to map the boundary of the proteolysis intermediate in the protein. The mass finger-printing analysis was carried out by the group of Dr. Paul Jenoe, Biozentrum, Basel.

2.8 Absorption spectroscopy

Protein concentrations were determined with absorption spectroscopy (Hewlett Packard 8453, HP845X UV-Visible system, Switzerland) at 280 nm. The extinction coefficients ε of the constructs were calculated from the amino acid sequence and listed in Table 1.

2.9 Analytical ultracentrifugation

Analytical ultracentrifugation analysis was performed for purified BepA_tr samples in the protein buffer. Sedimentation velocity (SV) measurements were carried out in a Beckman XLA analytical ultracentrifuge equipped with absorption optics. The SV runs

were performed at 54,000 rpm at 20^{o}C using a 12 mm double sector cell. All analytical ultracentrifugation experiments were performed by Ariel Lustig, Biozentrum, Univeristy of Basel.

2.10 His-tag cleavage using proteolytic enzymes

His-tag cleavage experiments were performed with two proteolytic enzymes; Thrombin (Amersham Biosciences, Switzerland) for BepA_tr and TAGZyme (Qiagen AG, Switzerland) for BepB.

2.10.1 His-tag cleavage of BepA_tr:

Purified BepA_tr was subjected to thrombin digestion (2 U – twice the concentration mentioned by the manufacture) in presence of 2 mM $CaCl_2$ for 6 hours at 4 ^{o}C. His-tag cleavage was observed by means of a gel shift procedure. For separating the uncleaved from cleaved BepA_tr, the digested sample was subjected to 2nd Ni-NTA column. Cleaved BepA_tr was subjected to pass through the Benzamidine sepharose^{TM} 4 Fast Flow column (Amersham Biosciences, Switzerland) to remove the thrombin. Cleavage was confirmed with the western blot method. The fractions were collected and the buffer exchange was carried out with protein buffer (20 mM Tris pH 8.0, 500 mM NaCl, 1 mM DTT) to remove the imidazole. This step was carried out in the concentrating unit and the protein was concentrated in the same way as before for crystallization.

2.10.2 His-tag cleavage of BepB:

Purified BepB from Ni-NTA column was subjected to buffer exchange by performing dialysis with TAGZyme Buffer (20 mM Potassium phosphate pH 7.2, 500 mM NaCl, 5 mM β-mercaptoethanol) before TAGZyme (DAPase : diamino peptidase) digestion. Digestion was performed with TAGZyme (100 mU, twice the concentration mentioned by the manufacture) for 5 hours at 4^{o}C with gentle agitation. Separation of cleaved BepB was done in a similar way as for BepA_tr. TAGZyme was efficiently removed during the 2nd nickel column, since it possess his-tag at the C-terminal end. Cleaved BepB was concentrated to 20 mg / mL and loaded on a Hiload^{TM} 16/60 Superdex^{TM} 75 (prep grade) gel filtration column (Amersham Biosciences Europe, Freiburg, Germany) pre-equilibrated with the gel filtration buffer (20 mM HEPES pH 7.5, 500 mM NaCl, 1 mM

DTT). The peak corresponding to cleaved BepB were pooled and concentrated to a final concentration of 16.2 mg/mL for crystallization trials.

2.11 Crystallization

Crystallization condition for cleaved BepA_tr was identified from the NCCR (National Center of Competence in Research; Prof. Markus Grutter's lab, NCCR groups) crystallization robotics technique (Cartesian dispensing systems - MicroSys 4000 from Genomic Solutions, U.K.) using the crystal farm (Crystal Farm Imaging System CF-400, U.S.A.) at 20^{o}C. The crystal plates were scanned using Plate Handler Server Crystal Farm (Version, v3.4) and images were viewed using the built-in software Crystal Navigator web based software (Imager server version 1.13). Crystals were obtained at room temperature by vapour diffusion using the hanging and sitting drop method. BepA_tr at a concentration of 17.5 mg/mL (assuming an ε_{280} of 23140 M^{-1} cm^{-1}) in 20 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM DTT was mixed with the reservoir solution (22 % polyethylene glycol 4000 (PEG 4K), 100 mM Tris pH 8.7, 10 mM $NiCl_2$) at a ratio of 1 : 1.5.

2.12 Heavy atom derivatives

Heavy atom soaking experiments were performed for cleaved BepA_tr crystals with 3 heavy atom derivatives (Potassium platinum chloride K_2PtCl_4 (Pt); methyl mercuric acetate $CH_3HgOOCCH_3$ (Hg) and samarium (III) acetate; $Sm(O_2C_2H_3)$ (Sm)) at 10 mM and 5 mM respectively, in presence of the stabilizing solution (24% PEG 4000, 100 mM Tris pH 8.7, $NiCl_2$ 10 mM) for a period of 5 to 6 hours. Fluorescent spectra of the crystals were performed on the X-ray diffractometer (SLS beamline, X06SA (PX), Villigen, Switzerland) measurements were made for all the soaks. Excitation spectrums of the heavy atom soaked crystals were taken to find the peak and inflation region for subsequent MAD experiment.

2.13 Data collection

Data collection was performed for the crystals without any cryoprotectant since the mother liquour has cryo protecting properties (22% PEG 4K & 500 mM NaCl). The

crystals were flash frozen in liquid nitrogen. Diffraction data were collected from a single native crystal (about 70 microns in diameter). Crystals from heavy atom soaks were mounted in a similar manner; a MAD experiment was performed for the crystal soaked with 5 mM K_2PtCl_4 at two different wavelengths (0.900 λ and 1.071 λ). All the measurements were performed at the beamline X06SA (PX) of the Swiss Light Source (SLS), Villigen, Switzerland.

2.14 Data processing

Diffraction data were processed with IPMOSFLM (version 6.2.4) / SCALA (version 5.0, CCP4 suite) [31]. A Self rotation function was performed using the programme POLARRFN (version 5.0) to validate the spacegroup. HKLVIEW 2.5 programme was used to identify the extinctions (systematic absences) along the *h*, *k*, *l* axis.

2.15 Bioinformatic analysis

Sequence analysis of Bep family of proteins and their individual domains was performed using ClustalW, a multiple sequence alignment programme [32]. Secondary structure prediction for the homologous BepA, B and C proteins was performed using PHD predict protein [33] and *JPRED* [34] and the output file showing secondary structural elements along with the multiple alignment was generated by CHROMA programme [35]. The NORSp (Predictor of Non-Regular Secondary Structure) program was used for predicting the non-regular structures in the protein sequence [36].

3.0 Results

3.1 Bep constructs

For the study of the Bep proteins from *B. henselae*, we have applied the recombinant method of expressing the protein of interest in *Escherichia coli* strain BL 21(DE3). Four different Bep proteins were chosen namely, the BepA, B, C & D and the individual BID domains of BepA & B. N-terminal hexahistidine tags (His_6-tag) were chosen for affinity chromatography purification. For the removal of the His-tag, a thrombin protease cleavage site was inserted between the His_6-tag and the protein sequence (Figure 11). All the Bep constructs were provided by Prof. C. Dehio's laboratory (Schmid. M and Guye. P, University of Basel) (Refer: Table 1).

3.2 BepA

3.2.1 Expression and Solubility test

A small scale culture of full-length BepA was tested for the expression level and soluble yield as a function of growth temperature, time of induction and duration of expression. The bacteria were grown at 30 ^{o}C and 37 ^{o}C, induced at optical density (OD) at 600 nm of 0.6 for 1, 2, 3, 4 and 5 hours. The cells were lysed and pelleted to estimate the expression level and the yield of soluble fraction. The best yield of soluble BepA was obtained when grown at 30 ^{o}C and induced at OD_{600} for 5 hours (Table 2).

3.2.2 Purification

BepA was purified using Ni-NTA (IMAC) affinity and size exclusion chromatography. The supernatant of the cell lysate was first purified on a nickel affinity column using the batch method. Elution of the protein was performed by stepwise gradient at 20, 50, 100 and 250 mM imidazole concentration.

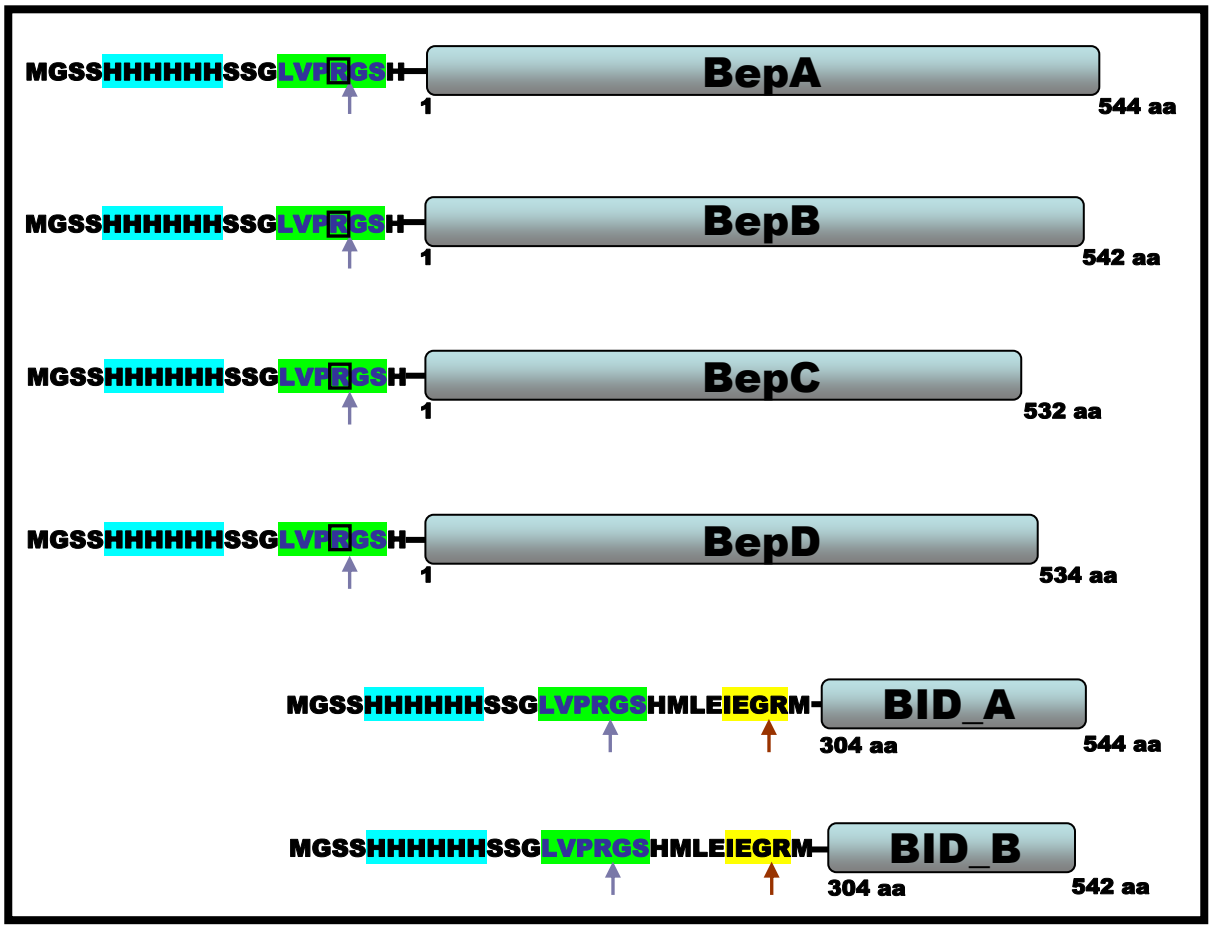


Figure 11: Design of N-His_6-constructs of BepA, B, C and D and BID domains of BepA and B; the hexahistidine tag is marked in blue, the thrombin cleavage site (violet arrows) and the Factor Xa cleavage site (red arrows) with their recognition sequence are marked in green and yellow respectively. The above constructs were obtained from Prof. C. Dehio's laboratory (Schmid M. and Guye P., University of Basel).

Constructs	Strain	Culture size	Induction temp. °C	Induction OD_{600} nm	IPTG conc.	Induction time	Yield / L culture
BepA	BL 21 (DE3)	2.5 L	30	0.54	200 µM	5 hrs	30 mg / L
BepB	BL 21 (DE3)	1.5 L	30	0.60	400 µM	4 hrs	35 mg / L
BepC	BL 21 (DE3)	1.0 L	30	0.57	400 µM	2 hrs	10 mg / L
BepD	BL 21 (DE3)	1.5 L	37	0.75	20 µM	6 hrs	40 mg / L
BID_A His_6	BL 21 (DE3)	1.5 L	37	0.55	500 µM	4 hrs	5 mg / L
BID_B His_6	BL 21 (DE3)	1.0 L	37	0.63	100 µM	4 hrs	20 mg / L

Table 2: An overall summary of expression studies on full length Beps and the BID domain constructs of BepA and BepB respectively. Cell cultures were grown in LB media in presence of ampicillin (antibiotic marker) at 37 ^{o}C.

BepA was eluted at 100 mM imidazole concentration. Unexpectedly, the SDS gel showed no band in the expected region for full length BepA (64 kDa) (Figure 12A). Instead, a strong band was found at the size of 40 kDa along with few contaminants. The Western blotting identified the His-tag fusion protein using the anti-histidine antibody. A strong signal was observed for the 40 kDa band which was identified as the truncated form of BepA (Figure 12B). In addition, a faint signal was found for the 35 kDa band (Figure 12B). The 35 kDa band that showed a weaker signal in western blot might be a degradation product. This result suggests that full-length BepA is expressed but cleaved during or after expression. In the following, this truncated version of BepA is called BepA_tr. In the second chromatography step, a Hiload 26/60 Superdex 75 gel filtration column was used for separating BepA_tr from the degradation product. This column has a fractionation range from 3 X 10^3 to 7 X 10^4 Da., and, thus, is well suited for resolving BepA_tr from the degradation product.

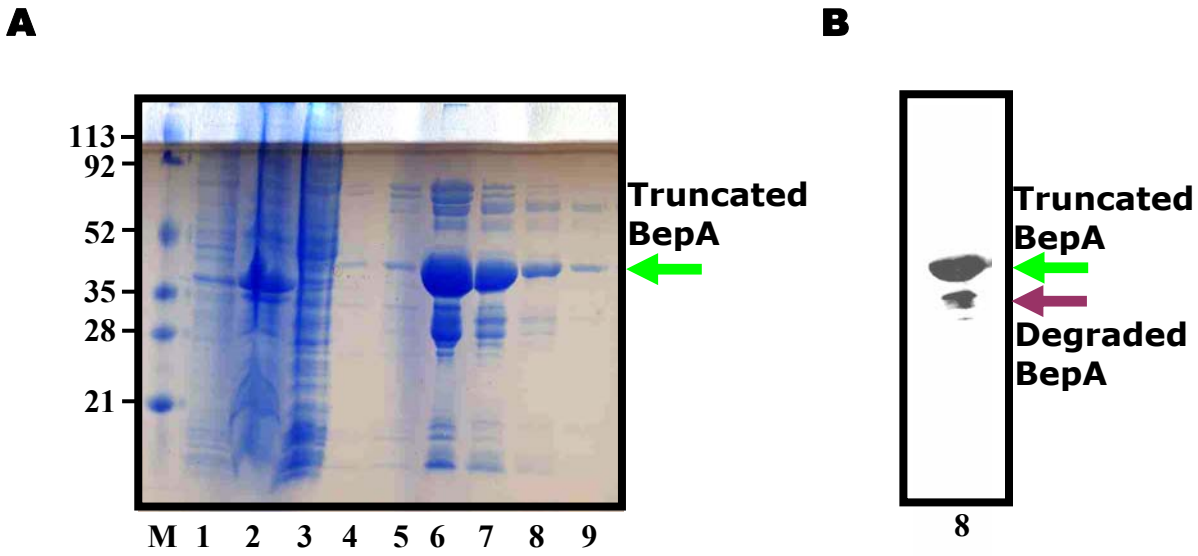


Figure 12: SDS-PAGE and Western blot analysis of BepA purification. (A) Cell lysate (1) was spun down to obtain the supernatant (2) and the pellet (3). The supernatant was purified on a nickel affinity column. Lane (4) and (5) show the wash with 20 mM and 50 mM imidazole, respectively, and lane (6) to (9) show the elution of the protein at 100 mM imidazole concentration. (B) Western blot (lane 8 from panel A) showed a positive signal (green arrow) for the main band. There was another signal below the main band which corresponds to a degradation product of BepA_tr (red arrow).

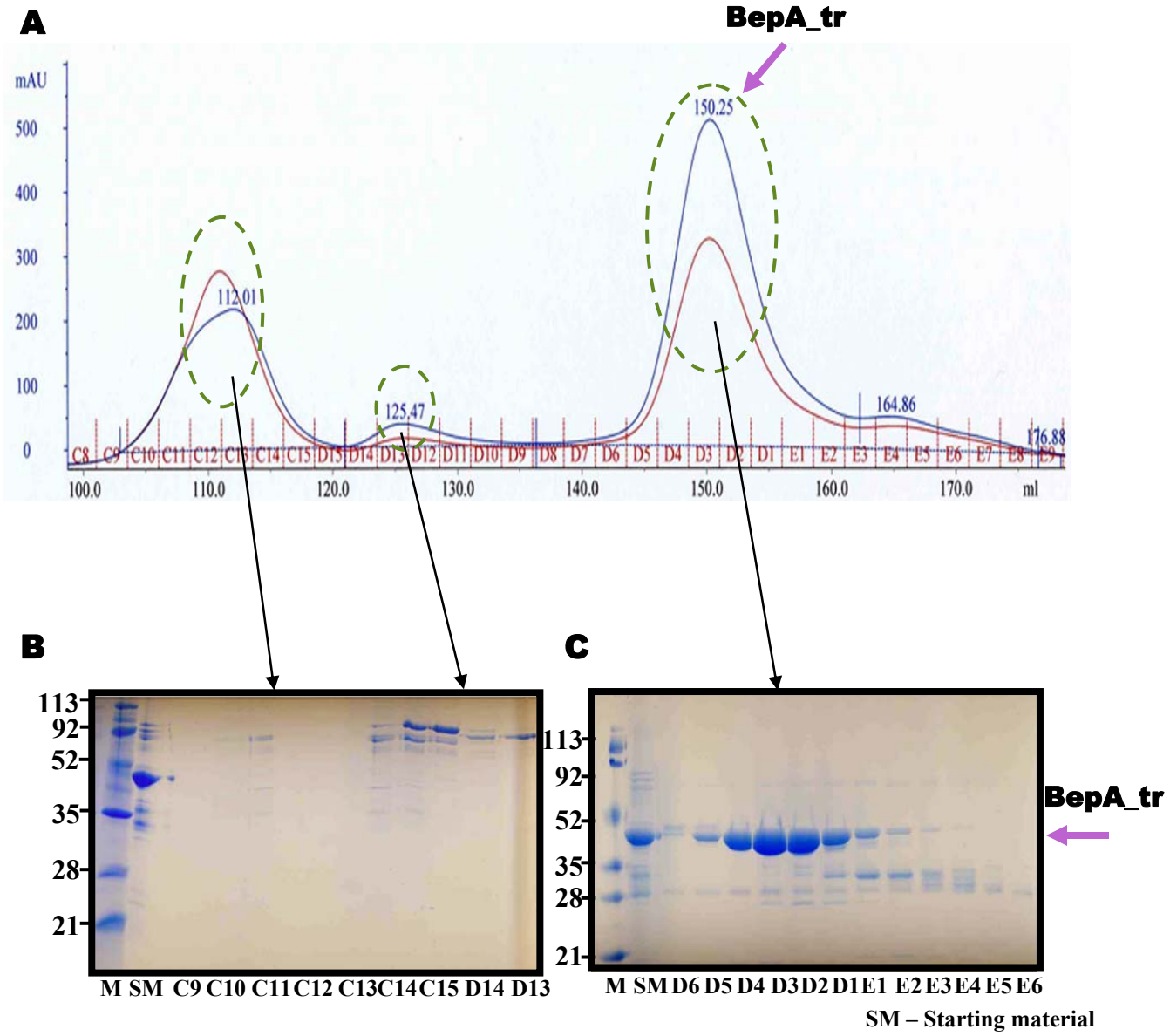


Figure 13: Purification of BepA_tr by size exclusion chromatography (A) Elution profile of BepA_tr on a Hiload 26/60 Superdex 75 gel filteration column (arrow). Elution peaks are labeled with their peak elution volume. (B) SDS-gel of the first and second elution peak fractions show only contaminants. (C) SDS-gel of the main peak fractions show highly purified BepA_tr (arrow).

BepA_tr was eluted as a single species as observed from the elution profile (Figure 13A) and SDS-gel showed 90 % purity (Figure 13C). The major peak fractions were checked by the Western blotting to observe the separation of BepA_tr from the degradation product (Figure 14). The fractions corresponding to pure BepA_tr devoid of degradation product were pooled and concentrated up to 20 mg/mL for the subsequent crystallization trials.

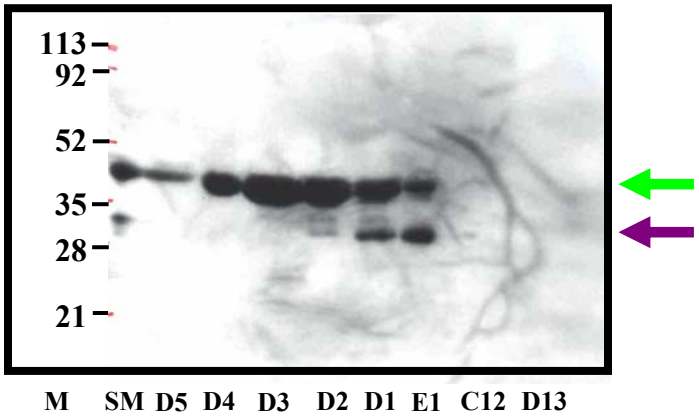


Figure 14: The Western blot analysis of purified BepA_tr. The Western blot picture shows the clear separation of the BepA_tr (green arrow) from the degraded protein (maroon arrow) in fractions D5, D4, D3 and D2.

3.2.3 Authentication and characterization of BepA_tr

Purified BepA_tr was confirmed using mass spectrometric (MS) analysis. Liquid chromatography-MS method was used for the molecular weight (MW) determination of BepA_tr which was found to be 39547 Da (Figure 15). Only a weak signal was obtained (Figure 15) not allowing for predicting the precise mass of BepA_tr. From the predicted mass, the sequence of the fragment was estimated by removing residues from the C-terminus of full length BepA until the theoretical mass matched the mass obtained from MS analysis (Figure 16).

BepA_tr was shown to be a monomer in solution from the purification step of gel filtration chromatography (Figure 13A). This was confirmed by analytical ultracentrifugation (AU) which showed that BepA_tr was homogenous with a calculated mass of around 37 kDa from sedimentation velocity measurements.

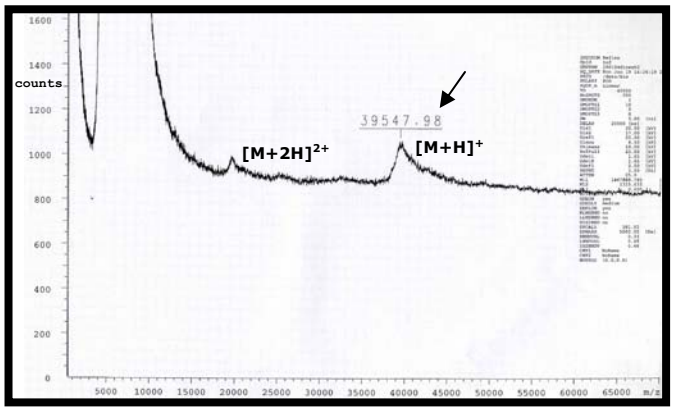


Figure 15: MALDI spectra of BepA_tr. The singly charged, protonated molecular ions, $[M + H]^+$ and the doubly charged ions, $[M + 2H]^{2+}$ are shown in the figure. The determined mass is indicated with an arrow.

3.2.4 Limited proteolysis

BepA_tr was subjected to limited trypsinolysis to check the domain stability. SDS-PAGE analysis of the digested samples shows a small shift of the band when compared to the control (Figure 15A). The corresponding Western blot of the SDS gel followed by immunological detection using anti-histidine antibodies showed no signal for the digested samples (Figure 15B). Thus, the result shows that the BepA_tr is stable upon tryptic digestion, which in turn indicates that the protein is folded and not accessible for proteolytic digestion. The observed gel shift of about 2 kDa is due to the action of the enzyme on the tryptic cleavage site (Figure 11) present in the N-terminal region of the protein. The tryptic digestion resulted in the removal of the His-tag as shown by a negative signal in the Western blot.

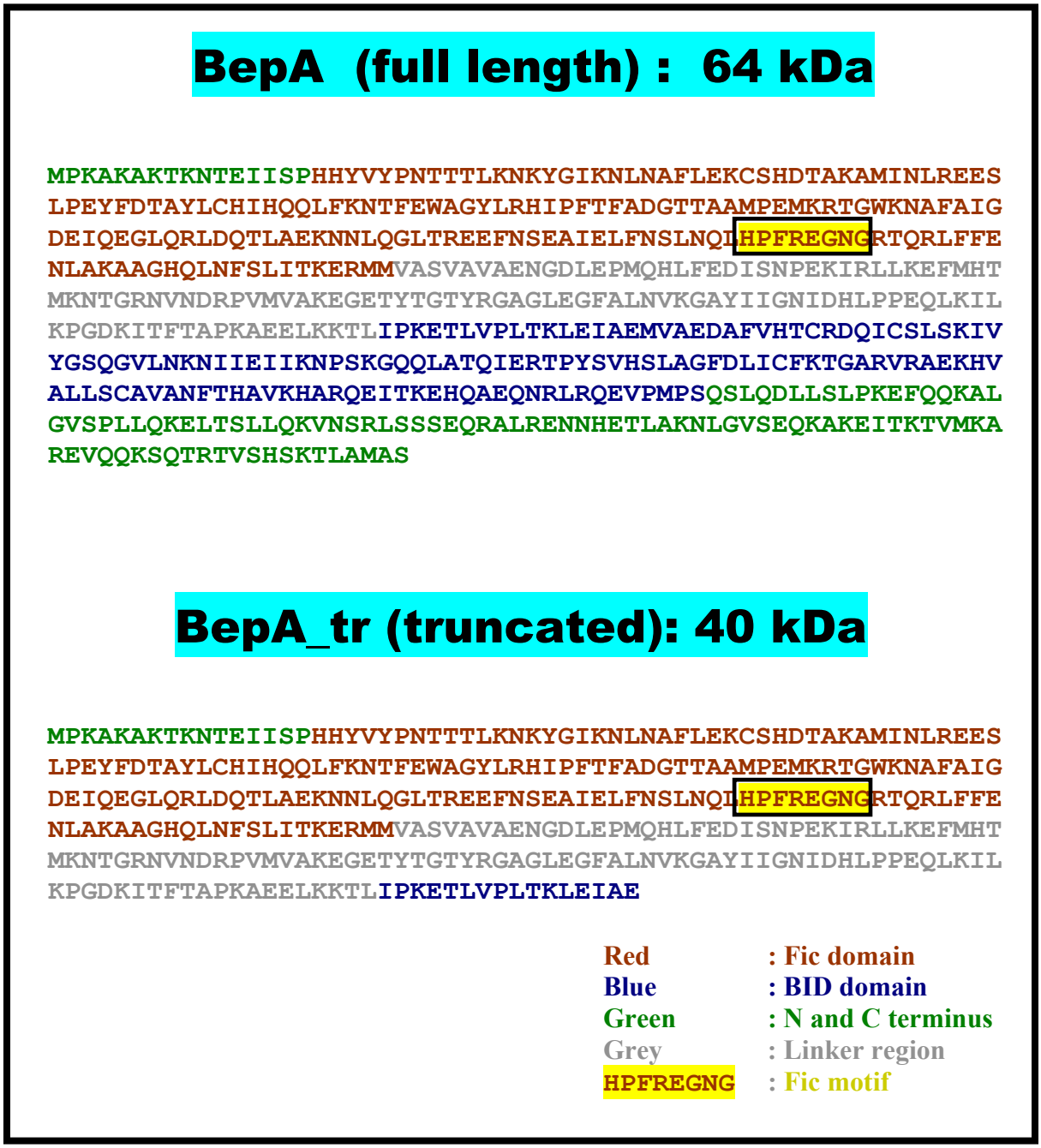


Figure 16: Sequence of BepA_tr derived by removing residues from the C-terminus of full length BepA until the theoretical mass matched the mass obtained from mass spectrometric analysis. The theoretical MW was calculated using the program ProtParam from the Expasy server [37]. The motif for the Fic domain is highlighted in yellow color.

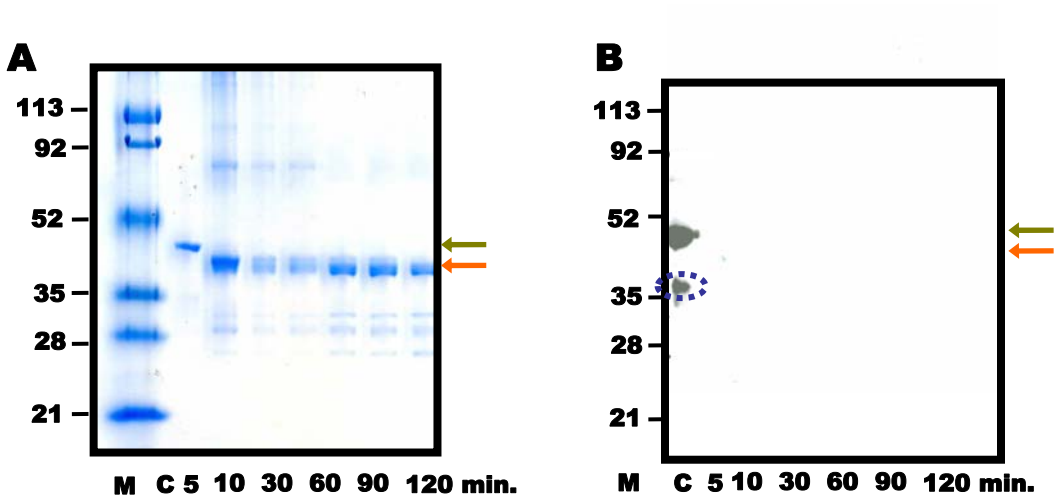


Figure 17: Domain folding studies on BepA_tr employing limited proteolysis. (A) SDS-PAGE, showing a time-course proteolysis. The green arrow indicates the position of untreated protein; the orange arrow indicates the position after tryptic digestion. (B) The Western blot of an SDS-gel equivalent to that shown in (A), showed signal only for the control (C: green arrow) and no signal was found for the tryptic digested samples (orange arrow). The signal for the degraded protein is shown in circle.

3.2.5 His-tag cleavage

From limited proteolysis study, we implicit the accessibility of the thrombin cleavage site for the His-tag cleavage. The His-tag was cleaved by thrombin treatment using two different concentrations of thrombin (1 and 2U respectively) of the enzyme. The time course was followed for 1 – 18 hours at 4 ^{o}C. SDS-PAGE analysis of the digested samples showed a shift by 2 kDa compared to the control (Figure 18A). Western blotting showed negative signal for the digested samples indicating the removal of the His-tag (Figure 18B). The digestion time was optimized to 6 hours which showed 90 % cleavage as observed from the SDS-gel (Figure 18C). The cleaved protein was not found in the flow through, later it was eluted with low concentration of imidazole (20 – 30 mM). This indicated that it was bound to the column non-specifically. The cleavage was confirmed by Western blotting. Thrombin was removed from the digested sample by passing through benzamidine sepharose beeds. Cleaved BepA_tr was then concentrated for several crystallization trials.

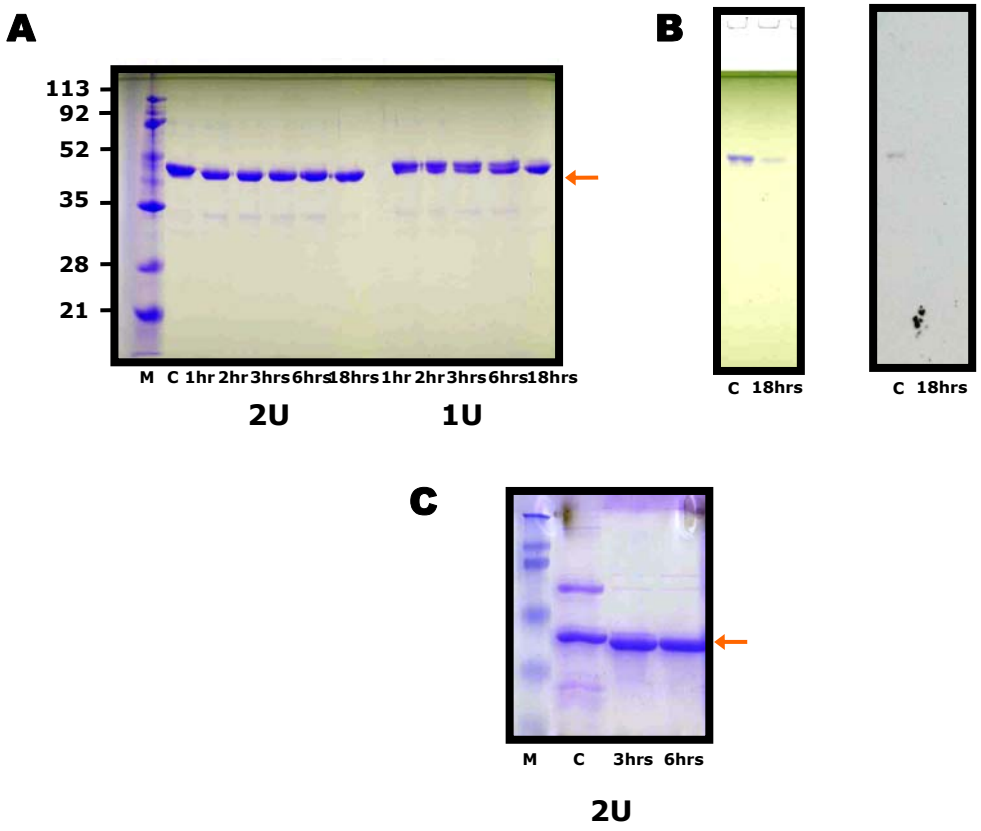


Figure 18: SDS-PAGE analysis of the thrombin digestion of BepA_tr samples. (A) A time course study (1 – 18 hrs) was carried out at 4 ^{o}C at two different concentrations (1 and 2 U) of thrombin. The gel shift of 2 kDa was observed during the course of digestion only when digested with 2U of thrombin (orange arrow). (B) The corresponding western blot of the digested sample was performed to confirm the cleavage of the His-tag. (C) SDS-gel shows the optimization of thrombin digestion to a period of 6 hrs at 2U of thrombin.

3.2.6 Crystals from cleaved BepA_tr

Several attempts to crystallize BepA_tr with and without his-tag were undertaken. Only the cleaved version of BepA_tr yielded crystals and the crystallization condition was obtained from the NCCR crystal farm, University of Zurich, Switzerland (Figure 19). The crystals grew to final size within 4 days and had bipyramidal or rod morphologies. The bipyramidal crystals grew to 100 microns in length and 25 microns in thickness. The crystallization conditions were optimized in house at 20 ^{o}C by both hanging and sitting drop methods (Figure 20).

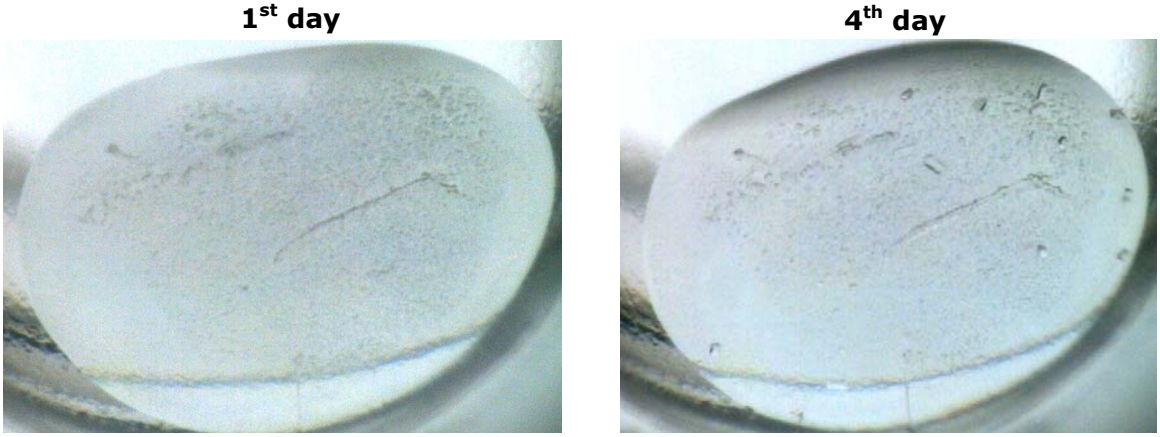


Figure 19: Demonstration of crystal growth for BepA_tr crystals on the 1st day and the 4th day (Crystals were obtained from the grid screen GS008 (Clear Strategy Screen™ I & II, pH 7.5 & 8.5) using the NCCR crystal farm facility). All the experimental set ups were performed at 20 ^{o}C. The crystallization condition was 20 % PEG 4000, 10 mM $NiCl_2$ and 100 mM Tris pH 8.5.

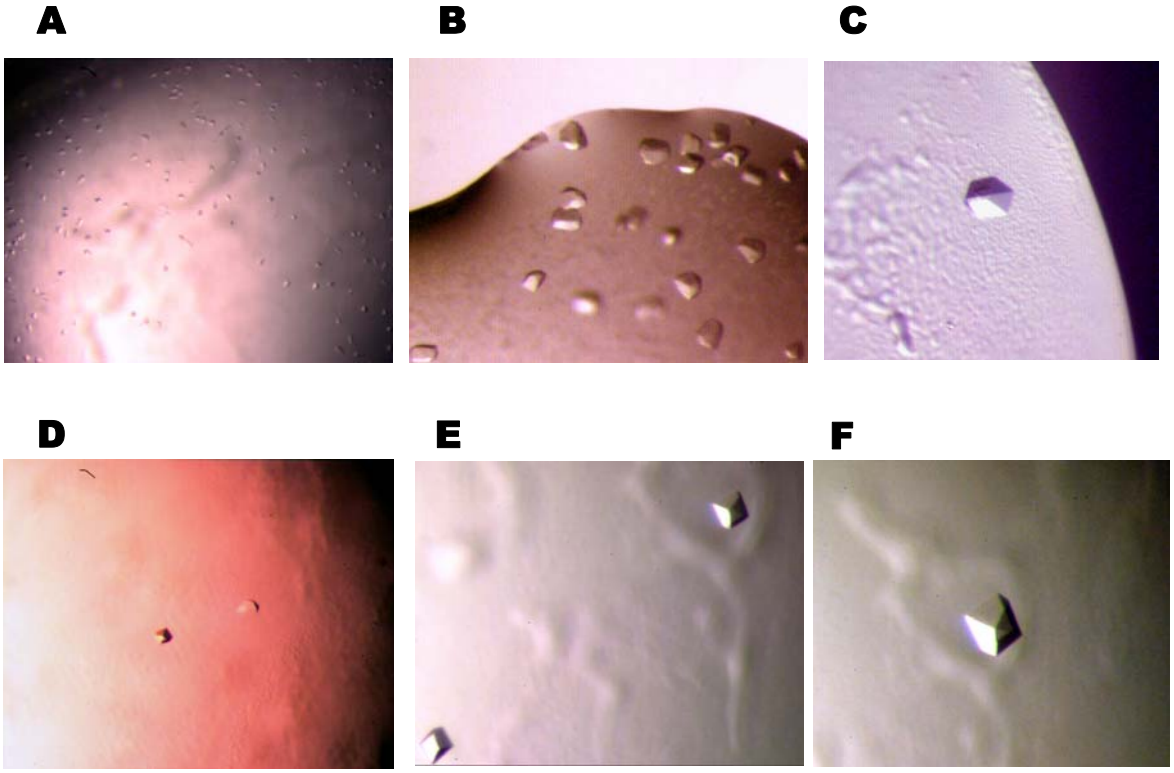


Figure 20: Crystals of BepA_tr grown using two different methods. Crystals A - C were grown using the hanging drop method and crystals D - F were grown using sitting drop method. 1 µl of protein solution was mixed with 1.5 µl of reservoir solution (20 % PEG 4000, 10 mM $NiCl_2$, 100 mM Tris pH 8.7). The reservoir contained 500 µl. The crystals grew within a week at 20 ^{o}C. (A) Showers of crystals obtained in the initial fine screening. (B) Crystal grew bigger upon an increase of the $NiCl_2$ concentration from 5 mM to 10 mM. The crystal doesn't show sharp edges. (C) Crystal size further increased by raising PEG 4K concentration from 20 % to 22%. The crystals were well formed and had sharp edges. (D) Crystals were very small and were stuck to the bottom of the sitting drop well. (E) Bigger crystals were obtained by changing the pH 8.5 to pH 8.7. (F) Crystals grew bigger up to 100 microns in diameter after a period of 3 weeks. All the crystal pictures were taken at the same magnification.

3.2.7 Data collection

BepA_tr crystals showed isotropic diffraction to about 3.3 Å resolution (Figure 21). Autoindexing clearly showed that they belong to the orthorhombic space group *P222* (Table 3). Extinctions were found along the *h00, 0k0* and *00l* axes were found indicating the presence of screw components along X, Y and Z (Figure 22A & B). The space group for BepA_tr crystal is therefore $P2_12_12_1$. A self rotation function was calculated to identify the orientation of two-fold axes for the authentication of the space group. The two-fold related peaks present along the X, Y, Z axes are shown in Figure 23. Assuming 1 monomer/asymmetric unit (a.u.), the Matthews coefficient V_m was determined to be 4.95 $Å^3$/Da using the following equation

$$\mathbf{V_m} = \frac{\textbf{Volume of unit cell } (Å^3)}{\textbf{Total weight of protein in the unit cell (Da)}}$$

$$= \frac{\boldsymbol{abc}}{\mathbf{mnZ}}$$

where ***a*** = 72.7 Å, ***b*** = 82.9 Å, ***c*** = 127 Å; **m** = 37678 Da, the molecular weight of cleaved BepA_tr; **n** is the number of molecules in the asymmetric unit (assuming 1), **Z** is the number of asymmetric units in the unit cell, 4 for $P2_12_12_1$.

The $\mathbf{V_m}$ value is very close to the upper border of the observed range for soluble protein (1.66 – 5.0 $Å^3$/Da) [38].

The solvent content was derived by the following equation:

$$\textbf{\% solvent content} = 1 - \frac{1.23}{\mathbf{V_m}}$$

$$= \mathbf{75\ \%}$$

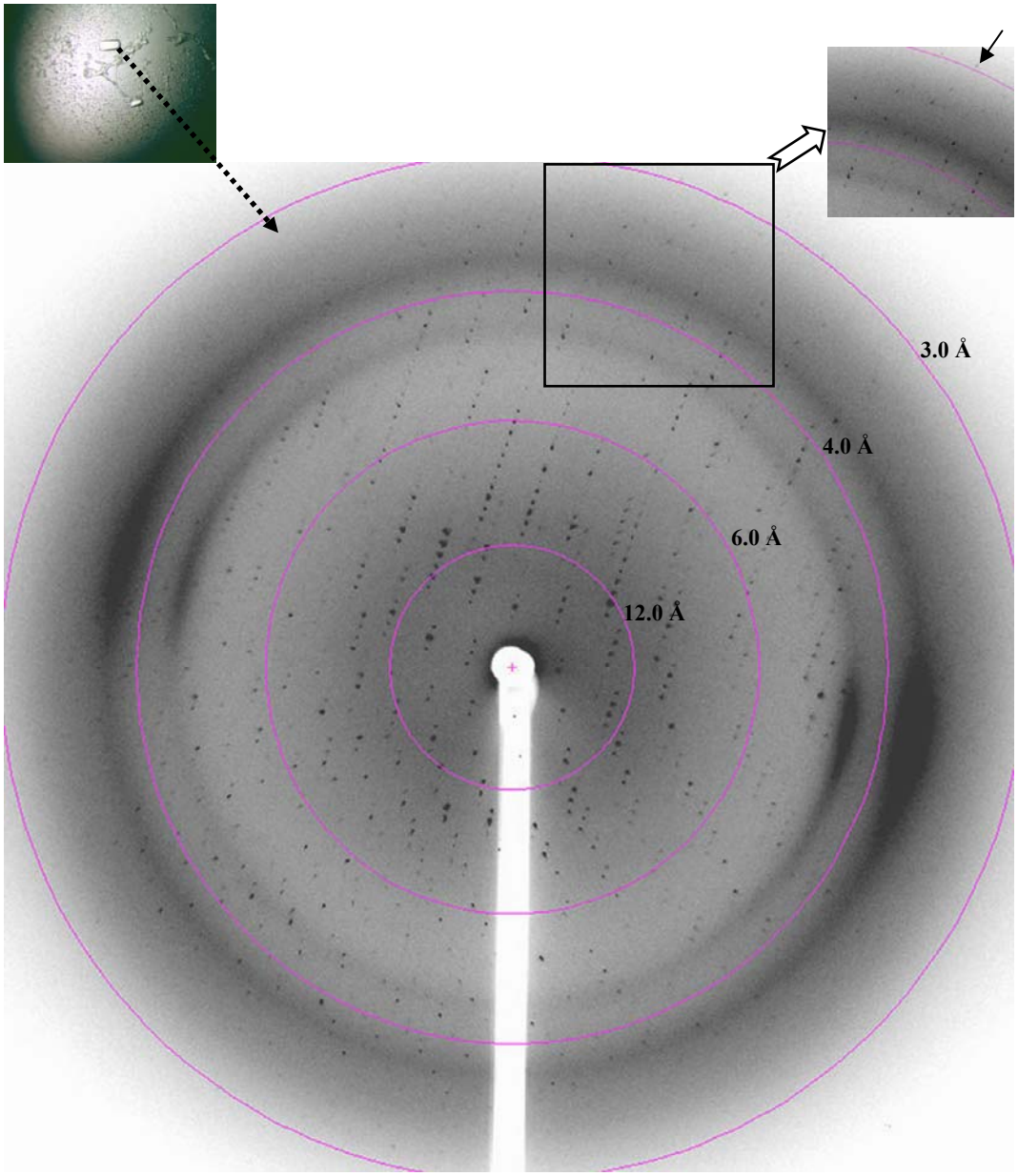


Figure 21: Diffraction pattern of a BepA_tr crystal. Data were collected at the Swiss Light Source (SLS), synchrotron source, Villigen, Switzerland. The images were recorded on a MAR-CCD image plate detector (Crystal-detector distance 300 mm, rotation angle 1.0°). The crystal showed isotropic diffraction, with spots diffracting beyond 3.0 Å which is demonstrated in the inset.

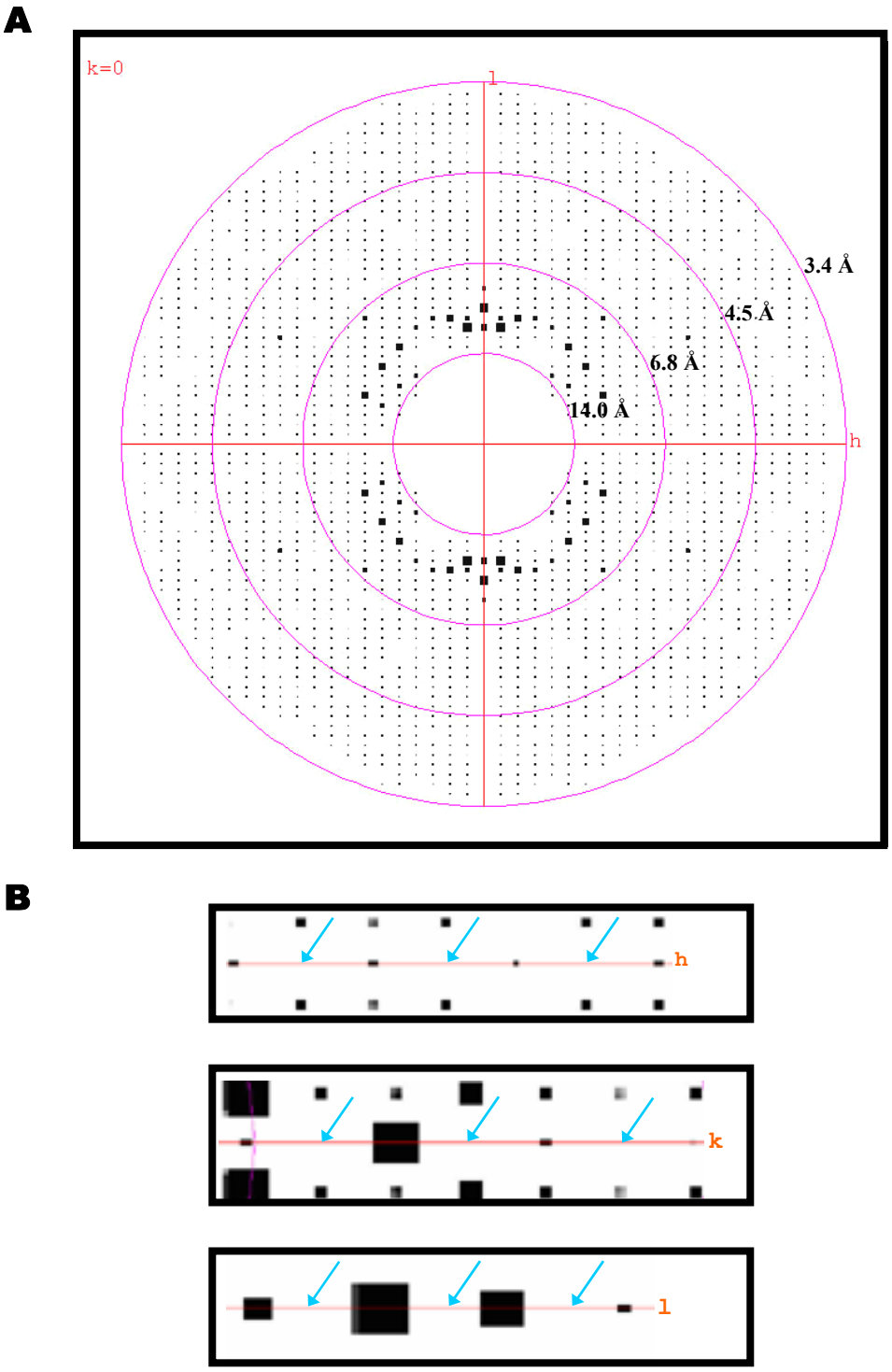


Figure 22: Integrated and scaled intensities of the reflections from BepA_tr crystal. Reflections with high intensities are shown as big spots. A, The diffraction pattern is isotropic up to a resolution of 3.3 Å. Data was processed in *P222* space group. B, Zooming in along the *h*, *k*, *l* axes shows the presence of systematic absences along *h00, 0k0* and *00l* (extinctions indicated with an arrow). Therefore, the space group of BepA_tr crystal is $P2_12_12_1$.

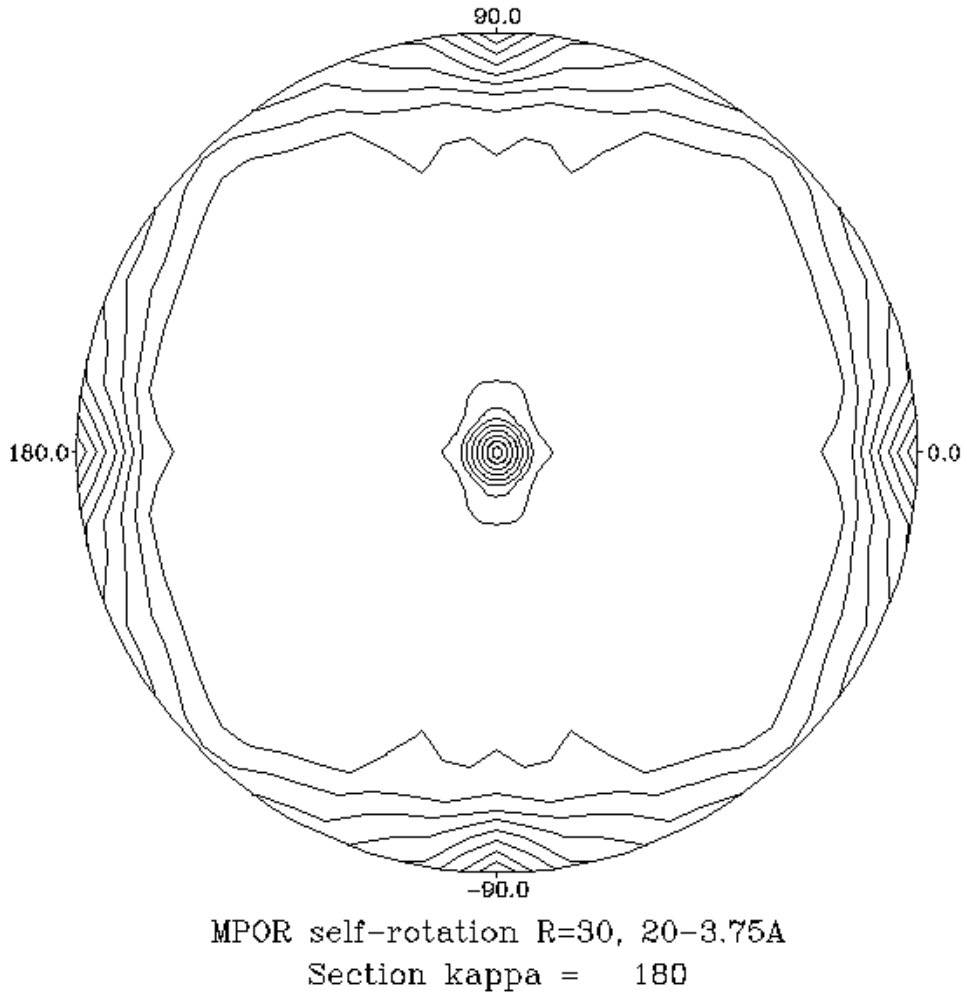


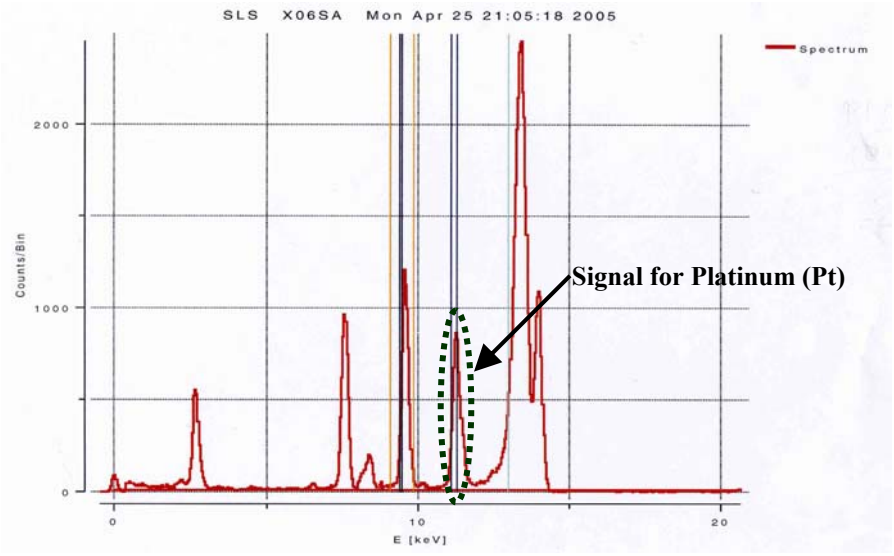
Figure 23: Self rotation diagram at section kappa 180^{o}, shows the crystallographic two-folds along X, Y, Z axes (*P222* space group) and no additional signal is found.

Table 3: Data collection and refinement statistics

Data Collection	
X-ray source; λ	**SLS, Villigen; 0.900 Å**
Detector	**MARCCD**
Temperature	**100 ^{0}K**
No. of crystals	**1**
Space group	***P2$_1$2$_1$2$_1$***
Unit cell dimensions	***a* = 72.7 Å *b* = 82.9 Å *c* = 127 Å** **α = β = γ = 90^o**
Resolution range (Å) [a]	**15.0 – 3.35 (3.53 – 3.35)**
Number of unique reflections [a]	**10106 (500)**
R_{sym} (%) [a, b]	**12.6 % (42 %)**
$\langle I \rangle / \langle \sigma(I) \rangle$ [a]	**5.3 (1.7)**
Completeness (%) [a]	**88 % (32 %)**
Multiplicity [a]	**3.5 (1.6)**

[a] Values in parentheses refer to the highest resolution shell.
[b] Rsym = ∑hkl ∑i (| I(hkl) - < I(hkl) > |) / ∑hkl ∑i < I(hkl) >, where I is the scaled intensity of a given reflection.

A



B

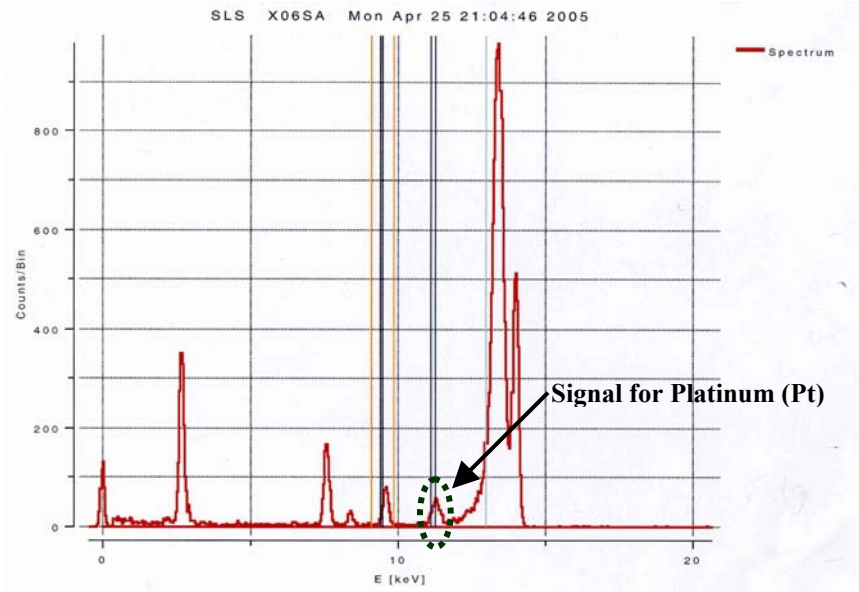


Figure 24: Fluorescence spectra of the BepA_tr crystal soaked in 10 mM $PtCl_4$ for 6 hours at room temperature. (A) The spectrum taken from the crystal. The signal obtained for Platinum (Pt) is indicated with black dots. (B) The spectrum taken from the solvent (region close to the crystal).

3.2.8 Heavy atom derivative

Heavy atom soaks of the native BepA_tr crystals with three heavy atom compounds, namely K_2PtCl_4 (Pt), $CH_3HgOOCCH_3$ (Hg) and $Sm(O_2C_2H_3)$ (Sm), were carried out for a period of 6 hours at two different concentrations (10 mM and 5 mM). Fluorescence spectra of the 5mM soaked crystal showed a strong signal for Pt (Figure 24A) but no signal for the other two derivatives. As a control, a spectrum was taken from the solvent which gave a very faint signal (Figure 24B). The peak (11.58087 KeV) and inflection edge of Pt (11.56700 KeV) were then identified (Figure 25). The soaked crystals showed diffraction up to 3.6 Å. A MAD experiment was performed at the peak and the inflection wavelengths (λ = 0.900 and λ = 1.071) for Pt. Identification of the heavy atom sites using Patterson search methods is in progress.

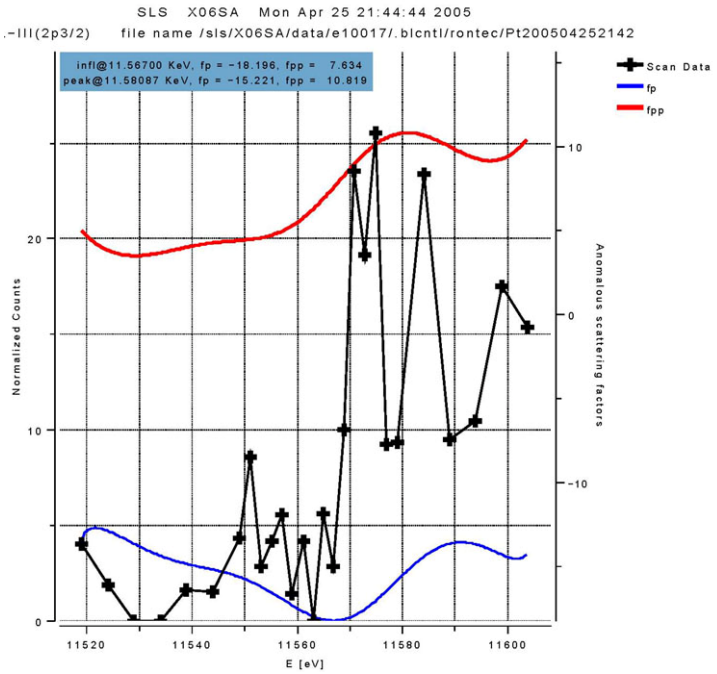


Figure 25: Excitation spectrum (emission energy at the Pt III edge) of the BepA_tr crystal soaked in 5 mM $PtCl_4$, showing the peak and the inflection points for Pt. The peak lies at 11.58087 KeV and the inflection point at 11.56700 KeV.

3.3 BID domain of BepA (BID_A)

Expression was conducted at 37°C for a period of 4 hours. The cells were lysed and pelleted to estimate the expression level and the yield of soluble fraction. Interestingly, the protein was found expressed and localized in the membrane fractions. In order to solubilise the protein from the membranes, solubilization experiments were conducted in the presence of detergents. Isolated membrane fractions were solubilised in the detergent n-decyl-β-D-maltoside (DM; 2.0 %) and purification was carried out with Ni-NTA beads in wizard midi columns. The detergent concentration was reduced during the washing step from 2.0 % to 0.3 % (three times the critical micellar concentration (CMC) value). After several washes with imidazole, the protein was eluted from the column at 200 mM imidazole concentration. SDS gel shows that the protein is soluble in the presence of the detergent (Figure 26A). Finally, the presence of BID_A was confirmed by western blot analysis (Figure 26B).

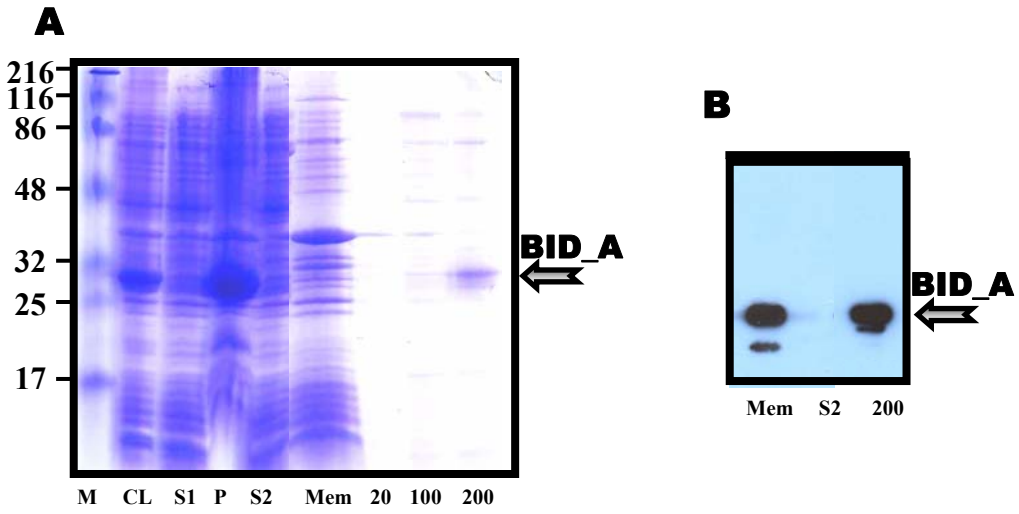


Figure 26: SDS-PAGE and the Western blot analysis of BID_A expression and purification. (A) Cell lysate was obtained by French press (CL) which were spun down at low spin to obtain the 1st supernatant (S1) and the pellet (P). The 1st supernatant was further centrifuged to separate the membrane fractions (Mem) from the cytosolic fractions (S2). The membrane fractions were solubilized using 2 % n-decyl-β-D-maltoside (DM) and purified on a nickel affinity column. The column was washed with 20 mM imidazole (20) and elution was carried out with 100 mM (100) and 200 mM (200) imidazole concentrations in presence of 0.3 % DM. BID_A was eluted at the 200 mM imidazole concentration (arrow). (B) The Western blotting showed positive signal for the membrane fraction (Mem) and 200 mM Ni-NTA fraction (200). There was no signal for the soluble fraction (S2).

3.4 BepB

3.4.1 Expression and Solubility test

Expression and solubility test for the full-length BepB were conducted in a similar way as explained for BepA. The best yield of soluble protein was obtained when BepB was grown at 30 ^{o}C and induced at OD_{600} for 4 hours (Figure 27A; Table 2).

3.4.2 Purification and characterisation of full length BepB

BepB was purified using Ni-NTA (IMAC) affinity and size exclusion chromatography. The supernatant of the cell lysate was first purified on a nickel affinity column using the batch method and BepB was eluted at 100 mM imidazole concentration. The Ni-NTA fractions showed 80 % purity by estimation from the SDS gel (Figure 27B). The Western blotting with an anti-his antibody showed signal for full length BepB (Figure 27C). In the second chromatography step, a Hiload 16/60 Superdex 75 gel filtration column was used and BepB was eluted in its monomeric range (calculated with the void volume (40 mL) from the manufacture) (Figure 28A). SDS-PAGE for the major peak fractions showed 95 % purity (Figure 28B). The fractions corresponding to pure BepB were pooled and concentrated up to 10 mg/mL for the subsequent crystallization trials.

3.4.3 His-tag cleavage

The His-tag was cleaved by TAGZyme (DAPase) treatment using 100 mU concentration of the enzyme. This enzyme is a dipeptidase and cuts every two amino acids from the N-terminus and stops its action at a specific amino acid (DAPase stop point: either K, R, P or Q). The enzyme has a C-terminal His-tag which can be exploited to remove the enzyme during the second Ni-NTA step. In the BepB construct, the first stop site of DAPase is found after the N-terminal His-tag at position R17, an Arginine was found in the BepB sequence (Refer; Figure 11). Digestion was performed for 6 hours at 4^{o}C. Second Ni-NTA column was used to separate cleaved from uncleaved BepB and TAGZyme. The cleaved protein was not found in the flow through, later it was eluted

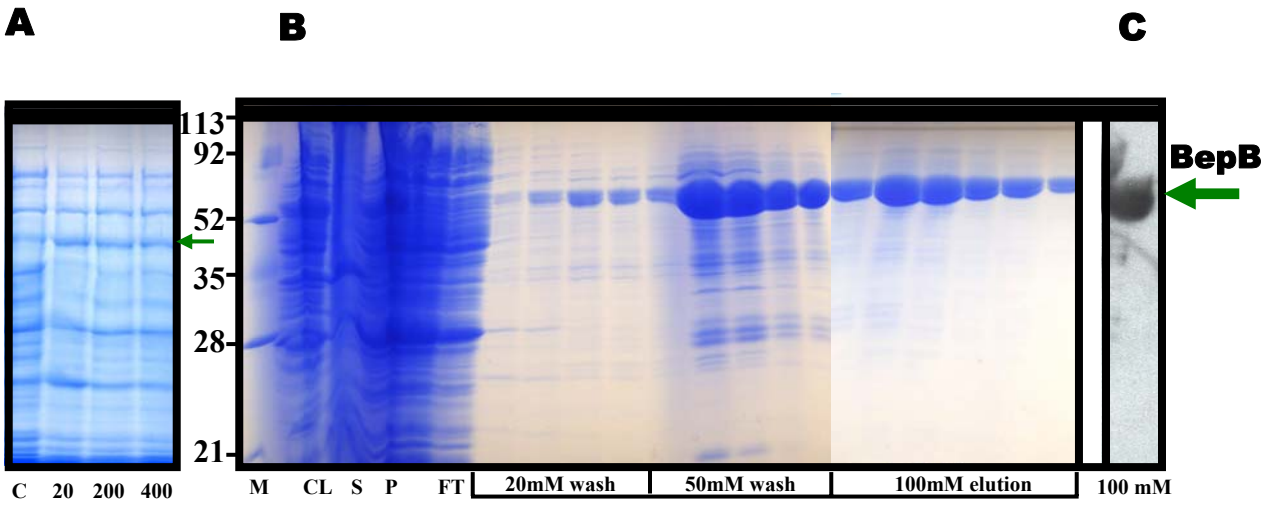


Figure 27: SDS-PAGE and Western blot analysis of BepB expression and purification. (A) Lane 1: Cells without IPTG induction (C - control), Lane 2: 20 µM IPTG induction, Lane 3: 200 µM IPTG induction and Lane 4: 400 µM IPTG induction. Expression of BepB was indicated by an arrow. (B) SDS-gel showing the steps involved for BepB purification on a nickel affinity column. Cell lysate was obtained by French press (CL) which were spun down to obtain the supernatant (S) and the pellet (P). BepB lysate supernatant was purified on a nickel affinity column. The flow through (FT), the washes with 20 mM and 50 mM imidazole, and the 100 mM imidazole elution of the protein were shown in the figure. (C) The Western blotting showed a positive signal for the main band (100 mM fraction) (arrow).

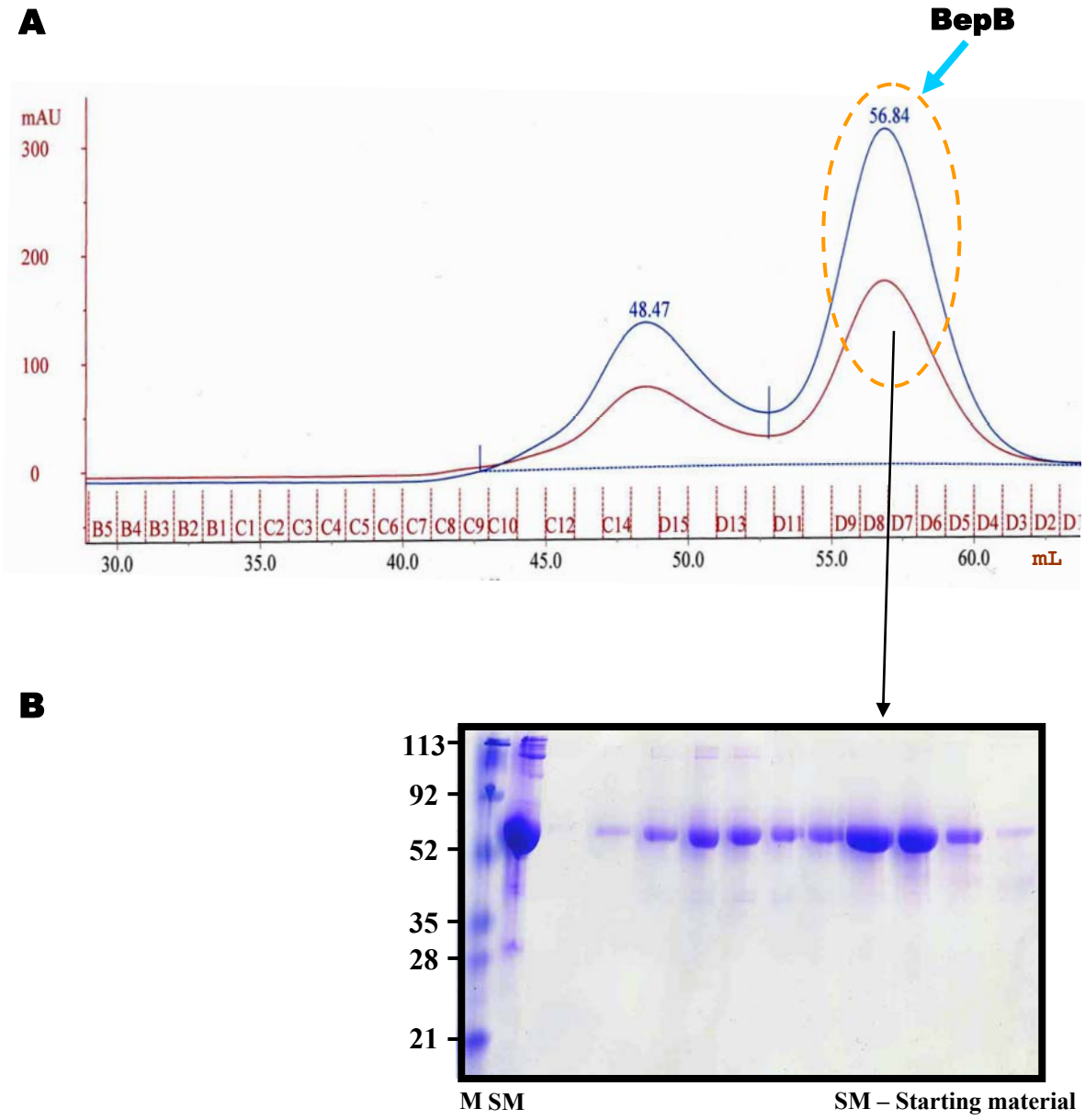


Figure 28: Purification of BepB by size exclusion chromatography (A) Elution profile of BepB on a Hiload 16/60 Superdex 75 gel filtration column (blue arrow). Elution peaks are labeled with their peak elution volume. (B) SDS-gel of the main peak fractions show highly purified BepB.

with low concentration of imidazole (40 mM). The Western blotting showed no signal for the cleaved BepB and signal was obtained for the uncleaved material (Figure 29). The

cleaved BepB was concentrated and several crystallization attempts were made. We haven't found any crystals so far from above trials.

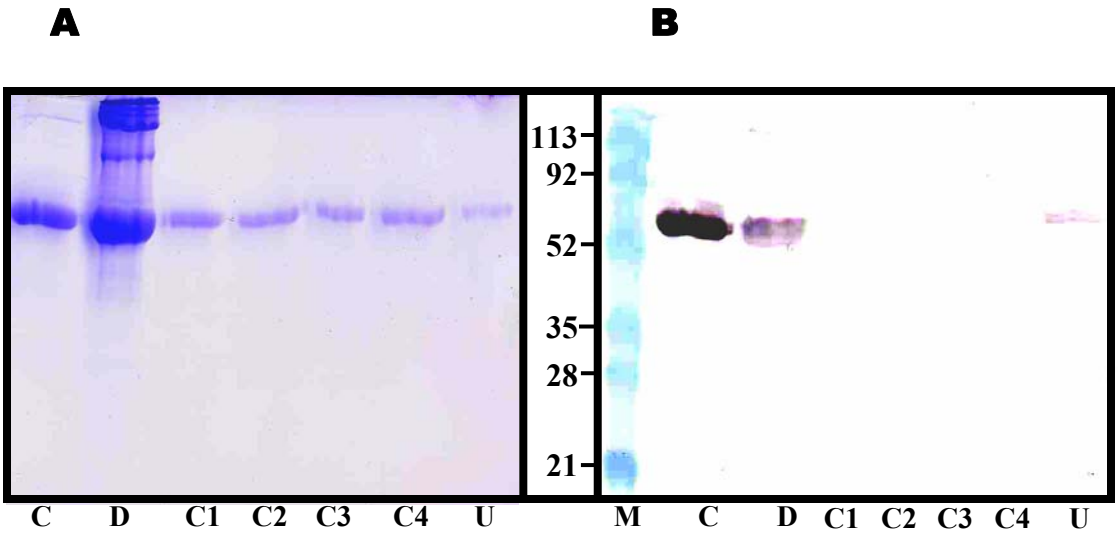


Figure 29: SDS-PAGE and the Western blot analysis of second Ni-NTA fractions obtained after TAGZyme digestion of BepB. The digestion was performed at 4^{o}C for a period of 6 hours. (A) Lane C: control (undigested sample), Lane D: Digested sample, Lane (C1-C4): cleaved BepB and Lane U: uncleaved BepB. (B) The corresponding western blot picture shows the signal for the samples (C), (D) and (U) and no signal was observed for samples (C1 – C4).

3.5 BID domain of BepB (BID_B)

Expression of BID_B behaved differently from BID_A. Expression was conducted at 37 ^{o}C for a period of 4 hours (Figure 30A; Table 2). The protein was found to be expressed and present in the soluble fraction. The supernatant of the cell lysate was loaded on a nickel affinity column using the batch method. BID_B was eluted at 150 mM imidazole concentration and they showed 95 % purity by estimation from the SDS-PAGE (Figure

30B). When the protein was subjected to the concentration step, the protein started to precipitate in the concentrating unit. This showed that the protein was not soluble upon concentration. The Ni-NTA fractions were pooled and dialyzed against the protein buffer containing 1M NaCl. In presence of 1 M NaCl, the sample was found to be soluble and concentrated up to 10 mg / mL. Concentrated BID_B was subjected to various crystallization trials. We haven't obtained crystals so far from the above trials.

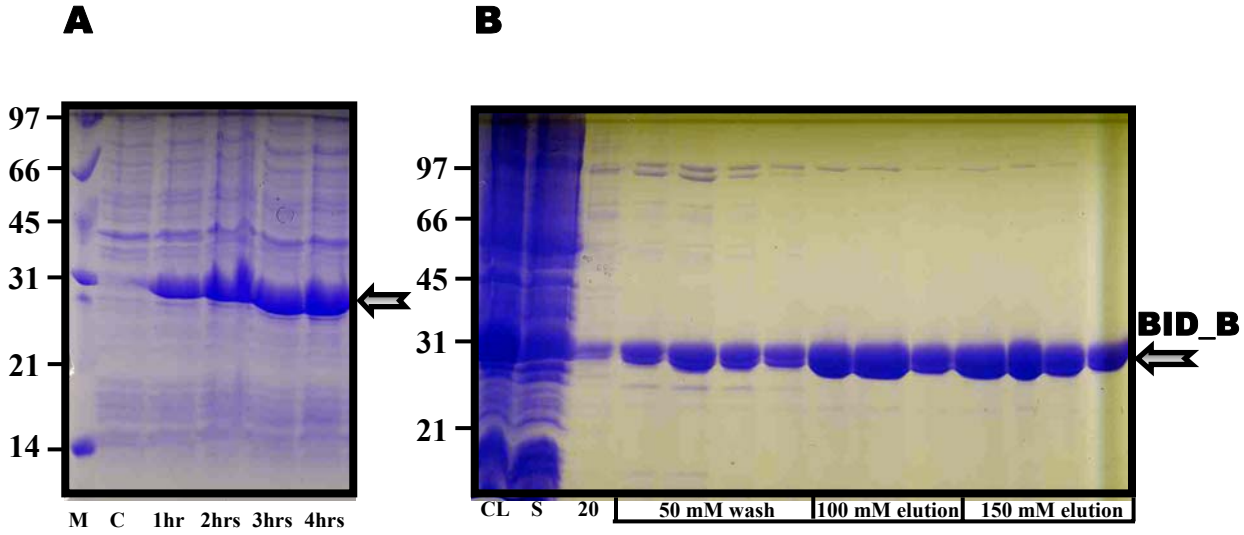


Figure 30: SDS-PAGE analysis of BID_B expression and purification. (A) Lane 1: Cells without IPTG induction (C - control), Lane 2: 1hour induction, Lane 3: 2 hours induction, Lane 4: 3 hours induction, Lane 5: 4 hours induction. Expression of BID_B was indicated by an arrow. (B) SDS-gel showing the purification of BID_B on a nickel affinity column. Cell lysate was obtained by French press (CL) which was spun down to obtain the soluble fraction (S). The soluble fraction was purified on a nickel affinity column. The column was washed with 20 and 50 mM imidazole concentrations (20 and 50 mM respectively) and elution at 100 and 150 mM imidazole concentrations (100 mm and 150 mM elution). The latter showed highly purified BID_B (arrow).

3.6 BepC: Preliminary result

3.6.1 Expression and localization studies

Expression and solubility test was conducted for the full-length BepC in a similar way as explained for BepA. The bacteria were grown at 30 ^{o}C and induced at OD_{600} nm of 0.57 for 3 hours (Table 2). The cells were lysed and pelleted to estimate the expression level and the yield of soluble fraction. Interestingly, the protein was not found in the soluble fraction as analyzed from the SDS-PAGE. The pellet fraction shows a strong band corresponding to the size of BepC. Our first interpretation was that the protein goes into inclusion bodies. But the procedure was modified as of for a membrane protein extraction by enforcing two centrifugation steps. Interestingly, BepC was found in the membrane fraction as observed on the SDS-PAGE (Figure 31). The above result was verified with Western blot and mass spectrometric analysis (data not shown).

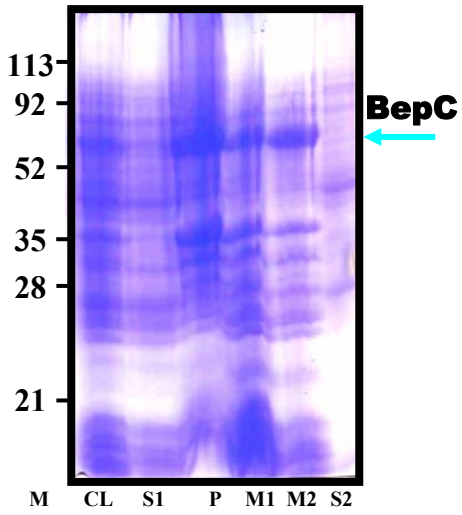


Figure 31: SDS-PAGE analysis of BepC expression and extraction from the membrane. Cell lysate was obtained by French press (CL) which were spun down at low spin to obtain the 1st supernatant (S1) and the pellet (P). The 1st supernatant was further centrifuged to separate the membrane fractions (M1 (1:10 dilution) and M2) from the soluble fractions (S2). BepC was localized in the membrane fraction (arrow).

3.7 BepD

3.7.1 Expression and Solubility test

Expression and solubility test for the full-length BepD were conducted in a similar way as explained for BepA. The best yield of soluble protein was obtained when BepD was grown at 37°C and induced at OD_{600} of 0.75 for 6 hours (Figure 32A; Table 2).

3.7.2 Purification of full length BepD

Purification of BepD was more difficult than that of BepA_tr and BepB, since it involved three chromatography steps, namely Ni-NTA (IMAC) affinity, ion exchange and size

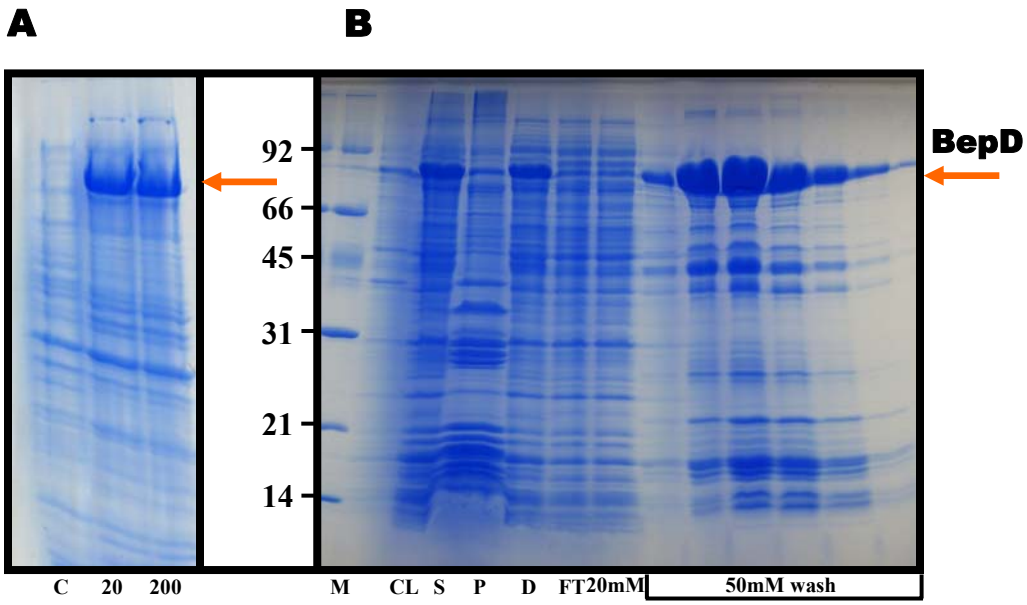


Figure 32: SDS-PAGE analysis of BepD expression and purification. (A) Lane 1: Cell lysate (c: control), Lane 2: 20 µM IPTG induction and Lane 3: 200 µM IPTG induction. Expression of BepD was indicated by an arrow. (B) SDS-gel showing the purification of BepD on a nickel affinity column. Cell lysate was obtained by French press (CL) which were spun down to obtain the supernatant (S) and the pellet (P). BepD treated with streptomycin sulphate was dialysed (D) with the Ni-NTA buffer (20 mM Tris pH 8.0, 300 mM NaCl, 5 mM imidazole and 5 mM β-mercaptoethanol) and loaded on a nickel column; the flow through (FT), wash with 10 mM and 50 mM imidazole were shown. BepD runs higher in the SDS-gel (~ 80 kDa) than the expected region of 63 kDa (arrow).

exclusion chromatography. The cells lysate was treated for DNA precipitation using streptomycin sulphate. Samples were spun down to pellet the DNA and subsequent dialysis step was performed to remove the streptomycin sulphate. The dialysed sample was first subjected to a nickel affinity column using the batch method. BepD was eluted at 50 mM imidazole concentration. In view of the fact that BepD eluted at very low concentration of imidazole, it had several contaminants as observed in the SDS-gel (Figure 32B). Interestingly, BepD was found to run as a band at a MW of 80 kDa than its expected MW of 63 kDa (Figure 32B).

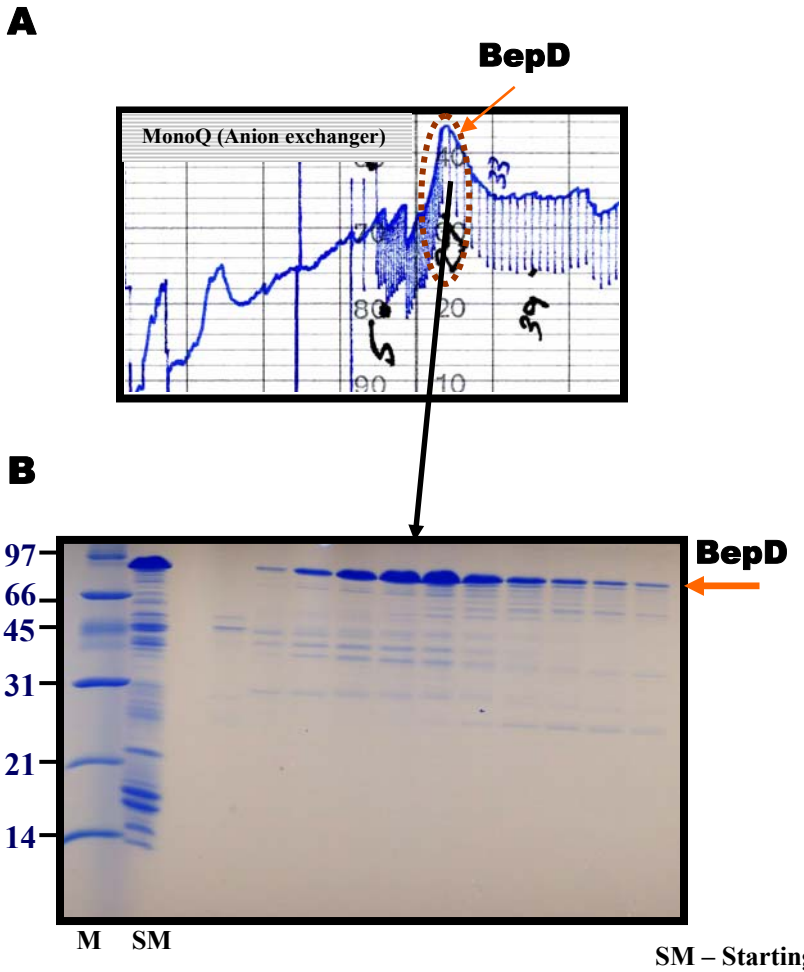


Figure 33: Purification of BepD by MonoQ (Anion exchanger) column using a linear gradient of NaCl (0 – 1 M) and SDS-PAGE analysis. (A) Elution profile of BepD eluting at 200 mM NaCl concentration (arrow). (B) SDS-gel of the main peak fractions show purified BepD.

In the next purification step, an ion-exchage column (MonoQ; anion exchanger) was used to separate BepD from the contaminants based on their ionic charges. Before loading onto the column, a desalting column (Sephadex25) was used to remove the salts from the Ni-NTA fractions.

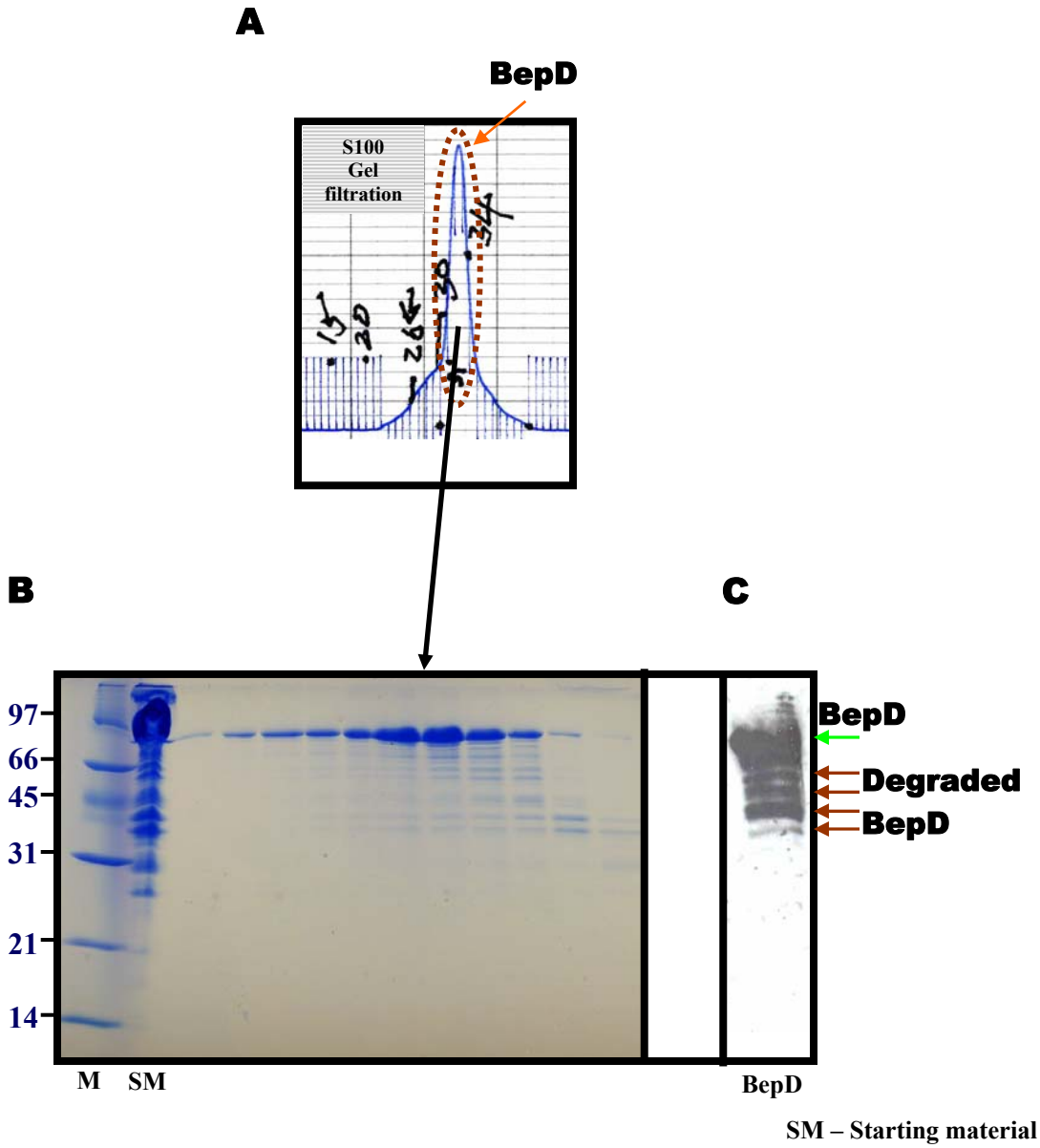


Figure 34: Purification of BepD by size exclusion chromatography and SDS-PAGE and western blot analysis of the purified fractions. (A) Elution profile of BepD on a Sephadex 100 gel filtration column. (arrow). (B) SDS-gel of the main peak fractions show purified BepD. (C) The Western blotting showed positive signal for BepD (green arrow). There were several signals present below the main band which corresponds to the degradation products of BepD (red arrows).

A linear gradient of NaCl (0 – 1 M) was used for eluting the protein. BepD eluted at a NaCl concentration of 200 mM (Figure 33A) and the sample was 70 % pure as judged by SDS-PAGE (Figure 33B). The final step of purification of BepD was achieved by loading the concentrated MonoQ fractions onto a Sephadex 100 gel filteration column. BepD eluted as a single peak (Figure 34A) and the fractions showed 90 % purity as estimated from the SDS-gel (Figure 34B).

3.7.3 Stability of BepD

The purified sample of BepD was tested for degradation by Western blot method using an anti-histidine antibody. The result showed that BepD forms a ladder of multiple bands from the Western blot analysis (Figure 34C.). This result indicates that this protein undergoes severe degradation from the C-terminus, which in turn implies that BepD is unstable during the course of purification.

4.0 Discussion

Structural investigations of the Beps of *B. henselae* have yielded valuable constructs and information that will allow us to solve the structure of some of these effector proteins. The results obtained from the constructs of Beps suggest that they are quite diverse from each other and their characteristic features are discussed in the following paragraphs.

4.1 Insights from the structural studies of Beps

4.1.1 BepA

Based on the results obtained for BepA, we have identified it as the most interesting protein among all the constructs. Full length BepA was expressed as a truncated protein (BepA_tr) which comprises of the Fic domain and the linker region.

4.1.1.1 BepA_tr

BepA_tr was found expressed in the cytosolic fractions along with a degradation product. By thorough examination, we managed to separate the degraded protein to obtain pure BepA_tr. Limited proteolysis treatment of BepA_tr indicated that the protein was stable and not accessible for protease action. Crystallization experiments showed no crystal formation for the His-tagged version of this protein. Upon cleaving the His-tag (cleaved BepA_tr), crystals were obtained. A control experiment was performed by testing the His-tag version of the same protein under the same crystallization condition. No crystals were grown in the control experiment, which clearly indicates that the His-tag was the hindrance for crystal formation. The cleaved BepA_tr crystals diffracted to a resolution of 3.3 Å. We have identified a potential heavy atom derivative which subsequently yielded MAD datasets for the phase determination. We are also in the process of optimizing the crystallization condition using additive screens for increasing the resolution of cleaved BepA_tr crystal. Recently, we have found few micro crystals being formed in presence of 2 mM L-cysteine mixed in the same crystallization condition.

Further optimization of the condition should improve the crystal size which we hope to show better diffraction for attaining high resolution data. Interestingly, we also identified a new condition from the NCCR crystal farm plate for cleaved BepA_tr. The appearance of the crystal was observed after a period of one month which is demonstrated in the Figure 35. The formulation of the new condition is 4 M sodium formate and 0.1 M Tris pH 7.5. The same kinetics of the crystal growth was observed under the same condition at pH 8.5. Ultimately, obtaining the structure of BepA_tr, might help us to elucidate their functional role in the eukaryotic cell.

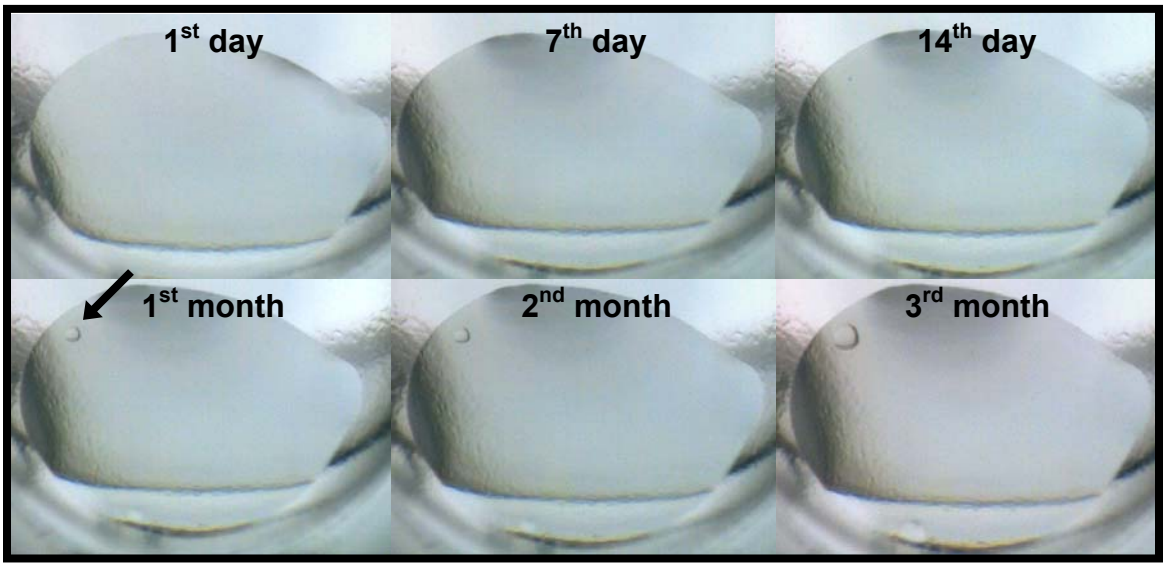
4.1.1.2 BID domain of BepA (BID_A)

Investigation of the BID domain of BepA (BID_A) showed that BID_A was overexpressed and found in the membrane fraction of the *E. coli*. Upon adding detergent, the protein was extracted from the membrane and found stable in solution. This result suggests that there are hydrophobic patches in this domain which can interact with the *E. coli* membranes. From these preliminary results, we are trying to role of BID_A in bacteria.

4.1.1.3 Hypothesis for truncation process in BepA expression

Based on the results obtained from the expression of BID_A, we tried to evaluate the expression of full length BepA in *E. coli* strain. On expressing the full length BepA, only the Fic domain and the linker region were found to be expressed but the BID domain was entirely absent. Since we found BID_A localized in the membrane fraction, we made a hypothesis for explaining the truncation process in BepA during expression. Our assumption is that BepA is expressed as a full length protein and gets associated with the membrane surface by means of its BID domain. Subsequently, it is proteolytically cleaved into two parts. The N-terminal part with the linker region segregates into the cytoplasm, whereas the BID domain will remain attached to the membrane (Figure 36). In order to confirm our hypothesis we need to make a C-terminal His-tag construct of BepA.

A



B

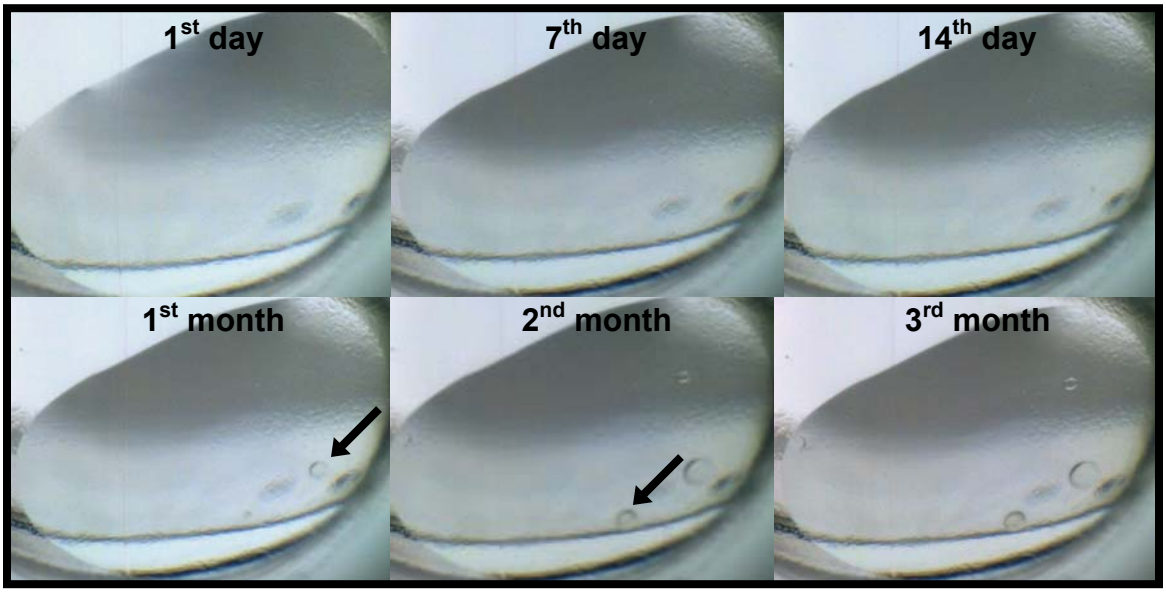


Figure 35: Demonstration of the crystal growth for BepA_tr crystals from 1st day till 3rd month using the NCCR crystal farm from the grid screen GS008 (Clear Strategy Screen™ I & II, pH 7.5 & 8.5). All the experimental trials were performed at 20 °C. Appearance of the crystal is indicated with an arrow mark. Images were obtained using the software provided from the crystal farm (A) and (B) The crystallization condition was 4 M sodium formate and 100 mM Tris pH 7.5 and 4 M sodium formate and 100 mM Tris pH 8.5 respectively.

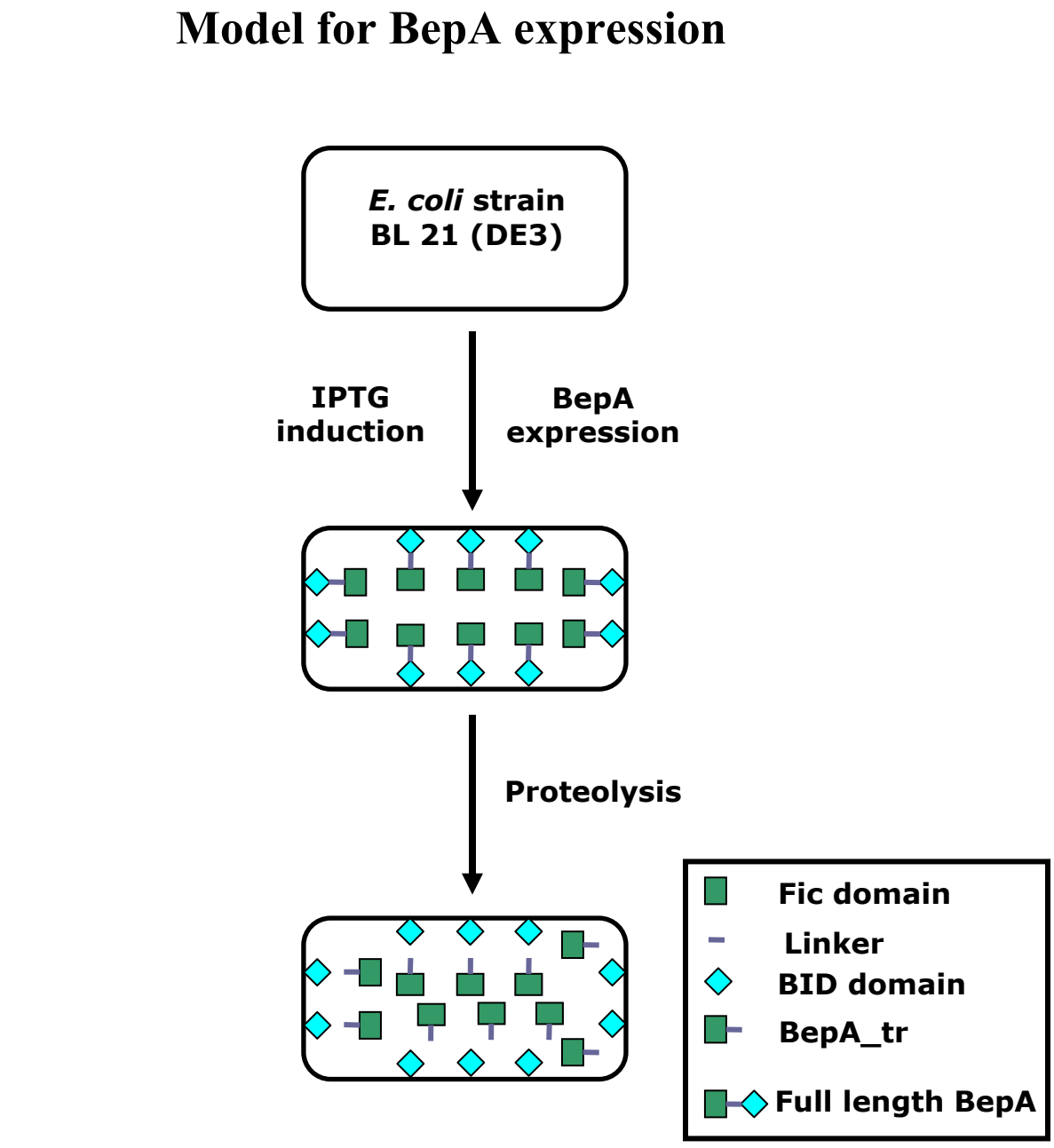


Figure 36: A schematic representation of the BepA expression in *E. coli*. BepA was expressed as a full length protein and gets associated with the *E. coli* membrane. Upon proteolysis, BepA is truncated into two parts; it forms BepA_tr (Fic plus the linker region) and the BID domain.

4.1.2 BepB

We worked extensively on this protein, since BepB was expressed in full length and showed no degradation in contrast to BepA. Purification of BepB was relatively easy compared to other Beps. It yielded 80 – 85 % pure protein from the nickel affinity column. Purification using a gel filtration column enabled us to obtain 95 % pure and homogenous protein sample. No crystal formation was observed in presence of the His-tag. We employed the same strategy used for BepA_tr by cleaving the His-tag by using TAGZyme. In the case of BepB, cleavage of the His-tag using thrombin was unsuccessful. We tried several crystallization experiements for the cleaved BepB. We haven't found any crystals so far from the above trials. Ultimately, obtaining the crystal for BepB will help us to elucidate the structure of a full length Bep protein of the Bep-ABC groups.

4.1.2.1 BID domain of BepB (BID_B)

Investigating the BID_B construct, it was overexpressed and present in the cytosolic fraction unlike BID_A. BID_B was purified by nickel affinity chromatography and the fractions showed 95 % purity. There were initial problems for concentrating the protein. The problem was solved by increasing the solubility of the protein by subjecting to high salt (1M NaCl) concentration. The protein was found soluble up to 10 mg/mL and was subjected for crystallization experiments. We haven't obtained crystals for this protein so far.

4.1.3 Assessment of His_6-tag cleavage for BepA_tr and BepB

His-tag cleavage experiments were made for BepA_tr and BepB. It worked for BepA_tr and BepB using the proteolytic enzymes, thrombin and TAGZyme, respectively. Unexpectedly, after cleavage, the protein was not found in the flow-through after the second Ni-NTA column although the cleavage was observed by gel shift method in the case of BepA_tr. On careful examination, we found that the cleaved protein was bound to the column, probably in a non-specific manner. In both the cases, the cleaved protein was eventually eluted from the column at low concentration of imidazole. The assessment of

the cleavage was easy in case of BepA compared to BepB, since the MW of BepA_tr was 40 kDa and a significant gel shift (2 kDa) was observed on SDS-PAGE analysis. Whereas, for BepB the MW was 64 kDa, and the gel shift procedure cannot be applied.

4.1.4 BepC

BepC was unique among all the Beps, since it was expressed and found in the membrane fraction of *E. coli*. This was proven by mass spectrometric and western blotting analysis (data not shown). The function of BepC is not known so far. The preliminary results obtained for BepC has to be thoroughly examined for further investigations. Obtaining the structure of the BepC can shed light into its function(s) in the eukaryotic cell.

4.1.5 BepD

BepD was the first protein among the Bep family of proteins where the functional phenotype was established in our collaborator's lab. Translocation of BepD into the host cells was effectively monitored by the CRAFT assay [26]. On the domain level, it possesses a different modular domain (pTyr) which gets tyrosines phosphorylated upon translocation into the host cells by src kinases. Purification of BepD was difficult since the protein bound to the nickel affinity column with very low affinity. In order to obtain pure BepD, two purification steps were employed. BepD was purified using ion exchange and gel filtration chromatography, the first step was very critical to remove a contaminant that absorbed signal at 260 nm. This may correspond to DNA or DNA-protein complexes. BepD also undergoes severe degradation at the C-terminal part which is seen from the Western blot experiment. BepD obtained from different purification steps were tested for phosphorylation by specific src kinases (Guye P., University of Basel). Interestingly, impure BepD obtained from the nickel column showed phosphorylation activity whereas, the pure BepD obtained after three chromatography steps was no longer phosphorylated and the sample was also found to be aggregated. We can suggest that BepD is highly unstable during the process of purification. All this information indicates that crystallization of BepD may be difficult. In addition, bioinformatic anaysis have predicted that BepD contains 40 – 50 % of non-regular secondary structures. This

suggests that the protein folding could be a problem, thus, it is accessible for severe proteolysis.

4.2 Role of BepA in eukaryotic cells

The functional role has been established for BepA [26, 27]. It performs an anti-apoptotic effect upon translocation into the eukaryotic host cell. Since BepA is composed of an N-terminal Fic domain and a C-terminal BID domain, their individual roles are discussed separately.

4.2.1 Putative role of the Fic domain

Fic proteins are predicted to be involved in the process of cell division in the synthesis of para-amino benzoate (PAB) or folate, indicating that the Fic protein and cAMP are involved in the regulation of cell division *via* folate metabolism [29]. This putative function was proposed for the *E. coli* Fic1 which showed this characteristic phenotype. The domain is present in many bacterial proteins and can be recognized by the HPFXXGNG motif. Fic homologues are found in closely related species *E. coli*, *Sinorhizobium* and in several other Gram-negative bacteria, but no function has been assigned to this gene in any of these systems. The localization of the Fic gene between loci encoding a presumable VirB / VirD4 T4SS suggest that the encoded protein may be a T4SS substrate or may have another function related to type IV secretion. The presence of the Fic domain in the secreted effector proteins BepA-BepC suggests a role in the eukaryotic host cell. Indeed, a search for Fic-like proteins in mammals retrieves the huntington-interacting protein E (HypE) [19]. The Fic domain is found to be annotated in the Pfam (PF02661) and the Interpro domain database (IPR003812). For BepA, the Fic region and the low complexity region near the C-terminus is highlighted from the Pfam database (Figure 37).

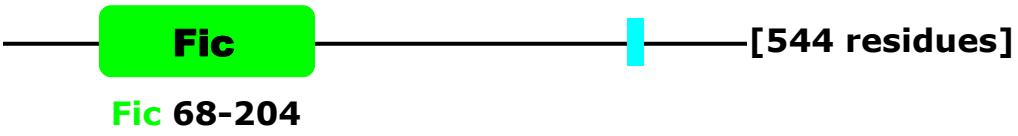


Figure 37: SwissPfam entry for Q6G2A9_BARHE (BepA) is PF02661. The presence of the Fic domain is shown in green colour for the putative cell filamentation protein BepA. Presence of low complexity region is indicated in blue color. The above information is obtained from Protein families database of alignments and HMMs, Wellcome Trust, Sanger institute.

4.2.2 Function of BID domain

The main function of BID domain appears to be to direct Beps into the host cell *via* VirB /VirD4 T4SS. Interestingly, results from our collaborator revealed that, BepA failed to perform the CRAFT assay. But using BepA-GFP fusion construct they have demonstrated that BepA gets translocated into the HUVEC but gets localized in the membrane regions of the cell. In addition, ectopic expression of BepA also confirmed its presence in the membrane region. By performing deletion constructs of BepA, our collobrators have identified that the BID domain of BepA is responsible for its membrane association. Apart from its translocation function, the BID domain of BepA was proven to elicit an anti-apoptotic response [30].

5.0 Conclusions

The expression studies on Beps (BepA-BepD) show that the effector proteins of the VirB/VirD4 T4SS are quite different from each other. They differ in their expression and localization properties in the *E. coli* system. Although BepA, B and C groups are highly homologues, they behaved differently from each other. Expression of BepD did not yield good results for structural studies. Among all the constructs, crystals were obtained only for the truncated BepA (BepA_tr) which subsequently yielded diffraction data for its structural investigation. Investigation on the new crystallization condition for BepA_tr is in progress. So far, no crystals were obtained for the other constructs. The progress of other constructs is under investigation. From the preliminary results obtained for BepA-BepD, we are trying to answer some of the questions which are given below.

Questions to be answered:

- What is the structure of the Fic domain?
- Can the structure of the Fic domain will help us to know its corresponding function in the target cells?
- What is the importance of the linker region present between the Fic and the BID domain?
- What is the structure of the BID domain? How does the localization of the BID domain in the membrane fraction of *E. coli* reveal their intrinsic nature in the host cells?
- What is the structure of full length BepA, B, C groups?
- How to avoid the degradation problem for further investigation of BepD?
- How does the structure of these effector proteins (BepA-BepD) will reveal the VirB/VirD4 mediated transport in *B. henselae*? What are the folding properties for Beps to mediate such a mechanism?

In order to answer these questions, we are in the process of investigating each of these Beps in detail. Currently, we are focusing on the BepA_tr, to solve the structure of the Fic domain and the linker region. Simultaneously, we are evaluating the BID domain of BepA in identifying the membrane binding region which is important for its localization in the eukaryotic cell. Interestingly, we have found that the BID is expressed as a single domain protein in *B. quintana*, closest homologue of *B. henselae*. We are expanding our work on this single domain protein for structural analysis. To attain insights into the BepD, short constructs of the pTyr domain or the homologues of BepD are under pipeline. On attaining the structure of the Beps, we can presumably analyze the translocation mechanism mediated through the VirB/VirD4 type IV pathway, their interacting partners and finally its subsequent interactions in the signaling pathways of the eukaryotic cells.

6.0 Perspectives

Structural investigation on the Type IV effector proteins (Beps) from *B. henselae* are fascinating subjects for study by virtue of their demonstrated contribution to the infection processes of several medically related diseases. The discovery of the structural components of the effector proteins and the cellular response of Type IV translocation mechanism will yield new information about signal transduction pathways activated during the course of infection. We have taken the first steps in performing the structural analysis on few of these Beps and preliminary insights into each of them have been established. Further studies will be continued on these proteins for obtaining the structure which may elucidate their pathogenic function in the target cells.

The structural biology of T4SSs is slowly advancing, with new X-ray structures and a multitude of biochemical data. The new challenge will be to visualize the complete secretion apparatus and to demonstrate their routes of entry into the hosts. Thus, the secretion system of the effector molecules increases the challenges to understand the process of secretion, the exact recognition signals for identifying the targets in the host cell and to decipher their effects on the biology of the recipient, at the cellular and the whole organism level. By learning more about the biology of the effector proteins of Type IV secretion system will increase the ability to exploit them as tools for biotechnology and as novel targets for antimicrobials.

Bibliography

1. Thanassi, D.G. and S.J. Hultgren, Multiple pathways allow protein secretion across the bacterial outer membrane. Curr Opin Cell Biol, 2000. **12**(4): p. 420-30.
2. Driessen, A.J., P. Fekkes, and J.P. van der Wolk, The Sec system. Curr Opin Microbiol, 1998. **1**(2): p. 216-22.
3. Cornelis, G.R., Yersinia type III secretion: send in the effectors. J Cell Biol, 2002. **158**(3): p. 401-8.
4. Chen, L., et al., A new type IV secretion system promotes conjugal transfer in Agrobacterium tumefaciens. J Bacteriol, 2002. **184**(17): p. 4838-45.
5. Cascales, E. and P.J. Christie, The versatile bacterial type IV secretion systems. Nat Rev Microbiol, 2003. **1**(2): p. 137-49.
6. Eisenbrandt, R., et al., Conjugative pili of IncP plasmids, and the Ti plasmid T pilus are composed of cyclic subunits. J Biol Chem, 1999. **274**(32): p. 22548-55.
7. Chai, Y., J. Zhu, and S.C. Winans, TrlR, a defective TraR-like protein of Agrobacterium tumefaciens, blocks TraR function in vitro by forming inactive TrlR:TraR dimers. Mol Microbiol, 2001. **40**(2): p. 414-21.
8. Kumar, R.B., Y.H. Xie, and A. Das, Subcellular localization of the Agrobacterium tumefaciens T-DNA transport pore proteins: VirB8 is essential for the assembly of the transport pore. Mol Microbiol, 2000. **36**(3): p. 608-17.
9. Christie, P.J., Agrobacterium tumefaciens T-complex transport apparatus: a paradigm for a new family of multifunctional transporters in eubacteria. J Bacteriol, 1997. **179**(10): p. 3085-94.
10. Christie, P.J. and J.P. Vogel, Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. Trends Microbiol, 2000. **8**(8): p. 354-60.
11. Burns, D.L., Type IV transporters of pathogenic bacteria. Curr Opin Microbiol, 2003. **6**(1): p. 29-34.
12. Farizo, K.M., et al., Membrane localization of the S1 subunit of pertussis toxin in

Bordetella pertussis and implications for pertussis toxin secretion. Infect Immun, 2002. **70**(3): p. 1193-201.

13. Odenbreit, S., et al., Translocation of Helicobacter pylori CagA into gastric epithelial cells by type IV secretion. Science, 2000. **287**(5457): p. 1497-500.
14. Nagai, H. and C.R. Roy, The DotA protein from Legionella pneumophila is secreted by a novel process that requires the Dot/Icm transporter. Embo J, 2001. **20**(21): p. 5962-70.
15. Foulongne, V., et al., Identification of Brucella suis genes affecting intracellular survival in an in vitro human macrophage infection model by signature-tagged transposon mutagenesis. Infect Immun, 2000. **68**(3): p. 1297-303.
16. Dehio, C., Bartonella interactions with endothelial cells and erythrocytes. Trends Microbiol, 2001. **9**(6): p. 279-85.
17. Karem, K.L., C.D. Paddock, and R.L. Regnery, Bartonella henselae, B. quintana, and B. bacilliformis: historical pathogens of emerging significance. Microbes Infect, 2000. **2**(10): p. 1193-205.
18. Dehio, C., et al., Interaction of Bartonella henselae with endothelial cells results in bacterial aggregation on the cell surface and the subsequent engulfment and internalisation of the bacterial aggregate by a unique structure, the invasome. J Cell Sci, 1997. **110** (Pt 18): p. 2141-54.
19. Schroder, G. and C. Dehio, Virulence-associated type IV secretion systems of Bartonella. Trends Microbiol, 2005.
20. Branley, J., et al., Prevalence of Bartonella henselae bacteremia, the causative agent of cat scratch disease, in an Australian cat population. Pathology, 1996. **28**(3): p. 262-5.
21. Foley, J.E., et al., Seroprevalence of Bartonella henselae in cattery cats: association with cattery hygiene and flea infestation. Vet Q, 1998. **20**(1): p. 1-5.
22. Karem, K.L., Immune aspects of Bartonella. Crit Rev Microbiol, 2000. **26**(3): p. 133-45.
23. Dehio, C. and A. Sander, Bartonella as emerging pathogens. Trends Microbiol, 1999. **7**(6): p. 226-8.
24. Dehio, C., Molecular and cellular basis of bartonella pathogenesis. Annu Rev

Microbiol, 2004. **58**: p. 365-90.

25. Dehio, C., Recent progress in understanding Bartonella-induced vascular proliferation. Curr Opin Microbiol, 2003. **6**(1): p. 61-5.
26. Schulein, R., et al., A bipartite signal mediates the transfer of type IV secretion substrates of Bartonella henselae into human cells. Proc Natl Acad Sci U S A, 2005. **102**(3): p. 856-61.
27. Schmid, M.C., et al., The VirB type IV secretion system of Bartonella henselae mediates invasion, proinflammatory activation and antiapoptotic protection of endothelial cells. Mol Microbiol, 2004. **52**(1): p. 81-92.
28. Kirby, J.E. and D.M. Nekorchuk, Bartonella-associated endothelial proliferation depends on inhibition of apoptosis. Proc Natl Acad Sci U S A, 2002. **99**(7): p. 4656-61.
29. Kawamukai, M., et al., Cloning of the fic-1 gene involved in cell filamentation induced by cyclic AMP and construction of a delta fic Escherichia coli strain. J Bacteriol, 1988. **170**(9): p. 3864-9.
30. Schmid, M. (2005). Ph.D thesis, University of Basel.
31. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr, 1994. **50**(Pt 5): p. 760-3.
32. Thompson, J.D., D.G. Higgins, and T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res, 1994. **22**(22): p. 4673-80.
33. Rost, B., C. Sander, and R. Schneider, Redefining the goals of protein secondary structure prediction. J Mol Biol, 1994. **235**(1): p. 13-26.
34. Cuff, J.A., et al., JPred: a consensus secondary structure prediction server. Bioinformatics, 1998. **14**(10): p. 892-3.
35. Goodstadt, L. and C.P. Ponting, CHROMA: consensus-based colouring of multiple alignments for publication. Bioinformatics, 2001. **17**(9): p. 845-6.
36. Liu, J. and B. Rost, NORSp: Predictions of long regions without regular secondary structure. Nucleic Acids Res, 2003. **31**(13): p. 3833-5.
37. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D.,

Bairoch *Protein Identification and Analysis Tools on the ExPASy Server;* (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005). pp. 571-607

38. Matthews BW. Solvent content of protein crystals. J Mol Biol. 1968 Apr 28;**33**(2):491-7.

Appendix

Curriculum Vitae

Dinesh Vellore Palanivelu

ADDRESS	Division of Structural Biology, Biozentrum, University of Basel, Klingelbergstrasse 50/70, CH – 4056 Basel, Switzerland
TELEPHONE	+41 61 267 2092
EMAIL	dinesh.palanivelu@unibas.ch
NATIONALITY	Indian

EDUCATION

2001 – Present	PhD in Biophysics, Prof. Tilman Schirmer's Group, Biozentrum, University of Basel, Switzerland **"Structural investigation into recombinant eye lens aquaporin (AQP0) and the effector proteins (BEPS) from *Bartonella henselae*"**
1998 – 1999	MPhil Biochemistry, Madras University, Chennai, INDIA
1998 – 1999	Post graduate diploma in computer application, Loyala College, Chennai, INDIA
1996 – 1998	MSc Biochemistry, Annamalai University, Chidambaram, INDIA
1992 – 1995	BSc Biochemistry, Madras University, Chennai, INDIA
1980 – 1992	Don Bosco Higher Secondary Matriculation School, Chennai, INDIA

PUBLICATION

Dinesh Palanivelu, Peter Agre, David Kozono, Ariel Lustig, Kitaru Suda, Andreas Engel and Tilman Schirmer. **Co-axial association of eye lens aquaporin AQP0.** (To be submitted)

RELEVANT EXPERIENCE

2004	Presented poster in the EMBO conference on "Structures in Biology", EMBL, Heidelberg, Germany.
	Presented poster in the "Biozentrum Symposium", Switzerland.
	Oral presentation in the "18th Regional Meeting on X-ray Crystallography of Biomacromolecules", Switzerland.
	Presented poster in the "22nd European Crystallographic Meeting", Budapest, Hungary.
2003	Presented poster in the "2nd International conference on Structure, Dynamics and Function of Proteins in Biological Membranes", Switzerland.
	Presented poster in the "Biozentrum Symposium", Switzerland.
2002	Oral presentation in the "Aquaporin Conference", Switzerland.
	Participitated in the "High-Throughput Techniques in Structural Biology" in EMBL, Heidelberg, Germany.
	Participitated in the "Membrane protein Symposium", Zurich, Switzerland.
2001	Participated in the Symposium and Workshop in "Current trends in Bioinformatics", Indian Institute of Science, Bangalore, INDIA.
1998	Attended the "All India science congress", Anna University, Chennai, INDIA.
1997	Oral presentation in the Symposium "SOCIETY OF BIOLOGICAL CHEMISTS", INDIA.

OTHER INVOLVEMENT IN SCIENCE

Present	Assist in drafting application proposals for synchrotron beam time.
	Organise the X-ray group seminar and discussion in the crystallography department.
2002 – 2003	Tutoring in "Basics of Biology" for final year school students.
2001 – Present	Instruct and supervise final year undergraduate students, and assess their performance in practical courses.

LANGUAGES Fluent English and Tamil

Basic German

AWARDS Awarded the best student in the academics in my Masters degree (1996 – 1998), Annamalai University, INDIA

REFERENCE

Prof. Tilman Schirmer, Division of Stuctural Biology, Biozentrum, University Of Basel Klingelbergstrasse 50-70, CH – 4056 Basel, Switzerland.

Phone : +41 61 267 2089

Email : tilman.schirmer@unibas.ch