MASARYKOVA UNIVERZITA
PŘÍRODOVĚDECKÁ FAKULTA
ÚSTAV BIOCHEMIE

Strukturní studie vybraných virových proteinů

Diplomová práce

Roman Baška

Vedoucí práce: Mgr. Pavel Plevka, Ph.D. **Brno 2019**

Bibliografický záznam

Autor:	Roman Baška Přírodovědecká fakulta, Masarykova univerzita Ústav biochemie
Název práce:	Strukturní studie vybraných virových proteinů
Studijní program:	Biochemie
Studijní obor:	Analytická biochemie
Vedoucí práce:	Mgr. Pavel Plevka, Ph.D.
Akademický rok:	2018/2019
Počet stran:	68
Klíčová slova:	fág; fágová terapie; *Staphylococcus aureus*; virus; proteinová krystalografie; strukturní

Bibliographic Entry

Author	Roman Baška Faculty of Science, Masaryk University Department of Biochemistry
Title of Thesis:	Structural study of selected virus proteins
Degree programme:	Biochemistry
Field of Study:	Analytical biochemistry
Supervisor:	Mgr. Pavel Plevka, Ph.D.
Academic Year:	2018/2019
Number of Pages:	68
Keywords:	phage; phage therapy; *Staphylococcus aureus*; virus; protein crystallography; structural

Abstrakt

S globálním nárustem rezistence bakterií vůči antibiotikům se nemoci způsobené zlatým stafylokokem stávají rostoucí hrozbou pro pacienty. Fágová terapie je alternativní léčbou k antibiotikům, s mnoha výhodnými vlastnostmi, jako vysoká specifita a nízká tvorba rezistence u bakterií. Fág phi812 a jeho mutanti infikují více jak 95 % testovaných kmenů *Staphylococcus aureus*, což z nich činí slibný léčebný prostředek. Nicméně, stále je nedostatek dat o struktuře tohoto fága, která by poskytla lepší porozumění jeho interakci s hostitelem a pomohla zavést fága phi812 do klinické praxe jako antibakteriální léčivo.

Tato práce se zabývá řešením struktury čtyř proteinů stafylokokového fága phi812, kombinací metod molekulární biologie a rentgenové krystalografie.

Abstract

With global increase in resistance of bacteria against antibiotics, diseases caused by staphylococcus are becoming an increasing threat to patients. Phage therapy is an alternative treatment to antibiotics, with many advantageous attributes, like high specifity and low probability of occurence of resistance in bacteria. Phage phi812 and its mutants are eliminating over 95 % of tested *Staphylococcus aureus* strains, making them promising therapy agents. However, there is lack of data on the phage structure, which would provide better understanding of its interaction with host and help to install phage phi812 in clinical practice as antibacterial therapeutics.

This thesis deals with structure determination of four proteins of staphylococcal phage phi812, by combining methods of molecular biology and X-ray crystallography.

MASARYKOVA UNIVERZITA
Přírodovědecká fakulta

ZADÁNÍ DIPLOMOVÉ PRÁCE

Akademický rok: 2016/2017

Ústav: Ústav biochemie

Student: Bc. Roman Baška

Program: Biochemie

Obor: Analytická biochemie

Ředitel *Ústavu biochemie* PřF MU Vám ve smyslu Studijního a zkušebního řádu MU určuje diplomovou práci s názvem:

Název práce: Strukturní studie vybraných virových proteinů

Název práce anglicky: Structural study of selected virus proteins

Oficiální zadání:
Jedna desetina produkce zemědělských plodin závisí na opylování hmyzem, které je prováděno především včelou medonosnou (Apis mellifera). Nicméně počet včelstev v USA a EU klesá. Virové infekce, způsobené viry z rodin Iflaviridae a Dicistroviridae způsobují kolapsy včelstev. V současné době neexistuje žádný lék na virové infekce včel. Student provede klonování, expresi proteinu, krystalizaci, sběr dat z rentgenové krystalografie a stanovení struktury 3C proteáz a 3D replikáz těchto virů. Struktury umožní vývoj protivirových látek. Následně bude student analyzovat komplexy těchto proteinů s inhibitory. Další částí výzkumného projektu bude určení struktur proteinů bakteriofága phi812, které jsou součástí kapsidy fága.

Jazyk závěrečné práce: angličtina

Vedoucí práce: Mgr. Pavel Plevka, Ph.D.

Konzultant: Mgr. Marta Šiborová

Datum zadání práce: 24. 2. 2017

V Brně dne: 23. 3. 2017

Souhlasím se zadáním (podpis, datum):

.....	.....	.....
Bc. Roman Baška	Mgr. Pavel Plevka, Ph.D.	doc. RNDr. Oldřich Janiczek, CSc.
student	**vedoucí práce**	**zástupce ředitele Ústavu biochemie pro pedagogické záležitosti**

Acknowledgement

First I would like to thank my supervisor Mgr. Pavel Plevka, Ph.D. for the opportunity to work in his research group and to gain valuable experience for further life.

I would also like to thank all my colleagues for their help and creating of friendly working environment. Especially to Marta, who took me under her wings and was great help ever since.

Many thanks belong to my family memebers for their support. Special thanks belong to Katherine and Patrik.

Prohlášení

Prohlašuji, že jsem svoji diplomovou práci vypracoval samostatně s využitím informačních zdrojů, ktoré jsou v práci citovány.

Brno 7. leden 2019 .....

Roman Baška

Table of contents

Introduction **10**

1 Theoretical section **11**

1.1 Viruses 11

1.1.1 Taxonomy of viruses 11

1.1.2 Prokaryotic viruses 12

1.1.3 Order *Caudovirales* 14

1.1.4 Staphylococcal phage phi812 16

1.1.5 Structure of phage phi812 17

1.2 X-ray crystallography 19

1.2.1 Crystals 19

1.2.2 X-ray diffraction by a crystal 20

1.2.3 Calculation of electron density distribution 21

1.2.4 Phase problem 22

1.2.5 Isomorphous replacement method 22

1.2.6 Anomalous scattering 23

1.2.7 Molecular replacement 23

1.2.8 Structure determination 23

2 Materials **24**

2.1 Chemicals 24

2.2 Equipments 25

3 Methods **26**

3.1 Cloning 26

3.1.1 Polymerase chain reaction (PCR) 26

3.1.2 Plasmid and DNA isolation 27

3.1.3 Agarose gel electrophoresis 27

3.1.4 Ligase independent cloning (LIC) 27

3.1.5 Transformation of competent cells 29

3.1.6 Colony PCR 29

3.2 Expression tests 30

3.2.1 Induction with IPTG 31

3.2.2 Induction by lactose 31

3.2.3 SDS-PAGE 33

3.3 Protein purification 34
3.4 Crystallization of purified proteins 35
3.5 Solving of protein structure 36
3.6 Model building and refinement 37
4 Results 38
4.1 BmpA 38
4.2 BmpB 45
4.3 BmpC 50
4.4 Prohead protease 56
5 Conclusions 62
6 Summary 65
7 References 66

Introduction

Antibiotics were introduced to human medicine in 1940s and played major role in fight against bacterial diseases [1]. After many decades of overuse and irresponsible prescribing in human and veterinary medicine, highly resistant strains of bacteria spread worldwide [1,2]. This represents a threat to patients, as two million health care-associated infections (HAIs) occur each year in United States alone, causing about hundred thousand casualties. Economic costs of the U.S. health care system caused by antimicrobial resistance in 2006 were estimated at $20 billion [3], while in the European union, hospitalization costs caused by Methicillin-resistant *Staphylococcus aureus* (MRSA) infections are estimated at €380 million annually [4]. The unpleasant situation could be handled by developing prevention strategies and novel therapeutics.

Staphylococcus aureus is bacterium associated with diseases like pneumonia and septicaemia [5]. As an alternative treatment of *S. aureus* infections, phage therapy seems to be suitable for its high specificity, slow development of resistance in bacteria and successful *in vivo* trials [6]. Suitable agent for phage therapy, which eliminates over 95 % of tested *S. aureus* strains, is phage phi812 [7]. However, there are many obstacles for phage therapy to be routinely used in human medicine, mainly due to regulatory concerns [8]. Detailed knowledge of virus structure, interaction with host and life cycle is important step towards introduction of phage therapy in clinical practice of western countries.

Aim of this study is to solve structure of selected proteins of *S. aureus* phage phi812. Enzyme prohead protease plays a role in capsid maturation. Proteins BmpA, BmpB and BmpC are putative building blocks of the phage base plate. Combination of methods of molecular biology and X-ray crystallography is used to produce proteins and solve their structures.

1 Theoretical section

1.1 Viruses

Viruses are obligatory intracellular parasites, which infect all cellular organisms. They are the most abundant live form in biosphere, with concentrations up to 10^8 per millilitre in marine waters [9]. Viruses play important role in global ecosystems and biogeochemical cycles [10]. Viruses reproduce within host cells and exist in extracellular environment as virions. Virions are non-cellular biomolecular assemblies containing genetic material and all the necessary apparatus for virus to infect its host. Virions contain either DNA or RNA, both in various forms as shown in Table 1. Viruses do not contain all the necessary machinery to reproduce on their own, therefore they need to use mechanisms available in host cell [11,12].

DNA genome		RNA genome	
dsDNA, circular	ssDNA, circular	dsRNA, linear	ssRNA, circular
dsDNA, linear	ssDNA, linear	ssRNA, linear	

Table 1. **Types of genomes present in viruses** [11].

Genetic material inside of virions is protected by protein structure - capsid. It is built of one or more types of proteins. Capsids usually exhibit icosahedral or helical symmetry and can be also covered by protective lipid membrane [11].

1.1.1 Taxonomy of viruses

Diversity among viruses renders it difficult to classify them in one common way. The International Committee on Taxonomy of Viruses (ICTV), assigns viruses a taxonomy group and sets rules for their nomenclature. To decide, which taxonomy rank virus belongs to, may depend on morphology of virion, type of nucleic acid, similarity between genome sequences or host preferences.

Useful tool to sort viruses is Baltimore classification. There are seven classes of viruses based on type of nucleic acid and replication strategy of viruses, thus using similarity features, which are not always reflected by the ICTV taxons.

	Class	Nucleic acid	Synthesis of mRNA	
DNA viruses	1	dsDNA	directly transcribed	mRNA
	2	(+) DNA or (-) DNA	opposite strand synthesised, then dsDNA transcribed	
RNA viruses	3	dsRNA	directly used in translation	
	4	(-) RNA	(+) RNA synthesised, then used in translation	
	5	(+) RNA	directly used in translation	
Reverse-transcribing viruses	6	(+) RNA	synthesised RNA/DNA double strand, then dsDNA and transcribed	
	7	dsDNA	directly transcribed; also (+) RNA synth., into (-) DNA, then dsDNA	

Fig. 1. **Baltimore classification.** Seven classes of viruses are established by different pathways of mRNA synthesis. Each class of viruses contains different type of nucleic acid in their virions. DNA viruses synthesise mRNA from dsDNA. RNA viruses need to synthesise (+) RNA strand or they already contain the plus strand within virion, thus directly using it for translation [11,13].

1.1.2 Prokaryotic viruses

Viruses infecting Bacteria and Archaea are the largest group of known viruses. Those infecting bacteria are also referred to as bacteriophages or phages. Majority of phages contain dsDNA, but members with ssDNA, ssRNA and dsRNA are also known [14]. There is large diversity among families regarding genomes sequences and virion morphology [15].

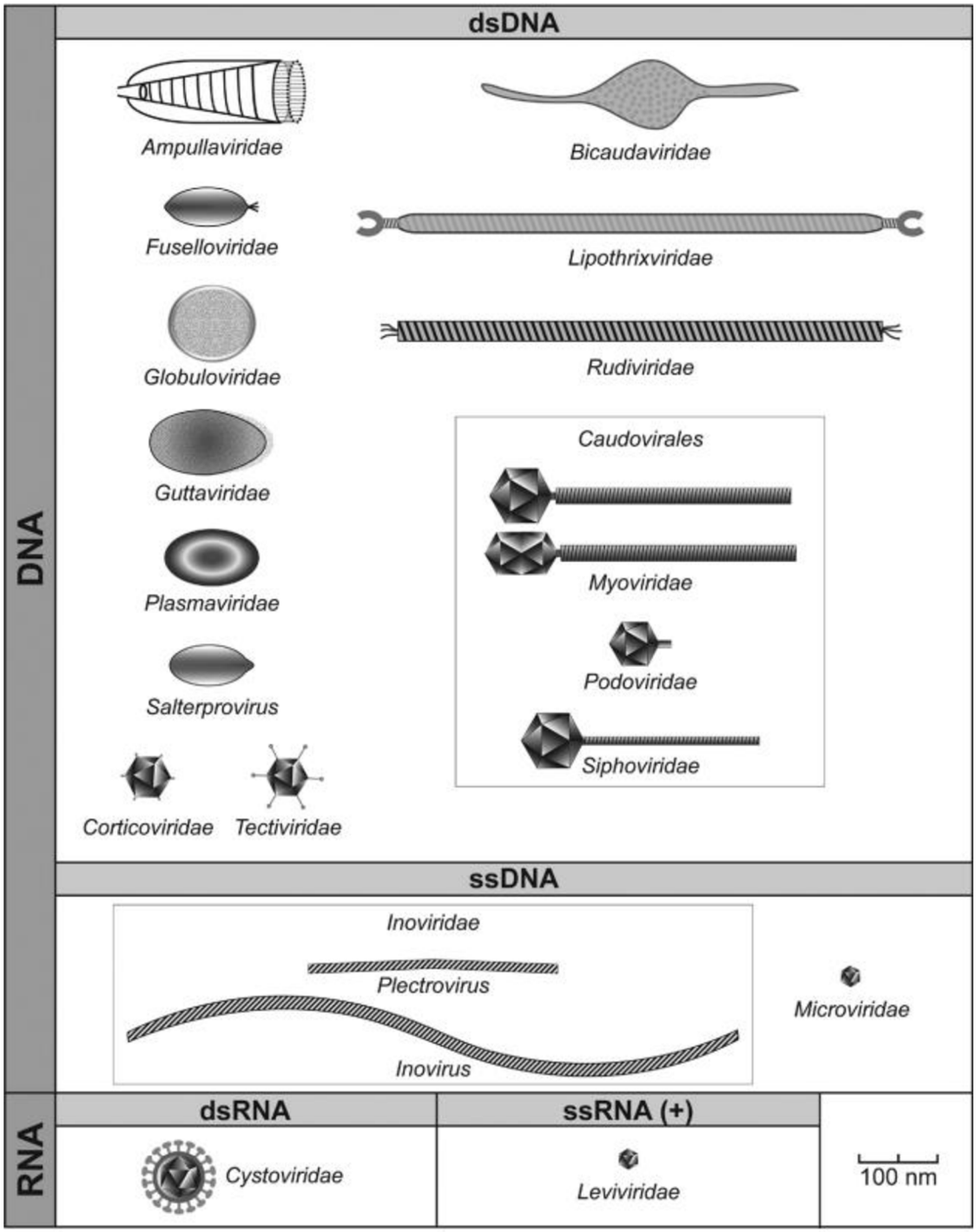

Fig. 2: **Morphology of *Bacteria* and *Archaea* infecting viruses.** *Plasmaviridae* contain lipoprotein envelope, *Fusseloviridae* have short spikes at one end, *Inoviridae* phages are long filaments, *Lipothrixviridae* are rod-like shaped and have lipoprotein envelope, *Rudiviridae* are straight rigid rods resembling tobacco mosaic virus. *Caudovirales* includes families of tailed phages. *Cystoviridae* has outer protein layer and inner lipid membrane. *Tectiviridae* has icosahedral capsid inside of lipoprotein vesicle with spikes; taken from [14].

1.1.3 Order *Caudovirales*

Caudovirales comprises four families of phages, which represent about 95 % of all known bacterial viruses [15]. The families were established by different morphology of virions, where *Myoviridae* contains phages with long contractile tails, *Siphoviridae* with long non-contractile tails and *Myoviridae* with short non-contractile tails [16]. Fourth family, *Ackermannviridae*, was established just recently [17] and comprise mostly former *Myoviridae* members, based on genomic similarities.

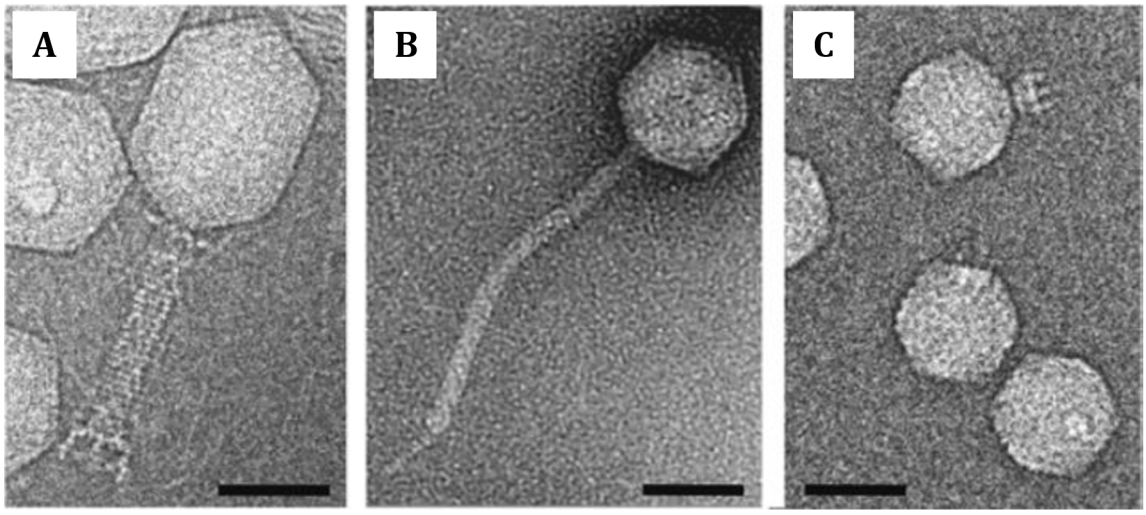

Fig. 3. **Phages from order *Caudovirales* under electron microscope.** A) Phage T4 (*Myoviridae*), B) SPP1 (*Siphoviridae*), C) P22 (*Podoviridae*); taken from [14].

Virions of viruses belonging to the *Caudovirales* family are non-enveloped, composed of capsid and tail. Length of virion is up to 400 nm. Proteins, which form head, are assembled with icosahedral symmetry, or in case of prolate heads, with symmetry derived from icosahedron. Head contains coil of dsDNA, which sequence contain large operons and highly clustered genes. Genomes are diverse, up to 400 kb long [16].

Tail is helical or composed of stacks of subunits, with tail tube exhibiting six-fold or three-fold symmetry. Other common tail parts are base plate and terminal fibers. Tail and capsid are bound via connector. Connector plays role in packaging of DNA into capsid, keeping DNA inside of capsid and enables its release out of capsid [18].

Virus begins infection of host cell by adsorption of tail to proteins or polysaccharides of cell surface. Virion then undergoes conformational changes, leading to ejection of genome into host cell. DNA is subsequently being replicated and transcribed into mRNA, which is further translated into virus proteins. These proteins ensure replication of phage DNA and assembly of novel virions. Separate pathways are usually needed for maturation of head and tail compartments. Head originates out of procapsid, built of capsid proteins and assembled with the help of scaffolding proteins. DNA is inserted into capsid by terminase through portal protein. Pressure inside of full capsid caused by stored DNA can reach up to 6 MPa. Progeny virions escape the cell after cell lysis, that is driven by cleavage of peptidoglycan by virus hydrolases [14,18].

Development of infection by virus depends on whether the virus is virulent or temperate. Virulent phages enter lytic cycle, which ends with cell death by lysis and release of novel virions. Temperate viruses can enter either lytic or lysogenic cycle. In lysogenic cycle, their genome is integrated into bacterial chromosome or stays in form of plasmid in cytoplasm. In this latency stage, virus is referred to as prophage and can stay within cell for many generations. Prophage may eventually enter the lytic cycle [18,19].

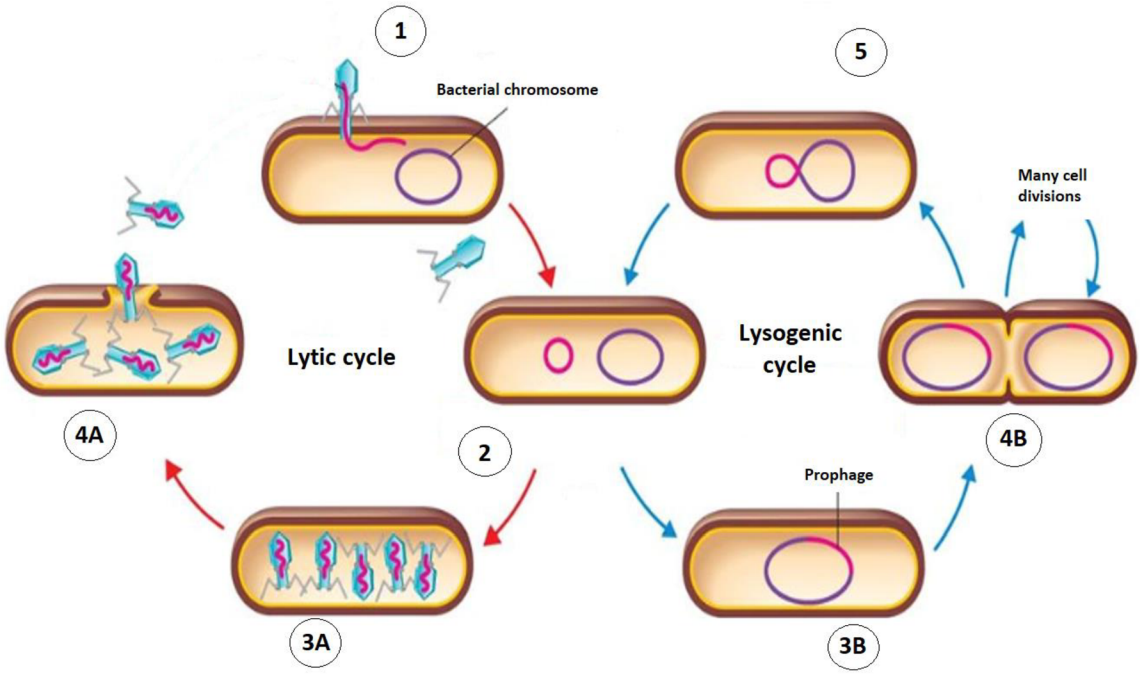

Fig. 4. **Scheme of lytic and lysogenic life cycle.** 1) Phage attaches to host cell and injects genome. 2) DNA enters either lytic or lysogenic cycle. 3a) New genome molecules and virus proteins are synthetized. Novel virions are assembled.

3b) Virus remains as prophage by incorporating its DNA into bacterial genome. 4a) Cell lysis and virion release. 4b) Bacterium undergoes division cycles. 5) Prophage excises from bacterial genome and enters lytic cycle; taken from [20].

1.1.4 Staphylococcal phage phi812

Staphylococcus aureus is gram-positive bacterium associated with wide range of human diseases, including skin infections, pneumonia or septicemia [21,22]. With global increase of bacterial resistance towards antibiotics, staphylococcus related infections are serious threat to patients, with eleven thousand deaths in the U.S. per year [2]. Novel therapeutics are developed to target this bacterium, including phage therapy. Phage phi812 and its mutants infect 95 % of tested *S. aureus* strains and are studied as potential phage therapy agents [7].

Phage therapy is based on the ability of phages to kill their natural hosts. In late stages of infection, phage encoded holins and lysins play role in degradation of bacterial cell wall with subsequent osmotic lysis of the cell and release of progeny virions [23].

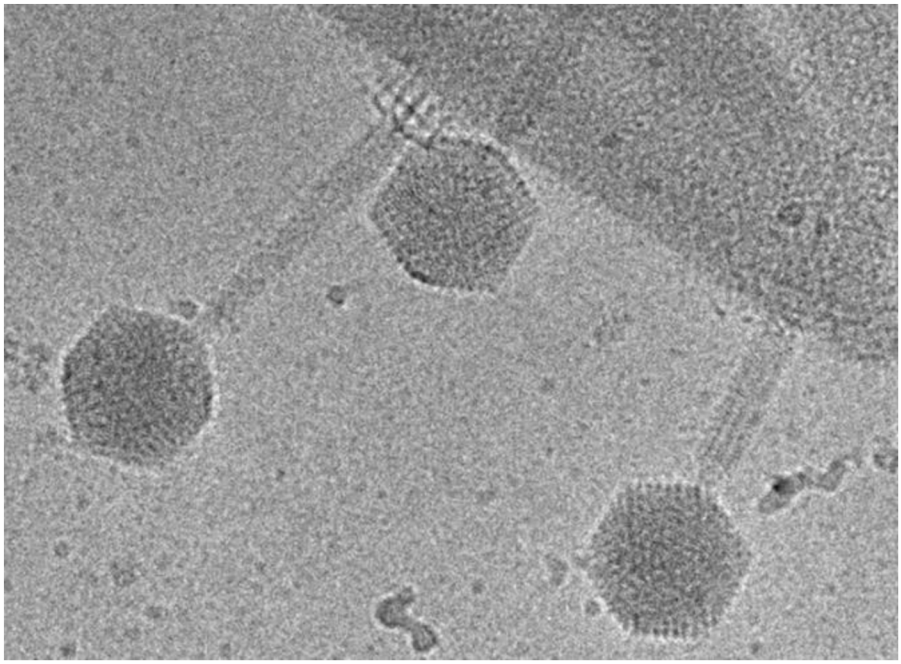

Fig. 5. **Cryo-EM image of staphylococcal phage phi 812 attached to the cell wall of *S. aureus*.** (Courtesy of M. Šiborová)

1.1.5 Structure of phage phi812

Virion head is isometric and 90 nm in diameter. Hexameric major capsid protein interacts with cement protein to stabilize the capsid. There are 16 major capsid proteins and five and one third of cement protein in icosahedral asymmetric unit of the capsid [24].

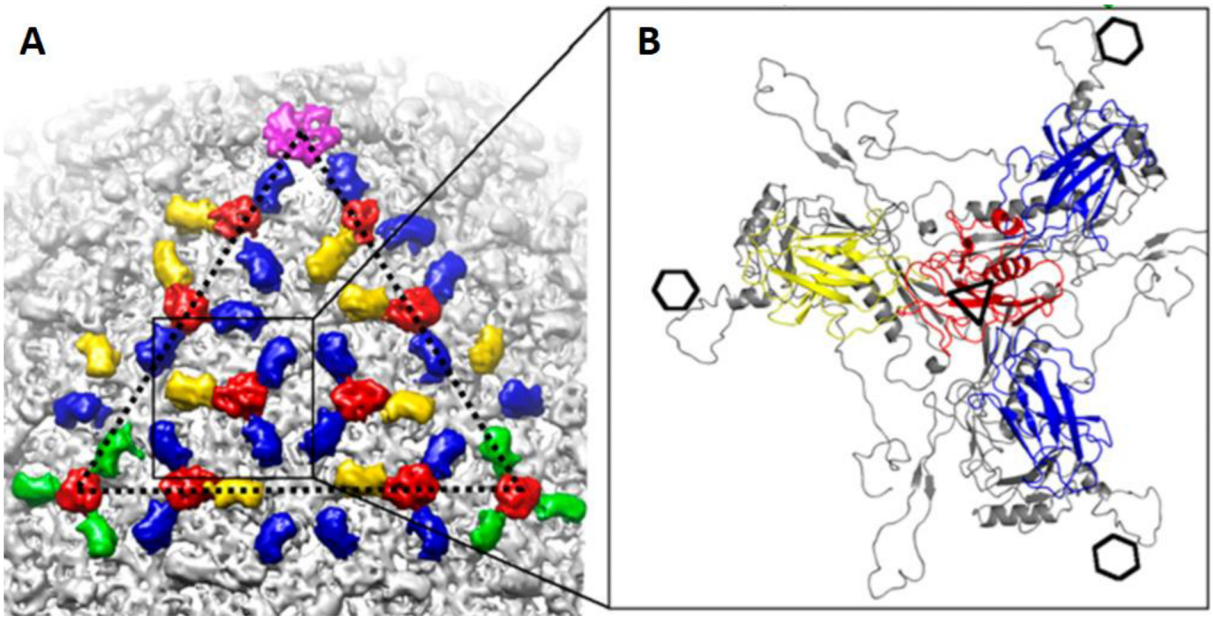

Fig. 6. **Phage phi812 and icosahedral asymmetrical unit of capsid.** A) Dotted triangle represents icosahedral asymmetric unit of the capsid. Cement protein is shown in red, other colors represent domains of major capsid protein. B) Cement protein forms connections to three major capsid proteins; taken from [24].

Tail of phi812 is 240 nm long with base plate 55 nm wide and 30 nm high. There are 312 copies of tail sheath protein along the tail, which are arranged in hexamer discs. The structure suggests that interaction with *S. aureus* cell wall begins with tail fibers. This triggers conformational changes upon which base plate no longer holds to tail tube and tail sheath proteins cause the tail to contract. After contraction tail sheath remains 40 % of original length and tail tube becomes exposed in order to penetrate into cell of *S. aureus*. Genome is not released upon conformational changes of tail sheath but subsequent conformational change inside of tail tube is necessary for genome release [24].

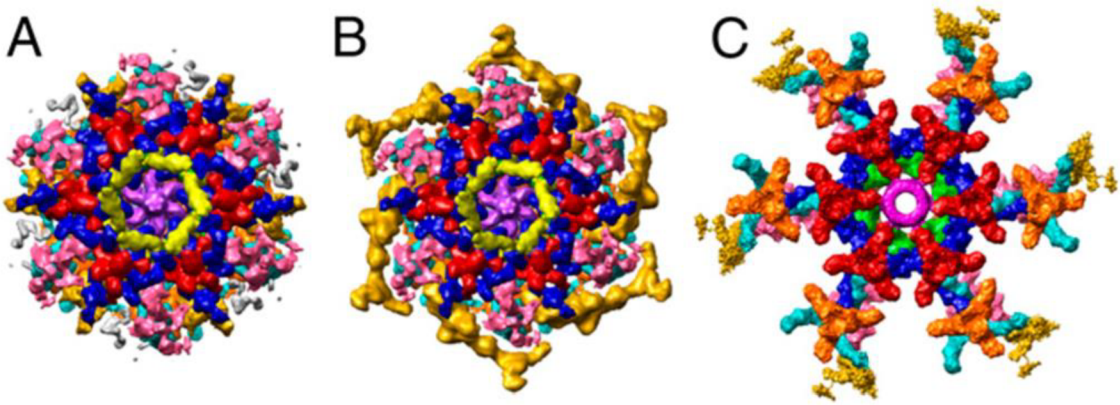

Fig. 7. **Baseplate of phage phi812 resolved by cryo-EM and bottom view at conformational changes during tail contraction.** A) Native base plate. B) Intermediate state. C) Base plate after tail contraction. Individual components of base plate are color coded; taken from [24].

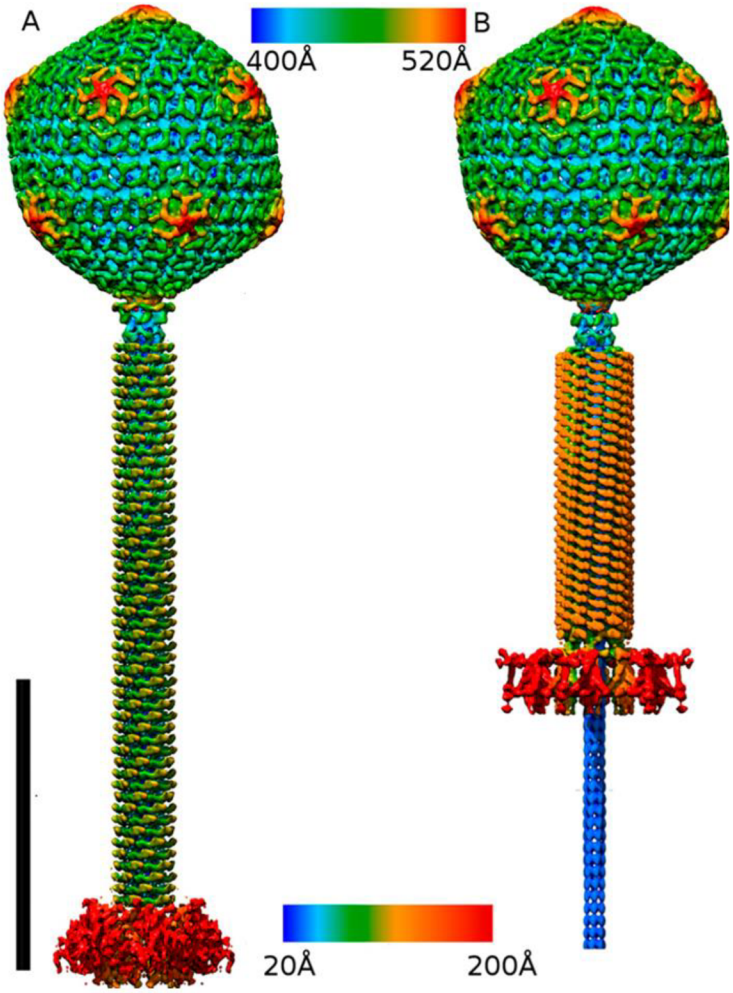

Fig. 8. **Cryo-EM recontruction of phage phi812 virions.** A) Virion in native state and B) after contraction of tail. Scale bar is 1 000 Å. Color bars represent distances of head from head centre and tail from tail six fold axis; taken from [24].

1.2 X-ray crystallography

Since 1914, when structure of several inorganic salts was determined using X-rays [25], crystallography has taken irreplaceable role in development of many scientific fields and determining structures from simple inorganic molecules to large biomolecular complexes. X-rays are electromagnetic radiation [26] with wave-length ranging from 0.01 to 10 nm [27]. As high energy photons of this radiation hit sample, they are elastically scattered mostly by electrons and after passing through the sample, scattered photons are detected. From obtained diffraction pattern, the original structure can be derived, since it represents the Fourier transform of electron density distribution. Molecules of the sample have to form well-ordered crystals. The reason to require crystals is that their inner symmetry enables constructive addition of waves in certain directions, thus creating a diffraction pattern [27].

The more electrons individual atoms have, the more likely the scattering event is. This brings various challenges to the measurement of biological samples. Crystals might not have sufficient quality and the scattering is in general weak, since the usual elements present in bio-samples are oxygen, nitrogen, carbon and hydrogen – all of which have small amount of electrons in their atomic shells. Therefore, highly coherent X-ray beams need to be used – commonly generated by synchrotrons [28].

1.2.1 Crystals

Crystals are objects with symmetric inner stucture. Whole structure can be represented by a unit cell, which can fill entire volume of crystal by a simple translation. Smallest part of the unit cell, that can fill entire unit cell by using symmetry operations, is asymmetric unit. Molecules are ordered in unit cell into motifs, and based on the motif symmetry, unit cell belongs to one of 230 space groups. The inner symmetry of crystal is important for structure determination, as it influences location of reflections in a diffraction pattern [26,28].

1.2.2 X-ray diffraction by a crystal

When X-ray beam passes through a crystal, photons are scattered by electrons. Photons travel as oscillating waves in space, which create new wavefronts upon scattering by crystal. Since many photons are part of the beam, these new waves interfere with each other. Total wave scattered by crystal is detectable primarily in directions where waves added together with the same phase. Eventually, they create a diffraction pattern on the detector. This event happens in cases (Fig. 9), which are described by Bragg's law:

$$2d \cdot \sin\theta = n\lambda,$$

where d is distance between Bragg planes, θ is scattering angle, n is order of reflection (an integer number) and λ is wavelength of radiation [26,29].

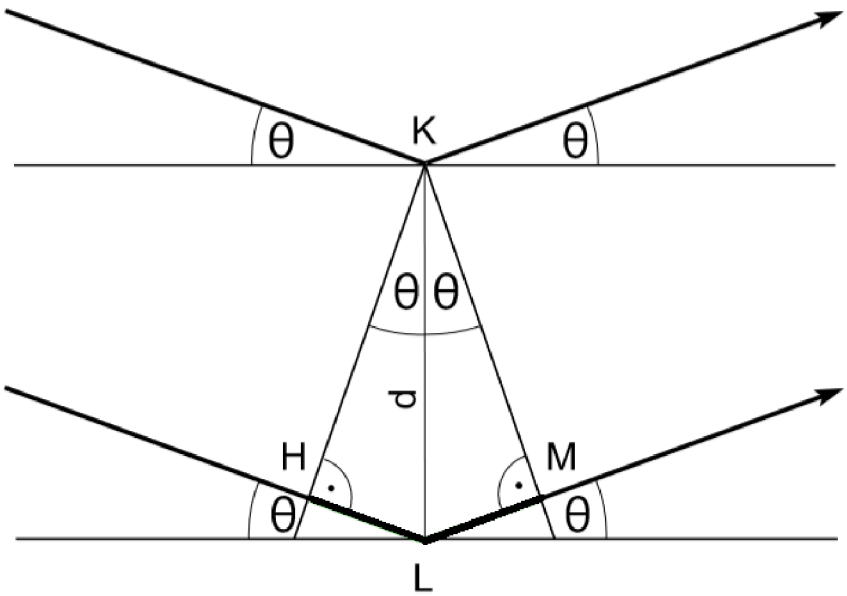

Fig. 9. **Incoming parallel waves with same direction are adding in phase.** Waves are scattered from the same crystal plane (K) or more parallel planes (K, L) with the same phase, when conditions described by Bragg's law are fulfilled. This means that pathlength difference of incident waves is equal integer number of wavelengths; taken from [31].

1.2.3 Calculation of electron density distribution

To solve structure of a molecule, its electron density distribution must be known, since it corresponds to the shape of molecule. To calculate electron density distribution, it is necessary to mathematically describe the scattering of X-rays by crystal. Useful tool to calculate wave addition is Argand diagram. It is 2D axial system where amplitudes of waves are vectors and phase shifts are angles between these vectors.

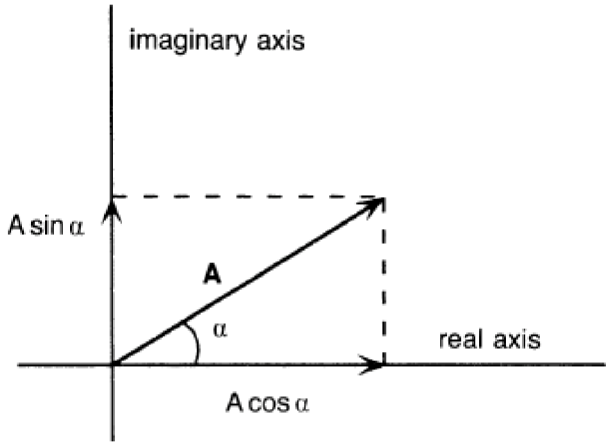

Fig. 10. **Argand diagram.** Vector A equals aplitude of wave and α represents phase shift with respect to the real axis; taken from [27].

To calculate scattered wave by a crystal, each element needs to be assigned a scattering factor. Value of scattering factor depends on amount of electrons in atomic shell and scattering angle of the beam. Scattering by unit cell can be then written as summation of scattering factors of all its atoms and their phase shifts, with respect to the origin of unit cell:

$$\mathbf{F}(\mathbf{S}) = \sum_{j=1}^{n} f_j \ \exp\left[2\pi i \mathbf{r}_j \cdot \mathbf{S}\right].$$

The F(S) is called structure factor and represents the sum of scattering by all atoms in unit cell. F(S) can be also written as an integration over unit cell:

$$F(h\ k\ l) = V \int_{x=0}^{1} \int_{y=0}^{1} \int_{z=0}^{1} \rho(x\ y\ z) exp[2\pi i(hx + ky + lz)]\ dx\ dy\ dz.$$

Structure factors are functions of electron density and by inverse Fourier transformation we obtain formula for calculating electron density distribution:

$$\rho(x\ y\ z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l}|F(h\ k\ l)|\ exp[-2\pi i(hx + ky + lz) + i\alpha(h\ k\ l\ )].$$

Structure factors can be measured from intensities of diffraction spots, while indexes (*hx*, *ky*, *lz*) are obtained from positions of the spots in the diffraction patterns. However, information about phases is lost during measurement and needs to be obtained with different approach [27].

1.2.4 Phase problem

As photons are detected, they only provide information about how many were scattered from particular direction, since detector only measures intensities of diffraction spots, but not the relative phases of incident photons. Various approaches were developed to determine the phase shifts between scattering centres in a molecule, thus enabling the calculation of electron density distributions [32].

When intensities of diffraction spots are Fourier transformed with zero phase angles, Patterson function is obtained:

$$P(\ u\ v\ w) = \frac{1}{V}\sum_{hkl}|F(h\ k\ l)|^2\ cos[2\pi(h\ u + k\ v + l\ w)].$$

The Patterson map represents all interatomic vectors in a molecule with each atom being in the origin. For small molecules the structure can be determined from this map without knowing phases, but for large molecules this map is not interpretable, however, it is still a useful tool in protein phase determination [33].

1.2.5 Isomorphous replacement method

In this method, native protein crystal is modified with heavy atoms. These atoms contain high number of electrons that strongly affect intensity of reflections. Modified crystal has therefore different structure factors than native crystal. By obtaining reflection intensities from native and modified crystals, coordinates of heavy atoms in unit cell can be calculated with Patterson function. With these data,

protein phases can be calculated geometrically with Harker construction. It is necessary that unit cell size and conformation of molecules are the same in native and modified crystals [34]. Heavy atoms are attached to crystal molecules by soaking method or by co-crystallization. Commonly used compounds to modify a crystal contain lanthanide ions, lead, platinum [35], iodine [36] or selenium [37].

1.2.6 Anomalous scattering

Anomalous scattering occurs when electron absorbs energy from photon, which is scattered inelastically. Anomalous scattering creates difference in structure factor amplitudes between pairs of reflections, which are scattered from the same crystal plane, but in opposite directions. These mirror reflections are Friedel pairs. Unlike in isomorphous replacement method, one crystal that contains suitable heavy atom is enough to determine protein phase angles, as reflections can be collected under different wavelengths. This enables to determine coordinates of heavy atoms in unit cell and subsequently protein phase angles [26, 37].

1.2.7 Molecular replacement

This method can be applied in case when similar protein structure has been previously determined. Assumption is that proteins with similar amino acid sequence have similar structures. The initial phase estimates of the new structure are provided by the already known structure to solve the phase problem. Model structure is first rotated and then translated in unit cell, to find the best orientation of model to match the observed diffraction [39].

1.2.8 Structure determination

After calculating electron density map of molecule, we build model of desired structure into this map. While building, several rounds of refinement are needed to improve protein phases. Correctness of model is indicated by R-factor:

$$R = \frac{\sum_{hkl} \left| |F_{obs}| - k|F_{calc}| \right|}{\sum_{hkl} |F_{obs}|},$$

where structure factors observed (F_{obs}) and structure factors from model (F_{calc}) are compared for agreement; k is the scaling factor [26,39].

2 Materials

2.1 Chemicals

Agarose	Sigma-Aldrich
Acrylamide mix	Sigma-Aldrich
Coomassie G-250	Sigma-Aldrich
Dithiothreitol (DTT)	Sigma-Aldrich
Deoxynucleotide triphosphates (dNTPs)	Sigma-Aldrich
Glycerol	Penta
Glycine	Sigma-Aldrich
Tris(hydroxymethyl)aminomethane (Tris)	Sigma-Aldrich
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	Thermo Fisher Scientific
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich
Lithium chloride	Carl Roth
Magnesium chloride	NEB
Midory green	Elisabeth Pharmacon
Octylthioglucoside (OTG)	Gold Biotechnology
Polyethylenglycol (PEG) 3350	Sigma-Aldrich
Polyethylenglycol (PEG) 4000	Sigma-Aldrich
Phenylmethylsulfonyl fluoride (PMSF)	Sigma-Aldrich
Sodium dodecyl sulphate (SDS)	Serva
TAE buffer	VWR prolabo chemicals
Tetramethylethylenediamine (TEMED)	Sigma-Aldrich
Tris(2-carboxyethyl)phosphine (TCEP)	Gold Biotechnology

Other used chemicals were in p. a. quality.

2.2 Equipments

ÄKTA Pure FPLC system	GE Healthcare Life Sciences
ÄKTA Prime FPLC system	GE Healthcare Life Sciences
Autoclave EL D-line 2840	TUTTNAUER
Centrifuge 4-16KS	SIGMA
Centrifuge 5424 R	Eppendorf
Centrifuge Z 32HK	HERMLE
Flowbox Bio II Advance	Tellstar
Heater DKT 200	MiuLab
Lamp CLS 150LED	Leica
Microscope S8AF	Leica
Microspin 12	Biosan
Power supply Enduro VE10 PAGE	Labnet
3D Rotatory Mixer	MiuLab
Shaker Innova 44R	New Brunswick Scientific
Sonicator Q700	Qsonica
Spectrometer NanoDrop 2000c	Thermo Fisher Scientific
Thermo Cycler typeE	GeneQ
Transilluminator	UVITEC Cambridge

3 Methods

3.1 Cloning

3.1.1 Polymerase chain reaction (PCR)

PCR was used to amplify sequence of DNA, which encodes gene of a studied protein. Specific pairs of primers were designed to enable amplification of target DNA sequences and subsequent insertion of PCR products into expression vectors. High fidelity DNA polymerase Phusion was used during PCR for its high accuracy, speed and robustness. PCR was carried out with reagents in Table 2, in total volume of 25 µl.

ddH2O		17 µl
HiFi Phusion	5×	5 µl
template DNA	30 ng/µl	1 µl
dNTP's	10 mM	0.5 µl
MgCl2	50 mM	0.5 µl
Phusion Polymerase	100 U/µl	0.5 µl
Forward primer	100 µM	0.25 µl
Reverse primer	100 µM	0.25 µl

Table 2. **Reagents for PCR reaction.**

Thermocycler program for PCR is shown in Table 3.

Initial denaturation	1 minute	98°C
Segment	30×	
Denaturation	10 seconds	98°C
Annealing	15 seconds	60°C
Extension	40 seconds	72°C
End of the segment		
Final extension	5 minutes	72°C
Final hold		4°C

Table 3. **Thermocycler program for PCR.**

3.1.2 Plasmid and DNA isolation

Isolation of DNA from bacterial cells was performed with High-Speed Plasmid MiniKit (Geneaid). Cleaning of PCR products was performed with Gel/PCR DNA Fragments Extraction Kit (Geneaid). All procedures were done as recommended by the manufacturer.

3.1.3 Agarose gel electrophoresis

Agarose gel electrophoresis was used to separate DNA molecules according to their length. Method is based on electrophoretic migration of charged molecules, which are separated by different speed of migration through a gel. Agarose gel was prepared by dissolving 0.5 g of agarose in 50 ml of TAE buffer (1×). Midori Green was added (2.5 µl) to intercalate into double stranded DNA molecules and visualize them under UV light. Electrophoresis ran for 25 minutes under constant voltage of 160 V.

3.1.4 Ligase independent cloning (LIC)

Vectors pTriExTi (high copy) and pET22T (low copy) were linearized by restriction endonuclease BseRI. Reaction volume was 20 µl and mixture (Tab. 4) was incubated for 3 hours at 37°C.

vector (pTriExTi/pET22T)	100 ng/µl	17 µl
CutSmart Buffer	10×	2 µl
BseRI	5 U/ml	1 µl

Table 4. **Composition of mixture for vector linearization.**

Linearized vectors and cleaned PCR product were incubated with T4 polymerase. The 3'-5' exonuclease activity of T4 polymerase was used to modify sticky ends of DNA molecules. Reaction mixture for PCR product modification was prepared according to Table 5. Composition of mixture for vector modification, is shown in Table 6.

PCR product		1 µl
NEB2	10×	2 µl
dGTP	100 mM	0.5 µl
DTT	100 mm	1 µl
T4 DNA pol.	3 U/ µl	0.4 µl
MilliQ H_2O		15.1 µl

Table 5. **Composition of reaction mixture for sticky ends modification of PCR product.**

vector		18 µl
NEB2	10×	2.8 µl
dCTP	100 mM	0.7 µl
DTT	100 mM	1.2 µl
T4 DNA pol.	3 U/ µl	0.6 µl
MilliQ H_2O		6.7 µl

Table 6. **Composition of reaction mixture for sticky ends modification of linearized vector.**

Reaction mixtures were incubated in thermocycler under following program:

30 minutes	22°C
20 minutes	72°C
final hold	4°C

Table 7. **Thermocycler program for sticky end modification.**

PCR product was mixed with vector molecule (Tab. 8) and incubated 15 minutes at room temperature. Extended sticky ends annealed together, and new vector molecule was transformed into competent cells, leaving out ligation step. Despite vector and insert molecules were not covalently bonded after incubation, extended sticky ends were annealed together during transformation with enough cohesion,

and were covalently connected after transformation by intracellular machinery of expression cells.

Vector	1 µl
Insert	2 µl
EDTA (25 nM)	1 µl

Table 8. **Composition of transformation mixture.**

3.1.5 Transformation of competent cells

Competent cells *Escherichia coli* TOP10 were first incubated on ice for 15 minutes. Transformation mixture was added, and cells were again incubated for 15 minutes on ice. Cells were heat-shocked at 42°C for 60 seconds and returned on ice for 2 minutes. 400 µl of SOC medium were added and bacterial suspension was incubated at 37°C for 1 hour. Bacterial suspension was gently spread over LBA plates with corresponding antibiotic.

3.1.6 Colony PCR

Colonies grown on LBA plates were screened for the presence of expression construct by using colony PCR. One milliliter of LBA medium was inoculated with single colony and incubated at 37°C for 1 hour. Mixture for colony PCR was prepared according to Table 9. The colony PCR ran under program shown in Table 10.

H_2O		9.1 µl
Kapa 2G Buffer	5×	2.5 µl
forward primer	100 µM	0.25 µl
reverse primer	100 µM	0.25 µl
dNTPs	10 mM	0.25 µl
Kapa 2G Polymerase	5 U/ µl	0.06 µl
cell culture		0.5 µl

Table 9. **Composition of reaction mixture for colony PCR.**

Initial denaturation	4 min	95°C
Segment	30×	
Denaturation	15 s	95°C
Annealing	15 s	55°C
Extension	15 s	72°C
End of segment		
Final extension	4 min	72°C
Final hold		4°C

Table 10. **Thermocycler program for colony PCR.**

Colony PCR products were examined by agarose gel electrophoresis to verify presence of transformed expression vector. Five milliliters of LBA medium were inoculated with positive colony containing high copy vector plasmid and 20 ml LBA medium were inoculated with positive colony containing low copy vector plasmid. Cells were incubated 12 hours at 37°C, 220 rpm. Cells were harvested, and plasmid DNA isolated and sequenced.

3.2 Expression tests

Vector plasmid was transformed into a suitable strain of expression cells. Different expression conditions were tested to obtain high yield of soluble protein. Optimized expression parameters were temperature, length of expression and type of induction. Range of tested expression temperatures was between 10°C - 37°C. Range of tested expression times was between 6 - 48 hours. Expression in cells was induced by adding IPTG in LB medium or by lactose in auto-induction medium. IPTG binds and releases the repressor of *lac* operon, what causes triggering of transcription of genes under this operon. It is not metabolized by the cell, therefore its concentration during experiment remains the same [41]. In case of lactose, after bacteria run out of preferred carbon sources, present lactose is enzymatically turned into allolactose. Allolactose induces expression by binding to the same repressor as IPTG [42]. Yields of expressed proteins were analyzed by polyacrylamide gel electrophoresis with sodium dodecyl sulphate (SDS-PAGE).

3.2.1 Induction with IPTG

Five milliliters of LB medium with antibiotic were inoculated by expression cells stored in glycerol stock and incubated 12 hours at 37°C, 250 rpm. Twenty microliters of grown bacterial culture was used to inoculate 21 ml of LB medium with antibiotic. Expression cells were further incubated at 37°C, until optical density of suspension was 0.4 – 0.6. Bacterial culture was then incubated 15 minutes at chosen expression temperature. One milliliter of uninduced culture was taken as control sample. Bacterial culture was induced by IPTG with concentration between 0.05 mmol - 1 mmol and further incubated. Four 1 ml samples were taken continuously during expression. All taken samples were centrifuged at 20 000 × g for 1 minute and pellet kept at -20°C.

3.2.2 Induction by lactose

Solutions and media prepared for lactose induction are shown in Table 11.

Z4	1 l	950 ml H_2O; 5 g yeast extract; 10 g tryptone
P20×	100 ml	90 ml H_2O; 17,8 g Na_2HPO_4; 11,3 g KH_2PO_4; 6,6 g $(NH_4)_2SO_4$
505	100 ml	50 g glycerol; 57 ml H_2O; 15 g glucose
505 2	100 ml	25 g glycerol; 73 ml H_2O; 2.5 g glucose; 10 g lactose monohydrate
Non-induction medium (NI)	100 ml	93 ml Z4; 5 ml P20×; 6 ml 505; 100 µl 2M $MgCl_2$; 100 µl ampicillin
Auto-induction medium (AI)	100 ml	93 ml Z4; 5 ml P20×; 2 ml 5052; 100 µl 2M $MgCl_2$; 100 µl ampicillin

Table 11. **Composition of solutions and media for lactose induction.**

Five milliliters of LB medium with antibiotic were inoculated by expression cells and incubated 12 hours at 37°C, 250 rpm. Hundred microliters of grown culture were used to inoculate 20 ml of NI and AI media, which were further incubated at either 37°C, 30°C or 24°C for 24 hours at 220 rpm. After 24 hours of incubation, OD_{600} was measured in 30 minutes intervals until it stopped increasing. Samples of

1 ml were taken from bacterial cultures in NI and AI media. Samples were centrifuged at 20 000 × g for 1 minute and pellet kept at -20°C.

Samples, which were collected during expression, were prepared for SDS-PAGE by resuspending bacterial pellets in 50 µl - 170 µl of lysis buffer. Composition of lysis buffer was the following:

Lysis buffer	10% (v/v) Solution 1; 90% (v/v) Solution 2
Solution 1	0.5 g 5% OTG; 100 mg lysozyme; 0.5 ml 1M Tris; 200 µl 5M NaCl; 2 ml 20% glycerol; 10 µl turbonuclease; H_2O (to 10 ml)
Solution 2	50 mM Tris (pH = 8); 150 mM NaCl

Table 12. **Composition of lysis buffer.**

Resuspended pellets were incubated on rotatory mixer at room temperature for 20 minutes. Ten microliters were taken out of each incubated sample as total fraction and mixed with 2.5 µl of protein loading buffer (PLB) (5×). PLB was prepared according to Table 13. All incubated samples were centrifuged at 20 000 × g for 1 minute. Supernatants represent soluble fractions and 10 µl of each supernatant was mixed with 2,5 µl of PLB. Samples mixed with PLB were incubated at 98°C for 3 minutes and loaded on polyacrylamide gel.

PLB (5×)	
1M Tris ph 6,8	1.25 ml
10% SDS	4 ml
glycerol	2.0 ml
bromphenol blue	2.0 mg
DTT	0.31 g
ddH_2O	to final volume 10 ml

Table 13. **Composition of PLB.**

3.2.3 SDS-PAGE

In this method, proteins are denatured by heat and covered by SDS, thus giving protein molecules negative surface charge, which is proportional to their molecular mass. Proteins are loaded on polyacrylamide gel and separated according to their mass.

Samples were loaded into wells of stacking gel in volume of 10 µl and separated under constant voltage of 150 V for 75 minutes in the presence of running buffer. Composition of used gels and buffers is shown in Table 14.

Stacking gel	
H_2O	0.68 v/v
30% acrylamide mix	0.17 v/v
1.0 M Tris pH=6.8	0.13 v/v
10% SDS	0.01 v/v
10% APS	0.01 v/v
TEMED	0.001 w/v
Separation gel	
H_2O	0.33 v/v
30% acrylamide mix	0.40 v/v
1.0 M Tris pH=6.8	0.25 v/v
10% SDS	0.01 v/v
10% APS	0.01 v/v
TEMED	0.0004 w/v
Running buffer (5×)	
Tris	30.28 g
glycine	144.13 g
SDS	10 g
ddH_2O	filled to 2 liters

Table 14. **Composition of gels and buffers used for SDS-PAGE.**

Resulting gels were stained with solutions containing Coomassie G250. Protein bands, which mostly corresponded to the mass of the expressed protein, were analyzed by MALDI-MS to verify the presence of expressed protein in the gel.

3.3 Protein purification

After high-yield expression conditions were found, protein was expressed in large volume of growth medium. Twenty milliliters of LB medium with antibiotic was inoculated by expression cells and incubated at 37°C, 220 rpm for 12 hours. One liter of LB medium with antibiotic was inoculated by 10 ml of prepared bacterial suspension, and incubated at 37°C, 220 rpm until OD_{600} = 0.4 – 0.6. Expression cells were incubated 20 minutes at chosen expression temperature. Protein expression was induced by addition of IPTG. Cells were further incubated at expression temperature, 220-250 rpm. Cell were after expression centrifuged at 5 000 × g for 15 minutes at 4°C, and pellet resuspended with resuspension buffer (RB). Cells were sonicated and 10 µl of sonicate taken as total fraction for SDS-PAGE. Sonicate was centrifuged at 20 000 × g for 30 minutes. Supernatant was collected, kept on ice and 10 µl taken as soluble fraction for SDS-PAGE.

Protein was first purified by metal ion affinity chromatography (IMAC) [43]. Supernatant was loaded on HisTrap HP 5 ml column. Column contains immobilized chelating groups charged with Ni^{2+} ions, which create complexes with histidine-tagged proteins and enable preparative purification. Protein was eluted from column by buffer with high concentration of imidazole. Eluted fractions were collected and analyzed with SDS-PAGE.

Proteins containing His-tag followed by Smt3 domain were cleaved by Ulp1-ac protease and incubated at 4°C with 5 mM TCEP (reducing agent) and 1 mM PMSF (serine protease inhibitor). To remove imidazole from protein sample, buffer was exchanged by size-exclusion chromatography, using HiPrep 26/10 desalting column.Fractions containing protein were collected and loaded on HisTrap HP 5 ml column.

Flow-through fractions with protein were collected and loaded on HiLoad Superdex 75 pg column to purify protein with high-resolution preparative size exclusion chromatography [44].

3.4 Crystallization of purified proteins

Purified proteins were screened for suitable crystallization conditions at Biomolecular Interactions and Crystallization Core Facility at CEITEC MU. Sitting drop vapor diffusion method was used for the screening. List of used crystallization screens is shown in Table 15.

Manufacturer	Screen
Qiagen - Nextal	Classics Suite
	Classics II Suite
	Classics Lite Suite
	PACT Suite
	AmSO4 Suite
	ComPAS Suite
	PEGs Suite
	PEGs II Suite
	pH Clear Suite
	pH Clear II Suite
Molecular Dimensions	Clear Strategy I
	Clear Strategy II
	JCSG-plus
	SG1
	Structure I + II
Hampton Research	Additive Screen

Table 15. **List of crystallization screens.**

Protein solution was mixed with screen conditions in ratios 1:2; 1:1 and 2:1, in total volume of 0.3 µl. Sitting drop method [45] was used for screening with solutions incubated at 4°C or 20°C. Sitting drops were automatically photo-documented for four months under visible and UV light. Conditions, in which protein formed crystals, were further optimized. Hanging drop method was used and 2 µl droplets were prepared manually and monitored under light microscope. Crystals of suitable size were tested for diffraction at Biological Small Angle X-ray Scattering Core Facility at CEITEC MU to verify, that salts present in crystallization

conditions do not form observed crystal. Phasing kits for preparation of heavy atom derivatives of native protein crystals are shown in Table 16.

Heavy atom kit	Compound	Manufacturer
Heavy atom screen M1	13	Hampton research
Heavy atom screen M2	2	
Heavy atom screen Hg	9	
Heavy atom screen PT	11	
Lanthanide phasing kit	1 8 9	Jena Bioscience
JBS Magic triangle	1	

Table 16. **List of phasing kits used to prepare heavy atom derivatives of native protein crystals.**

Hanging drop method [45] was used to prepare heavy atom derivatives of native crystal by cocrystallization. Suitable crystallization conditions were modified by adding heavy atom kit compounds, in concentration range from 1 mM to 50 mM. Another method to incorporate heavy atoms into native crystal was by soaking [36] already grown native crystals with kit compounds in concentrations recommended by manufacturer. Compound was added directly to hanging drop with native crystal, or native crystal was fished out and incubated in solution consisting of mother liquor and kit compound. Soaking time varied from several seconds to tenths of minutes. Native and derivatised crystals were fished out and immediately stored in liquid nitrogen.

3.5 Solving of protein structure

Measuring of diffraction data, indexing, integration, scaling and phasing were performed at Synchrotron SOLEIL, France. Structure was solved by using SHELXD program [46].

3.6 Model building and refinement

Application *Coot* was used for model building [47]. Maximum likelihood refinement was done with program REFMAC [48]. Validation of model was performed by web service MolProbity [49].

4 Results

4.1 BmpA

Gene of protein BmpA (gp118) is 705 bp long. The gene was successfully amplified by PCR, which was confirmed by agarose gel electrophoresis (Fig. 11). Amplified sequence was 734 bp long. Primers were designed as shown in Table 17. Both forward and reverse primer had 5'- ends designed to be suitable for LIC. Vector plasmids were pTriExTi (high copy) and pET22T (low copy). New expression vectors BmpA_pTriExTi and BmpA_pET22T were successfully transformed into competent cells, which was confirmed by colony PCR (Fig. 12 and 13) and sequencing of the inserts. BmpA_pTriExTi contains gene of BmpA with N-terminal His-tag followed by S-tag and Smt3 domain. BmpA_pET22T contains gene of BmpA with N-terminal His-tag followed by Smt3 domain.

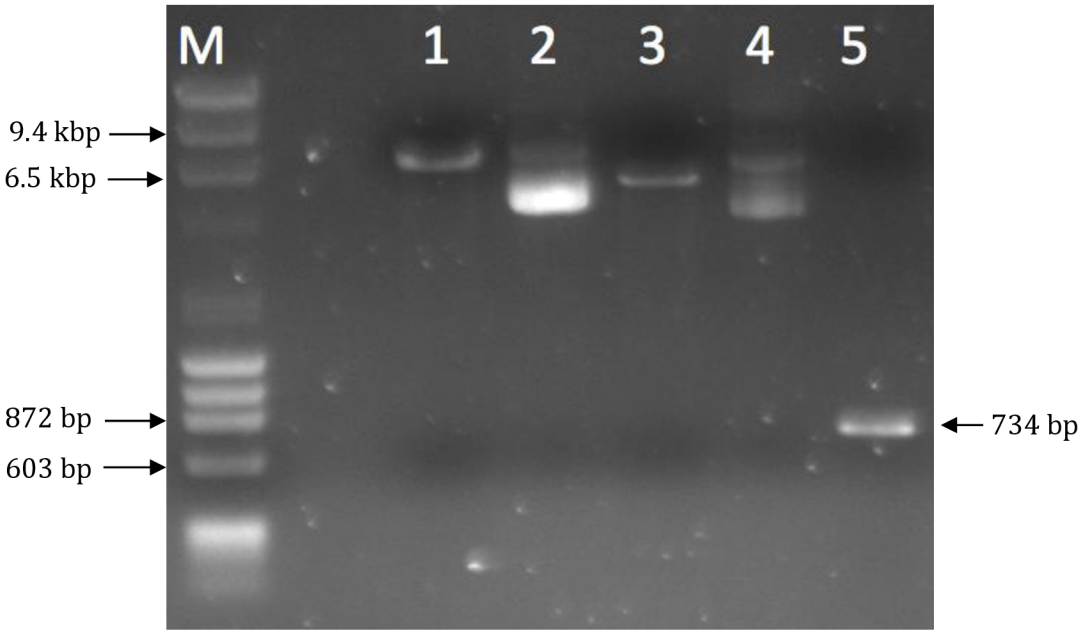

Fig. 11. **Agarose gel electrophoresis of PCR product and linearized vectors.** M: DNA marker; 1: linearized pTriExTi (6.9 kbp); 2: uncleaved pTriExTi; 3: linearized pET22T (5.8 kbp); 4: uncleaved pET22T; 5: PCR product.

Primer	Sequence (5'-3')
Forward	AGATTGGTGGCATGAGATTTAAGAAGCACGTAGTTC AACATGAAGAAACG
Reverse	GAGGAGAGTTAAGACTTACTATTCAAATAAAGCAAA AATTCCTTCCTCATCTTGTCC

Table 17. **Sequences of primers for LIC of BmpA.**

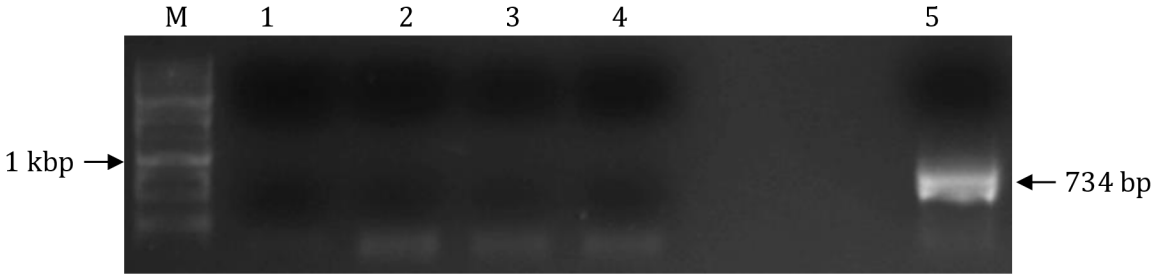

Fig. 12. **Agarose gel electrophoresis of colony PCR products for vector BmpA_pTriExTi.** M: 1kbp DNA ladder; 1 – 4: bacterial colonies without PCR product in vector plasmid; 5 - bacterial colony containing BmpA_pTriExTi.

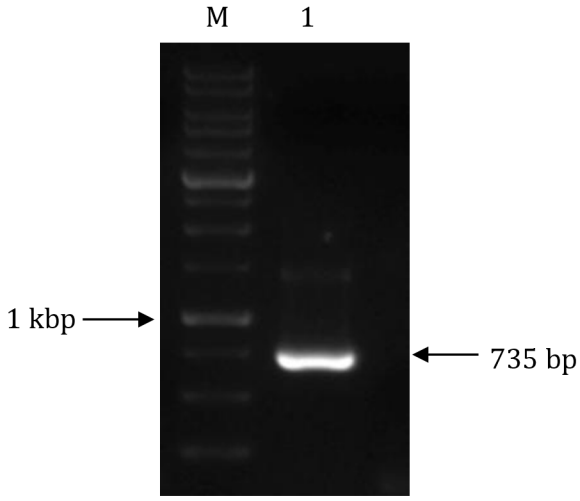

Fig. 13. **Agarose gel electrophoresis of colony PCR product for vector BmpA_pET22T.** M: 1kbp DNA ladder; 1: bacterial colony containing transformed vector plasmid.

Isolated expression vectors BmpA_pTriExTi and BmpA_pET22T were transformed into expression cells *E. coli* ArcticExpress, *E. coli* Tuner and *E. coli* BL21 (DE3). List of expression tests, performed to find expression conditions for high yield of soluble BmpA, are shown in Table 18. The highest yield was produced under conditions in Table 19. SDS-PAGE gel of this conditions is shown in Fig. 14.

BmpA_pTriExTi		**BmpA_pET22T**	
***E. coli* BL21 (DE3)**			
18°C	0.05/0.1/0.3 mM IPTG	18°C	0.1/0.5/0.7/1 mM IPTG
24°C	IM	24°C	IM, 0.7 mM IPTG
30°C	IM	30°C	IM, 0.1/0.5/1 mM IPTG
37°C	IM, 0.1/0.3 mM IPTG	37°C	IM, 0.1/1 mM IPTG
***E. coli* Tuner**			
18°C	0.05/0.1/0.3 mM IPTG	18°C	0.05/0.1/0.3 mM IPTG
37°C	0.1/0.3 mM IPTG	37°C	0.1 mM IPTG
***E. coli* ArcticExpress**			
10°C	0.1/0.5/1 mM IPTG	10°C	1 mM IPTG

Table 18. **List of tested expression conditions for BmpA_pTriExTi and BmpA_pET22T.**

Vector	Cells	Temperature	Induction	Time
BmpA_pTriExTi	*E. coli* Tuner	18°C	0.1 mM IPTG	17 hours

Table 19. **Expression conditions with highest yield of soluble BmpA.**

For successful purification of BmpA, composition of used buffers had to be optimized (Table 20). Expression cells were harvested from 2 liters of LB medium.

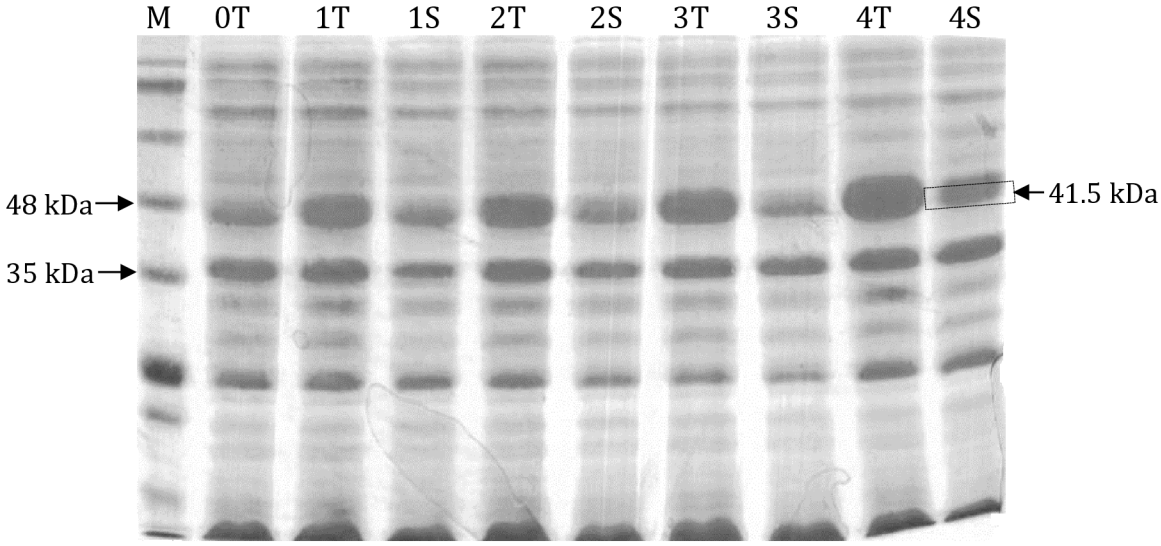

Fig. 14. **SDS-PAGE gel of BmpA expression test with highest yield of soluble protein.** M: marker; T: total fraction; S: soluble fraction; 0: before induction; 1: two hours after induction; 2: five hours after induction; 3: ten hours after induction; 4: seventeen hours after induction. Bracket shows bend of soluble BmpA.

Resuspension buffer	50 mM HEPES (pH=8); 400 mM NaCl; 1% glycerol; 5 mM β-mercaptoethanol
HisTrap loading buffer	50 mM HEPES (pH=8); 400 mM NaCl; 1% glycerol; 20 mM imidazole
Histrap elution buffer	50 mM HEPES (pH=8); 400 mM NaCl; 1% glycerol; 400 mM imidazole
Negative HisTrap elution buffer	50 mM HEPES (pH=8); 400 mM NaCl; 1% glycerol; 5 mM imidazole
Superdex loading buffer	50 mM HEPES (pH=8); 400 mM NaCl; 1% glycerol

Table 20. **List of buffers used for purification of BmpA.**

BmpA was first purified with IMAC on HisTrap column (Fig. 15 and 16). N-terminal His-tag and Smt3 domain were cleaved by Ulp1-ac protease (Fig. 17) and imidazole was removed from sample by buffer exchange. Negative HisTrap yielded very low amount of BmpA in flow-through, since BmpA bound strongly to the column even without His-tag.

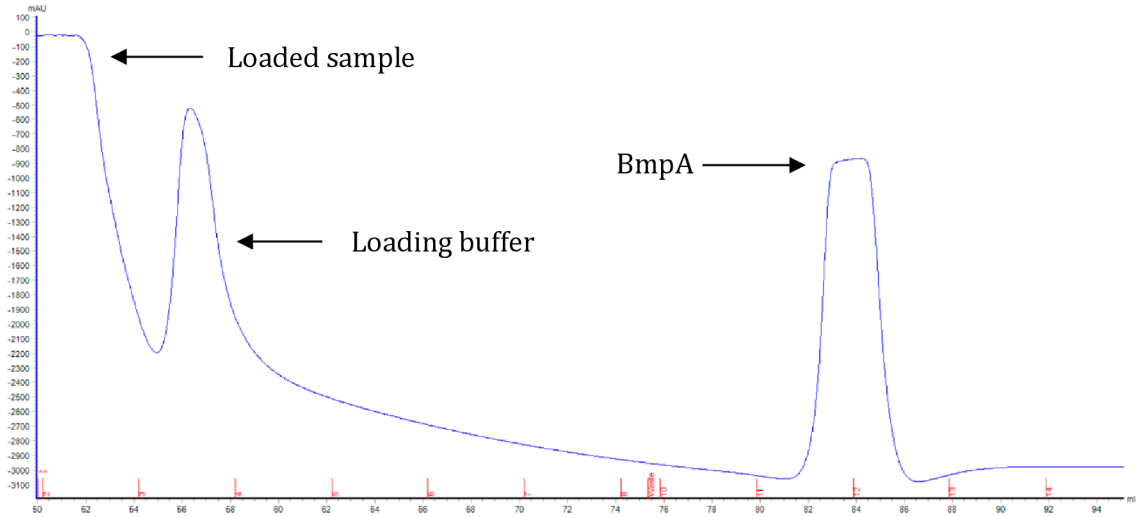

Fig. 15. **Chromatogram of purification of BmpA with HisTrap column.**

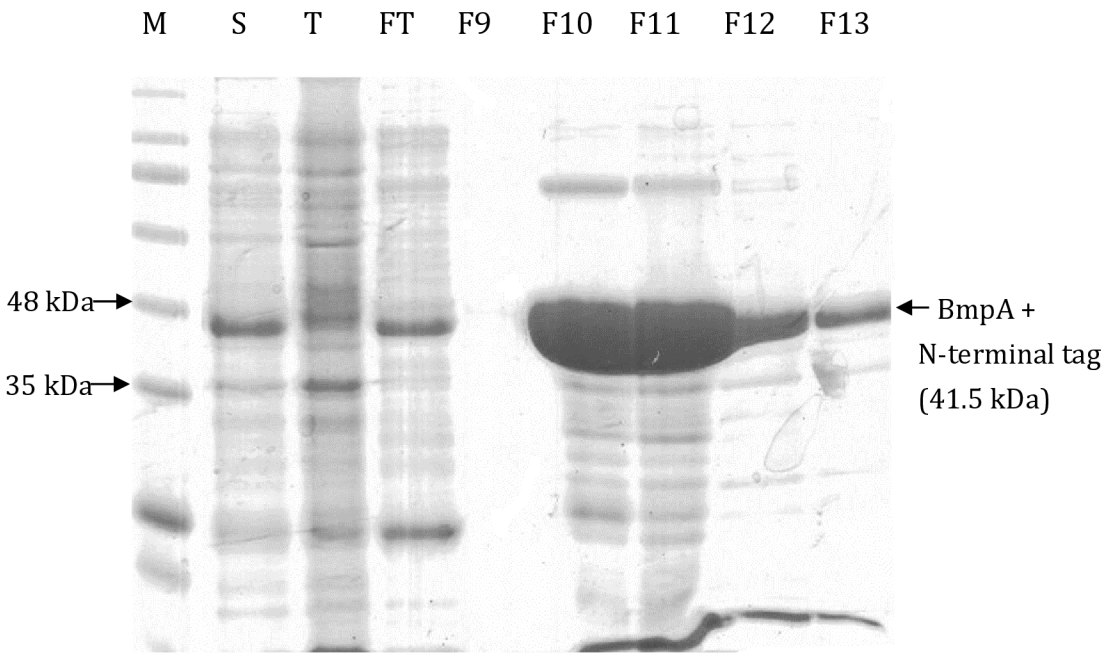

Fig. 16. **SDS-PAGE gel of BmpA after purification with HisTrap column.** M: marker; S: soluble fraction; T: total fraction; FT: flow through; F9-F13: elution fractions.

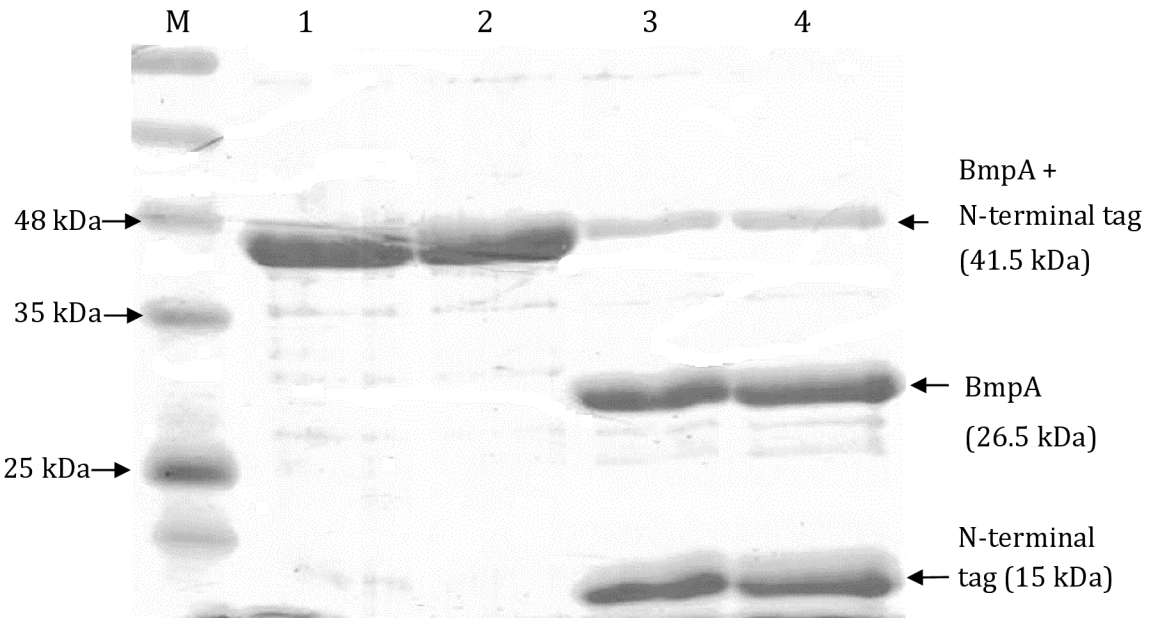

Fig. 17. **SDS-PAGE gel of BmpA after cleavage with Ulp-1ac protease.** M: marker; 1: control sample without protease stored at -20°C; 2: control sample without protease stored at 4°C; 3: total fraction at 4°C; 4: soluble fraction at 4°C.

Elution of BmpA from HisTrap column was performed with low concentration of imidazole, so that peptide with His-tag remains bound to the column (Fig. 18).

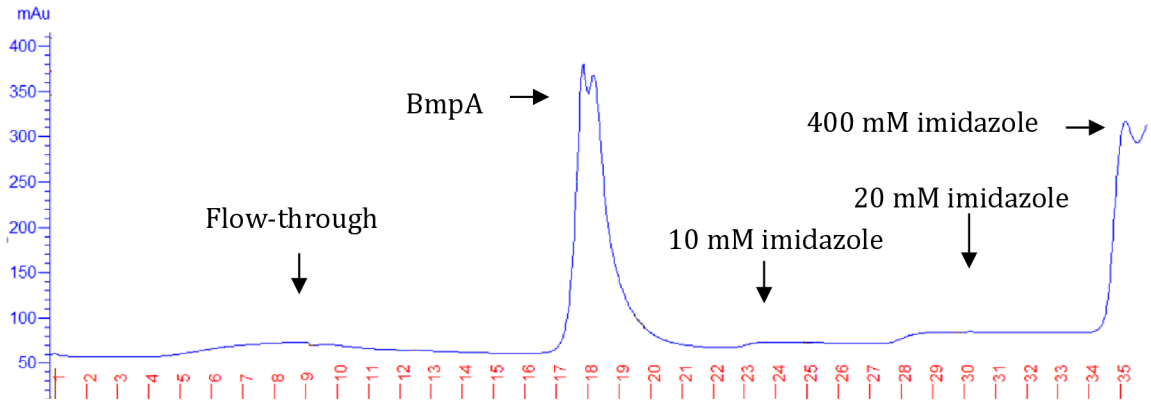

Fig. 18. **Negative HisTrap and elution of BmpA.** Fractions in flow-through (1-15) were eluted without imidazole. Fractions with BmpA (16-21) were eluted with 5 mM imidazole. Fractions 22-26 were eluted with 10 mM imidazole. Fractions 27-33 were eluted with 20 mM imidazole.

Eluted sample containing BmpA was further purified by size exclusion chromatography (Fig. 19 and 20). Eluted samples containing BmpA were concentrated to final concentration 10.9 mg/ml. Thirteen crystallization screens were set up. No crystals emerged after fourth month of monitoring the sitting drops.

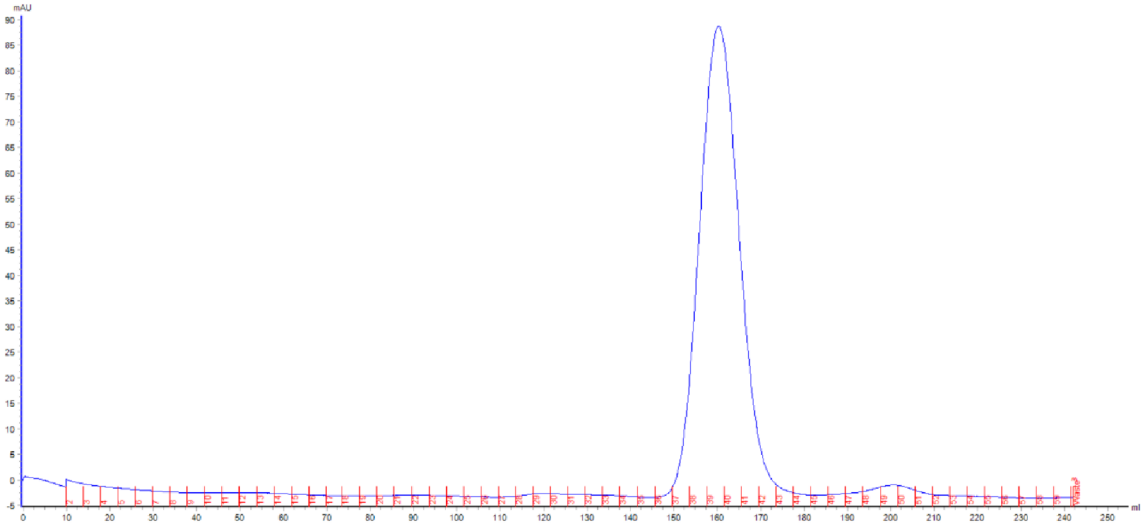

Fig. 19. **Chromatogram of size exclusion chromatography of BmpA on Superdex 75 column.** Retention time of eluted peak suggests, that BmpA was in solution in monomeric state.

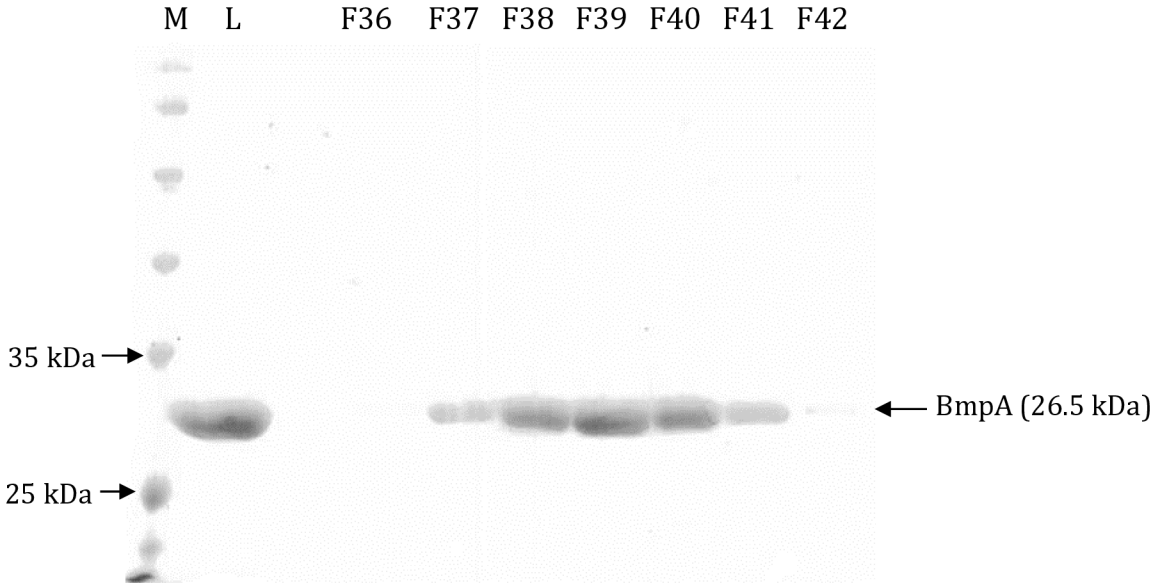

Fig. 20. **SDS-PAGE gel of BmpA after size-exclusion chromatography on Superdex 75 column.** M: marker; L: loaded sample; F36-F42: eluted fractions.

4.2 BmpB

I was provided with expression vector BmpB_pET-28a containing N-terminal His-tag followed by Smt3 domain and gene encoding BmpB, and with optimized purification protocol, both courtesy of Y. Levdansky. Plasmid was sequenced and transformed in expression cells *E. coli* BL21 (DE3). After induction with IPTG (0.4 mM), expression cells were incubated in LB medium with 1% glucose and 1 mM $MgCl_2$ for 16 hours at 18°C. Composition of optimized buffers was:

Resuspension buffer	50 mM HEPES (pH=8); 200 mM NaCl; 5 mM β-mercaptoethanol
HisTrap loading buffer	50 mM HEPES (pH=8); 200 mM NaCl; 40 mM imidazole
Histrap elution buffer	50 mM HEPES (pH=8); 200 mM NaCl; 500 mM imidazole
Superdex loading buffer	50 mM HEPES (pH=8); 200 mM NaCl

Table 21. **List of buffers used for purification of BmpB.**

Protein was purified with IMAC by HisTrap column (Fig. 21 and 22). Eluted fractions were collected and incubated with Ulp1-ac protease to remove N-terminal His-tag with Smt3 domain from BmpB protein (Fig. 23). Control samples showed, that the expressed protein is cleaved also without previously added Ulp1-ac protease. Results of MALDI-MS analysis suggested that protein was cleaved at two different sites, which were not exactly determined by MALDI-MS.

After buffer exchange and negative HisTrap, the protein sample was loaded onto Superdex column (Fig 24 and 25). Purified BmpB was concentrated to 9.4 mg/ml and analyzed by DLS (Fig. 26). Twenty crystallization screens were set up. No crystals were formed within four months.

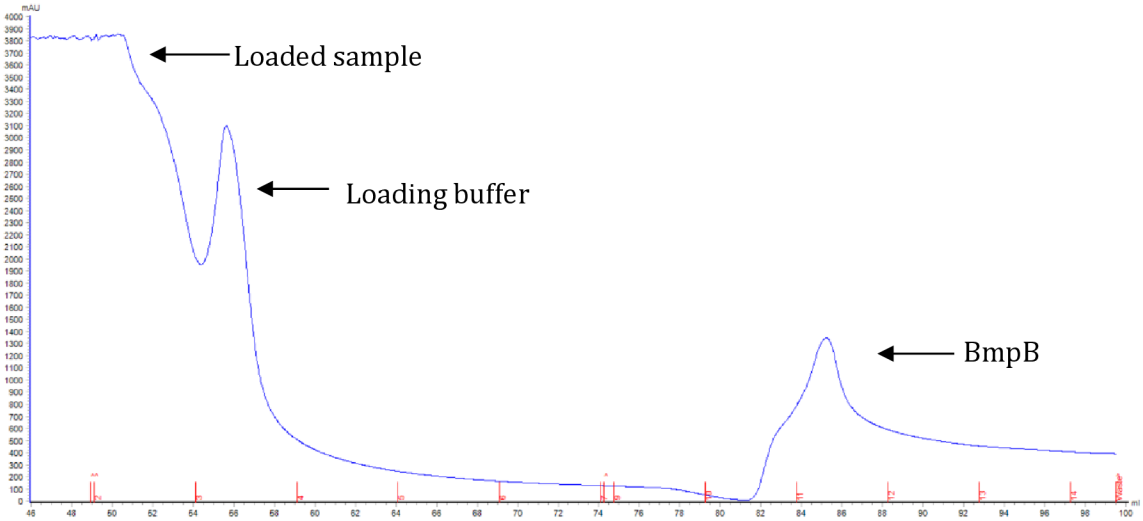

Fig. 21. **Chromatogram of IMAC purification of BmpB on HisTrap column.**

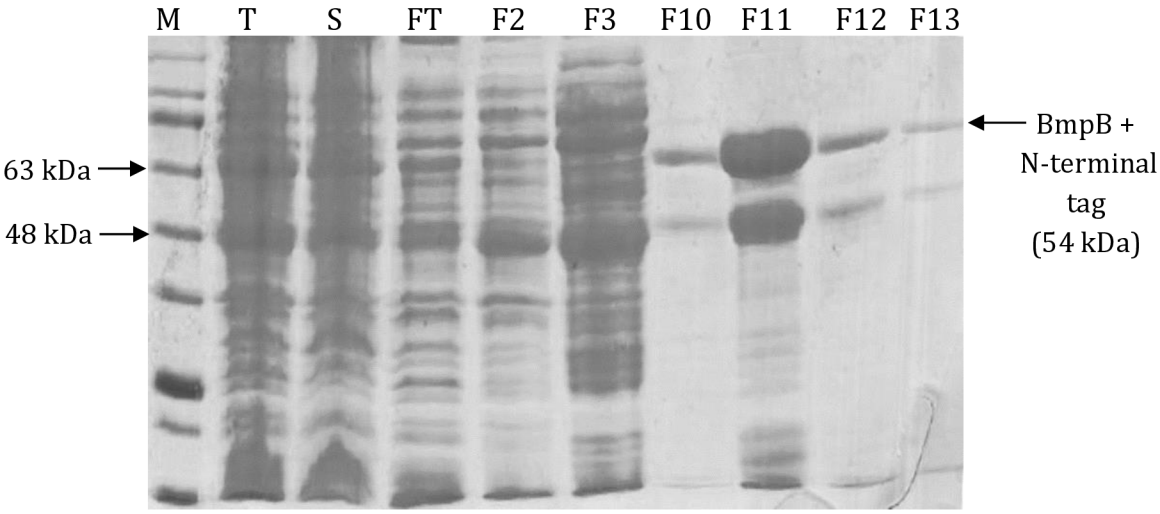

Fig. 22. **SDS-PAGE gel of purification of BmpB on HisTrap column.** M: marker; T: total fraction; S: soluble fraction; FT: flow through; F2-F13: eluted fractions.

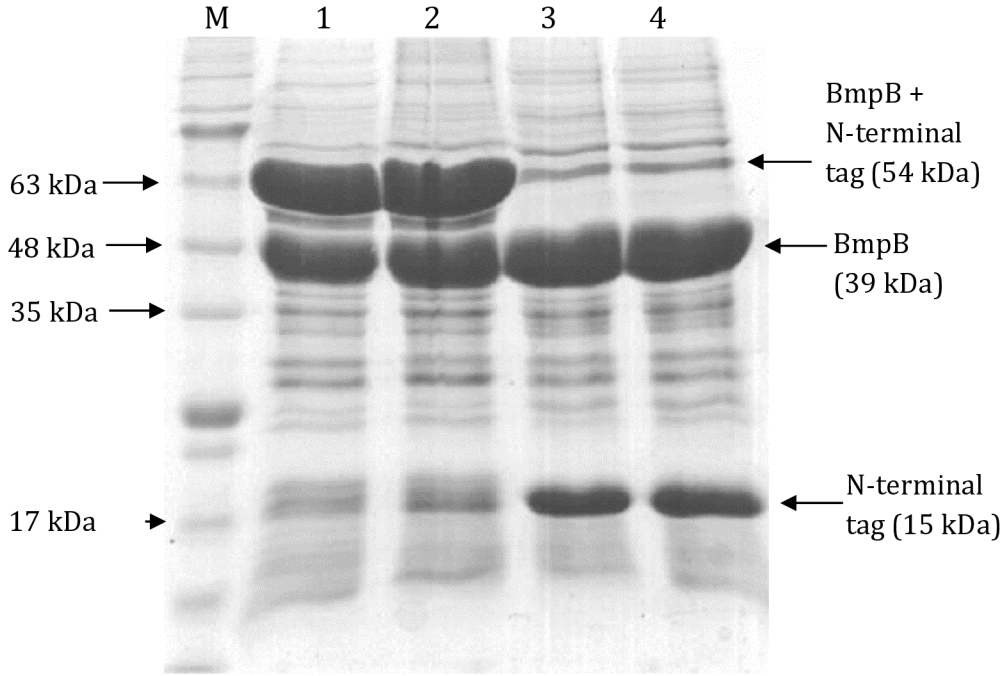

Fig. 23. **SDS-PAGE gel of BmpB after cleavage with Ulp1-ac protease.** M: marker; 1: control sample without protease, stored at -20°C; 2: control sample without protease stored at 20°C; 3: total fraction at 20°C; 4: soluble fraction at 20°C. Control samples indicated cleavage of approximately 15 kDa N-terminal part even without added protease.

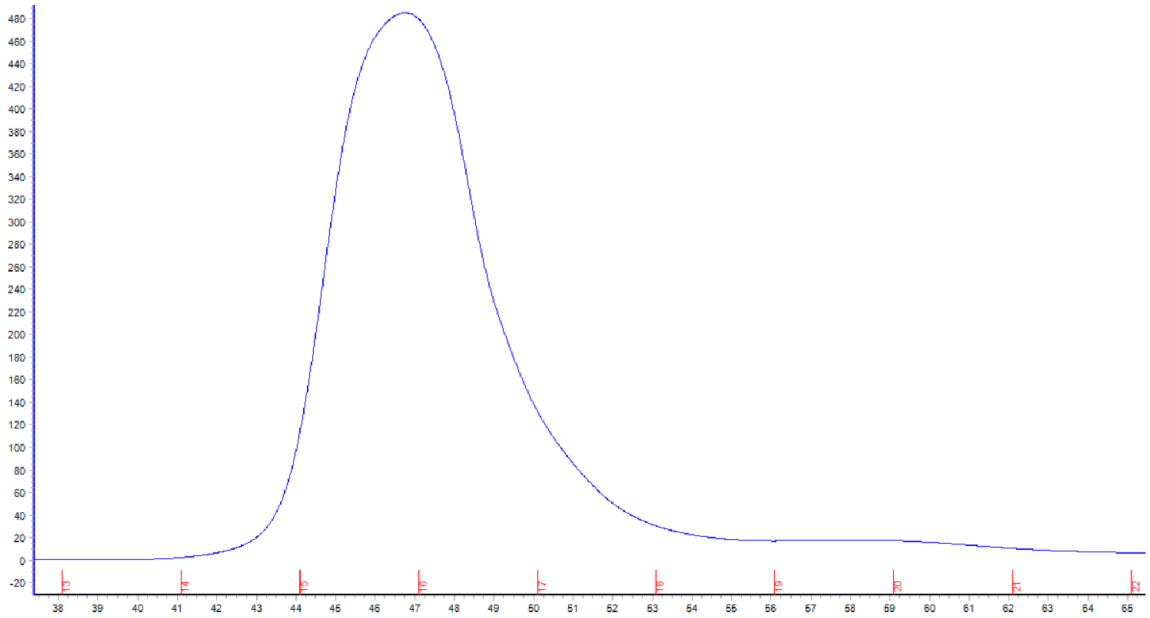

Fig. 24. **Chromatogram of size exclusion chromatography of BmpB on Superdex 75 column.** Peak retention time suggests, that BmpB was in solution as dimer.

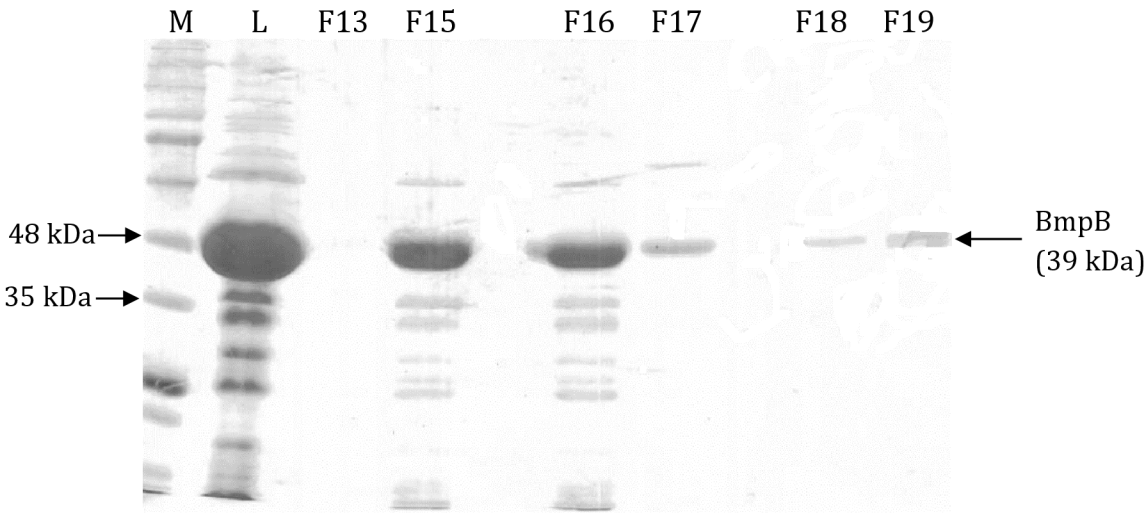

Fig. 25. **SDS-PAGE gel of BmpB after size exclusion chromatography on Superdex 75 column.** M: marker; L: loaded sample; F13-F19: eluted samples.

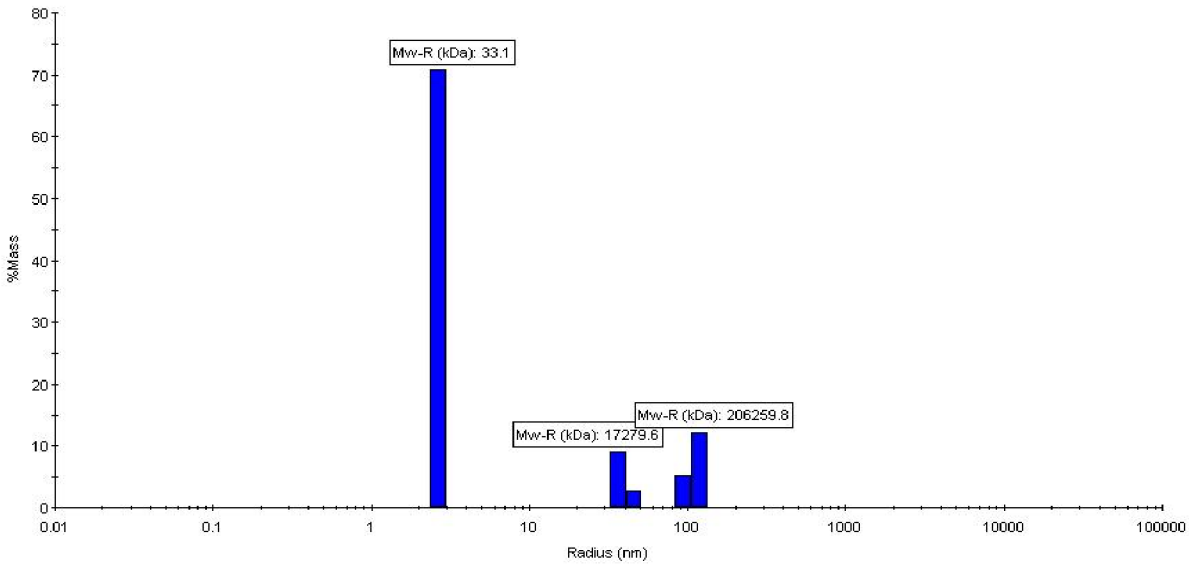

Fig. 26. **Graph of DLS analysis of solution of purified BmpB.** The majority component (33.1 kDa) was determined with relative molecular weight close to BmpB (39 kDa).

New sets of primers were designed to amplify gene of BmpB and to prepare new expression vectors (Table 22 and 23). The primers enabled insertion of BmpB gene into vectors pTriExTi, pET22T and pET-28a. Amplification of gene by PCR was successful for both sets of primers (Fig. 27).

Primer	Sequence (5'-3')
Forward	AGATTGGTGGCATGAAAACTAGAAAATTAACTAACATACTATCAAAATT AATAGATAAGACAATGGCAGG
Reverse	GAGGAGAGTTAAGACTTATTACTTTAGTTCTACTTTTATTTCTCCTGC

Table 22. **Primers for LIC of BmpB.**

Primer	Sequence (5'-3')
Forward	CCATGGTCAAAACTAGAAAATTAACTAACATACTATCAAAATTAATAGA TAAGAC
Reverse	CTCGAGTCTTTAGTTCTACTTTTATTTCTCCTGCTC

Table 23. **Primers for cloning of BmpB into pET-28a.**

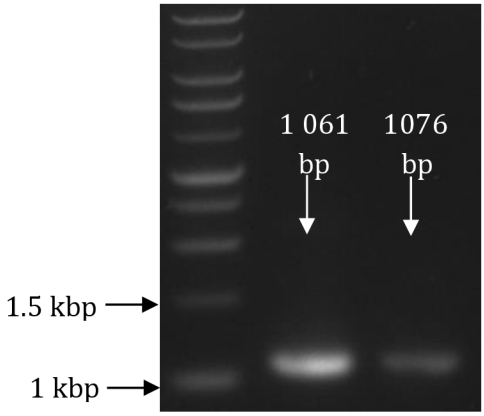

Fig. 27. **PCR products.** Gene of BmpB amplified with primer from Table 22 (1076 bp) and table 23 (1061 bp).

4.3 BmpC

I was provided with high copy vector plasmid pET-28b, containing gene BmpC (gp 121) with C-terminal His-Tag, in bacterial stock *E. coli* BL21 (DE3), together with purification protocol (courtesy of Y. Levdansky). Isolated plasmid was sequenced, and protein purified. Expression cells were induced by IPTG (0,4 mM) and incubated in LB medium with 1 mM $MgCl_2$ and 1% glucose for 17 hours at 18°C. Protein was first purified by IMAC on HisTrap column (Fig. 28 and 29).

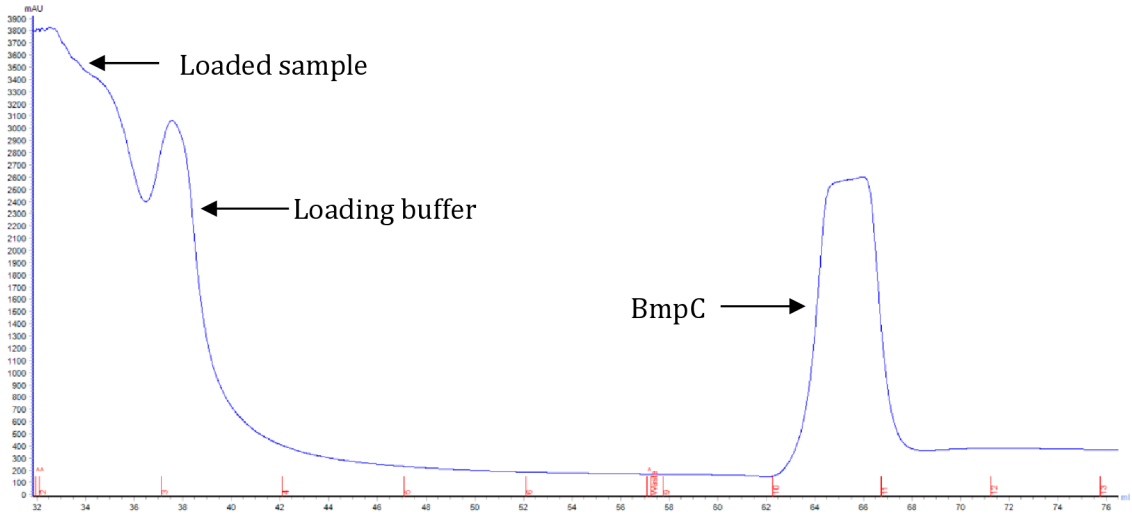

Fig. 28. **Chromatogram of BmpC purification with HisTrap column.**

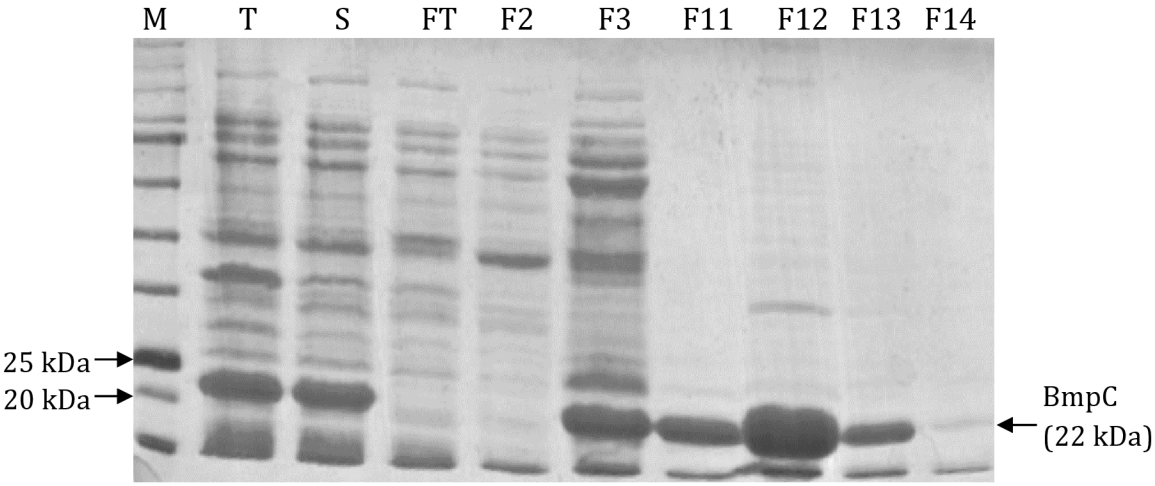

Fig. 29. **SDS-PAGE gel of BmpC after IMAC chromatography on HisTrap column.** M: marker; T: total fraction; S: soluble fraction; FT: flow through; F2-F14: eluted fractions; A: bend corresponding to BmpC with C-terminal His-tag.

Eluted fractions of BmpC were loaded onto Superdex column and purified with size exclusion chromatography (Fig. 30 and 31). Buffers used during purification are in Table 21.

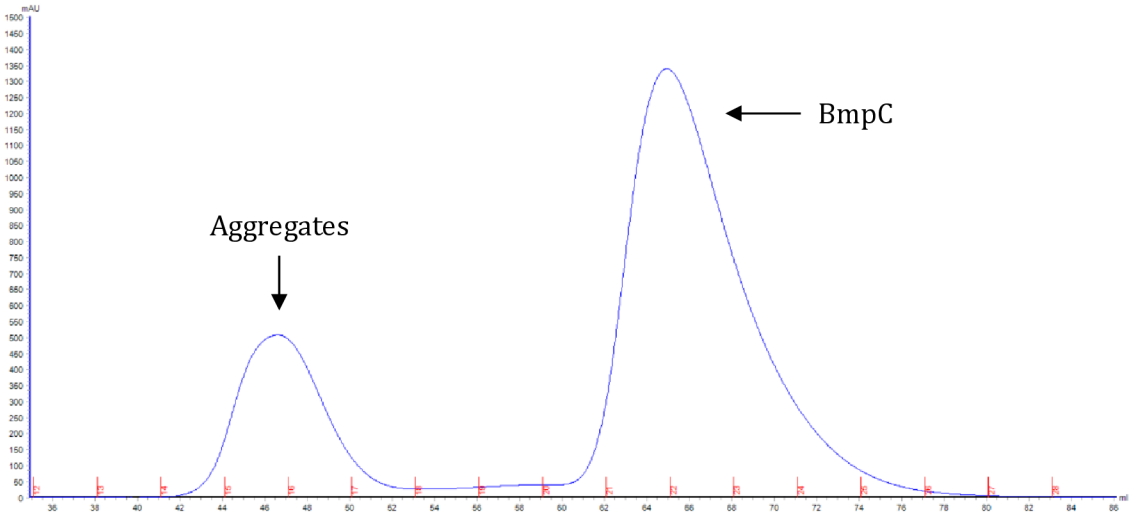

Fig. 30. **Chromatogram of size exclusion chromatography of BmpC on Superdex 75 column.**

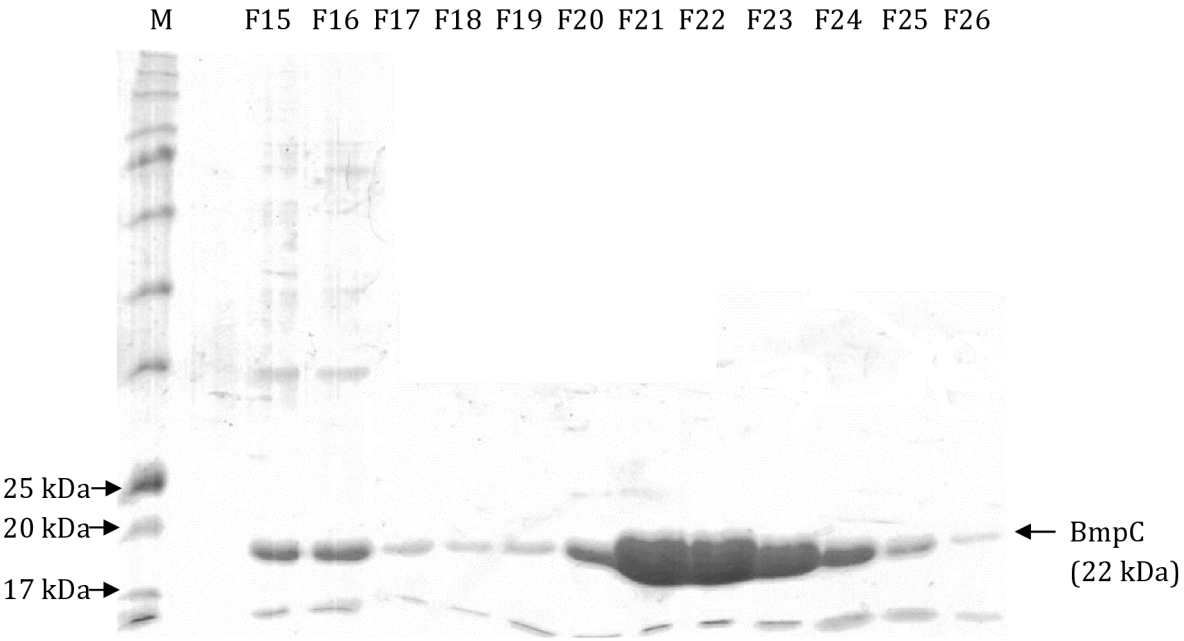

Fig. 31. **SDS-PAGE gel of BmpC after size exclusion chromatography on Superdex 75 column.** M: marker; F15-F26: eluted fractions.

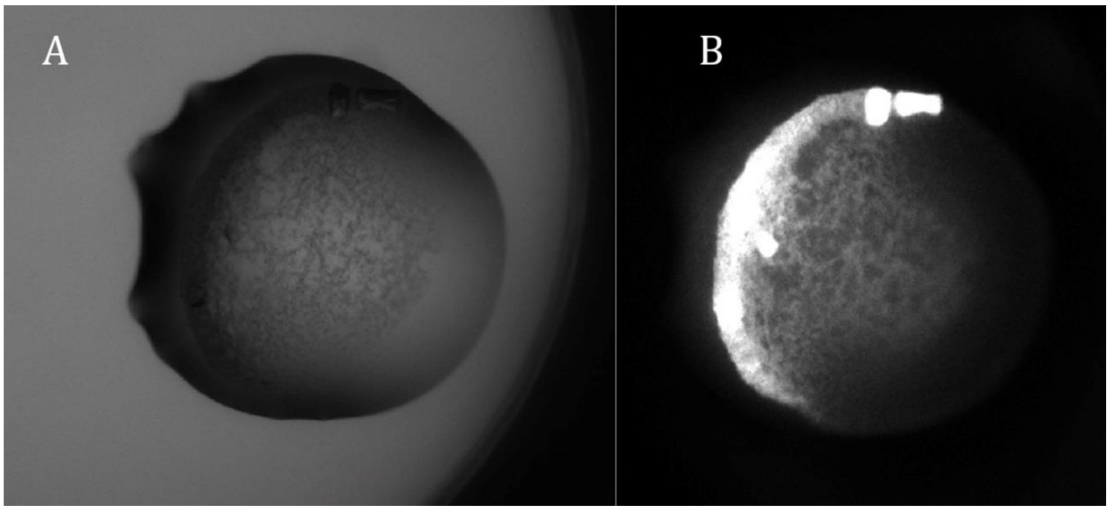

Fig. 32. **Inspection of crystallization screening.** Crystals of BmpC grown in sitting drop. A) crystals under visible light. B) crystals under UV light.

Purified BmpC was concentrated to 14 mg/ml. Twenty crystallization screens were set up. First crystals emerged after 16 days of incubation. Crystals grew under three different screening conditions. Hanging drops for vapor diffusion crystallization were prepared under each condition. To optimize size and shape of the crystals, 48 additional conditions with varying parameters were tested for hanging drop crystallization. Conditions, under which the largest and most compact crystals grew, were further optimized by changing concentration of salt, PEG, and pH of the buffer. Together, 72 new crystallization conditions were tried. Large and compact crystals grew in three conditions (Tab. 24). Three crystallization plates (each 24 wells) were filled with hanging drops under these conditions.

1.6 M LiCl 0.1 M Tris, pH=7.5 20.0 % w/v PEG 4 000	1.6 M LiCl 0.1 M Tris, pH=7.5 22.5 % w/v PEG 4 000	1.6 M LiCl 0.1 M Tris, pH=7.5 25.0 % w/v PEG 4 000

Table 24. **Optimized crystallization conditions of BmpC.**

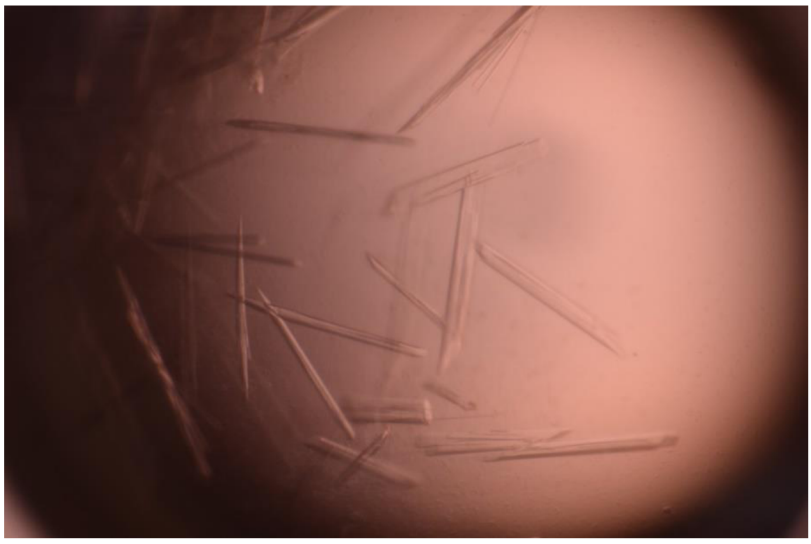

Fig. 33. **BmpC crystals.** Crystals grew under optimized conditions and were suitable for X-ray data collection. Volume of hanging drop was 2 µl.

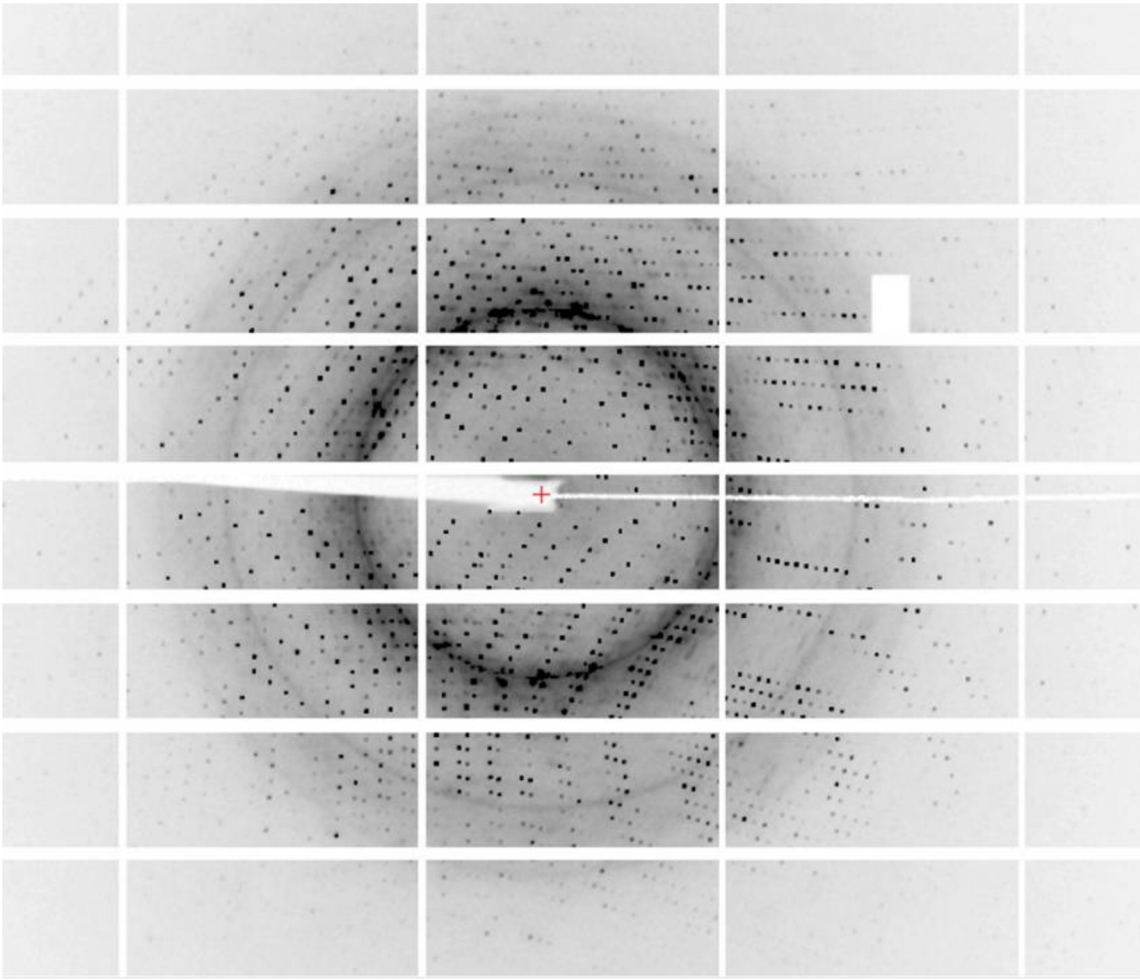

Fig. 34. **Diffraction pattern of BmpC crystal.** Measured at SOLEIL Synchrotron, France.

To prepare heavy atom derivatives of BmpC crystals, 8 phasing compounds were used for cocrystallization with the purified protein. No crystals of BmpC grew under the cocrystallization conditions. I3C phasing kit was as the only one used for soaking of native crystals of BmpC.

Structure of BmpC was solved by using multiple wavelength anomalous dispersion method. Diffraction datasets were collected for crystals soaked with I3C at 12.67 keV and 5.00 keV.

Parameters of crystal lattice calculated by CCP4 are shown in Table 25. Asymmetric unit and single chain of BmpC are shown in Fig. 35 and 36. Overall statistics of the model quality were calculated using MolProbity (Fig. 37).

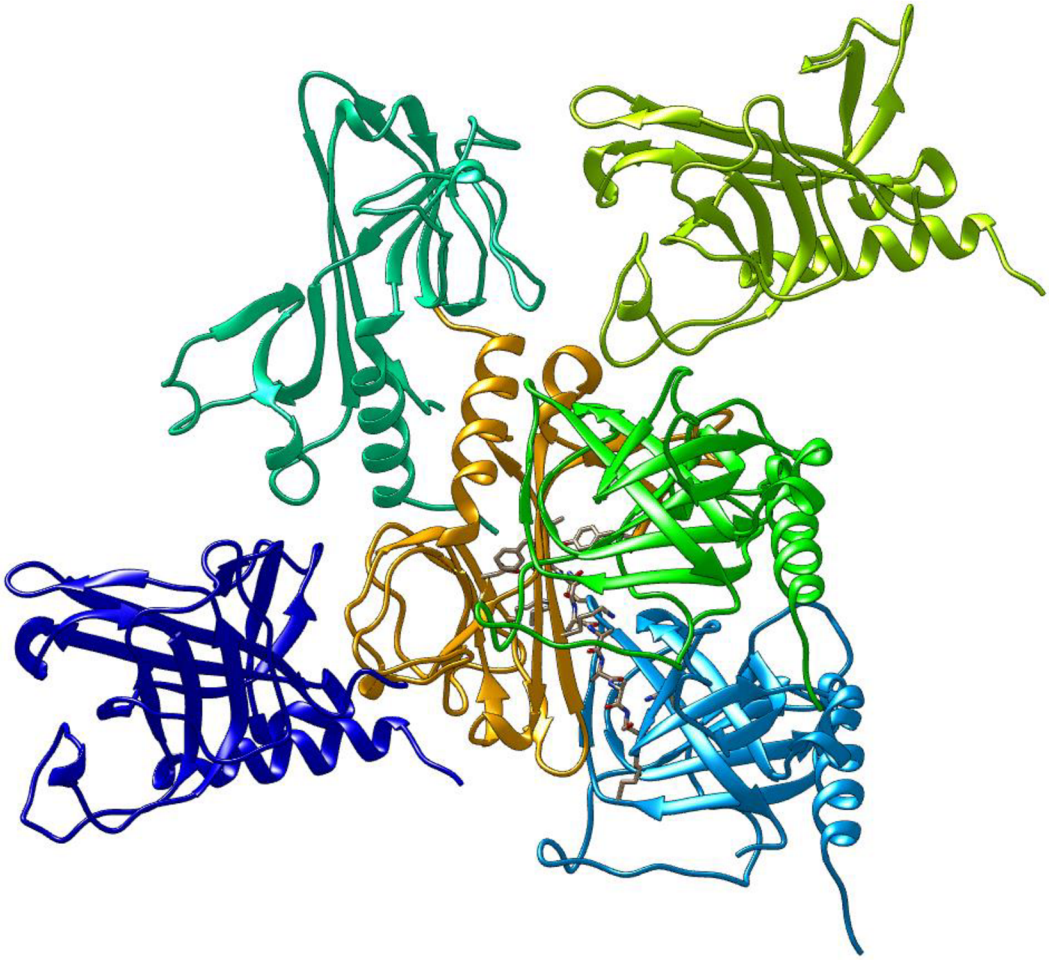

Fig. 35. **Asymmetric unit of BmpC crystal.** There are six BmpC molecules in one asymmetric unit. There is an unknown chain located in between blue and green molecule of BmpC.

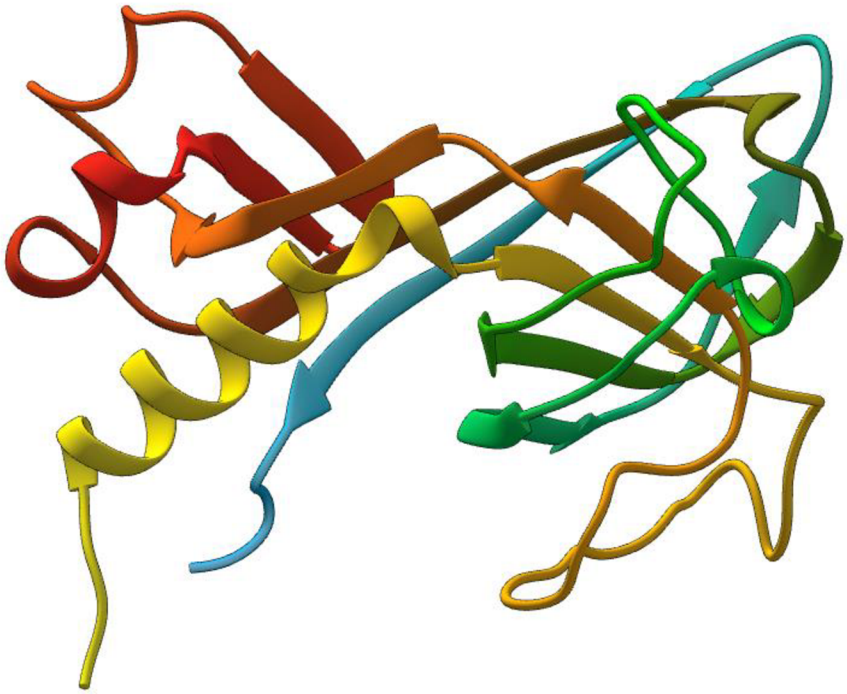

Fig. 36. **Single chain of BmpC.**

Unit cell parameters	a=92.93 Å; b=75.14 Å; c=94.45 Å; α=90°; β=109.87°; γ=90°
Space group	P 1 21 1
R-factor	0.2519
R-free	0.2926

Table 25. **Statistics of BmpC model calculated by CCP4.**

Clashscore, all atoms:	3	
Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.		
Poor rotamers	8	0.86%
Favored rotamers	858	92.76%
Ramachandran outliers	0	0.00%
Ramachandran favored	1007	97.48%
MolProbity score^	1.19	
Cβ deviations >0.25Å	1	0.10%
Bad bonds:	0 / 8364	0.00%
Bad angles:	2 / 11378	0.02%
Cis Prolines:	6 / 60	10.00%

Fig. 37. **Statistics of BmpC model.** Table was generated by MolProbity.

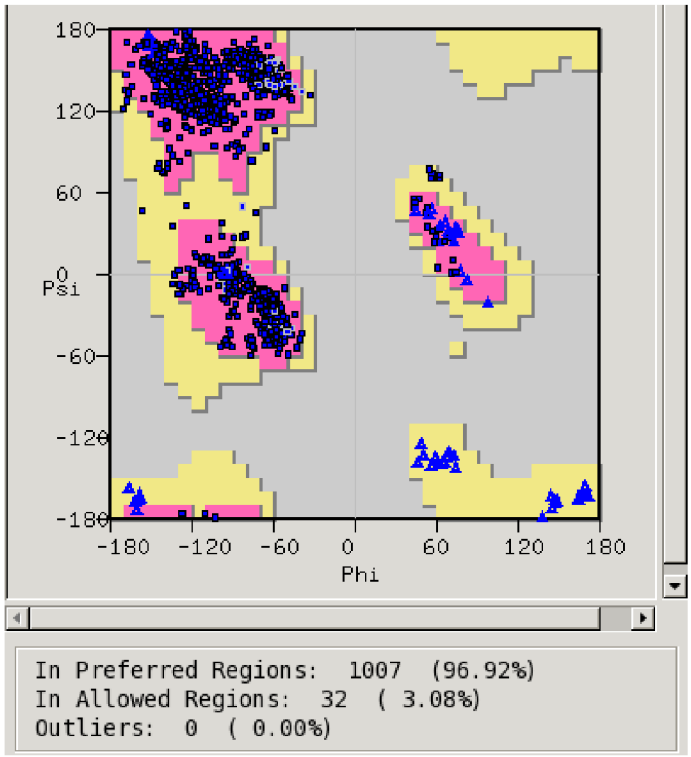

Fig. 38. **Ramachandran diagram of refined model of BmpC.** Diagram was generated in Coot.

4.4 Prohead protease

I was provided with 4 expression vector plasmids (courtesy of M. Šiborová). PP_pTriExTi contains gene of prohead protease with N-terminal His-tag followed by S-tag and Smt3 domain. PP_pET22T contains gene of prohead protease with N-terminal His-tag followed by Smt3 domain. Plasmids were stored in expression cells *E. coli* BL21 (DE3). Plasmid DNA was isolated and sequenced.

Expression vectors PP_short_pTriExTi and PP_short_pET22T contained shorter version of the original prohead protease gene. Shorter version of prohead protease gene was lacking 26 amino acids from N-terminal part and 8 amino acids from C-terminus. This was due to removal of flexible terminal parts of protein, which can hinder formation of protein crystals. The flexible regions were predicted by Phyre 2. Plasmids were sequenced and transformed into expression cells *E. coli* BL21 (DE3).

Expression tests were performed with all four vectors. For vectors PP_short_pTriExTi and PP_short_pET22T, no expressed protein was detected by MALDI-MS. MALDI-MS analysis confirmed the presence of expressed short version

of protein from PP_short_pTriExTi and PP_short_pET22T. Only vectors with short version of the protein were transformed into expression cells *E. coli* ArcticExpress and *E. coli* Tuner, and further used for expression tests (Tab. 26). Expression condition with highest yield of soluble protein is shown in Table 27.

PP_short_pTriExTi		**PP_short_pET22T**	
***E. coli* BL21 (DE3)**			
18°C	0.1/0.5/1 mM	18°C	0.1/0.5/1 mM
20°C	1 mM IPTG	20°C	1 mM IPTG
22°C	0.7 mM IPTG	22°C	0.7 mM IPTG
24°C	AI	24°C	AI
25°C	0.5/1 mM IPTG	25°C	0.5/1 mM IPTG
30°C	0.3/0.5/0.7/1 mM IPTG, AI	30°C	0.3/0.5/0.7/1 mM IPTG, AI
37°C	AI, 1 mM IPTG	37°C	AI, 1 mM IPTG
***E. coli* Tuner**			
18°C	0.05/0.1/0.3 mM IPTG	18°C	0.05/0.1/0.3 mM IPTG
37°C	1 mM IPTG	37°C	1 mM IPTG
***E. coli* ArcticExpress**			
10°C	0.1/0.5/1 mM IPTG	10°C	0.1/0.5/1 mM IPTG

Table 26. **List of expression conditions for vectors PP_short_pTriExTi and PP_short_pET22T.**

Vector	**Cells**	**Temperature**	**Induction**	**Time**
PP_short_pTriExTi	*E. coli* BL21 (DE3)	18°C	1 mM IPTG	7 hours

Table 27. **Expression conditions, which gave highest yield of soluble prohead protease.**

Prohead protease was purified from expression cells harvested from 3 liters of LB medium. Protein was first separated by IMAC with HisTrap column (Fig. 40 and 41). The protein was then cleaved by Ulp1-ac protease to remove His-tag with Smt3 domain (Fig. 42).

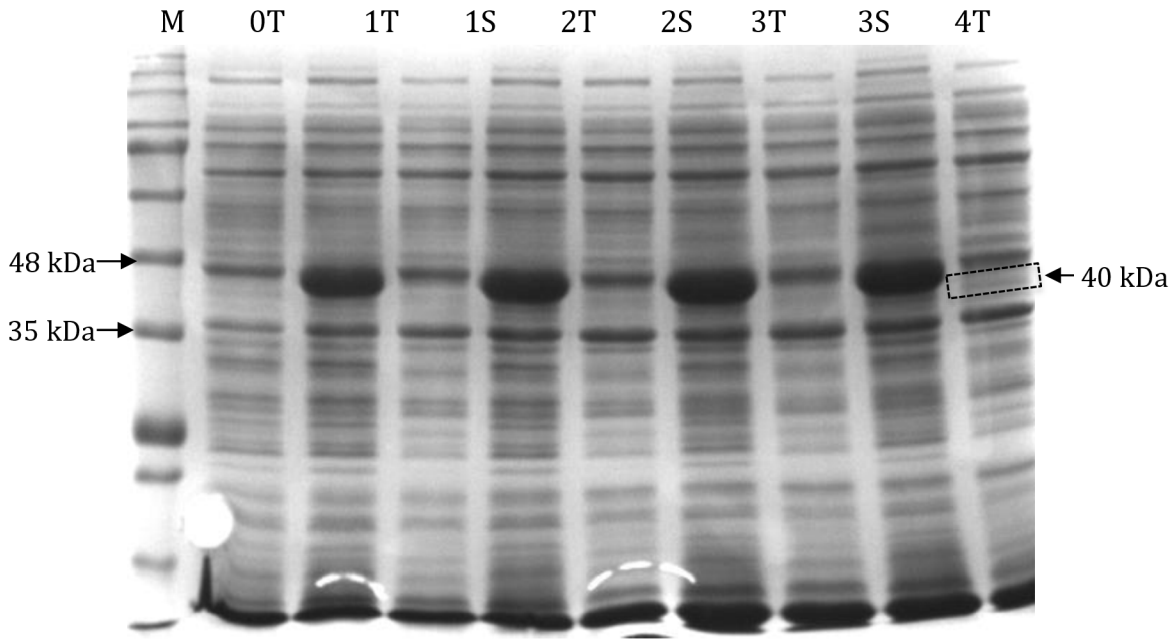

Fig. 39. **SDS-PAGE gel of expression condition with highest yield of soluble prohead protease.** Position of soluble prohead protease is highlighted by bracket.

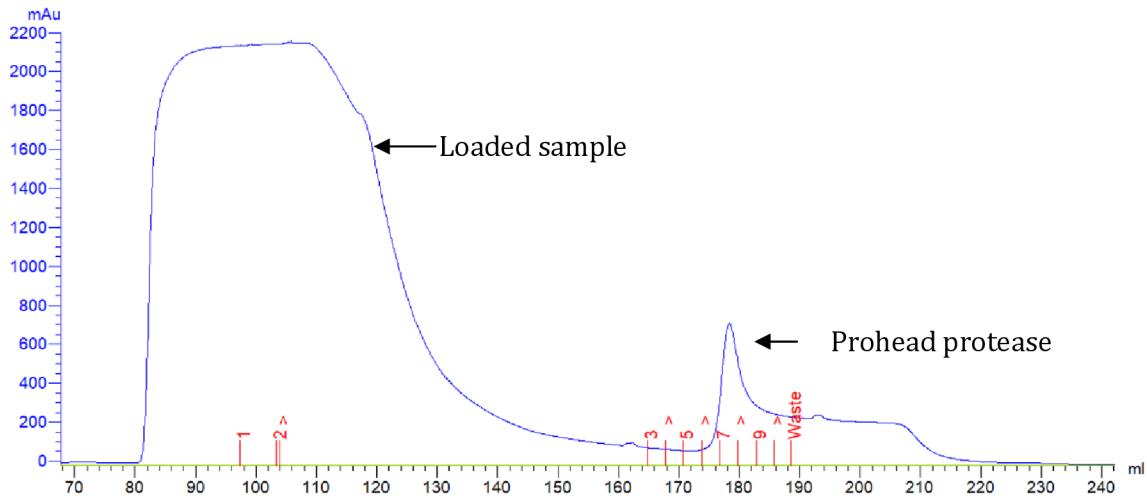

Fig. 40. **Chromatogram of prohead protease purification with IMAC on HisTrap column.**

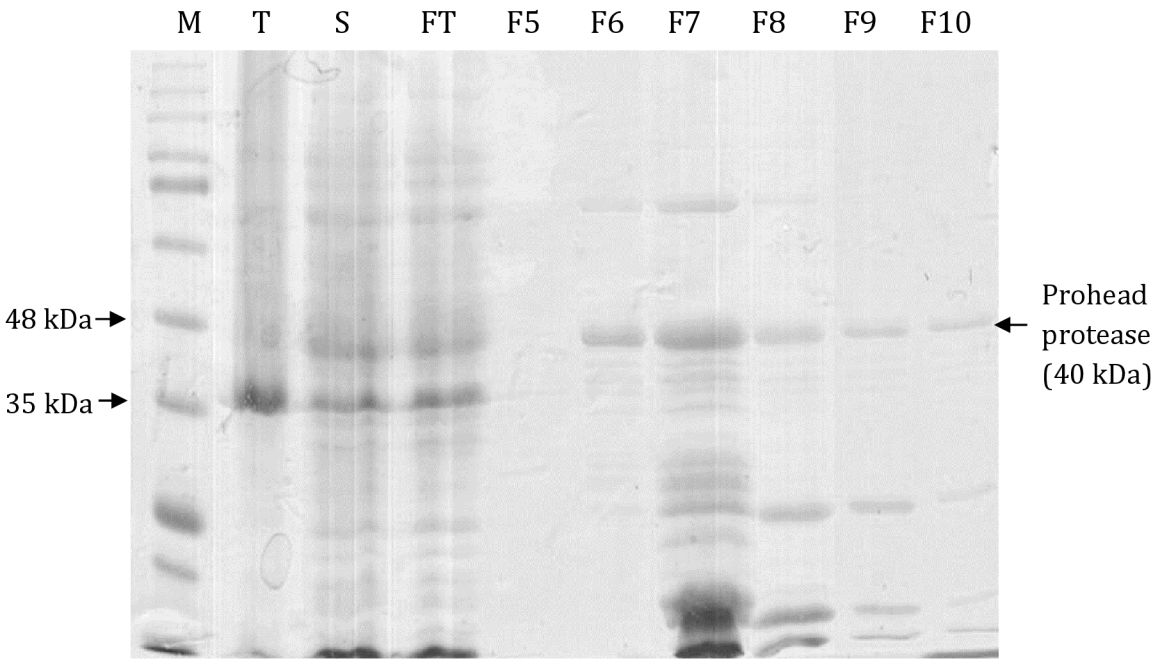

Fig. 41. **SDS-PAGE gel of purified prohead protease with IMAC on HisTrap column.** M: marker; T: total fraction; S: soluble fraction; FT: flowthrough; F5-F10: eluted fractions.

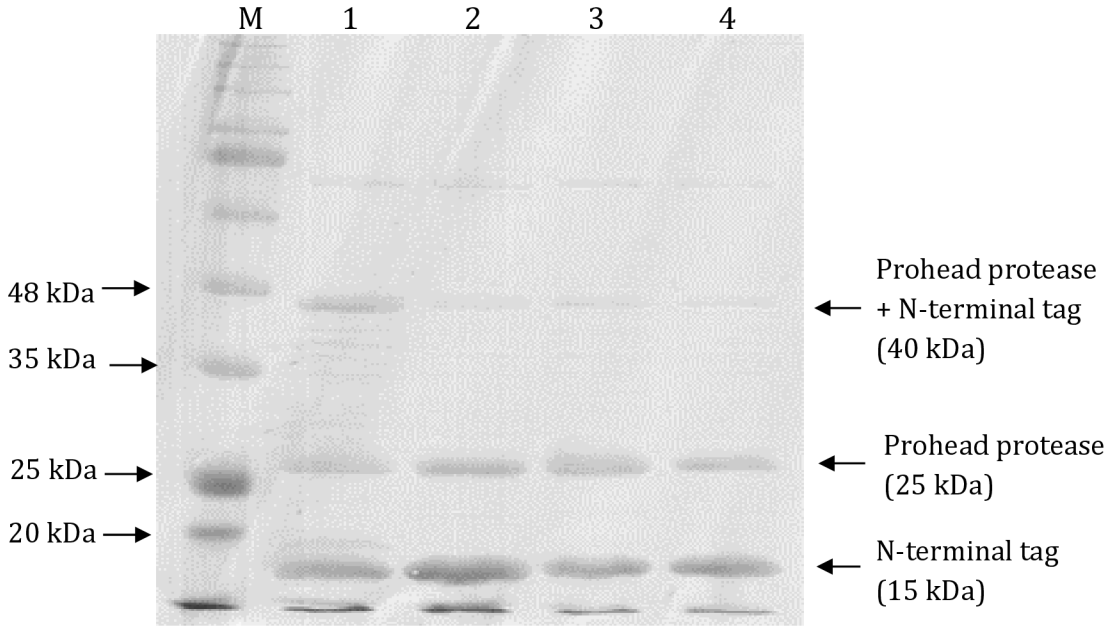

Fig. 42. **SDS-PAGE gel of prohead protease after cleavage with Ulp1-ac protease.** M: marker; 1: control sample without protease, stored at -20°C; 2: control sample without protease stored at 4°C; 3: total fraction at 4°C; 4: soluble fraction at 4°C. Control samples indicated cleavage of approximately 15 kDa of N-terminal part even without added protease.

After buffer exchange and negative HisTrap, sample containing soluble prohead protease was purified (Fig. 43 and 44) with Superdex column. Purified protein was concentrated to 11.3 mg/ml and measured for DLS (Fig. 45).

Resuspension buffer	50 mM Tris (pH=8), 200 mM NaCl, 5 mM β-mercaptoethanol
HisTrap loading buffer	50 mM Tris (pH=8), 200 mM NaCl, 20 mM imidazole
Histrap elution buffer	50 mM Tris (pH=8), 200 mM NaCl, 500 mM imidazole
Superdex loading buffer	50 mM Tris (pH=8), 200 mM NaCl

Table 28. **List of buffers used during purification of prohead protease.**

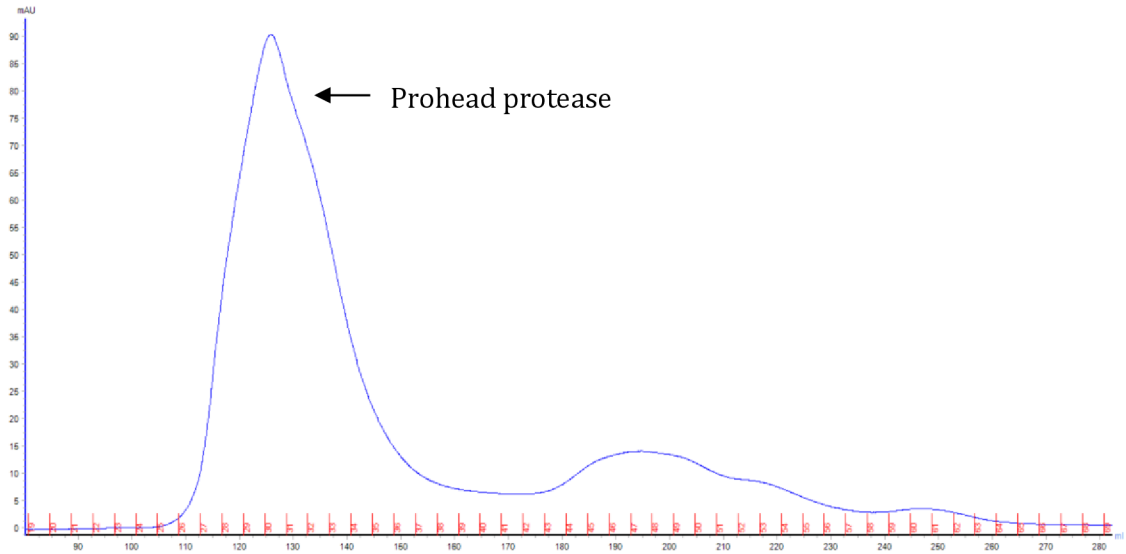

Fig. 43. **Chromatogram of size exclusion chromatography of prohead protease on Superdex 75 column.** Retention time of eluted peak suggests, that prohead protease was in solution in monomeric state.

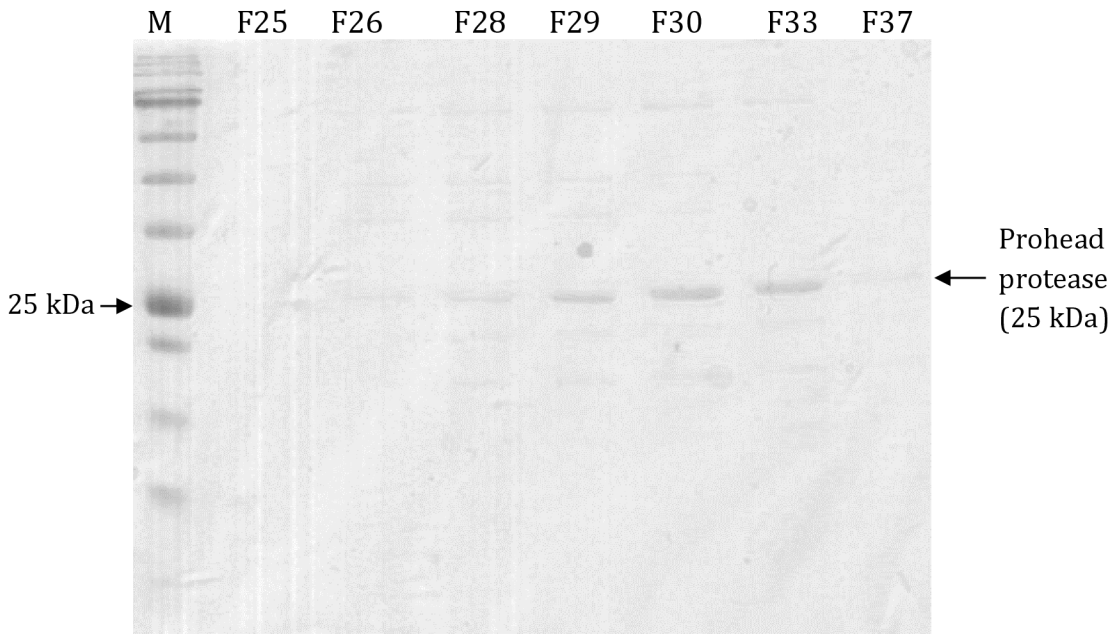

Fig. 44. **SDS-PAGE gel of prohead protease after size exclusion chromatography on Superdex 75 column.**

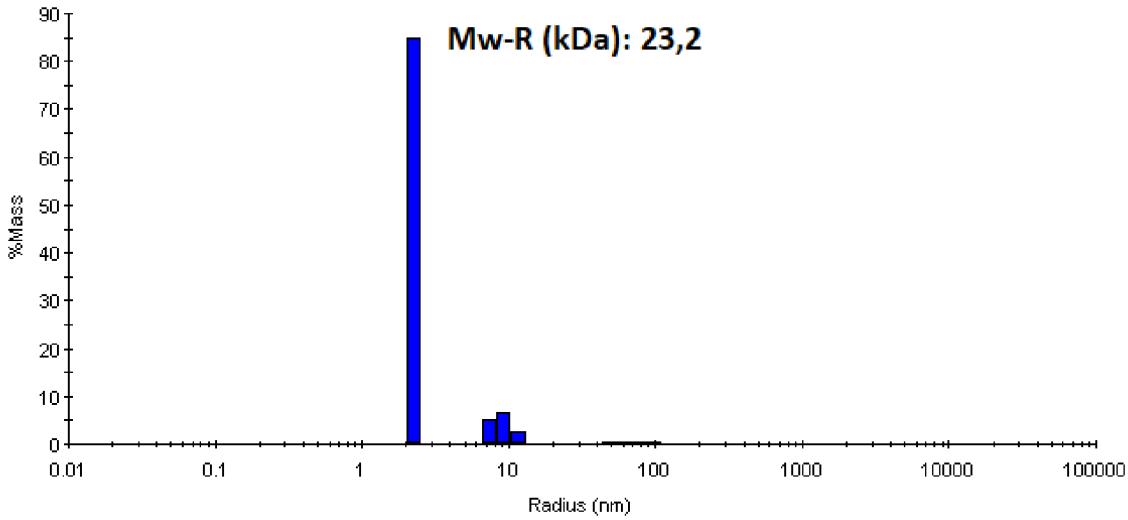

Fig. 45. **DLS analysis of purified prohead protease sample.** Component with relative molecular weight 23.2 kDa represent 84.3 % of scattering mass.

Eight crystallization screens were prepared, using the sitting drop vapor diffusion method. No crystals were observed during first week after screening.

5 Conclusions

The main goal of this thesis was to determine structure of studied virus proteins by X-rax crystallography. This requires to clone gene of each protein into expression vector, to purify expressed proteins and to crystallize them. If protein crystals provide high resolution X-ray diffraction data and protein phases can be determined, initial protein model is obtained and further refined to build final model of structure.

For expression of protein BmpA, two expression vector plasmids were prepared by the ligase independent cloning. Both constructs were designed with N-terminal tag. Plasmids were transformed into three strains of expression cells. Multiple expression tests were carried out to find the best conditions for protein expression in small volume of growth medium. Highest yield of expressed protein was observed in *E. coli* Tuner cells, with high copy BmpA_pTriExTi plasmid. Corresponding expression conditions were further used to produce protein in large volume of growth medium with subsequent purification. After Smt3 domain was cleaved off, precipitation of BmpA was observed progressively with each further purification procedure. Therefore, composition of buffers used for the purification was optimized to keep BmpA in soluble form. Since BmpA was attached to HisTrap column even after cleavage of His-tag, elution with low concentration of imidazole was necessary. Purified BmpA was concentrated to 10.9 mg/ml and screened for crystallization. No large protein crystals were observed after four months of inspecting the crystallizations screens. Structure of protein BmpA was not solved, since BmpA did not form suitable crystals for data collection. There are several reasons, which could have restrained the crystallization, like protein concentration, screen condition and temperature. Future experiment will be aimed to purify BmpA from larger volume of expression cells. The reason is to prepare purified sample with higher concentration of BmpA. This will enable testing of more crystallization conditions and temperatures. If no crystals appear under new conditions, construction of new vector plasmid will be required.

Protein BmpB was expressed with N-terminal His-tag followed by Smt3 domain. Expression and purification were carried out as recommended by previously obtained protocol. After N-terminal part with Smt3 domain was cleaved off, results

of SDS-PAGE showed, that whole construct was partially cleaved prior to addition of the protease. MALDI-MS analysis results suggest that the protein was cleaved in two different sites and BmpB is in solution in two forms, which differ in their N-terminal part. BmpB was purified with size exclusion chromatography and concentrated to 9.4 mg/ml. No crystals appeared after four months of incubation of crystallization screens. Structure of BmpB was not solved because no crystals for data collection formed in the crystallization screens. One of the reasons for unsuccessful crystallization of BmpB could be insufficient concentration of protein. Therefore, purification of BmpB from larger volume of harvested cells will be considered. However, main concern for restrained crystallization is possibility of BmpB being cleaved at two different sites. Two different N-terminal parts can interfere with regular packing of BmpB molecules into crystals. Therefore, new constructs are being currently prepared. Vectors, which are going to be used for the cloning are pTriExTi, pET22T and pET-28a.

Protein BmpC was designed with C-terminal His-tag. Expression and purification were carried out according to previously obtained protocol. BmpC was after purification concentrated to 14.0 mg/ml. Crystals of BmpC were observed in crystallization screens after two weeks of incubation. Conditions for crystal growth were successfully optimized, and large compact crystals of BmpC were prepared. Crystals were derivatized with iodine compound and structure solved at 1.8 Å by MAD method. Model of BmpC structure will be further refined. After the improvement of overall statistics of the model, we will fit the model into existing cryo-EM electron density map of phage phi812 base plate. We will also try to identify the unknown chain present in asymmetric unit.

Since BmpA, BmpB and BmpC were purified, cryo-EM will be considered for their structure solving in the future. Even though single molecules of these proteins are too small to be solved by cryo-EM, together in solution they might create regular complexes big enough to be examined by cryo-EM. The assumption is that phage proteins, which have similar function are clustered together in the genome, therefore are expressed one after another and used in virion assembly. By preparing mixture of the purified proteins in suitable ratios, complexes of these proteins might form, as they would form in phage-infected bacterial cell. However, BmpA (gp 118), BmpB (gp119) and BmpC (gp121) are in the genome separated by

gp120. Therefore, aim of future experiment is to purify gp120 and to prepare mixture of all four proteins in various ratios. If regular complexes of these four proteins assemble, they will be big enough to be examined by cryo-EM.

Four expression vectors were used for expression tests of prohead protease. Vector plasmid pTriExTi with short version of prohead protease gene produced the highest yield of expressed protein. SDS-PAGE result suggests that N-terminal part of protein, containing Smt3 domain was partially cleaved off prior to addition of protease. This brings similar concerns regarding crystallization, as in the case of BmpB. Different cleavage sites on the N-terminus of the prohead protease molecules might inhibit crystal formation. The protein was purified and concentrated to 11.3 mg/ml. Crystallization screens are currently incubated and no crystals have been observed. The crystallization screens will be further inspected. In case of successful crystallization of prohead protease, corresponding screen conditions will be used for optimization with hanging drop vapor diffusion method. If crystals do not appear after two months of screen incubation, Prohead protease will be purified from larger volume of expression cells and more concentrated sample for crystallization screens will be prepared. Cloning of new expression vector will also be considered.

6 Summary

Diseases related to *S. aureus* are a serious threat for patients world-wide. With decreasing number of effective antibiotics, novel therapeutics need to be developed. Phages, which infect *S. aureus*, are a promising option, but closer details regarding their structure are necessary for implementation of phages in human medicine.

Theoretical part of this thesis was focused on basic virus biology with emphasis on bacterial viruses. Theory of X-ray diffraction by crystal and calculation of electron density distribution were also briefly described.

The aim of this thesis in experimental part was to determine structure of four proteins of staphylococcal phage phi812 by combining methods of molecular biology and X-ray crystallography. BmpA, BmpB and BmpC are structural proteins and part of base plate. Prohead protease plays role in capsid maturation. All proteins were synthetized by using heterologous expression system of *E. coli*. Proteins were purified and tested for crystal formation with crystallization screens. Protein crystals were produced only by BmpC. After optimization of crystallization conditions and soaking of crystals with iodine solution I3C, structure of BmpC was solved with resolution 1.8 Å. In case of proteins, which did not crystallize, further approaches to determine structures were discussed.

7 References

[1] C. L. Ventola, "The Antibiotic Resistance Crisis," *Pharm. Ther.*, vol. 40, no. 4, pp. 277–283, Apr. 2015.

[2] M. Gross, "Antibiotics in crisis," *Curr. Biol.*, vol. 23, no. 24, pp. R1063–R1065, Dec. 2013.

[3] Z. Golkar, O. Bagasra, and D. G. Pace, "Bacteriophage therapy: a potential solution for the antibiotic resistance crisis," *J. Infect. Dev. Ctries.*, vol. 8, no. 02, pp. 129–136, Feb. 2014.

[4] I. M. Gould, J. Reilly, D. Bunyan, and A. Walker, "Costs of healthcare-associated methicillin-resistant Staphylococcus aureus and its control," *Clin. Microbiol. Infect.*, vol. 16, no. 12, pp. 1721–1728, Dec. 2010.

[5] M. J. Kuehnert, H. A. Hill, B. A. Kupronis, J. I. Tokars, S. L. Solomon, and D. B. Jernigan, "Methicillin-resistant–Staphylococcus aureus Hospitalizations, United States," *Emerg. Infect. Dis.*, vol. 11, no. 6, pp. 468–872, Jun. 2005.

[6] S. Chibani-Chennoufi, J. Sidoti, A. Bruttin, E. Kutter, S. Sarker, and H. Brüssow, "In Vitro and In Vivo Bacteriolytic Activities of Escherichia coli Phages: Implications for Phage Therapy," *Antimicrob. Agents Chemother.*, vol. 48, no. 7, pp. 2558–2569, Jul. 2004.

[7] R. Pantůcek *et al.*, "The polyvalent staphylococcal phage phi 812: its host-range mutants and related phages," *Virology*, vol. 246, no. 2, pp. 241–252, Jul. 1998.

[8] R. M. Carlton, "Phage Therapy: Past History and Future Prospects," *Arch. Immunol. Ther. Exp.*, vol. 47, pp. 267–274, 1999.

[9] Ø. Bergh, K. Y. BØrsheim, G. Bratbak, and M. Heldal, "High abundance of viruses found in aquatic environments," *Nature*, vol. 340, no. 6233, pp. 467–468, Aug. 1989.

[10] S. Jacquet, T. Miki, R. Noble, P. Peduzzi, and S. Wilhelm, "Viruses in aquatic ecosystems: important advancements of the last 20 years and prospects for the future in the field of microbial oceanography and limnology," *Adv. Oceanogr. Limnol.*, vol. 1, no. 1, pp. 97–141, Jun. 2010.

[11] J. B. Carter and V. A. Saunders, *Virology: Principles and Applications*. Wiley, 2013.

[12] E. C. Keen, "Tradeoffs in bacteriophage life histories," *Bacteriophage*, vol. 4, Feb. 2014.

[13] D. Baltimore, "Expression of animal virus genomes.," *Bacteriol. Rev.*, vol. 35, no. 3, pp. 235–241, Sep. 1971.

[14] A. M. Q. King, M. J. Adams, E. B. Carstens, and E. J. Lefkowitz, Eds., "Order - Caudovirales," in *Virus Taxonomy*, San Diego: Elsevier, 2012, pp. 39–45.

[15] H.-W. Ackermann and D. Prangishvili, "Prokaryote viruses studied by electron microscopy," *Arch. Virol.*, vol. 157, no. 10, pp. 1843–1849, Oct. 2012.

[16] H.-W. Ackermann, "Tailed Bacteriophages: The Order Caudovirales," in *Advances in Virus Research*, vol. 51, K. Maramorosch, F. A. Murphy, and A. J. Shatkin, Eds. Academic Press, 1998, pp. 135–201.

[17] E. M. Adriaenssens *et al.*, "Taxonomy of prokaryotic viruses: 2017 update from the ICTV Bacterial and Archaeal Viruses Subcommittee," *Arch. Virol.*, vol. 163, no. 4, pp. 1125–1129, Apr. 2018.

[18] M. H. V. van Regenmortel and B. W. J. Mahy, *Desk Encyclopedia of General Virology*. Academic Press, 2010.

[19] E. Kutter and A. Sulakvelidze, *Bacteriophages: Biology and Applications*. CRC Press, 2004.

[20] G. J. Tortora, B. R. Funke, and C. L. Case, *Microbiology: an introduction*, 12th ed. USA: Pearson Publish, 2016.

[21] F. D. Lowy, "Staphylococcus aureus Infections," *N. Engl. J. Med.*, vol. 339, no. 8, pp. 520–532, Aug. 1998.
[22] "National Nosocomial Infections Surveillance (NNIS) System Report, Data Summary from January 1992-June 2001, Issued August 2001," *Am. J. Infect. Control*, vol. 29, no. 6, pp. 404–421, Dec. 2001.
[23] M. J. Loessner, "Bacteriophage endolysins — current state of research and applications," *Curr. Opin. Microbiol.*, vol. 8, no. 4, pp. 480–487, Aug. 2005.
[24] J. Nováček, M. Šiborová, M. Benešík, R. Pantůček, J. Doškař, and P. Plevka, "Structure and genome release of Twort-like Myoviridae phage with a double-layered baseplate," *Proc. Natl. Acad. Sci.*, vol. 113, no. 33, pp. 9351–9356, Aug. 2016.
[25] Bragg William Henry and Bragg William Lawrence, "The reflection of X-rays by crystals," *Proc. R. Soc. Lond. Ser. Contain. Pap. Math. Phys. Character*, vol. 88, no. 605, pp. 428–438, Jul. 1913.
[26] B. R. Wheaton, *The Tiger and the Shark: Empirical Roots of Wave-Particle Dualism*. Cambridge University Press, 1991.
[27] J. Drenth, *Principles of Protein X-Ray Crystallography*, 3rd ed. New York: Springer-Verlag, 2007.
[28] G. P. Williams, "A general review of synchrotron radiation, its uses and special technologies," *Vacuum*, vol. 32, no. 6, pp. 333–345, Jan. 1982.
[29] M. S. Smyth and J. H. J. Martin, "x Ray crystallography," *Mol. Pathol.*, vol. 53, no. 1, pp. 8–14, Feb. 2000.
[30] C. G. Pope, "X-Ray Diffraction and the Bragg Equation," *J. Chem. Educ.*, vol. 74, no. 1, p. 129, Jan. 1997.
[31] L. Žídek, "Strukturní biochemie - Skripta k přednášce C9530 Přírodovědecké fakulty Masarykovy univerzity." 2014.
[32] G. Taylor, "The phase problem," *Acta Crystallogr. D Biol. Crystallogr.*, vol. 59, no. 11, pp. 1881–1890, Nov. 2003.
[33] A. L. Patterson, "A Direct Method for the Determination of the Components of Interatomic Distances in Crystals," *Z. Für Krist. - Cryst. Mater.*, vol. 90, no. 1–6, pp. 517–542, 2015.
[34] J. C. KENDREW, G. BODO, H. M. DINTZIS, R. G. PARRISH, H. WYCKOFF, and D. C. PHILLIPS, "A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis," *Nature*, vol. 181, p. 662, Mar. 1958.
[35] A. C. W. Pike, E. F. Garman, T. Krojer, F. von Delft, and E. P. Carpenter, "An overview of heavy-atom derivatization of protein crystals," *Acta Crystallogr. Sect. Struct. Biol.*, vol. 72, no. Pt 3, pp. 303–318, Mar. 2016.
[36] Z. Dauter, M. Dauter, and K. R. Rajashankar, "Novel approach to phasing proteins: derivatization by short cryo-soaking with halides," *Acta Crystallogr. D Biol. Crystallogr.*, vol. 56, no. 2, pp. 232–237, Feb. 2000.
[37] H. Walden, "Selenium incorporation using recombinant techniques," *Acta Crystallogr. D Biol. Crystallogr.*, vol. 66, no. Pt 4, pp. 352–357, Apr. 2010.
[38] W. A. Hendrickson, "Determination of macromolecular structures from anomalous diffraction of synchrotron radiation," *Science*, vol. 254, no. 5028, pp. 51–58, Oct. 1991.
[39] P. Evans and A. McCoy, "An introduction to molecular replacement," *Acta Crystallogr. D Biol. Crystallogr.*, vol. 64, no. Pt 1, pp. 1–10, Jan. 2008.
[40] A. Wlodawer, W. Minor, Z. Dauter, and M. Jaskolski, "Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures," *FEBS J.*, vol. 275, no. 1, pp. 1–21, Jan. 2008.
[41] A. Marbach and K. Bettenbrock, "lac operon induction in Escherichia coli: Systematic comparison of IPTG and TMG induction and influence of the transacetylase LacA," *J. Biotechnol.*, vol. 157, no. 1, pp. 82–88, Jan. 2012.
[42] Y. Zhang, X. M. Qu, and S. L. Yang, "[The influences of lactose as an inducer on the expression of the recombinant proteins in Escherichia coli BL21 (DE3)]," *Sheng Wu Gong Cheng Xue Bao Chin. J. Biotechnol.*, vol. 16, no. 4, pp. 464–468, Jul. 2000.

[43] J. A. Bornhorst and J. J. Falke, "[16] Purification of Proteins Using Polyhistidine Affinity Tags," *Methods Enzymol.*, vol. 326, pp. 245–254, 2000.
[44] A. Madadlou, S. O'Sullivan, and D. Sheehan, "Fast protein liquid chromatography," *Methods Mol. Biol. Clifton NJ*, vol. 681, pp. 439–447, 2011.
[45] A. Hampel *et al.*, "Single Crystals of Transfer RNA from Formylmethionine and Phenylalanine Transfer RNA's," *Science*, vol. 162, no. 3860, pp. 1384–1387, Dec. 1968.
[46] T. R. Schneider and G. M. Sheldrick, "Substructure solution with SHELXD," *Acta Crystallogr. D Biol. Crystallogr.*, vol. 58, no. 10, pp. 1772–1779, Oct. 2002.
[47] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan, "Features and development of Coot," *Acta Crystallogr. D Biol. Crystallogr.*, vol. 66, no. Pt 4, pp. 486–501, Apr. 2010.
[48] A. A. Vagin *et al.*, "REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use," *Acta Crystallogr. D Biol. Crystallogr.*, vol. 60, no. 12, pp. 2184–2195, Dec. 2004.
[49] V. B. Chen *et al.*, "MolProbity: all-atom structure validation for macromolecular crystallography," *Acta Crystallogr. D Biol. Crystallogr.*, vol. 66, no. Pt 1, pp. 12–21, Jan. 2010.