Alina Vitaliyivna Kulakova

BSc in Biochemistry

Production of membrane histidine kinases and structural studies on Thiosulfate dehydrogenase

Dissertation for the Master degree in Biochemistry for Health

Supervisor: Margarida Archer Frazão
Co-supervisors: Ana Lúcia Rosário, José Artur Brito

October 2015

Alina Vitaliyivna Kulakova

BSc in Biochemistry

Production of membrane histidine kinases and structural studies on Thiosulfate dehydrogenase

Dissertation for Master degree in Biochemistry for Health

Supervisor: Dr. Margarida Archer Frazão
Co-supervisors: Dr. Ana Lúcia Rosário, Dr. José Artur Brito

President: Dr. António Sebastião Rodrigues
Arguer: Dr. Ana Luísa Moreira de Carvalho
Vowel(s): Dr. Maria Teresa Catarino, Dr. Margarida Archer Frazão

Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa

October 2015

Production of membrane histidine kinases and structural studies on Thiosulfate dehydrogenase

Acknowledgments

I would like to express my acknowledgments to Dr. Margarida Archer Frazão for accepting me in her laboratory and giving me opportunity to perform this work, for support and constructive criticism that allowed me to learn and develop my capabilities.

I want to acknowledge my colleagues from the Membrane Protein Crystallography Laboratory. Very especial thanks to Ana Lúcia Rosário for all that she taught me, for all the help and support during this year. In most difficult and critical moments of my work she was always ready to help and I am very happy to have met someone like her. I want to thank José Artur Brito for teaching me crystallization, data collection and structure resolution, for all help and criticism that helped me to learn. Many thanks for Ana Maria for teaching me the cloning, giving me advices, for help and support, even when she was already part of another research group. Many thanks for my colleague, Diogo Athayde, for his help, for work discussions and advices. Thank you all for the good atmosphere of trust and friendship in the laboratory.

I also would like to acknowledge all members of the Macromolecular Crystallography unit at ITQB for all tips and advices that helped me in my work.

Gostaria de agradecer a minha família por toda ajuda e o apoio que me deram. Obrigada aos meus pais, Leonid e Lyudmyla, por todas as dificuldades que passaram para tornar os meus sonhos realidade. Por eu estar aqui, neste país, a estudar e a trabalhar no que mais gosto. Obrigada à minha irmã, Esmeralda, por estar sempre ao meu lado e ser a melhor irmã do mundo. Obrigada aos meus avós, por tudo o que me ensinaram e por tudo o que fizeram por mim. Obrigada a todos os meus amigos e ao meu companheiro, Maik. Obrigada por acreditarem em mim, sem vocês não seria possível estar onde estou hoje.

Abstract

This thesis is organized in three chapters. Chapter 1 comprises a brief introduction on the methodology used throughout the work herein presented. Chapter 2 describes the production of membrane histidine kinases from *Staphylococcus aureus* and *Clostridium difficile* and chapter 3 the structural studies of a thiosulfate dehydrogenase from *Campylobacter jejuni*. The experimental work was performed in the Membrane Protein Crystallography Laboratory at Instituto de Tecnologia Química e Biológica, and also included a visit to SOLEIL synchrotron, Paris, France, to collect X-ray diffraction data.

Histidine kinases (HKs) are part of two-component signal transduction systems (TCS) that are integrated in a large variety of cellular signalling circuits, *e.g.* response to environmental changes, drug resistance and virulence. TCS are common in bacteria but absent in mammals, an evidence of their potential use as targets for the development of antimicrobial therapies. The aim of this work was the structural and functional characterization of purified histidine kinases. This work started by cloning several membrane sensor HKs from *S. aureus* and *C. difficile,* SrrB, AgrC, ArlS, BceS SaeS and NreB were selected for expression and solubilization tests. Generally, all targets showed a good expression in *Escherichia coli,* but membrane extraction was a problematic step in this study. Due to the low protein quantity in the soluble fraction, it was not possible to obtain sufficient amount of pure protein to further proceed for functional and structural studies. Therefore, we decided to perform structural studies on a soluble protein, a thiosulfate dehydrogenase from *Campylobacter jejuni*.

C. jejuni is a microaerophile mucosal intestinal pathogen that causes acute bloody diarrhoea in humans. Such as many of sulfur bacteria, *C. jejuni* uses thiosulfate dehydrogenase (TsdA). TsdA was shown to be a bifunctional enzyme displaying both tetrathionate reductase and thiosulfate dehydrogenase activities. Depending on the organism needs, TsdA displays more of one of the activities. Since tetrathionate can be produced in the intestine during inflammation process, the ability of *C. jejuni* to reduce tetrathionate promotes its growth in the host organism gut. The aim of this work was the crystallization, data collection and model building for several *C. jejuni* TsdA variants, namely the "as isolated" form and the C138G and N254G mutants. All variants were crystallized by the vapour diffusion method and X-ray diffraction data was collected at a synchrotron source to 1.95, 2.37 and 1.72 Å resolutions, respectively. The structures were determined by molecular replacement using *Allochromatium vinosum* TsdA as a template. Similar to *A. vinosum* TsdA, *C. jejuni* TsdA is a heart-shaped molecule composed by two domains and containing two hexacoordinated heme molecules. The proximal axial ligand is His^{99} and distal ligand is Cys^{138} for heme 1 and His^{207} and Met^{255} for heme 2.

Keywords: Histidine kinases, membrane protein production and purification, X-ray crystallography, thiosulfate dehydrogenase, model building.

Resumo

Esta tese está organizada em três capítulos. O capítulo 1 inclui uma breve introdução da metodologia utilizada. No capítulo 2 está descrita a produção das cinases de histidinas membranares de *Staphylococcus aureus* e *Clostridium difficile*; e no capítulo 3 são descritos os estudos estruturais da proteína tiossulfato desidrogenase de *Campylobacter jejuni.* O trabalho experimental foi realizado no Laboratório de Cristalografia de Proteínas Membranares no Instituto de Tecnologia Química e Biológica, e também inclui uma visita ao sincrotrão do SOLEIL em Paris, França.

As cinases de histidina (HK) fazem parte do sistema de transdução do sinal de dois componentes (TCS) e estão integradas em inúmeros circuitos de sinalização, tais como resposta a alterações ambientais, resistência a antibióticos e virulência. TCS são comuns em bactérias e ausentes em mamíferos, o que leva à sua potencial utilização na terapia antimicrobiana. Este trabalho tinha como objetivo os estudos estruturais e funcionais de amostras puras de proteínas - cinases de histidina. O trabalho experimental começou com a clonagem de HK de *S. aureus* e *C. difficile.* AgrC, ArlS, BceS, SaeS e NreB foram selecionados para prosseguir com expressão e solubilização. Na generalidade, todas as proteínas alvo têm uma boa expressão em *Escherichia coli,* mas a extração a partir da membrana total foi um passo problemático neste estudo. Devido a uma baixa quantidade da proteína na fração solúvel, não foi possível obter proteína pura suficiente para prosseguir com os estudos funcionais e estruturais. Assim foi decidido escolher uma proteína solúvel para estudos estruturais – a tiossulfato desidrogenase de *Campylobacter jejuni*.

C. jejuni é um patogénio microaerófilo da mucosa intestinal que causa diarreia sanguinolenta aguda em humanos. Tal como muitas bactérias sulfurosas, *C. jejuni* utiliza a tiossulfato desidrogenase (TsdA). TsdA mostra ser uma enzima bifuncional, exibindo as atividades de tiossulfato desidrogenase e tetrationato redutase. As bactérias desempenham mais de uma atividade do que da outra, dependendo do tipo e necessidades do organismo. Uma vez que tetrationato pode ser produzido no intestino durante os processos inflamatórios, a habilidade de reduzir o tetrationato leva ao crescimento do *C. jejuni* no intestino do hospedeiro. O objetivo deste trabalho foi a cristalização, recolha dos dados e construção do modelo cristalográfico para diferentes variantes de TsdA *C. jejuni,* nomeadamente nativa e mutantes C138G e N254G. Todos os variantes foram cristalizados pelo método de difusão de vapor e os dados de difração de raios-X foram recolhidos em sincrotrão a resoluções 1.95, 2.37 e 1.72 Å, respetivamente. Todas as estruturas foram determinadas por substituição molecular usando como modelo a TsdA do *Allochromatium vinosum.* À semelhança da TsdA de *A. vinosum,* a TsdA de *C. jejuni* é uma molécula em forma de coração, constituída por dois domínios que contêm duas moléculas de hemo hexacoordenadas. O ligando proximal do hemo 1 é His^{99} e o ligando distal é Cys^{138}, tal como His^{207} e Met^{255} para o hemo 2.

Palavras-chave: Cinase de histidina, produção e purificação das proteínas membranares, cristalografia raios-X, tiossulfato desidrogenase, construção do modelo.

Table of contents

Index of Figures

Index of Tables

Abbreviation

A. vinosum	*Allochromatium vinosum*
Å	Ångström
ADP	Adenosine diphosphates
AIP	Autoinducing Peptide
ATP	Adenosine triphosphates
bp	Base pair
BSA	Bovine serum albumin
C. difficile	*Clostridium difficile*
C. jejuni	*Campylobacter jejuni*
CA	Catalytic ATP-binding domain
CAMPs	Cationic Antimicrobial Peptides
CB	Carbenicillin
CD	Circular Dichroism
CM	Chloramphenicol
CMC	Critical Micellar Concentration
Cymal5	5-cyclohexyl-1-pentyl-β-D-maltoside
Da	Dalton
DDM	*n*-dodecyl-β-D-maltoside
DHp	Dimerization and Histidine phosphotransfer domain
DLS	Dynamic Light Scattering
DM	Dodecyl-β-maltoside
DNA	Deoxyribonucleic acid
E.coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
GOI	Gene of interest
HEGA	Cyclohexylethanoyl-N-Hydroxyethylglucamide
HK	Histidine Kinase
HPLC	High Performance Liquid Chromatography
HPt	Histidine phosphotransferase
IMAC	Immobilized Metal-Affinity Chromatography
IPTG	Isopropyl-β-D-thiogalactopyranoside
LDAO	n-Dodecyl-N,N-Dimethylamine-N-Oxide
MAD	Multiple Anomalous Dispersion
MKH_2	Menaquinol pool
MP	Membrane Protein
MRSA	Methicillin-resistant *S. aureus*
Ni-NTA	Nickel-nitriloacetic acid
NMR	Nuclear Magnetic Ressonance
OD	Optical Density
PAGE	Polyacrylamide gel electrophoresis
P_i	Inorganic phosphate
PBS	Phosphate buffer solution
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEG	Polyethylene Glycol
RNA	Ribonucleic acid
Rpm	Revolutions per minute
RR	Response Regulator
SAD	Single Angle Dispersion
SAXS	Single Angle X-ray scattering
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
S. aureus	*Staphylococcus aureus*

TCS	Two-component signal transduction system
Tris	Tris(hidroxymethyl)-aminomethane
Triton X-100	Nonaethylene glycol octylphenol ether
TsdA	Thiosulfate dehydrogenase
w/v	Weight by volume
v/v	Volume by volume
2D	Bi-dimensional
3D	Three-dimensional

Chapter 1

Introduction for Materials and Methods

1. Introduction for materials and methods

In this chapter we will make a brief description of the techniques that were used on the production of membrane sensor histidine kinases and structural studies on thiosulfate dehydrogenase.

1.1 Recombinant DNA technology

Protein structure and function studies require large quantities of pure protein. Many proteins are normally expressed at very low amounts, precluding their isolation and purification processes. In order to produce larger quantities of a target protein, recombinant DNA technology is used. This process starts with DNA cloning by recombinant DNA methods. Molecular cloning is a set of experimental methodologies used to assemble recombinant DNA molecules and to direct their replication within a specific organism. This process consists in a DNA fragment insertion into a vector that is introduced into a host cell. The vector replicates and generates a large number of identical DNA molecules. In *Escherichia coli* (*E. coli*) organism, widely used for recombinant protein expression, plasmids are the most commonly used vectors. Plasmids are circular, double-stranded DNA molecules which naturally occur in bacteria and in lower eukaryotic cells. This DNA duplicates before every cell division, like a chromosomal DNA, and those copies are transferred to the daughter cells. Such process assures vector propagation through cell generation. Different plasmid types were engineered in order to optimize their use as vectors in DNA cloning. DNA fragment insertion into the plasmid requires two types of enzymes: restriction enzymes and DNA ligase. Restriction enzymes are endonucleases, which recognize specific sequences (restriction sites) and cleave DNA strands – DNA restriction. Restriction of specific gene of interest (GOI) allows the isolation of the gene encoding for the target protein. The GOI can be inserted into the plasmid by DNA ligase, which covalently join the complementary ends of the fragment and plasmid DNA.

Cell transformation

Cell transformation process consists in introducing a plasmid into a cell. Transformation can occur when the recombinant vector and host cells are mixed, under certain conditions. Normally, 1 cell in about 10000 is transformed and must be selected. This selection is possible because plasmids carry DNA fragments that confer resistance to specific antibiotics, such as ampicillin, chloramphenicol and others. After the transformation process all cells are placed in a solid growth medium (usually Luria Bertani[1] broth with agar) with the appropriate antibiotic and only transformed cells will survive and grow (figure 1.1).

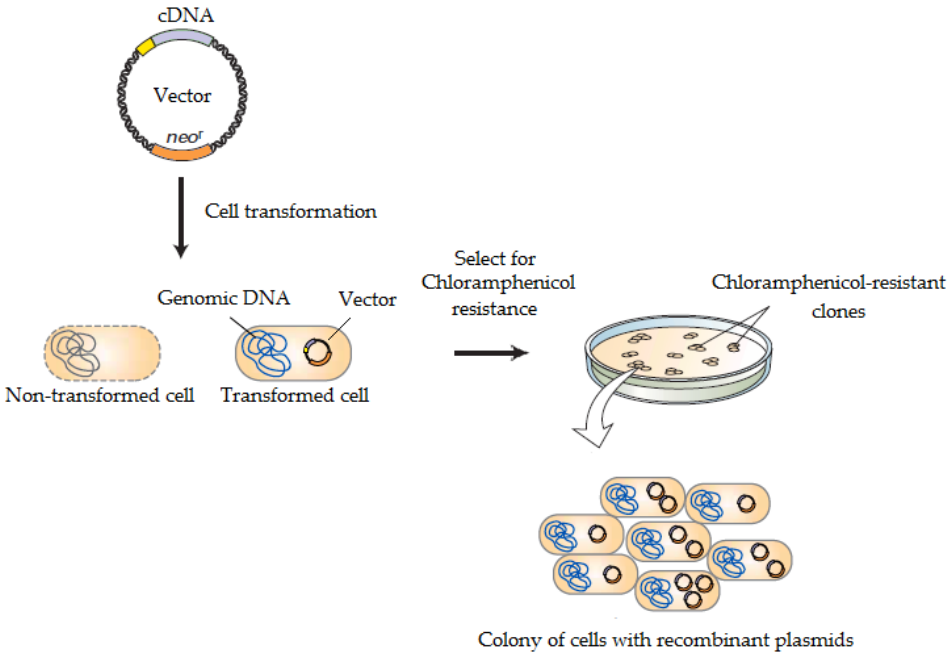

Figure 1.1 - Plasmid insertion and GOI propagation. The vector is inserted in a prokaryotic cell that will be selected by chloramphenicol resistance (Adapted from Molecular Cell Biology[2]).

1.2 Protein production

Vectors engineered for protein production are called expression vectors. These specific vectors contain a promoter region – DNA sequence that allows the transcription of the inserted GOI. The promoter is located at the adjacent region upstream of the target gene, so in the presence of an inductor, transcription occurs. One of the most common promoters is the *lac* promoter. Transcription only occurs in the presence of lactose or lactose analogues, such as isopropyl β-D-1-thiogalactopyranoside (IPTG).

Most soluble proteins can be overexpressed to high yields in terms of protein quantities. This is in sharp contrast to what normally happens with membrane proteins. The crucial and often limiting step is the protein insertion and folding within the membrane. Incorrect insertion or folding produces large quantities of inactive and/or aggregated protein.

Prokaryotic and eukaryotic expression systems can be used for membrane protein production. Both systems present several advantages and disadvantages. The bacterial expression system is cheaper and faster compared to an eukaryotic system, but the bacterial machinery does not allow for protein post-translational modifications (*e.g.* phosphorylation and glycosylation), which is crucial for the proteins that require it. However, post-translational modification introduces heterogeneity to the protein solution, which can be harmful for the crystallization of some proteins. The most frequently used bacteria expression system is *E. coli*[2].

Solubilization and purification of membrane proteins

Before proceeding with the purification process, the membrane protein needs to be solubilized from the membrane. Integral membrane proteins are extracted from the cell membrane with the help of detergents. Detergents are amphiphilic molecules that can form micelles with membrane proteins allowing their hydrophobic regions to be shielded in the interior of the micellar complex, while their hydrophilic regions are exposed to the solvent. These detergent-membrane protein micellar complexes are soluble in aqueous solution. Detergents can be classified as ionic (anionic and cationic), non-ionic and zwitterionic. Non-ionic detergents, particularly polyethylene and sugar detergents, are most often used in membrane protein solubilization and purification processes. Polyethylene detergents contain a polymeric chain of variable length that is usually used at neutral pH and at peripheral membrane protein extraction. Sugar detergents contain an hydrophobic part that is linked to a carbohydrate, such as glucose, maltose and sucrose, by a glycosidic bond. Those detergents are commonly used at integral membrane protein extraction.

The critical micelle concentration (CMC, minimal detergent concentration for micelle formation) is an important parameter to take in consideration when choosing a suitable detergent for membrane extraction (Figure 1.2).

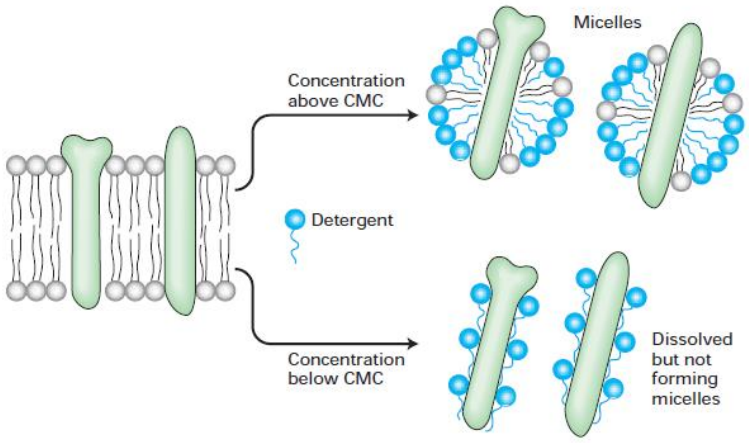

Figure 1.2 - Solubilization of membrane protein. Concentration below CMC value allows dissolution of membrane protein, but not micelle formation. In order to obtain micelles, concentration above CMC value is necessary (adapted from Molecular cell biology[2]).

Recombinant proteins can have peptide affinity tags or fusion proteins in their composition. These affinity markers are attached to the target protein by gene fusion to the amino(N)-terminal or carboxyl(C)-terminal. The presence of affinity markers allows for easy detection, immobilization and protein purification by affinity chromatography. Different affinity tags may be added to the target protein, such as *His-tag*, *Strep-tag* and *Flag-tag*. The most used one is the *His-tag*, which have 6-10 consecutive histidine residues allowing the protein to be purified by immobilized metal-affinity

chromatography (IMAC). Metal-chelating matrixes are loaded with transitional metal ions, such as cobalt (Co^{2+}), nickel (Ni^{2+}), copper (Cu^{2+}) and zinc (Zn^{2+}). The target protein is able to bind such matrices because of the interaction between the metal ion and the histidine residues. After purification, the protein quality needs to be assessed. Various techniques, such as SDS-PAGE, western blots, dynamic light scattering (DLS) and others can be used to determine the protein purity, homogeneity and stability. Afterwards, several techniques, such as X-ray crystallography, small angle X-ray scattering (SAXS), nuclear magnetic resonance (NMR) and circular dichroism (CD), may be used to obtain protein structure information. Usually, this information is complemented with specific functional assays to determine the activity of a pure protein in solution.

1.3 X-ray Crystallography

X-ray crystallography is the most used technique to determine three-dimensional structures of biological macromolecules, such as proteins, nucleic acids and viral particles. Structure determination requires single crystals of the macromolecule under study[3].

Crystallization is the most critical step in protein crystallography. Since it is not possible to know in advance in which conditions a specific protein will crystallize, many different conditions have to be tested, such as protein concentration, precipitation agents, temperature, pH, buffers, ionic strength and others. For this purpose crystallization screens are performed in order to test different conditions. Several commercial kits are available that include conditions in which other proteins have crystallized. Many laboratories use crystallization robots for crystallization screens, because it uses nanolitre volume drops, reducing the amount of protein needed to test an extensive number of conditions[3].

Different crystallization methods can be used, such as vapour diffusion, microbatch, microdialysis and free-interface diffusion[3]. The most used technique is the vapour diffusion that consists in mixing the precipitant and protein solutions in a droplet, which will be placed in an enclosed space with more concentrated precipitant solution (reservoir). In order to achieve equilibrium in this closed system, water transfer occurs from the drop to reservoir, which leads to an increase of protein and precipitant concentration in the drop (in diagram B, Figure 1.3, this process correspond to the passage from the pink to the blue region). The crystallization process is composed by two phases: nucleation and growth (Figure 1.3, diagram A). The nucleation process consists in the formation of molecule aggregates that will become crystals. When the protein and precipitant concentrations correspond to nucleation phase, protein nuclei are formed and free protein concentration decreases that correspond to the grow phase (in diagram B, this process corresponds to the passage from blue to the green region). The protein in a supersaturated state may lead to nucleation conditions, where small crystals appear or to an amorphous solid (precipitated protein). So, it is necessary to choose appropriate protein/precipitant concentrations, to allow transition from nucleation to growing phase where suitable crystals may form.

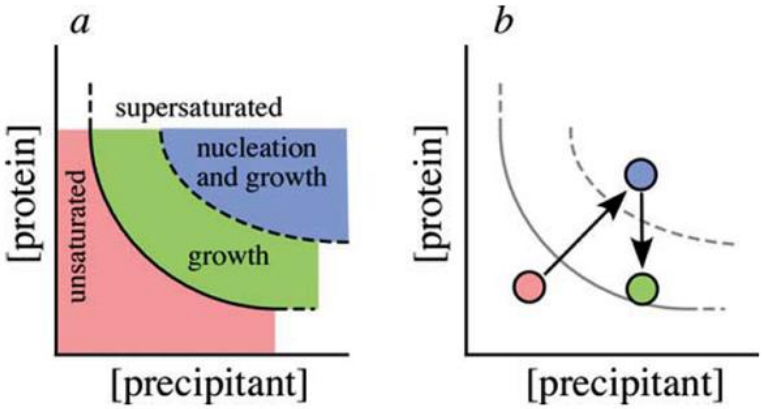

Figure 1.3 - Phase diagram. Pink region corresponds to unsaturated zone, green region to growth phase and blue to nucleation. The ideal strategy is to start in the blue region conditions of diagram and, when the nuclei are formed, pass to the green region that corresponds to the growing phase without any additional nucleation process[4]

The vapour diffusion technique is commonly applied using the sitting and hanging drop methods (Figure 1.4). Choosing the crystallization method is important, because the drop shape, and surface area and tension affects the equilibration rate and hence the number of nuclei[4].

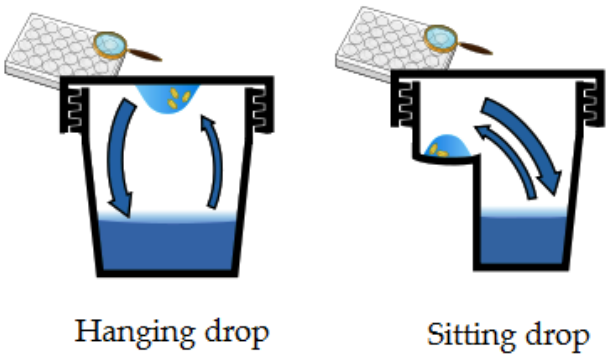

Figure 1.4 - Hanging and sitting drop crystallization methods. In the hanging drop method, a drop (usually 1-2 µl) of the protein solution is placed on an inverted cover slip and mixed with reservoir solution. In the sitting drop method, drop is placed on a pedestal that is separated from the reservoir[4].

The crystal, is then transferred from the crystallization drop and mounted on the diffractometer in order to assess diffraction quality (resolution limit and diffraction pattern). Crystals are frequently pooled using tiny loops. During mounting and data collection, the crystal must be surrounded by mother liquor, which will prevent it to dry out and degrade.

During X-ray exposure in data acquisition radicals are formed that start damaging the crystals and affecting its diffraction pattern. One of the strategies used to prevent crystal degradation is cryo-cooling the crystals. The crystal is maintained at cryogenic temperatures (~100K) in a gas steam of liquid nitrogen during data collection. Crystals usually have a solvent content of 30 to 50%, the water molecules have to be replaced by cryoprotectant solutions, which contain glycerol, glucose and other sugars otherwise, upon freezing, ice will be formed inside the crystal thus destroying its internal order[4].

X-ray diffraction

A crystal is a highly ordered microscopic structure that is composed by unit cells – smallest repeating units (figure 1.5).

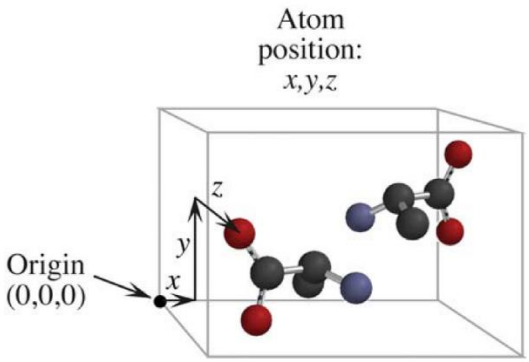

Figure 1.5 - Unit cell. The position of an atom in the unit cell can be defined by spatial coordinates x,y,z[5].

Unit cell contains the asymmetric unit – smallest fraction that can be rotated and translated. After obtaining crystals with the desired quality, these are mounted between an X-ray source and X-ray detector. Several X-ray sources are available: X-ray sealed tube, rotating anode X-ray tube, liquid anode and synchrotron X-ray source. Synchrotron radiation sources provide X-ray beams with high intensity, low divergence and a wide range of energy (figure 1.6). High intensity increases the signal of data acquisition, low divergence reduces overlap between adjacent reflections and wide range of energy enable the selection of specific wavelength[3].

Figure 1.6 - SOLEIL synchrotron. Saint-Aubin, France. (adapted from - http://enygf2015.org/program/technical-visits/saclay/)

When the X-ray beam hits the crystal, it is diffracted only it certain directions. Wave's constructive interference reinforces the wave amplitude, and destructive interferences cancel each other. Constructive interference gives rise to small dark spots with different intensities, *i.e.* reflections that constitute the

diffraction pattern. These are recorded on a detector (figure 1.7). Intensities of reflections are dependent from the photon number that a given crystal produces in the diffraction pattern. The position and intensities of reflections contain information needed to determine the molecular structure.

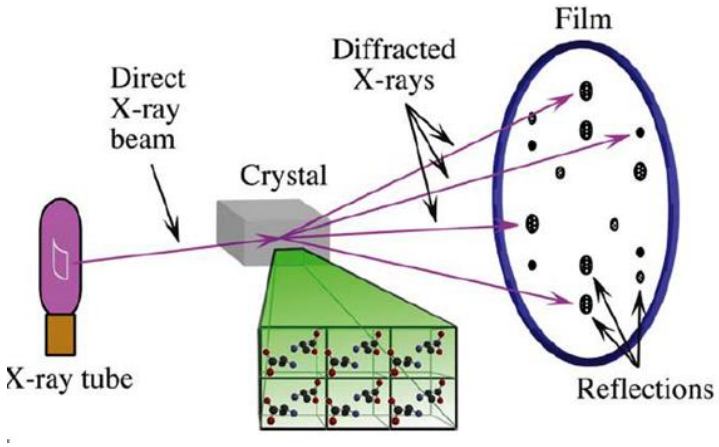

Figure 1.7 - X-ray data collection process. The crystals diffract the source beam into many discrete beams that produce distinct spots, i.e. reflections, on the detector[5].

An optical scanner transmits the spots positions and intensities in digital form to a computer for analysis. The usual strategy is to take few test images at 0 and 90º apart about an axis perpendicular to the X-ray beam and index these. Indexing consists in determining the reciprocal unit cell, namely parameters and the direction of its axis. This step allows determining the Bravais lattice that is used to define the crystalline arrangement. Bravais lattice allows the identification of Laue symmetry and space group, so a data collection strategy can be devised to obtain a data set as complete as possible using the least scanning width possible. After data collection, all images are integrated and intensities are extracted and scaled. Several diffraction patterns are obtained in different crystal orientations, which make it possible to construct a three-dimensional (3D) data set from bi-dimensional (2D) diffraction patterns.

The atomic position in a crystal structure is defined by three-dimensional coordinates (x,y,z) and individual diffracted beams are defined by Miller indices (h,k,l). Diffracted beams correspond to the sum of all atoms in unit cell that are characterized by an amplitude $|F_{hkl}|$ and a phase α_{hkl}. The amplitude is related with the number of atoms and the phase is derived from the atom position within the unit cell[3]. However, the phase information is lost during data collection (see below).

Structure factor, F_{hkl}, is a mathematical description of radiation scattering and is represented by $F_{hkl}=\sum_h\sum_k\sum_l |F_{hkl}| \exp(\alpha_{hkl})$. The structure factors $|F_{hkl}|$ are proportional to the square root of the intensity of diffracted beams, which are experimentally measured. The phase (α_{hkl}) is lost during the experiment and several methods can be used to estimate these lost phases. In order to calculate an electron density map that is defined by $\rho(x,y,z)=1/V \sum_h\sum_k\sum_l |F_{hkl}| \exp(i\alpha_{hkl}) \exp[-2\pi i(hx_j+ ky_j+ lz_j)]$ phase information is absolutely required[6].

Phase determination

Several direct and indirect methods can be chosen to overcome the phase problem. *Ab initio* is a direct method that requires one good-quality data set at atomic resolution (<1.2 Å). Indirect methods comprise molecular replacement, anomalous dispersion and isomorphous replacement. Molecular replacement is a method that is frequently used, in which a known structure is used to estimate the phases for an unknown structure. The model protein must have structural homology with the unknown protein, such as in the case of mutated proteins, complexes with ligands or structurally related proteins. When the structural similarity is unknown, molecular replacement is used in the base of the sequence identity, which must to be at least 30-40%. Anomalous dispersion is a method that is based on the heavy atoms ability to scatter X-rays anomalously near their absorption edges, which leads to changes in amplitudes and phases. The resulting changes might be used to estimate the lost phases. In this method one or more wavelengths might be used. Single anomalous dispersion (SAD) uses one wavelength that corresponds to the peak of the heavy atom. Multiple anomalous dispersion (MAD) use two or more wavelengths that include peak of the heavy atom and might include low energy remote, inflection point and/or high energy remote wavelengths. Isomorphous replacement method also requires heavy atoms that induce measurable changes in diffraction pattern. This changes, in relation to the "native" crystals, can be used to deduce the positions of the heavy atoms, which provide the estimates of the protein phase angles. Both methods might be used for metalloproteins, which contain one or more endogenous heavy atoms (e.g. iron (Fe), cupper (Cu), molybdenum (Mo), zinc (Zn) and nickel (Ni)). For the other proteins, crystals might be soaked with several metals, such as mercury (Hg), platinum (Pt), gold (Au), lead (Pb), silver (Ag), or protein incubated with them and then co-crystalized.

After phase estimation, an electron density map can calculated and an initial model built. The model must be refined against the experimental data in order to optimize the agreement between observed and calculated structure factors amplitudes $|F_{hkl}|^3$. .

Model validation

Several validation tools were developed to evaluate the quality of the experimental data, the refined model and agreements between them. Several statistical parameters must be reported, such as resolution, completeness, multiplicity, signal-to-noise ratio and merging R factor, which allows assessing the quality of the diffraction data.

Resolution (Å) determines the level of detail of an electron density maps. Completeness is the relation between measured crystallographic reflections and the total number of theoretically possible unique reflections. Completeness should be as high as possible (usually aiming at 100% completeness). Redundancy (or multiplicity) is defined to be the ratio between the total number of measured reflections and the number of unique reflections present. This calculation gives the average number of independent measurements for each reflection, so the average value is more precise. Signal-to-noise ratio ($I/\sigma(I)$)

indicates how many times above the noise the intensities were measured. R_{merge} value measures the agreement between the multiple independent observations of the same reflection: a good data set should have a low R_{merge} and high values might indicate problems with the data collection/processing. In the highest resolution shell, R_{merge} may be as high 30-40%, although values in order of 60-70% are still reasonable for high-symmetry space groups. R_{factor} or R_{work} measure the agreement between experimental structure factor amplitudes (F_{obs}) and calculated from the model (F_{calc}). R_{free} is the R_{factor} that use structure factor amplitudes from a "test set", which contains data not used in refinement, and is useful for cross validation.

Model geometry can be validated using different programs that analyze stereochemistry, local chemistry environments and others. Other properties, such as bond lengths, bond angles, φ ψ torsion angles, side chain torsion angles, peptide flips, clashes, Cβ derivations, asparagine/histidine/glutamine side chain flips, local environment profiles, Ramachandran and rotamer outliers, can also be checked[3].

Thesis objectives

The aims of this work were histidine kinases production followed by functional and structural studies. Production included four steps: cloning, expression, solubilization and purification. Solubilization and purification steps were critical steps in this work and it was not possible to obtain a pure protein for the functional and structural studies. So, it was decided to choose a soluble protein to proceed with structural characterization. Thus, several variants of thiosulfate dehydrogenase were selected for the structural studies.

Chapter 2

Production of membrane histidine kinases

2. Introduction

2.1 Two-component Systems and Bacterial Antibiotic Resistance

The two-component signal transduction systems (TCS) enable organisms to sense, respond and adapt to a wide range of environments or intracellular states, *e.g*. stresses and growth conditions. These signalling systems are integrated into a large variety of cellular signalling circuits, *e.g.* carbohydrate metabolism, quorum sensing, iron regulation, drug resistance and virulence[7].

These systems involve two components: histidine kinase (HK) and an associated response regulator (RR). TCS are found in all three kingdoms of Life: Bacteria, Archaea and Eukarya[8]. For the unicellular organisms HKs are related with nutrient sensing, chemoattractants, osmotic conditions and others. In Eukarya, HKs are found in plants and free-living organisms, such as yeast, fungi and protozoa and are involved in regulation of hormone-dependent processes. No HK genes are present in the *Caenorhabditis elegans*, *Drosophila melanogaster* and *Homo sapiens*, as these enzymes are absent in animal kingdom[9].

Structure and Function of Histidine Kinases

The HK family is extremely diversified, resulting from a simple combination of sensing, catalytic and auxiliary domains. HKs molecular weight usually range from <40 to >200 kDa and the structural architecture was adapted to the specific needs of signalling system[10]. A basic HK structure contains sensing domain, linker domain and kinase catalytic core (figure 2.1). When an environmental stimulus is detected directly or indirectly by the N-terminal sensing domain, it changes conformation. The sensing domain is connected to the kinase catalytic core by the linker domain, which is composed by several helices and cytoplasmic linker. The dynamic nature of these helices allows TCS to perceive any conformational changes of the sensing domain and to communicate these changes to the catalytic core[10]. Kinase catalytic core is composed by the catalytic and ATPase domain (CA domain) followed by the dimerization and histidine phosphotransferase domain (DHp). The CA domain is responsible for ATP binding and autophosphorylation of conserved histidine residue found in DHp domain. The phosphorylated histidine residue is the RR phosphodonor[7].

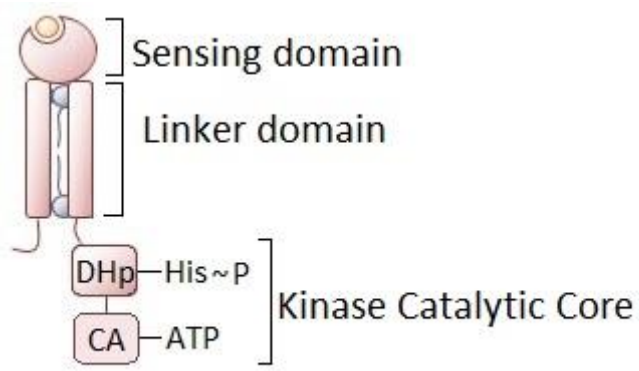

Figure 2.1 - Structure of basic Histidine Kinase. Basic HK structure is composed by Sensing domain, Linker domain and Kinase catalytic core. Catalytic core possesses DHp and CA domains. (Adapted from Pikyee Ma PhD thesis (2010)).

Despite of their diversity, HKs can be divided in two classes: orthodox and hybrid kinases. Most orthodox HKs function as periplasmic membrane receptors, while some are soluble cytoplasmic HKs. Soluble HKs can be regulated by intracellular stimuli and/or interaction with other proteins. The more elaborated kinases, hybrid kinases, contain multiple phosphodonor and phosphoacceptor sites. Their complexity allows the integration of different inputs and outputs into a signalling pathway. Hybrid kinases are rare in prokaryotes, whereas eukaryotic HKs are almost exclusively hybrid kinases[8].

Structure and Function of Response Regulator

A typical response regulator contains two domains: one conserved N-terminal regulatory domain and one variable C-terminal effector domain. The regulatory domain functions as an activator of the effector domain. Many different types of effector domains exist in nature, where the most prevalent are involved in DNA binding, and their response is coupled to changes in transcription[7,10].

TCS Architecture

Orthodox and Hybrid kinases are involved in different TCS types, the phosphotransfer and the phosphorelay systems, respectively[7,10] (figure 2.2). Phosphotransfer systems have a more simplistic design, which is composed by three phosphotransfer reactions:

1. Autophosphorylation: HK-His + ATP ↔ HK-His~P + ADP
2. Phosphotransfer: HK-His~P + RR-Asp ↔ HK-His + RR-Asp~P
3. Dephosphorylation: RR-Asp~P +H_2O ↔ RR-Asp + P_i

Phosphorelay systems are involved in more elaborated multiple phosphotransfer pathways and incorporate five phosphotransfer reactions:

1. Autophosphorylation: $HK\text{-}His_I + ATP \leftrightarrow HK\text{-}His_I\sim P + ADP$
2. Phosphotransfer I: $HK\text{-}His_I\sim P + Rec\text{-}Asp_1 \leftrightarrow HK\text{-}His_I + Rec\text{-}Asp_I\sim P$
3. Phosphotransfer II $HK\text{-}Asp_I\sim P + HPt\text{-}His_{II} \leftrightarrow Rec\text{-}Asp_I + HPt\text{-}His_{II}\sim P$
4. Phosphotransfer III: $HPt\text{-}His_{II}\sim P + RR\text{-}Asp_{II} \leftrightarrow HPt\text{-}His_{II} + RR\text{-}Asp_{II}\sim P$
5. Dephosphorylation: $RR\text{-}Asp_{II}\sim P + H_2O \leftrightarrow RR\text{-}Asp_{II} + Pi$

Autophosphorylation is a bimolecular reaction between homodimers, in which one HK monomer catalyses the phosphorylation in the second monomer[10].

In the phosphotransfer system, an HK autophosphorylates a conserved histidine residue and transfer phosphoryl group to a response regulator. In the phosphorelay system, the autophosphorylation is followed by phosphate transfer to an aspartate residue of an additional C-terminal receiver domain. Then, a phosphate is transferred to the histidine residue of an Histidine Phosphotransferase (HPt). HPt transfers a phosphoryl group to aspartate residue of the soluble response regulator[7,10].

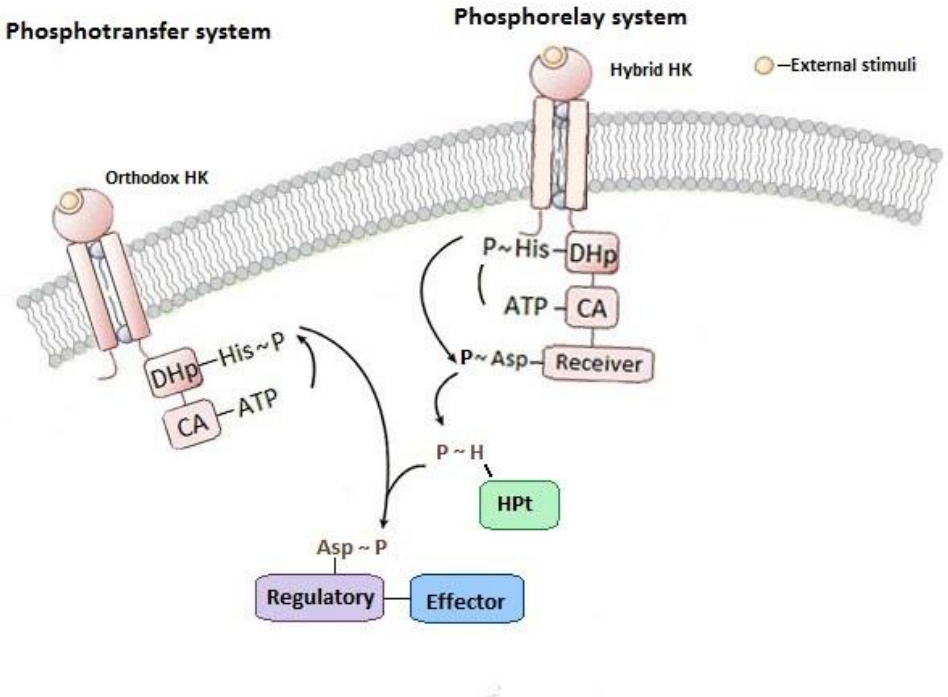

Figure 2.2 - **Phosphotransfer and Phosphorelay systems.** Phosphotransfer system mechanism: Orthodox HK CA domain transfers phosphate (P_i) to the conserved histidine residue of DHp domain. Then, P_i is transferred to the aspartate residue of RR regulatory domain and activation of effector domain occurs. Phosphorelay system mechanism: Hybrid HK CA domain transfers P_i to the conserved histidine residue of DHp domain. This process is followed by P_i transference to aspartate residue of receiver domain. Then, phosphate is transferred to HPt that transfers P_i to aspartate residue of RR regulatory domain. Regulatory domain phosphorylation activates the effector domain. (Adapted from Laub e Goulian, 2007)[7].

Two-component systems in prokaryotes

Bacteria can express a different number of TCS, *i.e.* ranging from 0 in *Mycoplasma genitalium* to 80 in *Synechocystis sp*[8]. The ability of bacteria to cause diseases is dependent on the expression of different cell wall-associated and secreted virulence factors. The expression of many virulence factors is controlled by several TCS, allowing bacteria to adhere and colonise host cells[11].

The *agr* operon is a global regulon that coordinates many critical virulence pathways. This operon consists of five genes (*agrB, agrD, agrC, agrA and hdl*) that encode two different transcripts (RNAII and RNAIII), whose synthesis is initiated by two different promoters (P2 and P3)[12]. Transcription at P2 results in the production of RNAII that encodes AgrB, AgrD, AgrC and AgrA. AgrC is a histidine kinase and AgrA is the respective response regulator. AgrD produces autoinducing peptides (AIP) that are secreted to the extracellular space by AgrB. AgrC recognizes the secreted AIP and, through phosphotransfer system, phosphorylates AgrA. This two-component system results in the production of RNAIII, an effector molecule that modulates transcription of numerous genes, *e.g*. several virulence factors[13].

The *agr* locus mediates ArlRS and SrrAB two-component systems. The ArlRS system, comprises a histidine kinase – ArlS, and a response regulator - ArlR, and is involved in the regulation of several virulence factors, such as protein A, α-toxins, β-haemolysin, coagulase, lipase, and others[12]. The staphylococcal respiratory response AB system (SrrAB) is involved in adaptation to anaerobic growth and in regulation of several virulence factors. SrrB is a histidine kinase and SrrA is the respective response regulator[14].

The NreABC is an oxygen sensing system that is involved in nitrogen regulation. Upon oxygen depletion, NreB (HK) autophosphorylation increases. NreB phosphorylates NreC (RR), which will bind to specific promoters of nitrate reductase and nitrite reductase operons and to *NarT* gene that is involved in nitrate and nitrite transport. NreA protein is also encoded by *nreABC* operon but its function is still unknown[15].

The TCS SaeR-SaeS is encoded by *SaeRS* locus and regulates the expression of genes that are important to initiate infection. SaeRS system is involved in the control of different virulence factors, such as α-toxins, β-haemolysin, coagulase and others. SaeRS system negatively affects AgrA expression, despite the fact that Agr systems acts upstream of Sae. This fact indicates that TCS systems interact with each other[14].

The glycopeptide resistance associated TCS GraS-GraR controls cationic antimicrobial peptides (CAMPs) resistance. CAMPs are small peptides (<100 amino acids) that exhibit a broad antimicrobial activity spectrum[16]. They are key components of host innate immune system that kill bacteria through membrane permeation, acting on DNA and RNA or by protein synthesis. GraSR system activates synthesis of specific enzymes that increase positive charges in bacterial cell surface leading to electrostatic repulsion of CAMPs[17].

In AlgB-KinB, the HK KinB controls several gene expressions involved in carbohydrate metabolism, quorum sensing, iron regulation, motility and virulence factors expression[18].

In SpaK-SpaR, the HK SpaK autophosphorylates in the presence of subtilin – an antimicrobial peptide antibiotic. SpaK phosphorylates SpaR, the RR, which will bind to specific promoters and induce a response[19]. SpaK-SpaR system from *Bacillus subtilis* regulates biosynthesis of subtilin, in order to inhibit growth of a broad range of pathogenic Gram-positive bacteria[20].

Some bacteria developed TCS systems that confer resistance to certain chemotherapeutics[7]. In another BceSR TCS, the BceS has the ability to sense bacitracin and the response regulator (BceR) positively regulates synthesis of transporters involved in bacitracin resistance process (BceAB and VraDE)[21,22].

In VanS-VanR TCS, VanS detects the presence of vancomycin and phosphorylates the respective response regulator. VanR (RR) induces expression of genes that contribute to vancomycin resistance[23,24].

Potential targets for antimicrobial therapy

Two component systems are potential targets for antimicrobial therapy, as they are used by pathogenic bacteria to control the expression of virulence factors[6]. Several reasons have led researchers to choose them as targets: TCS are common in bacteria but absent in mammalian organisms, general inhibitors of TCS would be broad-spectrum antibiotics while some could allow for a more selective inhibition.

Antimicrobial resistant infections

Antibiotic resistance has emerged as a global health problem in recent years. At least 50 000 deaths are reported each year in Europe due to antimicrobial resistant infections[21].

Staphylococcus aureus (*S. aureus*) is a gram-positive bacterium that commonly infects the skin, soft-tissues, bones, joints, lungs and central nervous system. This pathogen expresses many potential virulence factors, such as proteins that promote colonization on the host tissues, factors that inhibit the host phagocytosis, and toxins that damage the host tissues. *S. aureus* possesses 16 two-component signalling systems[21].

Different β-lactam antibiotics were developed against *S. aureus* infection. Cephalexin, dicloxacillin, vancomycin and methicillin are some of the developed antibiotics successively used against different bacterial infections. Despite their efficiency, different bacterial species have developed resistance to those antibiotics. Methicillin-resistant *S. aureus* (MRSA) is considered a multidrug resistant bacterium, because methicillin is a highly effective antibiotic against bacterial infections[25]. MRSA is one of the most common causes of healthcare-associated infections.

Clostridium difficile (*C. difficile*) is gram-positive spore forming bacterium that is transmitted by fecal-oral route. Colonization by *C .difficile* stimulates the release of toxins that induce inflammation and damage of tight junction between epithelial cells that is followed by loss of the epithelial barrier

function. In addition, toxins induce apoptosis and necrosis of epithelial cells that contribute to inflammatory response. This bacterium causes life-threatening diarrhoea and mostly occurs in hospitalized or recently hospitalized patients[26]. Although antibiotic resistance is not yet a problem for these bacteria, *C. difficile* spreads rapidly because it is naturally resistant to many drugs that are used to treat other infections.

3. Materials and methods

3.1 DNA cloning of HK targets in pET-52b(+) (Novagen)

General methodologies and reagents

For all solutions and buffer preparations pH was measured using a Sartorius pH meter. Analytic balance (KERN) and Precision balance (KERN) were used. All protocols that include cell manipulation were performed in Laminar flow chamber, Bio II model –A (Telstar). SDS-PAGE was performed using *Mini-PROTEAN*©® Tetra cell assembly (Bio-Rad). All restrictions enzymes and buffers were from New England Biolabs, unless otherwise stated. All absorbance measures at 600 nm (OD_{600}) were measured in a spectrophotometer (Ultrospect 10, Amersham) using non-inoculated media as blank.

Reagents

- Acetic acid (glacial) 95,5-100,5% (Carlo Erba)
- Acrylamide /Bis-Acrylamide (30%/0,8%w/v) (Alfa Aesar)
- Agar-Agar (ROTH)
- Agarose (CAMBREX)
- ATP (Jene Bioscience)
- Bovine serum albumin - BSA 100x – 10 mg/µl (Biolabs)
- Carbenicilin (ROTH)
- Cloramphenicol (ROTH)
- Cymal-4 (Affymetrix)
- C-Hega 8 (Affymetrix)
- DDM (Glycon)
- DM (Glycon)
- dNTP (X)
- DNA polymerase, Phusion High-Fidelity (Thermo Scientific)
- DNA purification Kit, Zymocean Gel DNA Recovery Kit (Zymo Research´s)
- D-Saccharose (ROTH)
- EDTA (Calbiochem)
- Ethanol absolute (ROTH)
- Gene Ruller 1Kb 0,5 µg/ml (Thermo Sientific)Sodium Chloride ≥99,5% D=2,17 g/ml (ROTH)
- Fos-Choline (Anatrace)
- Glycerine (ROTH)
- His-Probe HRP (Thermo scientific)
- Imidazole (Merk)
- IPTG (ROTH)

- LDAO (Anatrace)
- Lysozime (Sigma)
- Mega-8 (Anatrace)
- Methanol (ROTH)
- NEB buffer 4 (1x) – 50 mM Potassium Acetate, 20 mM Tris-acetate, 10 mM Magnesium Acetate, 1 mM DTT pH 7.9 ´
- NotI-HF (Biolabs)
- N-Octyl-β-D-Glucopyranoside (Affimetrix)
- N-Octyl- β-D-Thioglucopyranoside (Glycon)
- Page Ruller^{TM} Plus Prestained Protein (Thermo scientific)
- Phusion DNA polymerase (Thermo scientific)
- Phusion HF Buffer (5x) 75 mM $MgCl_2$ (Thermo scientific)
- Plasmid isolation kit, ZR BAC DNA Miniprep Kit (Zymo Research´s)
- Plasmid purification kit, Wizard SV Gel and PCR Clean-Up System (Promega)
- pET-52b(+) (Novagen)
- Potassium Chloride ≥ 99,0% D=1,98 g/ml (Sigma)
- Primary antibody: Anti-His IgG (Fc Specific)-Alkaine Phosphatase antibody) (GE Healthcare)
- SacI-HF (Biolabs)
- Sodium dodecyl sulphate SDS (ROTH)
- Secondary antibody: Anti-Mouse IgG (Fc Specific)-Alkaline Phosphatase conjugated (Sigma Aldrich)
- Supersignal West Pico Chemiluminescent Substrate (BioRAD)
- SYBR Safe (Invitrogen)
- Triptone (ROTH)
- Tris base ≥ 99,0% (Sigma)
- Triton X-100 (ROTH)
- Tween-20 (Merk)
- T4 DNA Ligase 1U/µl (Invitrogen)
- Vircon (Rely^{+}ON)
- Western Blot reagent – BCIP/NBT Liquid Substate System (Sigma Aldrich)
- XmaI (Biolabs)
- Yeast extract (ROTH)

Genomic DNA and targets for *Staphylococcus aureus* ad *Clostridium difficile*

Genomic DNA used for cloning was *Staphylococcus aureus COL* strain and *Clostridium difficile 630* strain. They were kindly provided by Bacterial Cell Biology laboratory (ITQB) and Microbiology Development laboratory (ITQB), respectively.

The selected *S. aureus* targets were *SrrB* gene (KEGG cede SACOL1534, 1752 bp), *AgrC* gene (KEGG cede SACOL2025, 1293 bp), *KinB* gene (KEGG cede SACOL0020, 1827 bp), *PhoR* gene (KEGG cede SACOL1739, 1665 bp), *GraS* gene (KEGG cede SACOL0717, 1041 bp), *Unk* gene (KEGG cede SACOL0202, 1557 bp), (attachment A table 5.1). The *C. difficile* targets were *VanS* gene (KEGG cede CD630_16250, 1143 bp), *SrrB* gene (KEGG cede CD630_14650, 1254 bp), *Unk* gene (KEGG cede CD630_16890, 1233 bp), *GraS* gene (KEGG cede CD630_15790, 1857 bp), *KinB* gene (KEGG cede CD630_17830, 1434 bp), *SaeS* gene (KEGG cede CD630_35990, 1194 bp), *BceS* gene (KEGG cede CD630_15300, 1005 bp), *AgrC* gene (KEGG cede CD630_06100, 1338 bp), (Appendix A, table 5.1).

PCR chain reaction of the *S. aureus* and *C. difficile* target genes

All targets were amplified by polymerase chain reaction (PCR) (MyCycler ™ Thermal Cycler, BioRAD) and cloned into pET-52b(+), which contains a N-terminal Strep•Tag II (cleavable by HRV 3C) and a 10x Histidine affinity tag in the C-terminus (cleavable by Thrombin). The reaction mix include 0.2 μM forward and reverse primers (Appendix A, table 3.2 and 3.3), 100 ng of template DNA, 200 μM dNTPs, 1x Phusion HF buffer and 2.5 U DNA polymerase. The PCR protocol for each target is described in Table 3.1. The size of the amplified genes were confirmed by 0.8% (w/v) agarose gel electrophoresis (Mini Electrophoresis system - VWR), visualized by Transiluminator Safe Imager ™ 2.0 (Invitrogen) and purified using Zymoclean Gel DNA Recovery Kit, (Appendix B, protocol nr 1).

Table 3.1 - PCR conditions used for *SrrB* and *AgrC* genes of MRSA

	Steps	Temperature	Time
1	**Initial denaturation**	98ºC	30 sec
2	**Amplification (30 cycles)**		
	Denaturation	98ºC	10 sec
	Annealing	65ºC	30 sec
	Extension	72ºC	1 min 30 sec
3	**Final extension**	72ºC	7 min
	Hold	4ºC	

Double digestion of pET-52b(+), MRSA *SrrB* and *AgrC*

The pET52-b(+) vector, and *SrrB* and *AgrC S. aureus* targets were double digested with specific restriction enzymes – XmaI and SacI-HF. *SrrB* double digestion reaction was prepared using 635 ng DNA insertion, 3 µg BSA, 10 U XmaI and 20 U SacI-HF in 1x NEB buffer 4. *AgrC* double digestion reaction comprised 419.25 ng DNA insertion, 3 µg BSA 10 U XmaI and 20 U SacI-HF in 1x NEB buffer 4. Double digestion reaction mix of pET-52b(+) comprises 1078,4 ng vector DNA, 3 µg BSA, 10 U XmaI and 20 U SacI-HF in 1x NEB buffer 4.

Samples were incubated in a dry bath (Labnet) for 3h at 37ºC, analyzed by 0.8% (w/v) agarose gel electrophoresis and purified with Zymoclean Gel DNA Recovery Kit. Samples were stored at -20ºC.

Ligation reaction (*S. aureus SrrB* and *AgrC* in pet52b(+))

For *SrrB* ligation reaction 191.5 ng digested DNA insert, 12.3 ng digested vector DNA, 1 U T4 ligase at 1x NEB buffer 4 were mixed. For AgrC ligation reaction 418 ng digested DNA insert, 12,3 digested vector DNA, 1 U T4 ligase at 1x NEB buffer 4 were mixed. Samples were incubated for 3h at 37ºC and transformed into *E.coli* XL-10.

Competent cell production protocol

One colony of *E. coli* was selected from the LB agar plate and was inoculated into 10 ml LB broth medium and incubated overnight at 37ºC in the shaker (MaxQ 4000 Barnstead lab-line) at 180 rpm. The pre-inoculum was added to 400 ml of fresh LB medium, incubated at 37ºC, with 180 rpm, until OD_{600} reached 0.6. Cell culture was placed on ice for 15 min and was centrifuged (JA-14 rotor, Allegra X-12R) at 4000 rpm for 8 min at 4ºC. Next, the pellet was gently ressuspended in 100 ml of ice-cold 100 mM potassium chloride (KCl), 50 mM manganese (II) chloride ($MnCl_2$), 30 mM Potassium acetate (CH_3COOK), 10 mM calcium chloride ($CaCl_2$) and 15% glycerol and centrifuged as before. After the second centrifugation, the pellet was gently ressuspended in 20 ml of ice-cold 10 mM 3-(*N*-morpholino) propanesulfonic acid (MOPS), 10 mM potassium chloride (KCl), 75 mM calcium chloride ($CaCl_2$) and 15% glycerol. Aliquots of 30x200 µl and 10x100 µl were prepared in sterile eppendorf tubes, flash frozen in liquid nitrogen and stored at -80ºC.

Cell Transformation

25 ng of previously prepared DNA (vector with putative insert) were mixed with 100 µl of several *E.coli* XL-10 competent cells and incubated on ice for 15 min. Samples were placed in 42ºC dry bath to perform heat shock, following an incubation on ice for 2 min. 500 µl of LB broth medium were added to the reaction and sample was incubated, to allow cell recovery. After 1 h at 37ºC in shaker at 180 rpm incubation, samples were centrifuged (Eppendorf 5424R) at 4000 rpm for 1 min and the precipitate was

recovered. 150 µl of supernatant were spread on a LB agar plate with a sterilized glass rod supplemented with 100 µg/mL carbenicillin and incubated overnight at 37ºC.

Sequencing

Isolated colonies were chosen for each target and inoculated into 5 ml of LB broth medium with 100 µg/mL of carbenicillin. After 16 h at 37ºC in shaker at 190 rpm, all samples were purified with Zymoclean Gel DNA Recovery Kit. Obtained DNA was cleaved with SacI-HF and XmaI restriction enzymes and analysed by agarose gel (Appendix B, protocol nr 1). Positive colonies were selected and forward and reverse sequencing reaction were performed with 20 µl of plasmid DNA, with a concentration of 50-100 ng/µl. For forward and reverse sequencing reaction, T7 forward (TAATACGACTCACTATAGGG) and T7 reverse (CTAGTTATTGCTCAGCGG) primer were used. STAB VIDA sequencing service was used.

3.2 Expression, solubilization and purification

Expression tests protocols

Plasmids pET52b(+)-SrrB and pET52b(+)-AgrC were transformed into *E.coli* C43, BL21, BL21 Star and BL21 Gold, using the same protocol described above. One colony of each transformed cell type were inoculated in 50 ml LB broth medium with 100 µg/mL of carbenicillin and incubated overnight at 37ºC at 180 rpm. 500 µl of overnight culture was transferred to 10 ml of media with 1 mg/l carbenicillin. Different media were tested - LB broth, TB, 2xYT and autoinduction medium. Cells were induced with 0.5 mM Isopropyl-β-D-thiogalactopyranoside (IPTG) when samples reached a specific OD_{600} (values depended on the media, TB: $OD_{600} \geq 1$; LB: OD_{600}=0.4-0.6; 2xYT: OD_{600}=1; M9: OD_{600}=0.7). Each sample was further incubated at different temperatures (37ºC, 28ºC and 21ºC) during different post-induction times (4h, 6h and overnight) in order to find the best expression conditions. Samples were centrifuged for 20 min at 8000 rpm (Eppendorf 5424R), precipitate was recovered and 40 samples were stored at -80ºC.

Extracting protein from bacteria

Membrane preparation using water lysis protocol (for small scale)

This protocol is used to disrupt the cell membrane in samples of small volumes, in the range of 1 – 5 ml. Cell pellet, obtained previously, was ressuspended in 1 ml 0.2 M Tris-HCl pH =8 and incubated for 20 min at room temperature. Then, 485 µl of S1 solution (1 M Sucrose, 0.2 M Tris-HCl pH 8, 1 mM EDTA) was added, and 6.5 µl of S2 (10 mg/ml lysozyme, 0.2 M Tris-HCl pH 8) was added after 90 s and the sample was incubated for 30 s. Immediately after, 960 µl of dH_2O were added and incubated for 20 min at room temperature. Sample was centrifuged (Allegra X-12R, JA-14 rotor) at 14 000 rpm for 20 min at 4ºC. Supernatant was discarded (periplasmic fraction) and pellet was homogenised in 2 ml of

dH_2O and incubated for 30 min at room temperature. The sample was centrifuged as before and the supernatant was discarded (cytoplasmic fraction). After that, pellet was homogenised in 1 ml of S3 (10 mM Tris-HCl pH 8, 5% glycerol, 1 mM mercaptoethanol).

The results were evaluated by SDS-PAGE and Western Blot (ChemiDoc MP Imager -BioRAD) or In Vision (Scanner Fujifilm fla 5100) techniques (Appendix B, protocol nr 2, 3 and 4).

Membrane preparation using cell disrupter protocol (for large scale)

In alternative to the above, this protocol can be used to disrupt sample of larger volumes, e.g. around 60 ml. Once the cell disruptor is refrigerated, two washing steps were performed, the first with 100 ml of dH_2O and the second with 100 ml of buffer, both at 10 kPsi. After that, homogenised cells were passed twice at 27 kPsi and stored on ice. Further washing steps included the passage of 100 ml buffer, followed by 100 ml 1:100 diluted Virkon detergent (1 tablet was dissolved in 500 ml), 100 ml 70% ethanol, 100 ml 20% ethanol and finally 100-200 ml of dH_2O addition.

Sample was centrifuged (JA-10/JA-14 rotor, Allegra X-12R) at 8 000 rpm for 20 min at 4ºC and the pellet was discarded. The supernatant was centrifuged (45Ti/70Ti rotor, Ultracentrifuge XL-10) at 40000 rpm for 2 h at 4ºC. The pellet, containing the cell membranes, was stored at -80ºC.

The results were evaluated by SDS-PAGE and Western Blot (ChemiDoc MP Imager -BioRAD) or In Vision (Scanner Fujifilm fla 5100) techniques (Appendix B, protocol nr 2, 3 and 4).

Solubilization tests

Solubilization of total cell membranes (10 mg/ml) was tested at different conditions. Three detergents, DM, DDM and Triton X-100 at 1.5% (w/v) were combined with various salt concentrations, NaCl at 250, 500 and 750 mM in buffer 50 mM Tris-HCl pH=7.5. Solubilization of 2 ml samples was performed during 2 h at 4ºC and, after, the samples were centrifuged (TLA-55 rotor, ultracentrifuge Optima TLX) 30 min at 40 000 rpm at 4ºC. Pellet was homogenised with 2 ml of 50 mM Tris-HCl pH 7.5. The presence of target protein in the pellet and soluble fraction were evaluated by SDS-PAGE and Western Blot (ChemiDoc MP Imager -BioRAD) or In Vision (Scanner Fujifilm fla 5100) techniques (Appendix B, protocol nr 2, 3 and 4).

Inner and Outer Membrane separation by Sucrose Gradient

Solutions with different sucrose concentrations (25, 30, 35, 40, 45, 50 and 55%) were prepared in 20 mM Tris-HCl, 50 mM EDTA, pH 7.5. 10 ml of each sucrose solution was slowly added to the centrifuge tube (figure 3.1). Total membrane was dissolved in 4 ml of 25% sucrose solution and added, slowly, to the top of the centrifuge tube. Samples were ultracentrifuged at 38 000 rpm at 4ºC overnight using a 45Ti rotor. As expected, the total membrane sample was separated into a white layer (outer membrane fraction at 50-55% sucrose) and yellow layer (inner membrane fraction at 35-40% sucrose).

After extracting separately the outer and inner membrane fraction, 20 mM Tris-HCl pH 7.5 was added to each sample in order to decrease EDTA traces. Samples were centrifuged for 2 h at 41 000 rpm at 4ºC using 70Ti rotor. The results were evaluated by SDS-PAGE and Western Blot (ChemiDoc MP Imager -BioRAD) or In Vision (Scanner Fujifilm fla 5100) techniques (Appendix B, protocol nr 2, 3 and 4).

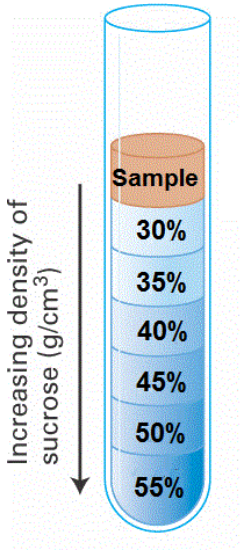

Figure 3.1 - Scheme of sucrose solutions addition in the sucrose gradient

Sucrose gradient solution: x% sucrose, 20 mM Tris-HCl, 50 mM EDTA pH=7.5 (x=55, 50, 45, 40, 35, 30, 25%)

Purification

Protein purification by HPLC

A 1 ml Nickel-nitriloacetic acid (Ni-NTA) column (GE Healthcare) was mounted on a AKTA purifier (GE Healthcare). The column was washed with water at a flow rate of 1 ml/min and equilibrated with binding buffer (specific to each experiment, e.g. for BceS purification 0.02% DDM and 250 mM NaCl in 50 mM Tris-HCl pH=7.5 was used). The previously solubilized membrane was introduced in the column (flow rate of 0.3 ml/min). Each sample flow-through was reserved. A washing step was performed by passing 5CV of binding buffer in the column. Protein was eluted with imidazole gradient from 0 to 1M imidazole. Finally, the column was washed with water and 20% ethanol. All samples were stored at 4ºC until processing. The results were evaluated by SDS-PAGE and Western Blot (ChemiDoc MP Imager -BioRAD) or In Vision (Scanner Fujifilm fla 5100) techniques (Appendix B, protocol nr 2, 3 and 4).

Batch purification using magnetic beads (His Mag Sepharose Ni)

An Eppendorf containing magnetic beads was placed in the magnetic rack (MagRack 6) to remove the storage solution. Next, equilibration step was performed by resuspending the beads in 500 µl of binding buffer and subsequent removing. Then 1 ml of solubilized total membrane was added, ressuspended and incubated for 30 min with slow mixing at room temperature. After that, the supernatant was removed and several washing steps were realised. Washing step consist in addition of 500 µl of binding buffer, mixing and subsequent removing. Than elution step were performed using elution buffer (50 mM Tris (Sigma) pH=7.5, 250/500 mM NaCl, 1 M Imidazole, 0.02% DM or DDM) with several imidazole concentrations (20, 40, 60, 80, 500 mM or 20, 100, 250, 500 and 1 000 mM imidazole). 500 µl of each elution buffer was added, mixed, removed and stored to evaluate by SDS-PAGE and Western Blot (ChemiDoc MP Imager - BioRAD) or In Vision (Scanner Fujifilm fla 5100) techniques (Appendix B, protocol nr 2, 3 and 4).

Batch purification

2 ml of Ni Sepharose 6 Fast Flow slurry (GE Healthcare) were placed into a 15 ml falcon tube and sedimented by centrifugation (Eppendorf 5424R) at 500 *xg* for 5 min at 4ºC. Supernatant was discarded, replaced with distilled water and centrifuged at 500 *xg* for 5 min at 4ºC. This step was repeated 2 times. After washing, the matrix were ressuspended in a specific binding buffer for equilibration. After centrifugation, total membrane fraction previously solubilized was added and incubated at 4ºC overnight under mild agitation.

Sample was transferred to the econo-pac column (BioRAD) and flow-through was collected. Column was washed with 1 ml binding buffer. The protein was eluted by passing 1 ml of solution with increasing imidazole concentration (20, 100, 250, 500 and 1000 mM, respectively). The results were evaluated by SDS-PAGE and Western Blot (ChemiDoc MP Imager -BioRAD) or In Vision (Scanner Fujifilm fla 5100) techniques (Appendix B, protocol nr 2, 3 and 4).

4. Results and Discussion

4.1 DNA cloning of HK targets in pET-52b(+)

Histidine kinases are complex and difficult proteins to study. To increase our chances of success, several HK targets were cloned in pET-52b(+).

Cloning of *AgrC* and *SrrB* from *S. aureus*

The first targets were *AgrC* and *SrrB* HK´s. PCR was performed has described in Materials and Methods, and the results were analysed by 0.8% agarose gel. Figure 4.1 shows single bands in the gel, indicative of the high specificity from the designed primers. *AgrC* gene size is 1250 bp, and from the band location in the agarose gel (between 1500 bp and 1000 bp of DNA ladder), we concluded that the observed band corresponds to *AgrC* target. The same conclusion was obtained for the *SrrB* gene, 1749 bp in size, the band is located between 2000 bp and 1500 bp. The chosen targets were successfully amplified.

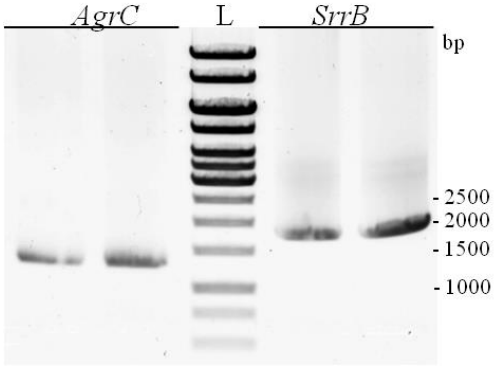

Figure 4.1 - 0.8% agarose gel. PCR results for *AgrC* and *SrrB* genes. Lanes: L - GeneRulerTM 1 kb DNA ladder; all bands correspond to 50 µl of putative amplified *AgrC* and *SrrB* genes.

DNA was extracted from the agarose gel and purified using a DNA gel purification kit (Zymo Research). The concentration measured was 21.3 ng/µl for *AgrC* and 16.4 ng/µl for *SrrB* target. We believe those low DNA concentrations was due to the little amount of *S. aureus* genomic DNA used in the PCR, approximately 15 ng was used. In order to increase DNA concentration of amplified targets, a 2nd PCR was performed, in which 1st PCR products were used as templates. PCR results were analysed by agarose gel (figure 4.2), where is observable an increase of band intensity, indicative of increased amplification of DNA target.

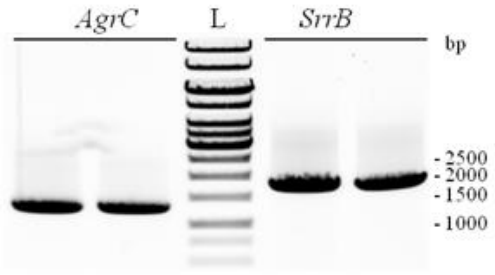

Figure 4.2 - 0.8% agarose gel. PCR results for *AgrC* and *SrrB* genes. Lanes: L - GeneRulerTM 1 kb DNA ladder; all bands correspond to 50 µl of putative amplified *AgrC* and *SrrB* genes.

Despite a stronger intensity in the gel, DNA concentration was lower than previous experiments. *AgrC* concentration was 10.5 ng/µl and *SrrB* was 10 ng/µl. Operator error or kit contamination are possible explanations for such lower concentration.

The following step consisted in the double digestion experiment of amplified DNA targets using the PCR product with higher concentration (*AgrC* 21.3 ng/µl and *SrrB* 16.4 ng/µl). Results are presented in a figure 4.3. *AgrC* and *SrrB* bands were extracted from the gel and DNA was purified. At the same time, pET-52b(+) double digestion was performed (figure 4.4).

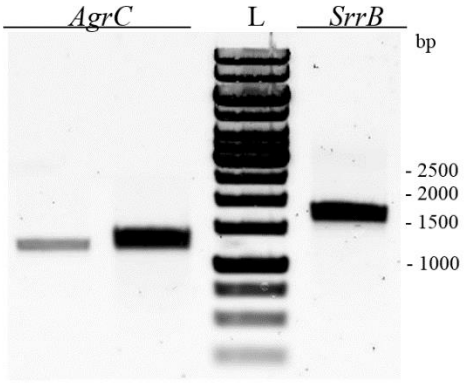

Figure 4.3 - 0.8% agarose gel. Double digestion results for *AgrC* and *SrrB* genes. Lanes: L - GeneRulerTM 1 kb DNA ladder; all bands correspond to 30 µl of putative amplified *AgrC* and *SrrB* genes.

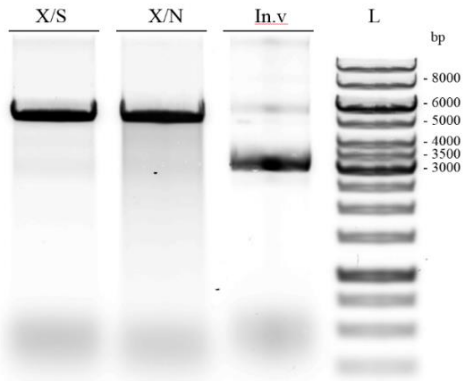

Figure 4.4 - 0.8% agarose gel. Results of pET-52b(+) double digestion. Lanes: X/S –digestion with XmaI and SacI-HF; X/N –digestion with XmaI and NorI; In.v – intact pET-52b(+), L- GeneRulerTM 1 kb DNA ladder.

pET-52b(+) digested with XmaI and SacI-HF, or with XmaI and NotI-HF, are located between 6 kbp and 5 kbp. Intact pET-52b(+) is located around 3 kbp band. When digested, the vector is an opened ring, and its conformation interferes with running through the agarose gel – running is slower comparing to a non-digested vector. Intact vector exhibits a more compact conformation (supercoiled) and runs faster. Such characteristic explains the difference around 3 kbp between digested (Figure 4.4, lanes X/S and X/N) and non-digested vector bands (Figure 4.4, lane In.V). The double digestion was successful.

After double digestion procedure, we proceeded to ligation reaction and transformation into *E.coli* XL-10. No colonies have appeared which means that cloning was unsuccessful. The problem may reside in the double digestion reaction – one of the enzymes might be nonfunctional. In order to check this hypothesis, pET-52b(+) was digested with XmaI and SacI-HF separately.

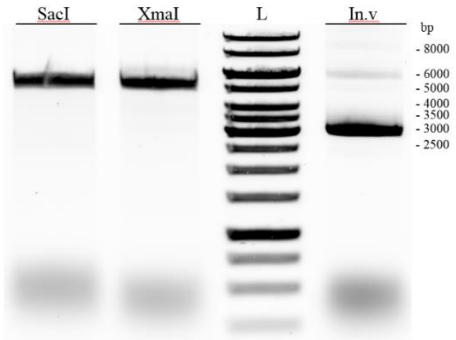

Figure 4.5 - 0.8% agarose gel. Results of pET-52b(+) double digestion. Lanes: SacI –digestion with SacI-HF; XmaI –digestion with XmaI; In.v – intact pET-52b(+), L- GeneRulerTM 1 kb DNA ladder.

Digested pET-52b(+) is located around 6 kbp and intact vector around 3 kbp. SacI-HF and XmaI samples present a band around 6 kbp (Figure 4.5), indicating a successful double digestion and good activity of the enzymes. The other reason might be contamination of the gel extraction kit, which also explains the low concentration of extracted DNA from stronger bands. All experiments were repeated with a new

Gel DNA Recovery Kit (Zymo Research). PCR experiments were performed using product from 1st PCR (Figure 4.6).

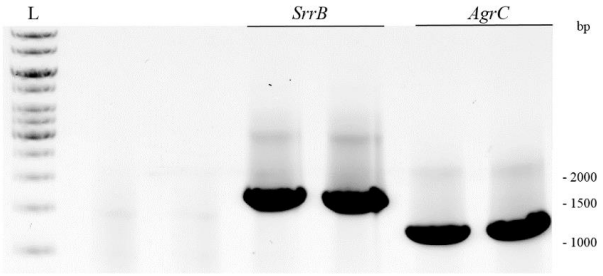

Figure 4.6 - 0.8% agarose gel. PCR results for *AgrC* and *SrrB* genes. Lanes: L - GeneRulerTM 1 kb DNA ladder; all bands correspond to 50 µl of putative amplified *AgrC* and *SrrB* genes.

DNA was extracted from the agarose gel and concentration was measured. *AgrC* concentration was 64.5 ng/µl and *SrrB* was 97.7 ng/µl, confirming that previous cloning experiments failed because of contaminated DNA gel extraction kit. Polymerase chain reaction was followed by double digestion (figure 4.7). After gel extraction, concentration of *SrrB* and *AgrC* targets were 38.8 ng/µl and 34.5 ng/µl, respectively.

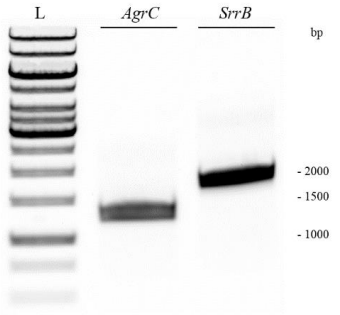

Figure 4.7 - 0.8% agarose gel. Double digestion results for *AgrC* and *SrrB* genes. Lanes: L - GeneRulerTM 1 kb DNA ladder; all bands correspond to 30 µl of putative amplified *AgrC* and *SrrB* genes.

Following ligation and transformation into the *E.coli* XL-10 and overnight growing, SrrB and AgrC colonies appeared. Four isolated colonies for each target were chosen. In order to confirm that chosen colonies hold pET-52b(+) with target genes, vector was isolated from each colony and digested with XmaI and SacI-HF enzymes. After *SrrB* result analysis (Figure 4.8), two positive colonies were selected: $SrrB_1$ and $SrrB_3$.

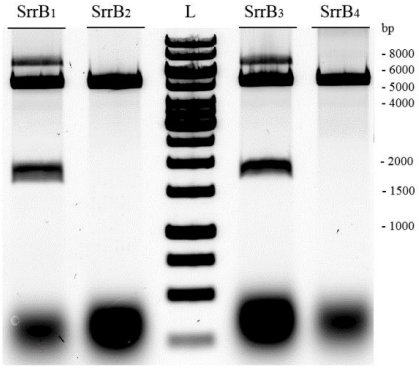

Figure 4.8 - 0.8% agarose gel. Results of *SrrB* vector double digestion. Lanes: $SrrB_1$ – colony nº1, $SrrB_2$ – colony nº2, L- GeneRulerTM 1 kb DNA ladder, $SrrB_3$ – colony nº3, $SrrB_4$ – colony nº4.

Positive colonies hold a band around 2 000 bp that corresponds to *SrrB* insert (1749 bp). All other two bands will correspond to pET-52b(+) (located between 8 kbp and 5 kbp ladder bands). Negative colonies ($SrrB_2$ and $SrrB_4$) present just one band (between 6 kbp and 5 kbp) that corresponds to the vector without DNA insert. In figure 4.9, *AgrC* double digestion results are presented. $AgrC_1$ and $AgrC_3$ colonies were selected, because they present a band located between 1500 bp and 1000 bp that corresponds to *AgrC* insert (1250 bp).

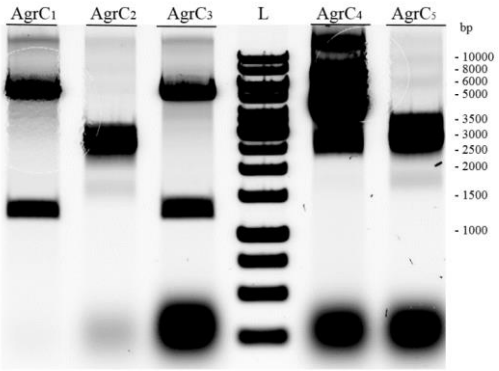

Figure 4.9 - 0.8% agarose gel. Results of *AgrC* vector double digestion. Lanes: AgrC1 – colony nº1, AgrC2 – colony nº2, AgrC3 – colony nº3, L- GeneRulerTM 1 kb DNA ladder, AgrC4 – colony nº4, AgrC5 – colony nº5.

Positive colonies $SrrB_1$, $SrrB_3$, $AgrC_1$ and $AgrC_3$ were sent to Stabvida (Monte da Caparica, Portugal) for sequencing. After sequencing results analysis $SrrB_3$ and $AgrC_1$ colonies were chosen for the following studies.

Cloning of other *S. aureus* and *C. difficile* targets

More HK targets were chosen for the cloning into pET-52b(+), *S. aureus targets* were: *KinB* (1827 bp), *GraS* (1041 bp), *PhoR* (1665 bp) and *Unk* (1557 bp); and *C. difficile* targets were: *SaeS* (1194 bp), *BceS* (1005 bp), *AgrC* (1338 bp), *SrrB* (1254 bp), *KinB* (1434 bp), *GraS* (1857 bp), *Unk* (1233 bp) and *VanS* (1143 bp). 1st PCR results are presented in the figure 4.10 and 4.11.

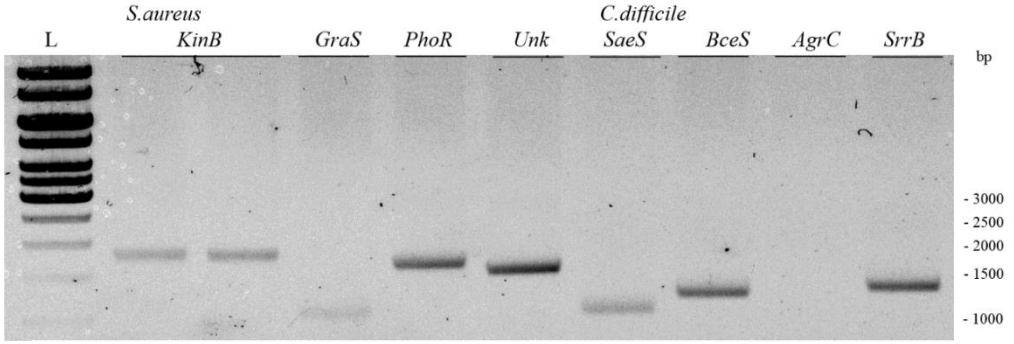

Figure 4.10 - 0.8% agarose gel. PCR results of *S. aureus (KinB, GraS, PhoR* and *Unk)* and *C. difficile (SaeS, BceS, AgrC* and *SrrB)* genes. Lanes: L- GeneRuler™ 1 kb DNA ladder; all band correspond to 50 µl of putative amplified genes.

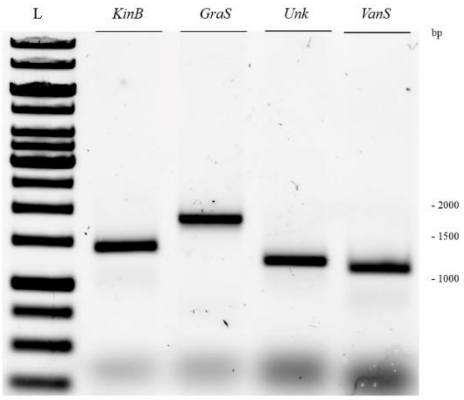

Figure 4.11 - 0.8% agarose gel. PCR results of *C. difficile (KinB, GraS, Unk* and *VanS)* genes. Lanes: L- GeneRuler™ 1 kb DNA ladder; all bands correspond to 50 µl of putative amplified genes.

All samples contain one single band, except *C. difficile AgrC* sample (Figure 4.10), where amplification was unsuccessful (no band). For targets with positive amplified genes, bands were extracted from gel as previously described and concentrations were measured. Concentrations of *S. aureus* target gene were [*KinB*]= 177.5 ng/µl, [*GraS*]=13.2 ng/µl, [*PhoR*]=74.4 ng/µl and [*Unk*]=56.9 ng/µl and concentrations of *C. difficile* targets were [*SaeS*]= 96.6 ng/µl, [*BceS*]=19.5 ng/µl, [*SrrB*]=65.2 ng/µl, [*KinB*]=93.3 ng/µl; [*GraS*]=4.8 ng/µl, [*Unk*]=36.4 ng/µl and [*VanS*]=107.7 ng/µl. Some targets, namely *S. aureus GraS* and *C. difficile BceS* and *GraS* present low DNA concentrations. *C. difficile AgrC* target didn't present any band, which means that problem might be in the primers or PCR protocol conditions; however, due to

the great number of selected targets we did not proceed with this one. In order to increase DNA concentration, a 2nd PCR was carried out using the amplified genes from 1st PCR as templates. Results were analysed by agarose gel (figure 4.12).

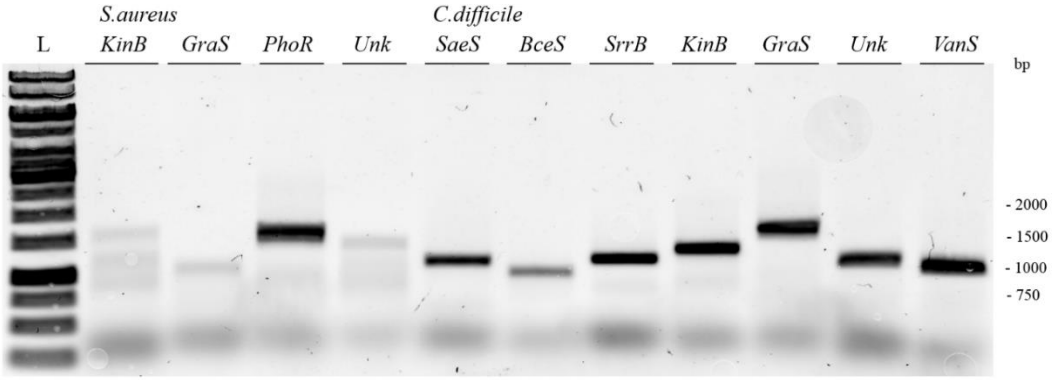

Figure 4.12 - 0.8% agarose gel. PCR results *S. aureus* (*KinB, GraS, PhoR* and *Unk)* and *C. difficile* (*SaeS, BceS, SrrB, KinB, GraS, Unk* and *VanS*) genes. L- GeneRuler™ 1 kb DNA ladder; all bands correspond to 50 µl of putative amplified genes.

Isolation of *S. aureus KinB, GraS* and *Unk* bands was not possible, because of the very weak intensity of the band. This fact might be due to the low quantity of DNA that was removed from 1st PCR tube. *S. aureus KinB* target presents 3 different bands. The reason might be in the pipetting of incorrect primer, because in the previous PCR reaction a single band was obtained. Concentrations that were obtained in the 2nd PCR were: *S. aureus* – [*PhoR*]= 125.6 ng/µl, *C. difficile* - [*SaeS*]= 80.1 ng/µl, [*BceS*]= 135.2 ng/µl, [*SrrB*]= 129.8 ng/µl, [*KinB*]= 18.1 ng/µl, [*GraS*]= 82.9 ng/µl, [*Unk*]= 157.7 ng/µl and [*VanS*]= 11 ng/µl.

In the following studies double digestion, ligation with pET-52b(+) and transformation into a *E.coli* XL-10 have been performed using more concentrated samples from 1st or 2nd polymerase chain reactions. Double digestion was done with XmaI and NotI-HF, for *S. aureus Unk* target, and with XmaI and SacI-HF, for the other targets. In case of successful cloning expression, solubilization tests have followed.

4.2 Expression test of the *S. aureus* targets

In addition to the cloned targets (*SrrB* and *AgrC*), several other *S. aureus* targets were already cloned. From those, *ArlS*, *SaeS* and *BceS* were chosen to proceed with expression tests.

Initial ArlS expression tests

Expression tests were started with ArlS target. Different media (LB broth, TB, 2xYT and Autoinduction medium) and *E. coli* strains (C43, BL21, BL21 Star and BL21 Gold) were used, at 37ºC

overnight. Results were analyzed by SDS-PAGE (figure 4.13), InVision (figure 4.14) and Western Blot techniques (figure 4.15).

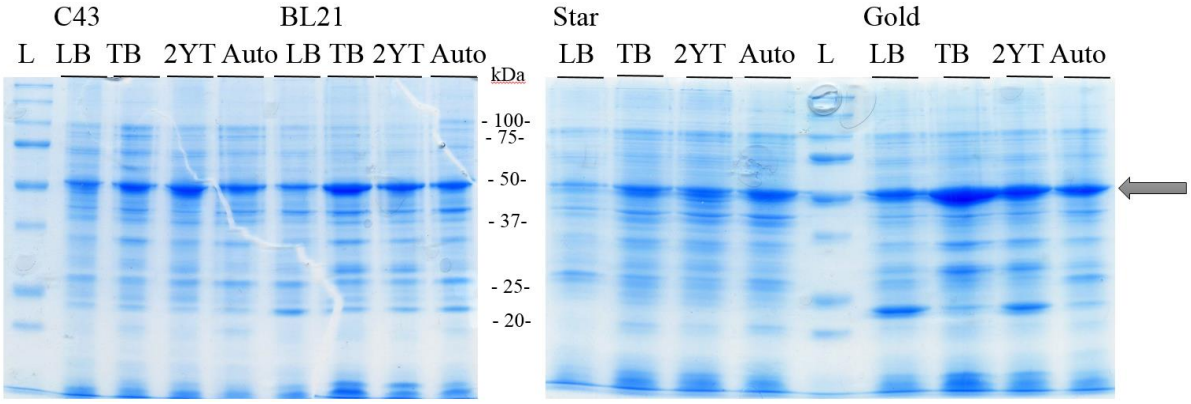

Figure 4.13 – 12% SDS-PAGE to analyse ArlS expression tests in *E. coli* at 37ºC. C43- C43 strain, BL21 – BL21 strain, Star - Bl21 Star strain, Gold - BL21 Gold strain; LB – LB broth medium, TB – TB medium, 2YT – 2xYT medium, Auto – Autoinduction medium, L - Precision Plus Protein Standards all blue ladder (BioRAD); Arrow – indicates to molecular weight where is supposed to appear bands of ArlS.

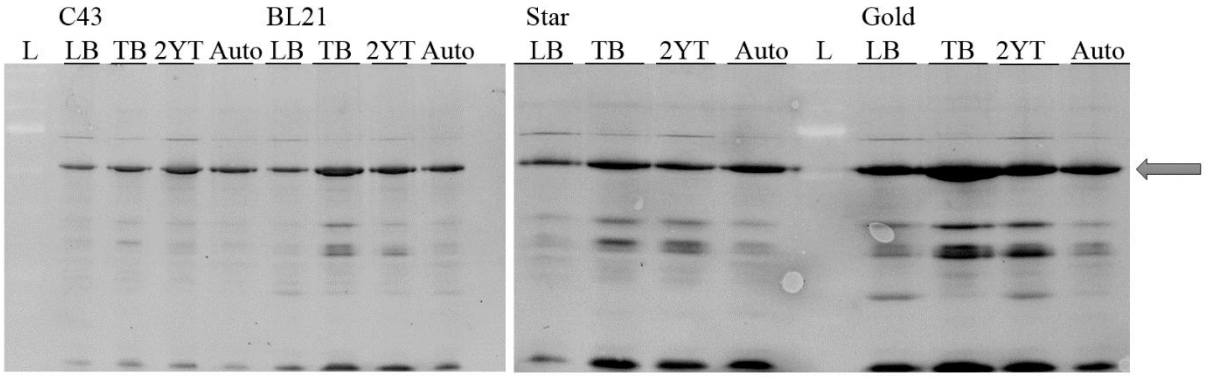

Figure 4.14 - InVision technique to access ArlS expression tests in *E. coli* at 37ºC. C43- C43 strain, BL21 – BL21 strain, Star - Bl21 Star strain, Gold - BL21 Gold strain; LB – LB broth medium, TB – TB medium, 2YT – 2xYT medium, Auto – Autoinduction medium, L - Precision Plus Protein Standards all blue ladder (BioRAD); Arrow – indicates to molecular weight where is supposed to appear bands of ArlS.

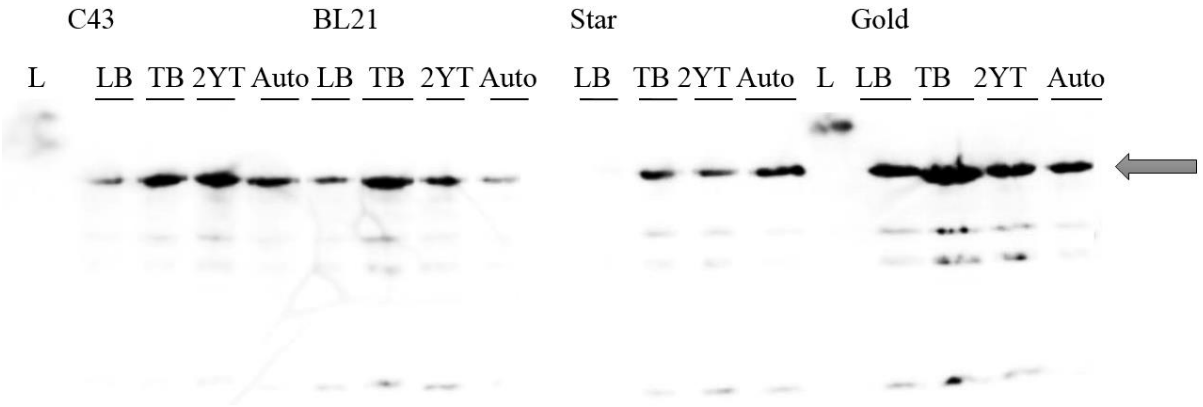

Figure 4.15 - Western Blot to evaluate ArlS expression tests in *E. coli* at 37ºC. C43- C43 strain, BL21 –BL21 strain, Star - Bl21 Star strain, Gold - BL21 Gold strain; LB – LB broth medium, TB – TB medium, 2YT – 2xYT medium, Auto – Autoinduction medium, L - Precision Plus Protein Standards all blue ladder (BioRAD); Arrow – indicates to molecular weight where is supposed to appear bands of ArlS.

The most intense bands can be observed around 50 kDa in the SDS-PAGE that correspond to ArlS molecular weight (52.4 kDa). This result is confirmed by Western Blot and In Vision techniques. The highest yield is obtained in *E.coli* BL21 Gold strain growth in TB media.

The same conditions (different media and *E.coli* strains) were tested at 28ºC overnight. Results were analyzed by SDS-PAGE (figure 4.16) and western blot techniques (figure 4.17).

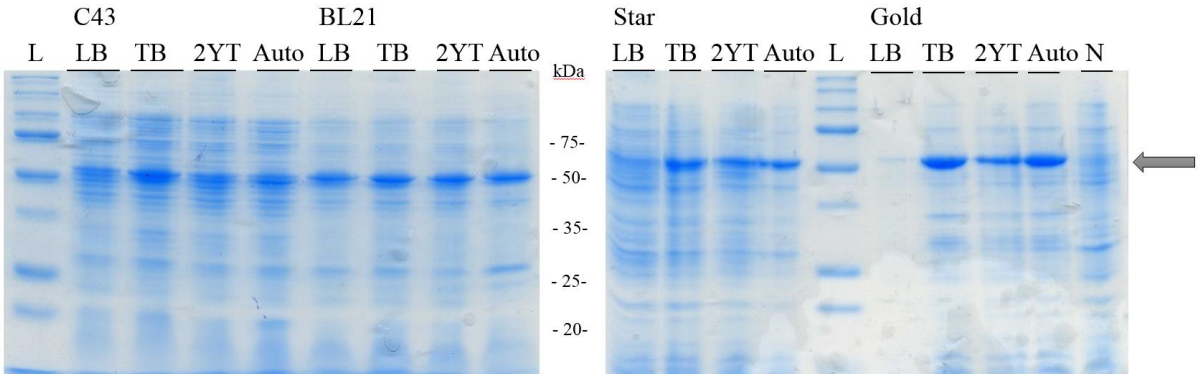

Figure 4.16 – 12% SDS-PAGE to analyse ArlS expression tests in *E. coli* at 28ºC. C43, BL21, BL21 Star and BL21 Gold – tested strains. LB broth, TB, 2YT, Autoinduction (Auto) – tested media; L - Precision Plus Protein Standards all blue ladder (BioRAD); Arrow – indicates to molecular weight where is supposed to appear bands of ArlS.

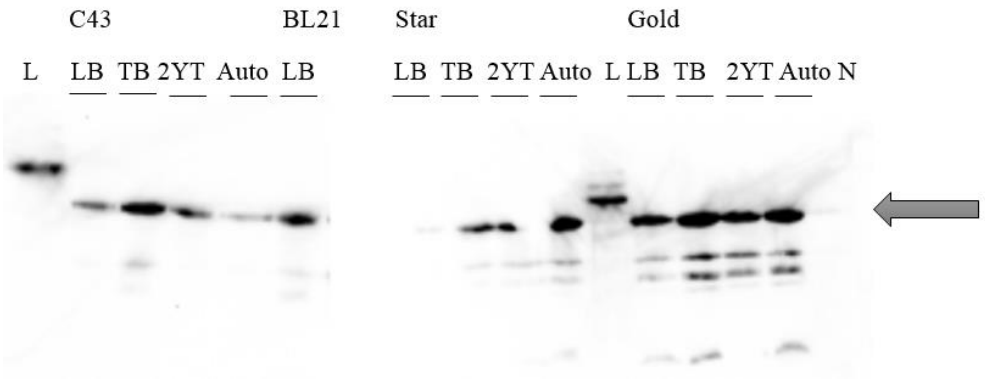

Figure 4.17 - Western Blot to evaluate ArlS expression tests in *E. coli* at 28ºC. C43, BL21, BL21 Star and BL21 Gold – tested strains. LB broth, TB, 2YT, Autoinduction (Auto) – tested media; L - Precision Plus Protein Standards all blue ladder (BioRAD); Arrow – indicates to molecular weight where is supposed to appear bands of ArlS.

More intense bands, such as in the previous expression test, appear around 50 kDa and correspond to ArlS target expressed by *E.coli* BL21 Gold in TB medium. Comparing with the previous expression test, 28ºC had lower band intensity, which means that 37ºC is the better condition for ArlS target expression.

ArlS target was further expressed in *E.coli* BL21 Gold in TB medium at 37ºC overnight and several solubilization tests were performed. Using a solubilization protocol[27], several detergents were tested: C-Hega 8, DDM, Cymal-4, LDAO, n-Octyl-β-D-Thioglucopyranoside, Fos-choline, Mega-8, DM and n-Octyl-β-D-glucopyranoside; for 2h, 4h and overnight (Appendix C, figure C.1, C.2 and C.3). In the soluble fraction no band corresponding to ArlS was detected, indicative of solubilization failure. Several factors influence protein solubilization, such as salt concentration or impropriate buffer. Different buffer and salt concentrations were tested in addition to LDAO and Fos-Choline detergents. These detergents were chosen because they represent two different detergents types: zwitterionic and non-ionic. In the soluble fraction no bands were detected around 50 kDa that could correspond to ArlS (Appendix C, figure C.4 and C.5). We concluded that these tested conditions were not suitable for the solubilization step. However, the membranes were obtained after chemical lyses, a method aggressive to the sample, which might influence the detergent capacity for solubilization. To overcome this problem, membrane disruption by pressure was performed (cell disruptor) and other solubilization conditions were tested, namely two different buffers (PBS and Tris), DDM detergent and 700 mM NaCl (Appendix C, figure C.6). Higher salt concentrations were used, which may positively influence protein solubilization. These conditions were chosen to assess whether the combination of buffer and detergents could influence solubilization.

No band in the soluble fraction corresponds to ArlS target. In the samples obtained by lyses in the cell disrupter, more intense ArlS bands were detected in the insoluble fraction, indicating that this membrane disruption process is more efficient, obtaining more protein quantity. Other parameter, that might influence protein solubilization, is temperature. Tris and PBS buffers in addition to Triton X-100

and DM were tested at room temperature (Appendix C, figure C.7). No bands around 50 kDa, which might correspond to the ArlS target were detected in soluble fraction, indicating that temperature has no significant effect on our target solubilization. So 4ºC is the best temperature to use, because it favors protein integrity and protects it from degradation. The other reason for solubilization failure might be in the 37ºC overnight expression. These conditions might lead to the production of large amounts of protein that may difficult the solubilization step. Inner and outer membrane separation was performed in order to remove other membrane proteins from the fraction containing the protein of interest. In the western blot several intense band are detected around 55 kDa in the white fraction that corresponds to the outer membrane, indicating that ArlS is located at the outer membrane (Appendix C, figure C.8). Outer membrane fraction was solubilized at different conditions, namely with two different buffers (PBS and Tris) and detergents (LDAO and DM) (Appendix C, figure C.9). No band in the soluble fraction was detected. In the last solubilization test more NaCl concentration were tested in combination 1% DM, DDM, LDAO and Triton X-100 (Appendix C, figure C.10 and figure C.11). No band is detected in the soluble fractions, which means that salt concentration is not relevant parameter in membrane extraction of ArlS. No positive results were obtained in these solubilization studies, and similar assays were performed for two other targets: BceS and SaeS.

ArlS, BceS and SaeS expression tests

As for the previous targets, BceS and SaeS expression tests were performed. In the first test, different media (LB broth, TB, 2xYT and Autoinduction) and *E.coli* strains (C43, BL21, BL21 Star and BL21 Gold) were used. BceS and SaeS results were evaluated by SDS-PAGE (figure 4.18 and 4.20) and Western blot techniques (figure 4.19 and 4.21)

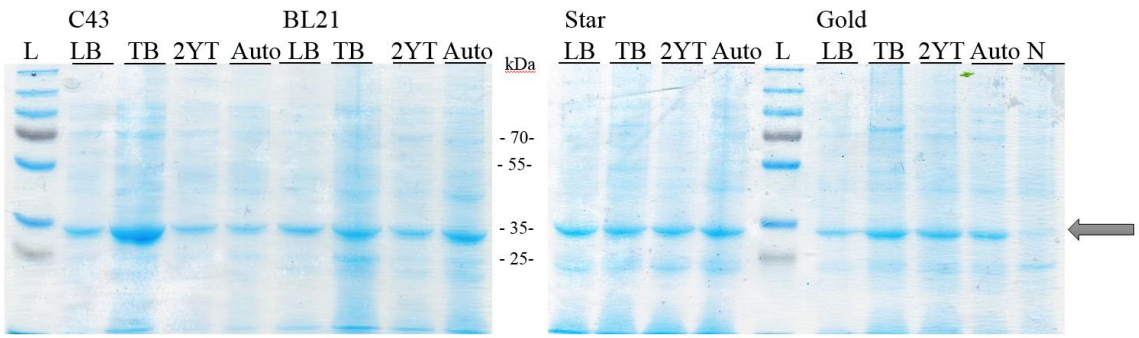

Figure 4.18 – 12% SDS-PAGE to analyse BceS expression tests in *E. coli* at 37ºC. C43, BL21, BL21 Star and BL21 Gold – tested strains. LB broth, TB, 2YT, Autoinduction (Auto) – tested media; L - Precision Plus Protein Standards all blue ladder (BioRAD); Arrow – indicates to molecular weight where is supposed to appear bands of BceS.

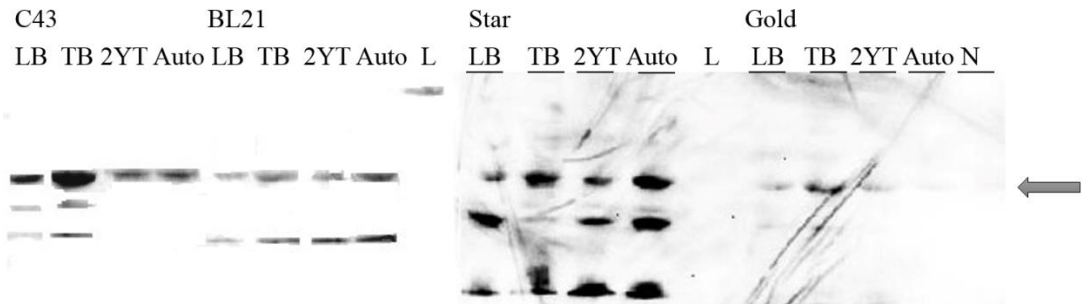

Figure 4.19 - Western blot to evaluate BceS expression tests in *E. coli* at 37ºC. C43, BL21, BL21 Star and BL21 Gold – tested strains. LB broth, TB, 2YT, Autoinduction (Auto) – tested media; L - Precision Plus Protein Standards all blue ladder (BioRAD). Arrow – indicates to molecular weight where is supposed to appear bands of BceS.

Several intense bands around 30 kDa are observed in the SDS-PAGE gel and western blot membrane belonging to the BceS protein (molecular weight is 34 kDa). The highest protein quantity (more intense band detected around 30 kDa) is expressed by *E.coli* C43 in TB media.

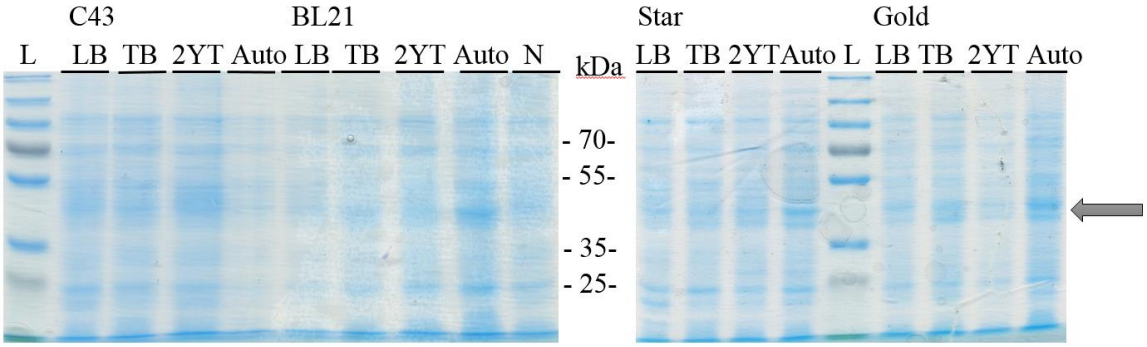

Figure 4.20 – 12% SDS-PAGE to analyse SaeS expression tests in *E. coli* at 37ºC. C43, BL21, BL21 Star and BL21 Gold – tested strains. LB broth, TB, 2YT, Autoinduction (Auto) – tested media; L - Precision Plus Protein Standards all blue ladder (BioRAD); Arrow – indicates to molecular weight where is supposed to appear bands of SaeS.

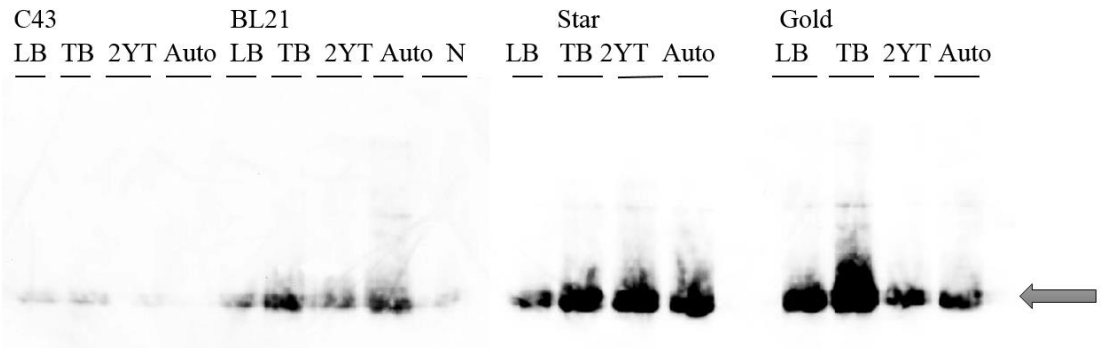

Figure 4.21 - Western blot to evaluate SaeS expression tests in *E. coli* at 37ºC. C43, BL21, BL21 Star and BL21 Gold – tested strains. LB broth, TB, 2YT, Autoinduction (Auto) – tested media; L - Precision Plus Protein Standards all blue ladder (BioRAD); Arrow – indicates to molecular weight where is supposed to appear bands of BceS.

Contrary to BceS target, SaeS protein production is very low in the *E.coli* C43 (no intense band was detected). However, in the western blot membrane, intense bands of expressed protein are observed for *E.coli* BL21 Star and Gold. The more intense band corresponds to the expression by *E.coli* BL21 Gold in TB medium. Despite the fact that western blot present very intense SaeS bands, in the SDS-PAGE these bands are not visible. The reason might be an unspecific binding of the chemiluminescence substrate.

From the previous results, the best medium and strain conditions were chosen for the BceS and SaeS targets and different temperatures and post-induction incubation times were tested. ArlS target were included for testing expression. The ArlS target production was tested in *E.coli* BL21 Gold in TB medium together with expression at 37ºC for 4h, 28ºC for 6h and 21ºC overnight. From previous results, 37ºC overnight expression was performed as reference. Results were evaluated by SDS-PAGE and Western Blot (figure 4.22) techniques.

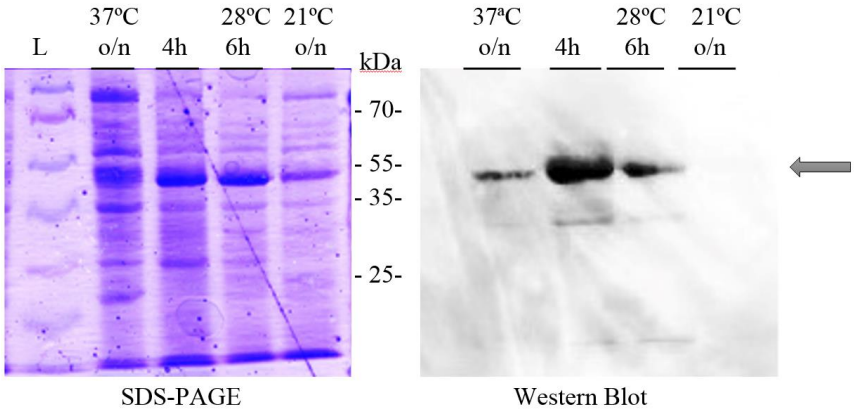

Figure 4.22 – 10% SDS-PAGE and Western Blot to analyse ArlS expression tests in *E. coli* at different temperatures and post-induction time. 37ºC, 28ºC and 21ºC – tested temperature, o/n – overnight expression, 4h – 4 hours expression, 6h – 6 hours expression ; L– Page Ruller™ Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear bands of ArlS.

The lowest ArlS band intensity was obtained at 21ºC, indicative of low protein expression under this condition. The most intense bands were observed when bacteria was grown at 37ºC for 4 hours or overnight. Expressing at 37ºC overnight presents additional bands, probably due to higher contamination with other *E.coli* proteins. Choosing 4 h expression at 37ºC might eliminate undesired proteins in the sample. The same conditions were tested for the BceS target in *E.coli* C43 strain in TB media. Samples were evaluated by SDS-PAGE and Western blot techniques (figure 4.23).

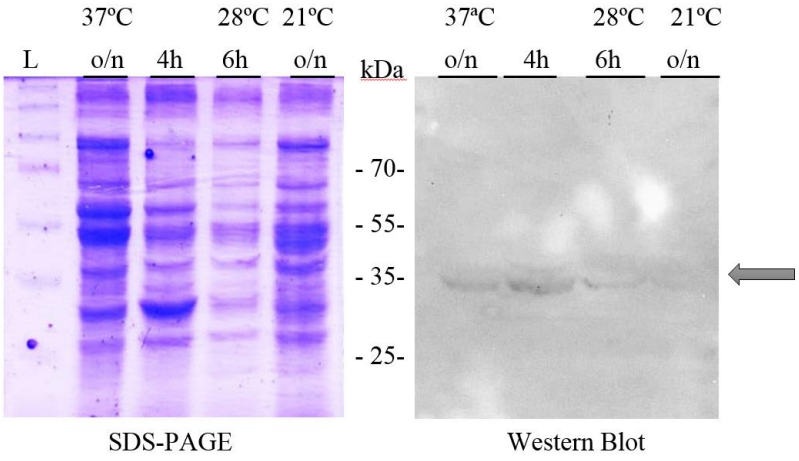

Figure 4.23 – 10% SDS-PAGE and Western Blot to analyse BceS expression tests in *E. coli* at different temperatures and post-induction time. 37ºC, 28ºC and 21ºC – tested temperature, o/n – overnight expression, 4h – 4 hours expression, 6h – 6 hours expression ; L– Page Ruller™ Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear band of BceS.

As in the previous study, expression overnight at 37ºC presents higher levels of *E.coli* contaminants. The Western blot result shows that higher expression (more intense band) is present when using 4 h expression at 37ºC. The same conditions were tested for the SaeS target in *E.coli* BL21 Gold and TB medium. Results were evaluated by SDS-PAGE and Western blot (figure 4.24)

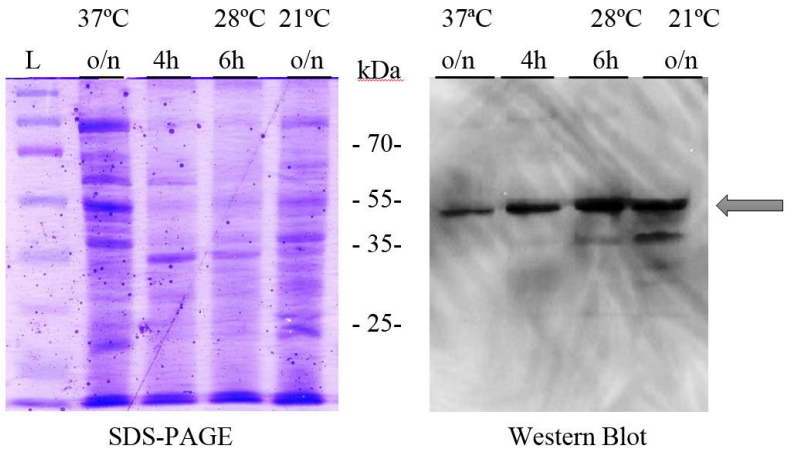

Figure 4.24 – 10% SDS-PAGE and Western Blot to analyse SaeS expression tests in *E. coli* at different temperatures and post-induction time. 37ºC, 28ºC and 21ºC – tested temperature, o/n – overnight expression, 4h – 4 hours expression, 6h – 6 hours expression ; L– Page Ruller™ Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear band of SaeS.

Such as in the previous tests, overnight expression at 37ºC presents higher contamination. Western blot analysis shows that expressing at 28ºC for 6h or 21ºC for overnight presents the same band intensity – indicating a similar SaeS production for both conditions. Two bands are detected in the Western blot membrane, which might correspond to different SaeS states/conformations. For the following studies

(crystallization and functional assays) homogeneity of the sample is very important so, 28ºC for 6h was chosen as the best condition for expression, because it contains less contaminants.

The growth volume is another parameter that might influence target protein expression. BceS expression in 1 L of TB medium was tested in three different conditions: the first one was divided in 4 parts (250 ml in each tube), second one was divided in two parts (500 ml in each tube) and third was intact (1 L) (Figure 4.25). The best strain (*E.coli* C43), temperature (37 ºC) and post-induction expression time (4h) were tested.

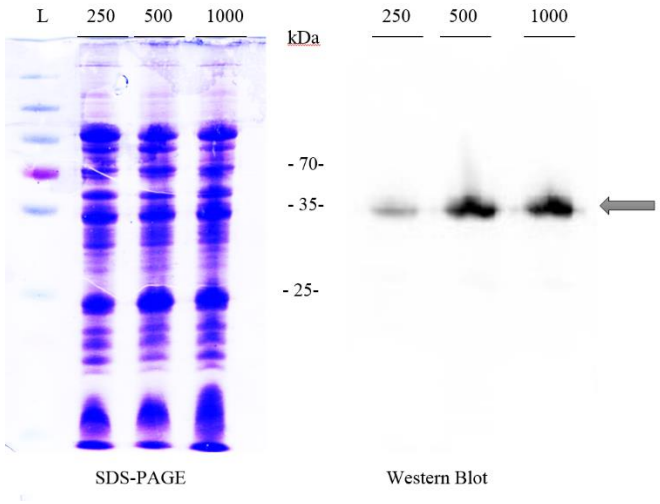

Figure 4.25 – 10% SDS-PAGE and Western Blot to analyse ArlS expression tests in *E. coli* at different volumes. 250 ml – 4x250 ml volume; 500 ml - 2x500 ml volume 1000 ml – 1 L volume ; L – Page Ruller^{TM} Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear band of ArlS.

The worst expression was obtained in the 4x250 ml. The 2x500 ml and 1 L volumes contain bands with same intensity, which indicates a similar BceS production. In order to spend less materials and time to centrifuge the samples in the large scale, 1L growing volume was chosen. Finally, best expression conditions were identified for each target. ArlS protein expression is best in *E.coli* BL21 Gold, TB medium at 37ºC for 4 h; BceS expression – in *E.coli* C43 and TB medium at 37ºC for 4 h; and SaeS expression – in *E.coli* BL21 Gold and TB medium at 28ºC for 6 h. The best medium volume was 1 L.

4.3 Solubilization tests

Knowing the best expression condition for each target, solubilization tests were followed. Different conditions were tested, namely different detergents and salt concentrations. Primary, ArlS solubilization tests were performed. Different detergents, Triton X-100, *n*-Decyl-β-D-maltopyranoside (DM) and *n*-Dodecyl-β-D-maltopyranoside (DDM) were tested in combination with several salt concentrations: NaCl at 250 mM, 500 mM and 750 mM. Results were evaluated by SDS-PAGE (figure 4.26) and In Vision technique (figure 4.27).

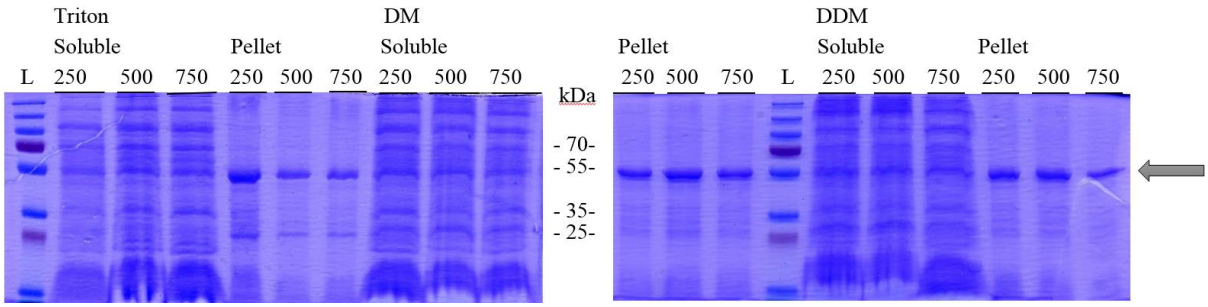

Figure 4.26 – 10% SDS-PAGE to analyse ArlS solubilization. Detergents: Triton X-100, DM and DDM, Soluble – Soluble fraction, Pellet – insoluble fraction, NaCl concentrations (mM): 250, 500 and 750; L – Page Ruller^{TM} Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear band of ArlS.

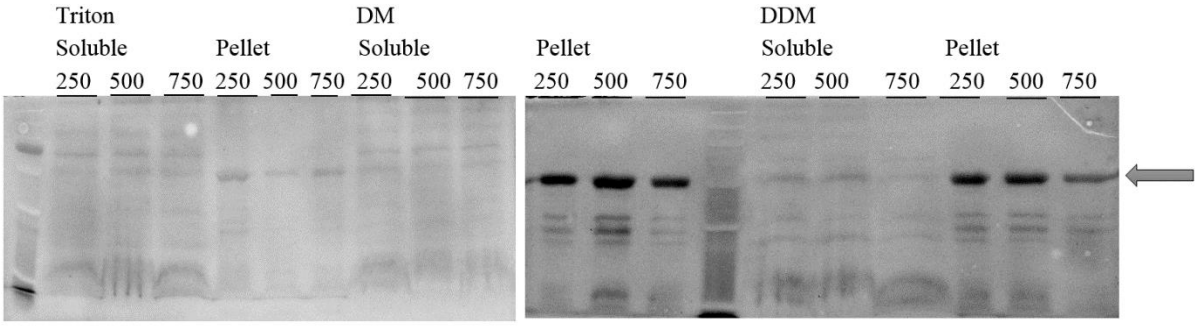

Figure 4.27 – In Vision to evaluate ArlS solubilization. Detergents: Triton X-100, DM and DDM, Soluble – Soluble fraction, Pellet – insoluble fraction, NaCl concentrations (mM): 250, 500 and 750; L – Page Ruller^{TM} Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear band of ArlS.

As can be observed in figure 3.28, the solubilized fraction contains a variety of *E.coli* proteins of different sizes. In the insoluble fraction an intense band is visible around 55 kDa that should correspond to ArlS protein. The same band is detected in the soluble fraction, but with much lower intensity. However, we chose 1.5% DDM at 500 mM NaCl condition to proceed with a purification test.

Then BceS target solubilization tests were done as for the previous target (Figure 4.28, 4.29 and 4.30).

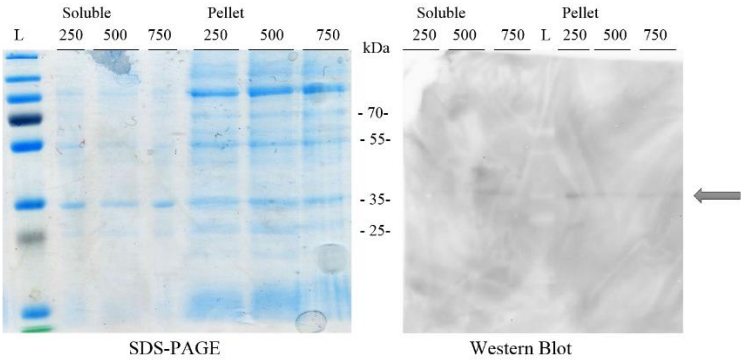

Figure 4.28 – 10% SDS-PAGE and Western blot to analyse BceS solubilization with 1.5% Triton X-100. Soluble – Soluble fraction, Pellet – insoluble fraction, NaCl concentrations (mM): 250, 500 and 750; L – Page Ruller^{TM} Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear band of BceS.

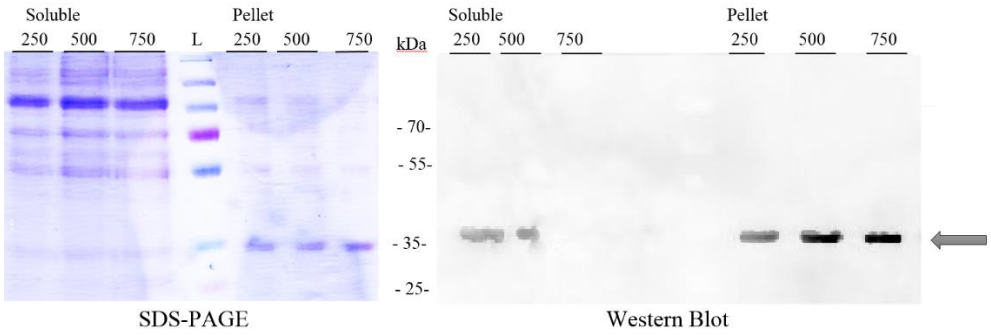

Figure 4.29 – 10% SDS-PAGE and Western blot to analyse BceS solubilization with 1.5% DM. Soluble – Soluble fraction, Pellet – insoluble fraction, NaCl concentrations (mM): 250, 500 and 750; L – Page Ruller™ Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear band of BceS.

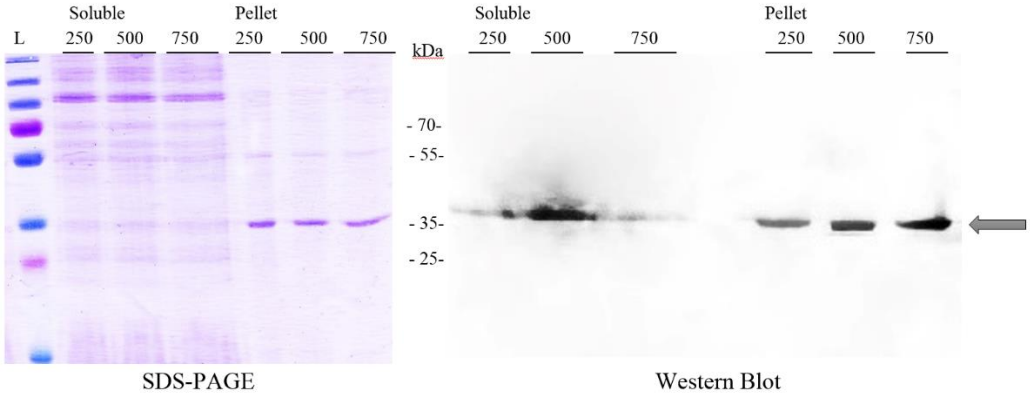

Figure 4.30 – 10% SDS-PAGE and Western blot to analyse BceS solubilization with 1.5% DDM. Soluble – Soluble fraction, Pellet – insoluble fraction, NaCl concentrations (mM): 250, 500 and 750; L – Page Ruller™ Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear band of BceS.

In the SDS-PAGE gel several bands around 35 kDa were detected in the insoluble fraction corresponding to the BceS target. The soluble fraction did not show any intense band around this value, however, in the western blot, a band at the correct position was visible for several of the conditions (Figure 4.29 and 4.30). 1.5% DDM with 500 mM NaCl and 1.5% DM with 250 mM NaCl are the samples containing the more intense BceS bands. This two solubilization conditions were chosen for the following purifications tests.

As for SaeS solubilization tests, the same conditions were tested (figure 4.31 and 4.32).

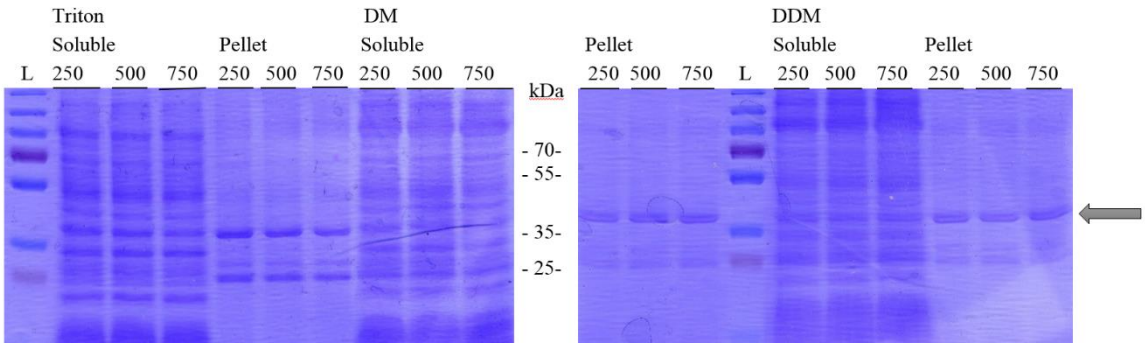

Figure 4.31 – 10% SDS-PAGE to analyse SaeS solubilization. Detergents: Triton X-100, DM and DDM, Soluble – Soluble fraction, Pellet – insoluble fraction, NaCl concentrations (mM): 250, 500 and 750; L – Page Ruller^{TM} Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear band of SaeS.

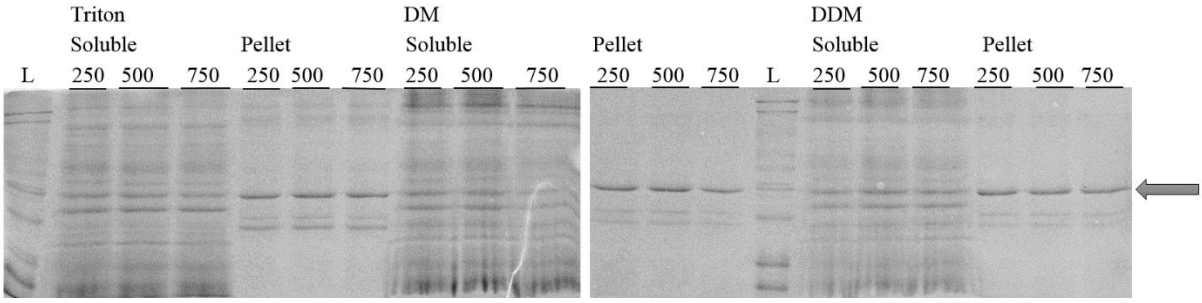

Figure 4.32 – In Vision to analyse SaeS solubilization. Detergents: Triton X-100, DM and DDM, Soluble – Soluble fraction, Pellet – insoluble fraction, NaCl concentrations (mM): 250, 500 and 750; L – Page Ruller^{TM} Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear band of SaeS.

Similar to the previous solubilization tests, several *E.coli* proteins are solubilized with the SaeS protein. An intense band appears in the insoluble fraction around 40 kDa that should correspond to SaeS protein. The same band was observed in the soluble fraction, but with lower intensity. 1.5% DDM at 500 mM NaCl condition was chosen to proceed with purification tests.

4.4 Purification tests: magnetic Beads

The first BceS purification test was performed using magnetic beads. Purification with magnetic beads requires very low membrane quantity and corresponds to small scale purification. Total membrane was solubilized with 1.5% DDM, 500 mM NaCl in 50 mM Tris-HCl pH 7.5. According to the manufacture´s protocol, different imidazole concentrations were chosen for elution of the protein.

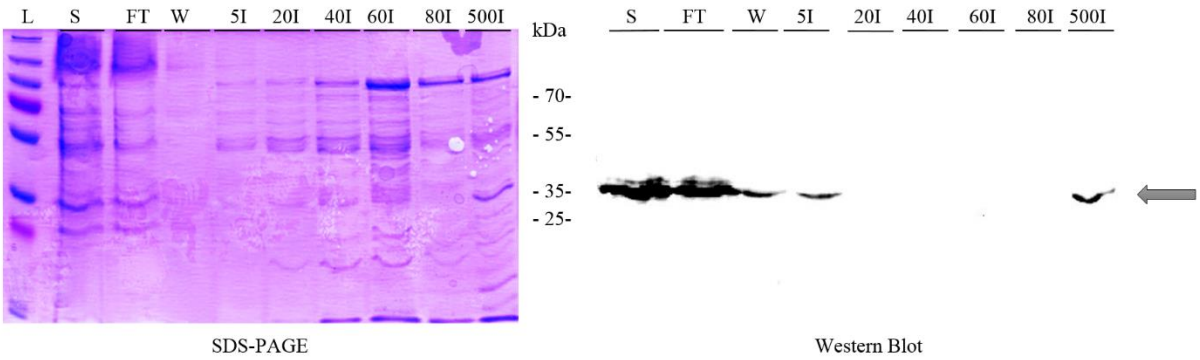

Figure 4.33 – 10% SDS-PAGE and Western blot to analyse BceS purification with magnetic bids. S – Solubilized membrane, FT – flow through, W – washing, I - Imidazole concentrations (mM): 5, 20, 40, 60, 80 and 500. L – Page Ruller^{TM} Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear band of BaeS.

The western blot shows an intense band in the flow-through, an indication of low binding affinity of the His-tag to the beads. The reason might be in the BceS tertiary structure: His-tag might be located inside the protein blocking the tag access to the magnetic beads. Smaller amounts of the target protein are present in the washing and 5 mM Imidazole eluted samples. Also, some protein is observed when eluted with 500 mM Imidazole which indicates that a portion of the protein was strongly attached to the beads. However, this purification has a very low yield, because many other proteins are present in the sample (Figure 4.33). Possibly, eluting by gradient increase in imidazole concentration could increase sample purity. In order to choose an adequate imidazole concentration for elution, HPLC purification was tried with Ni-NTA column where an imidazole gradient can be used for protein elution (Figure 4.34). The results were analyzed by SDS-PAGE (figure 4.35).

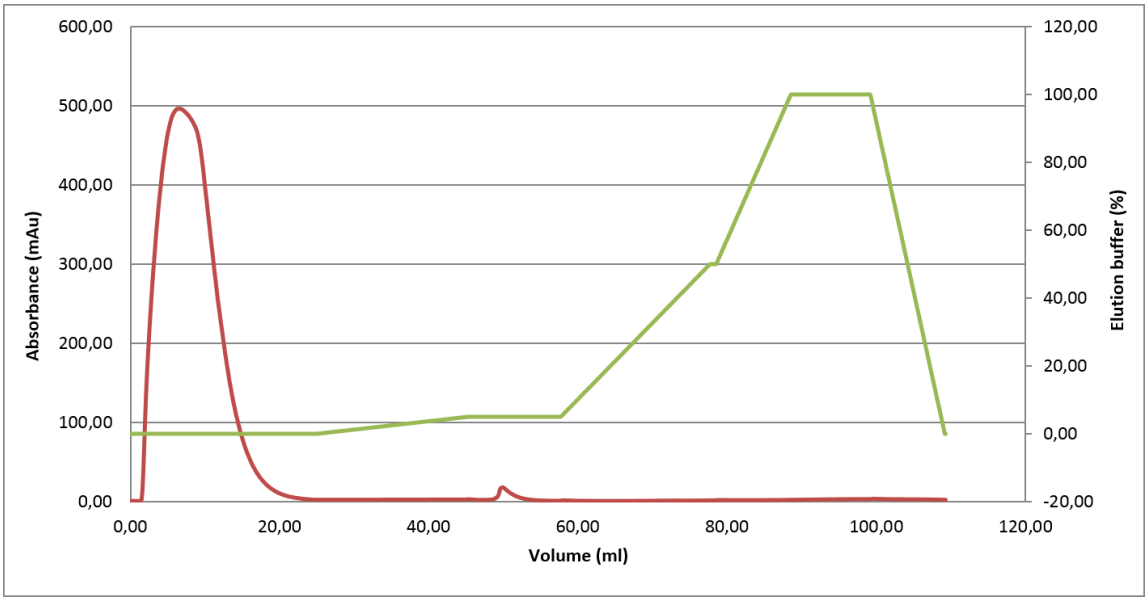

Figure 4.34 - Chromatogram of BceS purification by HPLC with Ni-NTA column. Red line – Absorbance of the protein in amount; Green line: Imidazole gradient percentage

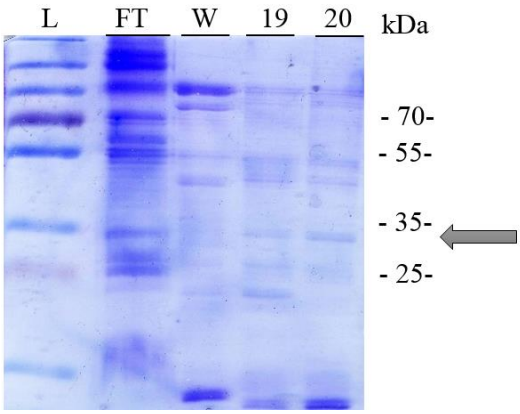

Figure 4.35 – 10% SDS-PAGE to analyse BceS purification by Ni-NTA column. FT – Flow through, W- washing step, 19 – 48-50 ml, 20 – 50-52 ml, L – Page RullerTM Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear band of BceS.

In the chromatogram (figure 4.34) we observe a large peak in the flow-through sample. At the washing step with 5 mM Imidazole, another peak is detected. On the elution step, very small peak is observed around 500 mM Imidazole. In the SDS-PAGE gel several bands are present around 35 kDa. The flow-through contains the more intense band that should correspond to the BceS target. At the washing step no intense band around 35 kDa is detected, such as in the elution step (48-52 ml), however other contaminants were also detected. This indicates the low affinity of the His-tag towards the column matrix. We concluded that purification in batch using magnetic beads was more successful, because intense higher quantity of BceS is purified (more intense band in the gel). For those reasons, purification in batch using Ni Sepharose 6 Fast Flow slurry in large quantity was performed. Since the previous utilized imidazole concentrations (60 and 800 mM) did not elute the protein, other concentrations were tested, namely 20, 100, 250, 500 mM and 1 M Imidazole. Each elution step was repeated twice in order to verify that all attached protein is eluted.

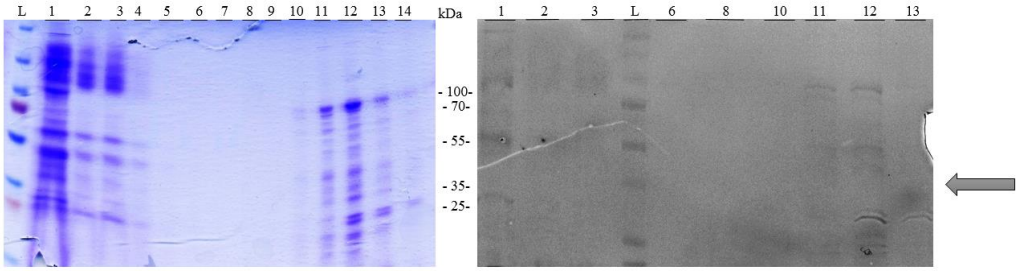

Figure 4.36 – 10% SDS-PAGE and In Vision to analyse BceS batch purification 1 – Soluble fraction, 2 – Flow through, 3 – First washing step; 4 – Sixth washing; Imidazole concentrations (mM): 5 – 20, 6 –20 2nd elution, 7 – 100, 8 –100 2nd elution, 9 – 250, 10 – 250 2nd elution, 11 – 500, 12 –500 2nd elution, 13 – 1000, 14 – 1000 2nd elution, L – Page Ruller^{TM} Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear band of BceS.

After SDS-PAGE and in vision analysis, BceS bands were detected in the flow-through, 500 mM and 1M imidazole eluted samples. In addition to BceS target, several contaminants were also eluted. This fact indicates that our purification was not efficient, as no pure protein was obtained.

We then proceeded with ArlS and SaeS purifications, performing purification in batch with Ni Sepharose 6 Fast Flow slurry.

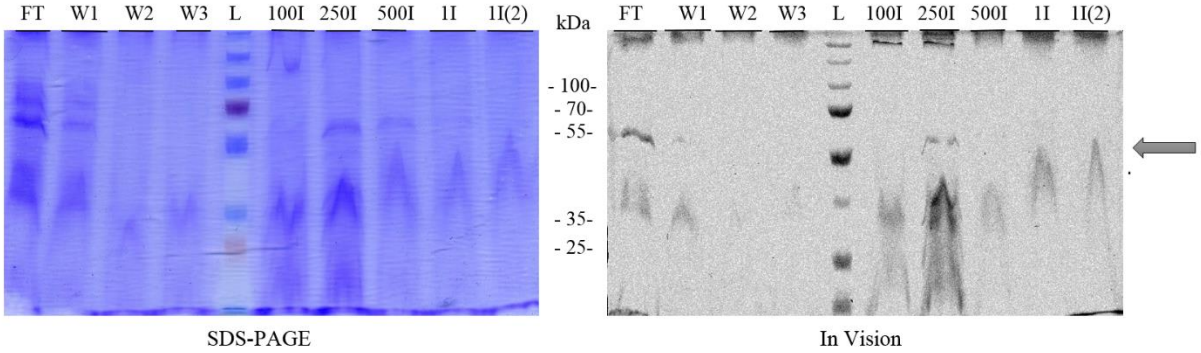

Figure 4.37 – 10% SDS-PAGE and In Vision to analyse ArlS batch purification. FT – Flow through, W1/W2/W3 – Washing steps, xI - Imidazole concentrations (mM): 100, 250, 500 and 1I - 1000, 1I(2) - 1000 2nd elution. L – Page Ruller™ Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear band of ArlS.

ArlS purification analysis by SDS-PAGE and In Vision technique (figure 4.37) shows intense bands around 35 kDa and 60 kDa. Molecular weight of ArlS target is 51 kDa, and the fact that membrane proteins have a tendency to run at lower molecular weights, led us to hypothesize that the band around 35 kDa might corresponds to the ArlS target. This target is eluted by 250 mM imidazole, but, similar to BceS target, the elution fraction contains many other *E.coli* proteins.

SaeS purification was also analyzed by SDS-PAGE and In Vision technique (figure 4.38).

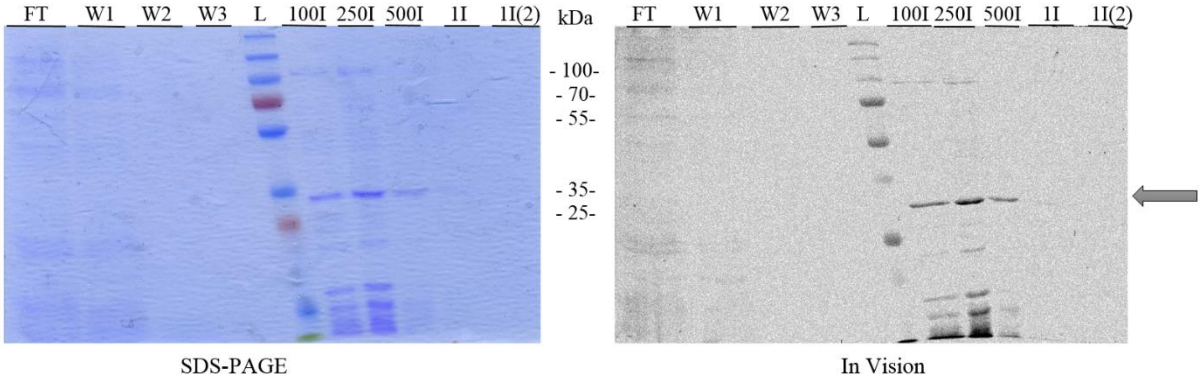

Figure 4.38 – 10% SDS-PAGE and In vision to analyse SaeS batch purification. FT – Flow through, W1/W2/W3 – Washing steps, xI - Imidazole concentrations (mM): 100, 250, 500 and 1I - 1000, 1I(2) - 1000 2nd elution. L – Page Ruller™ Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear band of SaeS.

One intense band is observed around 35 kDa in the samples eluted with 100, 250 and 500 mM of imidazole, corresponding to the SaeS target (39.7 kDa). Despite their protein contamination, 250 mM imidazole sample contains more SaeS protein quantity (band is more intense, defined and less

contaminated). This sample was chosen to estimate the amount of purified protein by densitometry and calculate the purification yield to determine whether it is profitable to move for a large-scale production. The band was quantified using ImageJ program. Several samples were quantified namely those eluted with 100 and 250 mM imidazole, together with lysozyme samples at 0.05, 0.1, 0.25 and 0.5 mg/ml for reference.

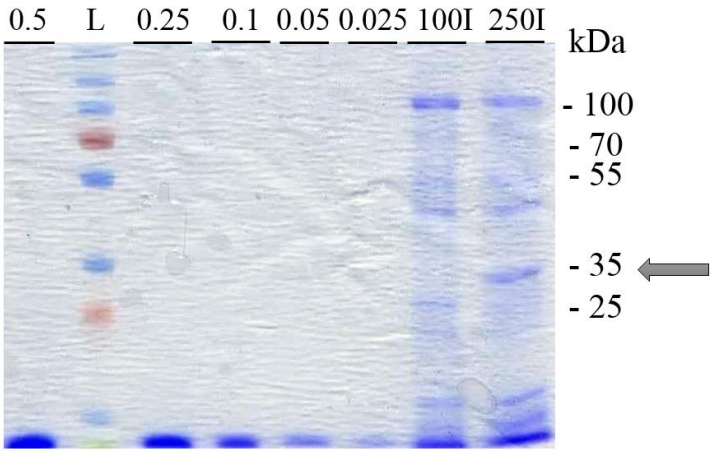

Figure 4.39 – 10% SDS-PAGE to analyse SaeS quantification. xI - Imidazole concentrations (mM): 100 and 250; Lysozyme concentration (mg/ml): 0.5, 0.25, 0.1, 0.05 and 0.025. L – Page Ruller^{TM} Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear band of SaeS.

Lysozyme bands corresponding to 0.25, 0.01, 0.05 and 0.025 mg/ml were measured with ImajeJ. The intensity of each band was plotted against the protein concentration and a correlation graphic was obtained. Each intensity and concentration was directly related with each other´s and a linear function was defined (figure 4.40).

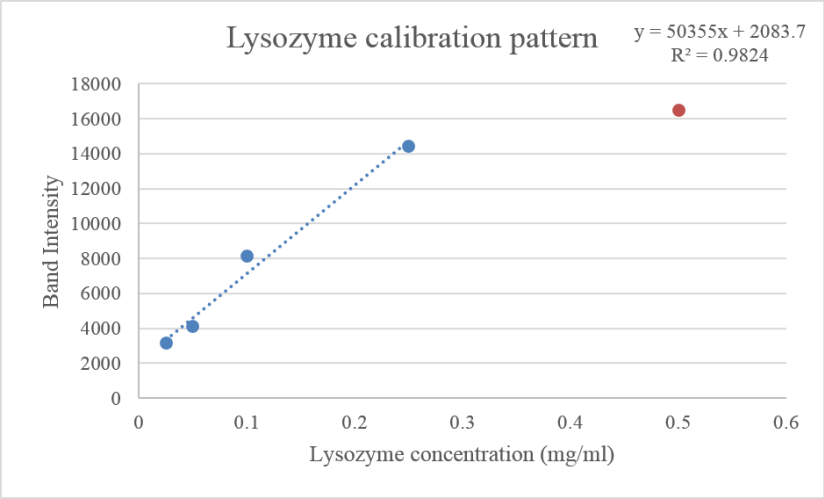

Figure 4.40 - Lysozyme calibration curve used to quantify SaeS in the purified fractions. Blue – points that were used to the calibration curve calculation; Red – point that were excluded from calibration curve.

Comparing all points we saw that 0.5 mg/ml (red point) shows a discrepancy with other points, and was excluded from the calculations. Four other bands, namely 0.25, 0.1, 0.05 and 0.025 mg/ml were used to

build our calibration curve: R^2 =0.982. The obtained function y=50355x + 2083.7 was used to estimate SaeS concentration. More intense and isolated SaeS band is located in 250 mM Imidazole sample. Through ImageJ program, intensity of the band was calculated (5526.93) that corresponds to 0.068 mg/ml. 0.05 mg/ml lysozyme band intensity is very similar to intensity of SaeS band in 250 mM imidazole fraction, indicating that our estimation is close to reality. Such quantity led us to produce the protein in the larger scale, so that we could increase protein quantity for crystallization and functional assays. The problem is the existent contamination with *E.coli* secondary products. To overcome this problem, we decided to perform a gel filtration to separate SaeS target from contaminants (figure 4.41).

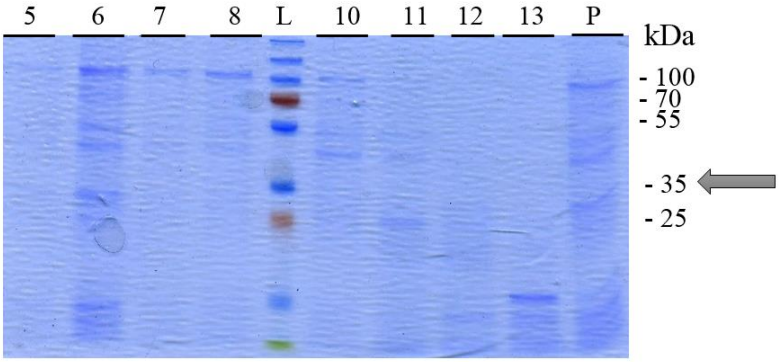

Figure 4.41 – 10% SDS-PAGE to analyse SaeS gel filtration. 5 – 5 tube, 6 – 6 tube, 7 – 7 tube, 8 – 8 tube, 10 – 10 tube, 11 – 11 tube, 12 – 12 tube, 13 – 13 tube, P – pellet, L – Page Ruller^{TM} Plus Prestained Protein ladder; Arrow – indicates to molecular weight where is supposed to appear band of SaeS.

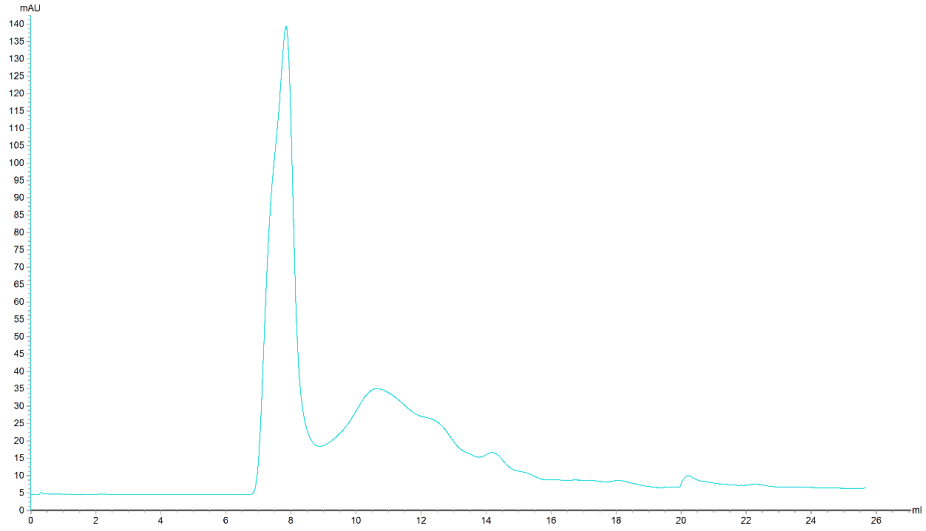

Figure 4.42 - SaeS gel filtration chromatography in superose 12 10/300 GL column

After SDS-PAGE analysis of gel filtration results, SaeS target might be detected in the pellet fraction and 18-20 ml of eluted volume. However, several contaminants are also present, indicating it was not

possible to completely separate SaeS target from other proteins. SaeS purification failed and we were not able to go for crystallization and functional assays. As such, other protein targets were chosen.

NreB, SrrB and AgrC targets were selected for expression tests. As observed in previous experiments, *E.coli* C43 and BL21 Gold were the strains that showed best expression yields in LB and TB media. The best temperature for expression are 37ºC or 28ºC for 4h and 6 h, respectively. Expression tests results were analyzed by SDS-PAGE (figure 4.43 and 4.44).

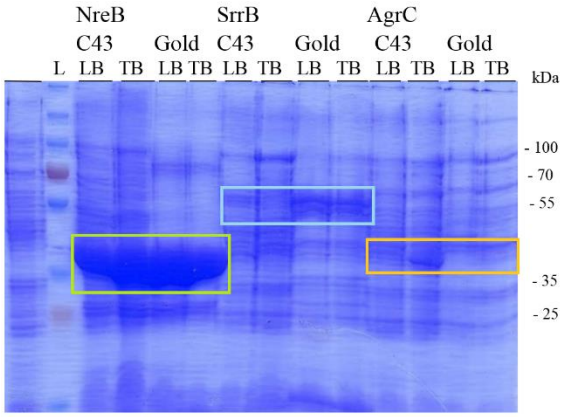

Figure 4.43 – 10% SDS-PAGE to analyse NreB, SrrB and AgrC expression in *E. coli* at 37ºC. C43- C43 strain, Gold - BL21 Gold strain; LB – LB broth medium, TB – TB medium, L - Precision Plus Protein Standards all blue ladder (BioRAD) Rectangle: Green – NreB bands; Blue – SrrB bands; Orange – AgrC bands;

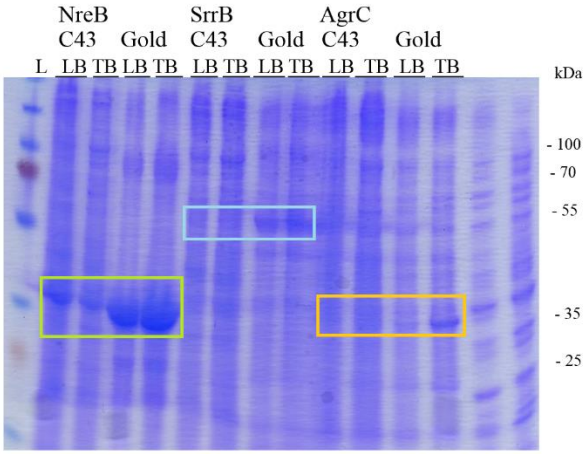

Figure 4.44 – 10% SDS-PAGE to analyse NreB, SrrB and AgrC expression in *E. coli* at 28ºC. C43- C43 strain, Gold - BL21 Gold strain; LB – LB broth medium, TB – TB medium, L - Precision Plus Protein Standards all blue ladder (BioRAD) Rectangle: Green – NreB bands; Blue – SrrB bands; Orange – AgrC bands;

Observing figure 4.47 and 4.48, SrrB and AgrC targets do not show any intense band around 66 and 49 kDa, indicating a low or absent expression of these selected targets. In the previous expression tests of SaeS SDS-PAGE gel (figure 4.20 and 4.24) did not have an intense band and expressed protein quantity was very low and contaminated to proceeds for large scale purification. That's why AgrC and SrrB samples were discarded. NreB sample has an intense band around 35 kDa, which corresponds to NreB HK (39 kDa). The more intense band is obtained at 37ºC 4 h post-induction and growing *E.coli* BL21

Gold strain in LB medium. This condition was used for 1 L expression of NreB target to produce protein sample for solubilization tests. In those tests, DM, DDM, Triton X-100 detergents were used in combination with different salts concentrations: 250, 500 and 750 mM NaCl. Results were analyzed by SDS-PAGE (figure 4.45).

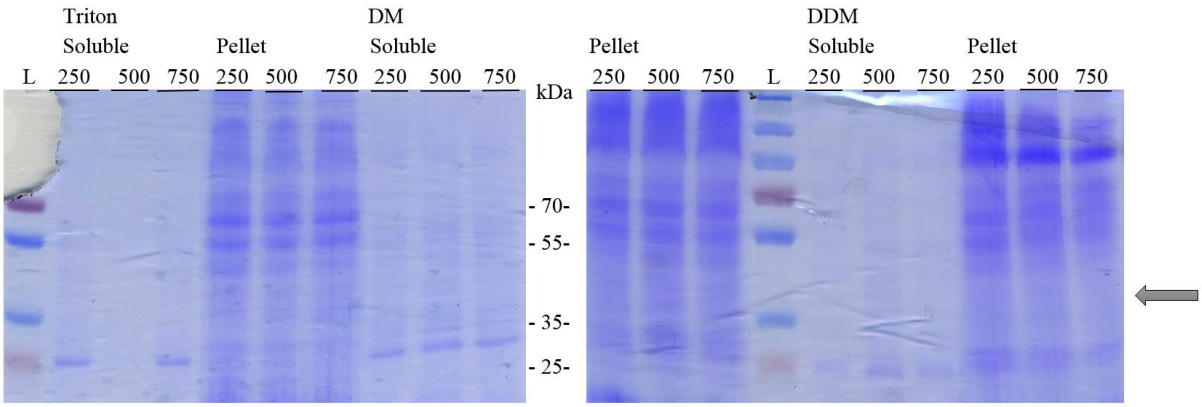

Figure 4.45 – 10% SDS-PAGE to analyse NreB solubilization. Detergents: Triton X-100, DM and DDM, Soluble – Soluble fraction, Pellet – insoluble fraction, NaCl concentrations (mM): 250, 500 and 750; L – Page Ruller^{TM} Plus Prestained Protein ladder; Arrow – indicates molecular weight where is supposed to appear a band of NreB.

No band was detected around 39 kDa that might correspond to the NreB target. That fact indicates that we failed to express the target. Solubilization test must be repeated with a new sample.

All results obtained in the cloning, expression, solubilization and purification experiments are summarized at the table 4.1 and 4.2.

Table 4.1 – Summary of all results obtained in the cloning, expression, solubilization and purification of *S. aureus* Histidine Kinases

MRSA (COL)	Cloning	Expression	Solubilization	Purification
ArlS	-	√ (High)	√	x
SaeS	-	√ (Low)	√	x
BceS	-	√ (High)	√	x
SrrB	√	√ (Low)	xx	xx
AgrC	√	√ (Low)	xx	xx
KinB	√	-	-	-
PhoR	√	-	-	-
GraS	√	-	-	-
Unk	√	-	-	-
NreB	√	√ (Very high)	-	-

Note: √ - positive result; x – negative result; xx – experiment not performed; - experiment to be performed

Table 4.2 Summary of all results obtained in the cloning, expression, solubilization and purification of *C. difficile* Histidine Kinases

C. difficile (630)	Cloning	Expression	Solubilization	Purification
VanS	√	-	-	-
SpaK	√	-	-	-
PhoR	√	-	-	-
SrrB	√	-	-	-
Unk	√	-	-	-
GraS	√	-	-	-
KinB	√	-	-	-
SaeS	√	-	-	-
BceS	√	-	-	-
AgrC	x	-	-	-

Note: √ - positive result; - experiment to be performed

As we can see in table 4.1 and 4.2, SrrB, AgrC, ArlS, BceS, SaeS and NreB were selected for expression and solubilization tests. Just two targets, namely ArlS and NreB had a good expression yield in *E. coli*. Solubilization was a problematic step in this study. Many different conditions were tested, such as different type of detergents, detergent concentration, buffers, pH, temperature, salts concentrations, time of solubilization and different cell lysis protocols. Despite of that, solubilization was not significantly improved. Purification trials were performed for the ArlS, BceS and SaeS proteins, but due to low protein quantity in a soluble fraction it was not possible to obtain sufficient amount of pure protein.

Chapter 3

Structural studies on Thiosulfate dehydrogenase

5. Introduction

5.1 Thiosulfate dehydrogenase

Sulfur (S) is the tenth most abundant element in the Universe and plays an important role in physiology being a constituent of vitamins, proteins and hormones. In nature sulfur can be found in different oxidation states ranging from +6 to -2, such as in sulfate and sulfide, respectively. Interconversion between various sulfur species constitute their global biogeochemical cycle (figure 5.1)[28].

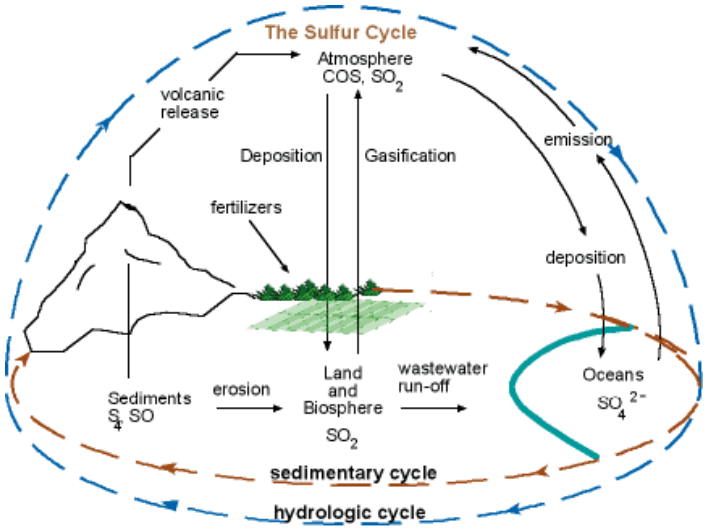

Figure 5.1 Schematic representation of the hydrologic and sedimentary sulfur cycle. (Adapted from http://environ.andrew.cmu.edu/m3/s4/cycleSulfur.shtml)

Elemental sulfur and reduced inorganic compounds, such as polysulfide, organic disulfides sulfate, sulfide and thiosulfate ($S_2O_3^{2-}$), are metabolized by large number of microorganisms[29]. Thiosulfate is an important intermediate in the biogeochemical sulfur cycle, which can be derived from sulfide soils or animal gut. Many sulfur bacteria use thiosulfate as a reductant agent, in the carbon fixation, or as an electron donor, in the energy-generating respiratory electron transport chain[30].

Bacterial pathways allow complete oxidation of thiosulfate to sulfate. Some sulfur bacteria oxidize thiosulfate to tetrathionate using thiosulfate dehydrogenase (TsdA), whose oxidation process is represented as

$$2\ S_2O_3^{2-} \rightarrow S_4O_6^{2-} + 2\ e^-$$

TsdA homologues are present in α-, β-, δ-, γ- and ε-Proteobacteria[31]. ε-Proteobacteria of genus *Campylobacter*, particularly *C. jejuni* and *C. coli*, cause a great number of food-borne human infections. *C. jejuni* is a microaerophilic mucosal intestinal pathogen in humans that causes acute bloody diarrhoea[6]. TsdA from *C. jejuni* was shown to be a bifunctional enzyme displaying tetrathionate reductase and

thiosulfate dehydrogenase activities (figure 5.2). When TsdA presents thiosulfate dehydrogenase activity (green pathway), two thiosulfate molecules give rise to one tetrathionate molecule and two electrons are transferred to a heme moiety. Then, electrons are further transferred to another heme, and then to a periplasmic cytochrome *c* (a membrane-bound *cb*-oxidase), for which the final acceptor is an oxygen molecule (O_2). On the opposite, when TsdA shows tetrathionate reductase activity (black pathway), one tetrathionate molecule is reduced to two thiosulfate molecules. The necessary electrons needed for tetrathionate reduction are provided from formate oxidation by formate dehydrogenase. Electrons are then transferred via the menaquinol pool (MKH_2) to cytochrome bc_1 complex, then to the periplasmic cytochrome and finally to a heme moiety in TsdA[6].

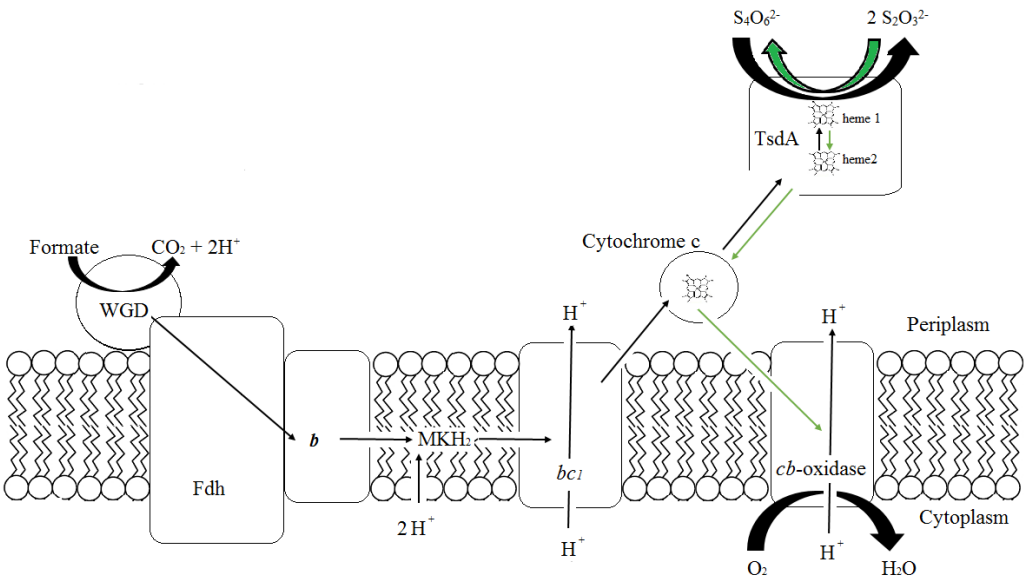

Figure 5.2 – *In vivo* model for bi-directional electron transfer between tetrathionate and thiosulfate catalysed by TsdA from *C. jejuni*. Electrons for the tetrathionate reduction (black pathway) are provided by formate via the formate dehydrogenase (Fdh), tungstopterin guanine dinucleotide (WGD), menaquinol pool (MKH_2), cytochrome bc_1 complex (bc_1) and periplasmic cytochrome c. Thiosulfate dehydrogenase activity yelds 2 electrons, for which the final acceptor is cb-type oxidase (green pathway). (Adapted from Liu et al., 2013)[6] .

The oxidation activity of *A. vinosum* TsdA for thiosulfate is over 40 times higher than that of *C. jejuni* TsdA. However, tetrathionate reduction is 1500 times higher for *C .jejuni* TsdA than that of *A. vinosum*. This fact indicates that *A. vinosum* TsdA is not a suitable enzyme for tetrathionate reduction, in contrast to *C. jejuni* TsdA[6], highlighting the importance of habitat adaptation for microorganisms.

Tetrathionate can be produced in intestinal mucosa from the oxidation of thiosulfate by reactive oxygen species during inflammation process. The ability of *C. jejuni* to use tetrathionate and thiosulfate in metabolism promotes its growth in the gut of the host organism[6].

6. Materials and methods

General methodologies and reagents

The pH of solutions and buffers was measured with a Sartorius pH meter. Analytic balance (KERN) and Precision balance (KERN) were also used. Crystals were observed using a Leica MZ16 stereomicroscope.

Reagents

- Ammonium Chloride (3.5 M Molecular Dimensions)
- CryoMix2 (CryoProtX screen)
- Imidazole (1 M – Molecular Dimensions)
- Malic Acid (1 M – Molecular Dimensions)
- Mount loop (MiTeGen)
- Nextal plates
- PEG 3350 (50% (*v/v*) - Molecular Dimensions)
- PEG 400 (50% (*v/v*) - Molecular Dimensions)
- PEG 4000 (50% (*v/v*) - Molecular Dimensions)
- Sodium borate (0.1 M – Molecular Dimensions)
- Sodium citrate (1.6 M – Hampton Research)

Crystallization and data collection

The "as isolated" TsdA at a concentration of 3.5 mg/ml in 50 mM BisTris pH 6.5 was crystallized using the hanging drop, vapour diffusion method. 3-µl drop, obtained by mixing 2 µl protein and 1 µl precipitant solution (18% (*v/v*) PEG 3350 and 0.2 M NH_4Cl (with the final pH set to 6.2 with HCl) were equilibrated against 500 µl of precipitant solution at 20ºC. Crystals appeared after 7 days and grew to maximum dimensions of ~0.3x0.1 mm. Before data collection, crystals were pulled from the drop using a micro-mount loop (MiTeGen) and transferred to a cryoprotective solution composed by 25% (*v*/v) PEG 3350, 0.2 M NH_4Cl and 5% (*v/v*) PEG 400.

The C138G TsdA mutant form concentrated at 3.3 mg/ml in 50 Mm BisTris pH 6.5 was crystallized in the same conditions as the "as isolated" TsdA and crystals were cryo-protected using a similar procedure as described above. Crystals appeared after 3 days and grew up to maximum dimensions of ~0.2x0.1 mm.

The N254G TsdA variant at of 4 mg/ml in 50 Mm BisTris pH 6.5 was crystallized using the hanging drop, vapour diffusion method. 2-µl drop, obtained by mixing of 1 µl protein and 1 µl precipitant solution (15% (*v/v)* PEG 4000 and 0.2 M imidazole/malate (with the final pH adjusted to 5.7 with HCl) were equilibrated against 500 µl of precipitant solution at 20ºC. Crystals appeared after 24 days and

grew up to ~0.2x0.2 mm. Crystals were cryoprotected with CryoMix2 solution from the CryoProtX screen, which is composed of 25 % (*v/v*) diethylene glycol 25 % (*v/v*) glycerol and 25 % (*v/v*) 1.2-propanediol.

Complete X-ray diffraction data sets were collected at a wavelength of 0.978 Å in beamline I04 at Diamond Light Source synchrotron (Didcot, Oxfordshire, UK). Data were indexed, integrated and scaled using *XDS* and converted to MTZ format with *XDSCONV*[32]. An R_{free} array of reflections corresponding to 5% of the data set were created at this stage for each data set.

Structure determination and refinement

The program *PHASER* as implemented in *PHENIX*[33] suite of programs was used to determine phases by molecular replacement using the *A. vinosum* TsdA model (PDB ID 4WQD) devoid of any solvent cofactors. Iterative manual model building was carried out with *COOT* and refined with *phenix.refine* until a complete model was built with highest refine convergence. The model was validated with the program *MolProbity*[34], as implemented in *PHENIX*. Since all TsdA crystals were isomorphous, the remaining structure determination of TsdA variants proceeded by an initial rigid body refinement with *BUSTER-TNT*[35]. Heme prosthetic groups and coordinating residues were removed from the model for this initial refinement. Program macro "Missing atoms" together with the "-L" flag (presence of an unknown ligand" in *BUSTER-TNT*), was used to render clear electron density in those regions. All figures were created with the program *PyMOL*[36].

7. Results and Discussion

The “as isolated”, C138G and N254G variants of TsdA from *Campylobacter jejuni* were crystallized using polyethylene glycol at pH around 6, as described in materials and methods. Crystals were subjected to data collection at Diamond synchrotron. The “as isolated”, C138G and N254G *C. jejuni* TsdA structures were obtained using molecular replacement strategy to solve the phase problem. The 3D-structure of the homologous enzyme from *Allochromatium vinosum* was used as search template (PDB entry 4WQD)[37] .

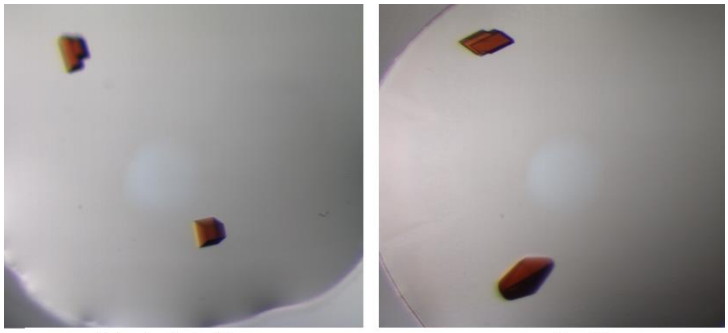

Figure 7.1 Crystals of "as isolated" and N254G variants

X-ray data collection, refinement statistics and model quality parameters are depicted in table 7.1.

Crystal of “as isolated” form belongs to the orthorhombic space group $P\ 2_12_12_1$, with two molecules in the asymmetric unit and unit cell dimensions a=294.1 b=44.9 and c=48.0 Å. Crystal solvent content is around 44% and Matthews coefficient[38] is 2.21Å^3 Da^{-1}. The structure was refined to 1.95 Å with R_{crys} of 17.5% and R_{free} of 22.1% and comprises Ser^{15}-Met^{321} residues, 2 hemes, 1 ethylene glycol molecule, 1 citrate ion and 185 waters. The citrate [$C_3H_5O(COO)_3^{3-}$], presented in the crystallization condition, is located between the two monomers of the asymmetric unit mediating contacts between them. This fact might indicate that citrate stabilizes the interface between the two monomers thus facilitating the protein crystallization.

The C138G crystal belongs to monoclinic space group $P\ 2_1$, with unit cell parameters a=71.2, b=47.1, c=90.5 Å and β=90.3º. Similarly to the “as isolated” form, the C138G variant contains a dimer in the asymmetric unit, that corresponds to Matthews coefficient of 2.10 Å^{3}Da^{-1} and a solvent content 42%. The structure was refined to 2.37 Å with R_{crys} of 18.6% and R_{free} of 21.7% and comprises Ser^{15}-Met321 residues, two hemes, one ethylene glycol molecule and 185 waters.

The N254G crystal belongs to tetragonal space group $P\ 4_1\ 2_1\ 2$, with unit cell parameters a=b=45.9 and c=293.02 Å. Contrary to the previous structures, N254G structure only contains one monomer in the asymmetric unit with a Matthews coefficient of 2.15 Å^3 Da^{-1} and solvent content of 43%. This structure was refined to 1.72 Å with R_{crys} of 16.6% and R_{free} of 18.7%. The model contains

Asp[17]- Gly[254] and Leu[262]-Ile[322] residues, two hemes, one ethylene glycol, one imidazole, one thiosulfate, one propylene glycol molecule and 210 waters. The ethylene glycol and polyethylene glycol molecules are present in the cryo solution, the imidazole was used during purification and the thiosulfate molecule was probably sequestered during expression. It was not possible to model the loop between Gly[255]-Asp[261]. The same observation was reported by Brito *et al.*[37] in some of *A. vinosum* TsdA structures: this loop is solvent exposed, highly disordered, and no electron density is observed for it.

Table 7.1 Data collection and refinement statistic for *C. jejuni* TsdA structures.

Data collection	**"As isolated"**	**C138G mutant**	**N254G mutant**
Synchrotron	Diamond Light Source		
Beamline	I 04		
Wavelength (Å)	0.978		
Resolution range (Å)	45.61-1.95 (2.02-1.95)	35.17-2.37 (2.45-2.37)	45.35-1.72 (1.78-1.72)
Space group	$P\ 2_1\ 2_1\ 2_1$	$P\ 1\ 2_1\ 1$	$P\ 4_1\ 2_1\ 2$
Unit cell parameters			
a, b, c (Å)	294.1, 44.9, 48.0	71.2, 47.1, 90.5	45.9, 45.9, 293.0
α, β, γ (degrees)	90, 90, 90	90, 96.3, 90	90, 90, 90
Total reflections	204650 (20336)	62706 (5294)	494378 (2618)
Unique reflections	47203 (4541)	23393 (2087)	34861 (3368)
Multiplicity	4.3 (4.4)	2.7 (2.4)	14.1 (8.0)
Completeness (%)	99 (100)	93 (92)	1.00
I/σ(I)	11.0 (1.79)	19.6 (4.89)	12.7 (2.1)
Wilson B-factor(Å^2)	25.9	30.7	20.1
R_{merge} (%)	9.1 (69.0)	4.0 (18.0)	12.0 (73.0)
R_{meas} (%)	10.2 (78.0)	5.0 (23.0)	12.8 (72.1)
$CC_{1/2}$	0.998 (0.80)	1.00 (0.94)	0.98 (0.78)
R_{work}	0.17 (0.28)	0.18 (0.22)	0.17 (0.27)
R_{free}	0.22 (0.32)	0.22 (0.32)	0.19 (0.28)
Refinement			
Number of atoms			
Non-hydrogen atoms	5356	5104	2679
Ligands	193	176	116
Protein residues	625	613	299
RMSD (bond) (Å)	0.01	0.003	0.01
RMSD (angles) (º)	1.33	1.05	1.27
Ramachandran favoured (%)	98	97	97
Ramachandran allowed (%)	1.9	2.6	2.7
Ramachandran outliers (%)	0	0	0
Average B-factor(Å^2)	34.71	35.35	26.82
Ligands	30.09	33.84	27.51
Solvent	35.28	31.42	35.13

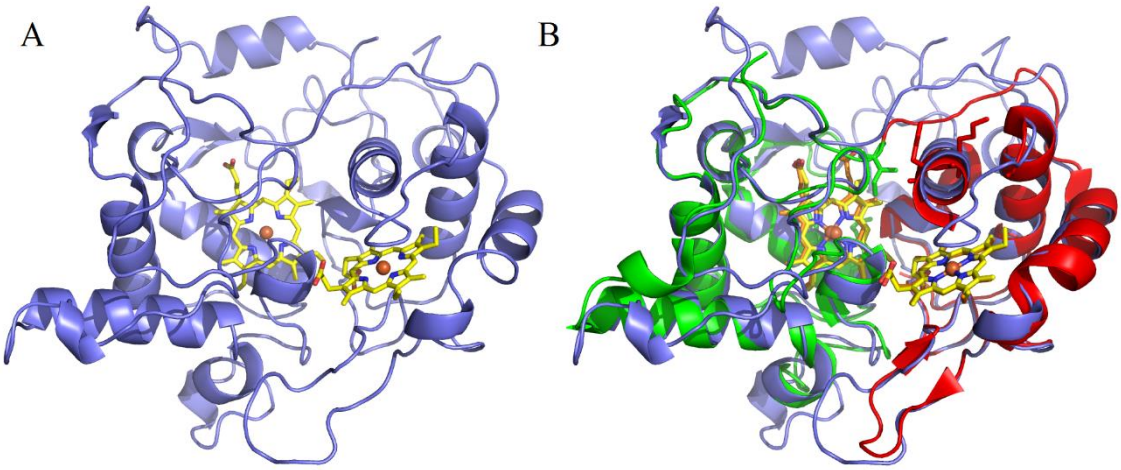

Figure 7.2 – Structural comparison of *C. jejuni* TsdA and *A. vinosum* TsdA. TsdA *C. jejuni:* Chain A is represented in blue; the heme prosthetic groups, are shown is sticks and coloured by atom type (yellow for carbon, blue for nitrogen, red for oxygen, and dark red for iron; this colour code will be used all figures). B, *C. jejuni* TsdA superposed with *A. vinosum* TsdA (PDB 4WQD), for which the N-terminal domain is coloured in green and C-terminal domain in red; heme colour code is the same as in A. The figures were created through PyMol program.

TsdA from *C. jejuni* (CjTsdA) is a heart-shaped molecule with approximate dimensions of 50x54x40 $Å^3$ (Figure 7.2 A). Each domain is composed by four α-helices that surround each heme molecule. The structure coordinates were submitted to the DALI server[39] that shows highest similarity with *Allochromatium vinosum* TsdA (AvTsdA). This TsdA structure (PDB entry 4WQD) was previously solved in our laboratory and shows 43% of sequence identity with CjTsdA and an RMSD of 1.2 Å for 223 aligned Cα atoms giving a Z-score of ~29 (two structures are considered similar with Z-scores above 10). Comparison of these structures (figure 7.2 B) shows that the α-helices around the hemes are highly superposed. This fact indicates that CjTsdA, such as AvTsdA, is also organized in two domains that are related by a pseudo-2-fold symmetry axis. The SoxAX complex from *Rhodovulum sulfidophilum* (PDB entry 2OZ1) and *Paracoccus pantotrophus* (PDB entry 2C1D) also shows to be similar with CjTsdA. These two structures show 22% of sequence identity with CjTsdA and an RMSD of 2.3 and 2.8 Å for 103 and 104 aligned Cα atoms with giving a Z-score of ~11 and ~7, respectively.

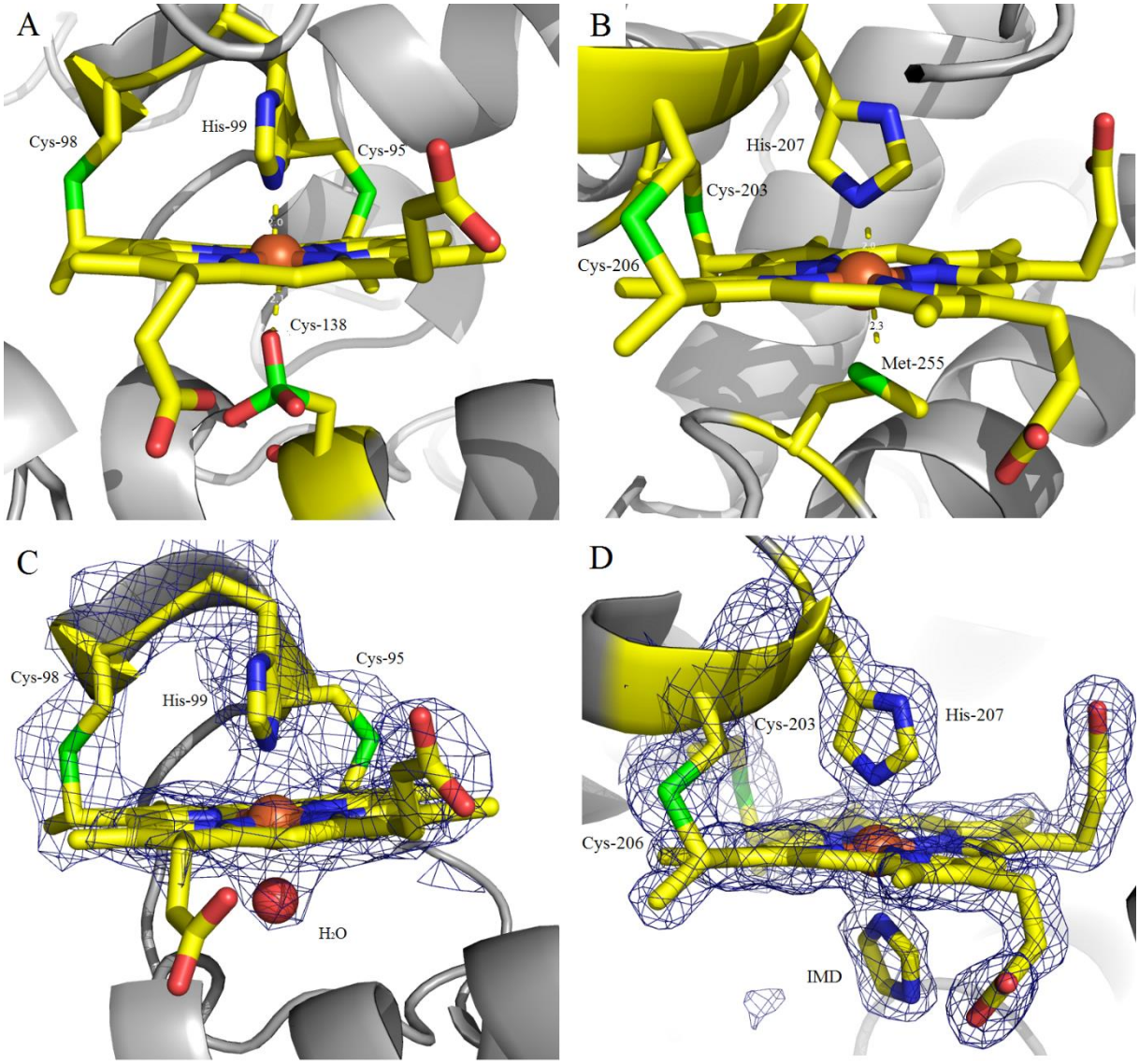

Figure 7.3 - Heme coordination of the "as isolated", C138G and N254G TsdA structures. A, heme 1 in "as isolated" is coordinated by His^{99} and oxidised Cys^{138}. B, "as isolated" heme 2 is coordinated by His^{207} and Met^{255}. C, C138G heme 1 is coordinated by His^{99} and a water molecule. D, N254G heme 2 is coordinated by His^{207} and an imidazole molecule. The figures were created through PyMol program.

C. jejuni TsdA contains two heme groups that are covalently bound to the polypeptide chain. The distance between the two heme irons is around 15.3 Å. Heme 1 is covalently bound to Cys^{95} and Cys^{98} and heme 2 to Cys^{203} and Cys^{206} by thioether bonds. Both heme irons are hexacoordinated. The proximal axial ligand of "as isolated" structure heme 1 iron is His^{99} and distal ligand is Cys^{138}. Cys^{138} shows a triple oxidation state in the Chain A (Figure 7.3 A) and double oxidation state in the Chain B. Cysteine oxidation is an irreversible process that occurred in the absence of reducing agents, e.g. TCEP. Heme 2 is axially coordinated by His^{207} and Met^{255}. In the C138G mutant Cys^{138} the cysteine distal ligand is substituted by a water molecule. In N254G mutant an imidazole molecule is coordinating the heme 2 iron rather than Met^{255}.

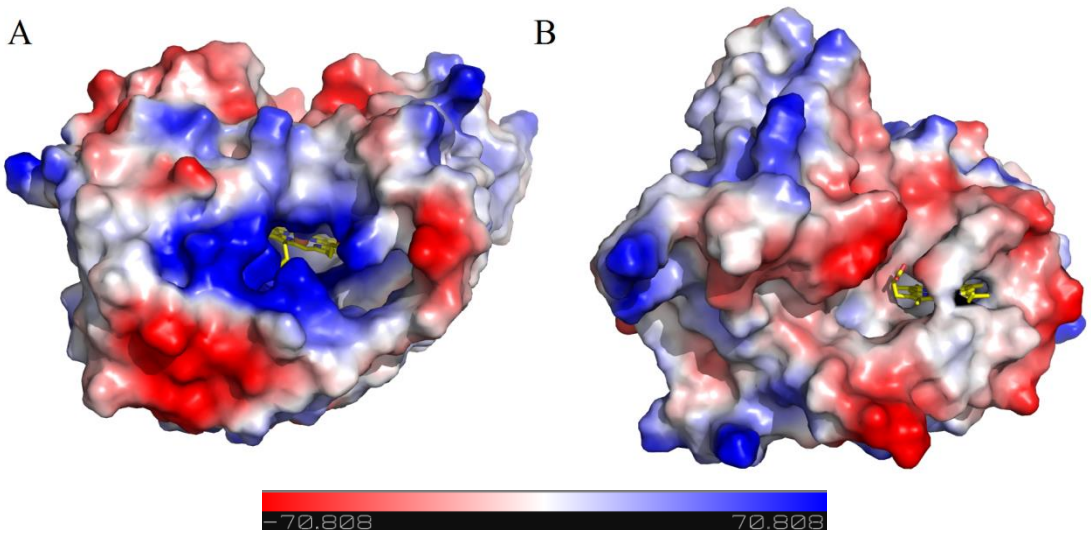

Figure 7.4 - TsdA electrostatic surface potential. A, electrostatic surface potential around heme1, B, electrostatic surface potential around heme 2. Positive charge is represented in blue and negative charge in red.

The electrostatic surface potential of CjTsdA was calculated with the program *PyMol* (figure 7.4). Heme 1 lies in the bottom of a positively charged cavity and is slightly buried inside the protein; in contrast, heme 2 is more solvent exposed and displays a positively charged patch surrounding more hydrophobic core close to the plane of the heme. One possible explanation for these observations is that the positively charged residues around heme 1 cavity attract the negatively charged substrate molecule $S_2O_3^{2-}$ and help to release the negatively charged product $S_4O_6^{2-}$; the same is valid for the reverse tetrathionate reductase reaction. In base of AvTsdA mechanism[37], we suppose that two thiosulfate molecules enter the active site inducing the Cys^{138} to tilt away from iron coordination sphere. Two electrons are released during the direct reaction between two thiosulfate molecules that reduce heme 1 and heme 2. The resultant tetrathionate molecule is released, and the hemes are oxidized. In case of tetrathionate reductase activity, two electrons that are provided from periplasmic cytochrome reduce heme 2 and heme 1. Tetrathionate molecule enters into the active site and receives two electrons that reduce tetrathionate to two thiosulfate molecules. After tetrathionate release, Cys^{138} gets back to the iron coordination sphere. Due the absence of electron density in heme 2, it was not possible to model Met^{255} - Asp^{261} and we cannot draw conclusions about iron coordination in the heme 2 during the reaction.

The aim of this work was achieved. It was possible to model three variants of CjTsdA. “As isolated”, C138G and N254G mutants were crystallized by the vapour diffusion method and X-ray diffraction data was collected at a synchrotron source to 1.95, 2.37 and 1.72 Å resolution.

Conclusion

It is important to mention that almost all histidine kinase targets were successfully cloned and showed a good expression level. The solubilization step was a problematic step in this study. Despite the innumerous conditions that were tested, solubilization was not significantly improved. Purification tests were conditioned by the low amount of protein available in the solubilized membrane fraction. Most of the protein did not presented affinity to the Ni-NTA column, which leads to a low quantity of the target protein eluted along with other *E. coli* proteins contaminating the sample. It was not possible to obtain pure protein for the functional and structural studies. In future work, other *S. aureus* and *C. difficile* targets should proceed with expression, solubilization and purification tests.

In the second part of the work, the aims were successfully achieved. It was possible to model three different variants of thiosulfate dehydrogenase. "As isolated", C138G and N254G TsdAs were crystallized by the vapour diffusion method in different conditions. X-ray diffraction data were collected at a synchrotron source to 1.95, 2.37 and 1.72 Å resolution, respectively. In the future work several crystallization trials in the presence of ligands will be performed, such as the optimization of the M255G mutant crystallization. These experiments will provide the information that is necessary to propose *C. jejuni* TsdA mechanism of reaction.

References

1. Bertani, G. Lysogeny at Mid-Twentieth Century: P1, P2, and Other Experimental Systems. *J. Bacteriol.* **186,** 595–600 (2004).
2. Darnell, H. L. / A. B. / P. M. / C. A. K. / M. K. / M. P. S. / L. Z. / J. *Molecular Cell Biology By Lodish, Berk, etc.* (WHFreeman, 2004).
3. Brito, J. A. & Archer, M. in *Practical Approaches to Biological Inorganic Chemistry* (ed. Louro, R. R. C. O.) 217–255 (Elsevier, 2013).
4. Rhodes, G. in *Crystallography Made Crystal Clear (Third Edition)* (ed. Rhodes, G.) 31–47 (Academic Press, 2006). at <http://www.sciencedirect.com/science/article/pii/B9780125870733500052>
5. Rhodes, G. in *Crystallography Made Crystal Clear (Third Edition)* (ed. Rhodes, G.) 7–30 (Academic Press, 2006). at <http://www.sciencedirect.com/science/article/pii/B9780125870733500040>
6. Liu, Y.-W., Denkmann, K., Kosciow, K., Dahl, C. & Kelly, D. J. Tetrathionate stimulated growth of Campylobacter jejuni identifies a new type of bi-functional tetrathionate reductase (TsdA) that is widely distributed in bacteria. *Mol. Microbiol.* **88,** 173–188 (2013).
7. Laub, M. T. & Goulian, M. Specificity in two-component signal transduction pathways. *Annu. Rev. Genet.* **41,** 121–145 (2007).
8. Falke, J. J., Bass, R. B., Butler, S. L., Chervitz, S. A. & Danielson, M. A. The two-component signaling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes. *Annu. Rev. Cell Dev. Biol.* **13,** 457–512 (1997).
9. Wolanin, P. M., Thomason, P. A. & Stock, J. B. Histidine protein kinases: key signal transducers outside the animal kingdom. *Genome Biol.* **3,** REVIEWS3013 (2002).
10. Stock, A. M., Robinson, V. L. & Goudreau, P. N. Two-component signal transduction. *Annu. Rev. Biochem.* **69,** 183–215 (2000).
11. Liang, X. *et al.* Global regulation of gene expression by ArlRS, a two-component signal transduction regulatory system of Staphylococcus aureus. *J. Bacteriol.* **187,** 5486–5492 (2005).
12. Fournier, B., Klier, A. & Rapoport, G. The two-component system ArlS-ArlR is a regulator of virulence gene expression in Staphylococcus aureus. *Mol. Microbiol.* **41,** 247–261 (2001).
13. Sakoulas, G., Moellering, R. C. & Eliopoulos, G. M. Adaptation of methicillin-resistant Staphylococcus aureus in the face of vancomycin therapy. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* **42 Suppl 1,** S40–50 (2006).
14. Liang, X. *et al.* Inactivation of a two-component signal transduction system, SaeRS, eliminates adherence and attenuates virulence of Staphylococcus aureus. *Infect. Immun.* **74,** 4655–4665 (2006).
15. Schlag, S. *et al.* Characterization of the Oxygen-Responsive NreABC Regulon of Staphylococcus aureus. *J. Bacteriol.* **190,** 7847–7858 (2008).
16. Pasupuleti, M., Schmidtchen, A. & Malmsten, M. Antimicrobial peptides: key components of the innate immune system. *Crit. Rev. Biotechnol.* **32,** 143–171 (2012).
17. Falord, M., Karimova, G., Hiron, A. & Msadek, T. GraXSR Proteins Interact with the VraFG ABC Transporter To Form a Five-Component System Required for Cationic Antimicrobial Peptide Sensing and Resistance in Staphylococcus aureus. *Antimicrob. Agents Chemother.* **56,** 1047–1058 (2012).
18. Damron, F. H. *et al.* Analysis of the Pseudomonas aeruginosa Regulon Controlled by the Sensor Kinase KinB and Sigma Factor RpoN. *J. Bacteriol.* **194,** 1317–1330 (2012).

19. Kleerebezem, M. Quorum sensing control of lantibiotic production; nisin and subtilin autoregulate their own biosynthesis. *Peptides* **25,** 1405–1414 (2004).
20. Chakicherla, A. *et al.* SpaK/SpaR two-component system characterized by a structure-driven domain-fusion method and in vitro phosphorylation studies. *PLoS Comput. Biol.* **5,** e1000401 (2009).
21. Kawada-Matsuo, M., Yoshida, Y., Nakamura, N. & Komatsuzawa, H. Role of two-component systems in the resistance of Staphylococcus aureus to antibacterial agents. *Virulence* **2,** 427–430 (2011).
22. Hiron, A., Falord, M., Valle, J., Débarbouillé, M. & Msadek, T. Bacitracin and nisin resistance in Staphylococcus aureus: a novel pathway involving the BraS/BraR two-component system (SA2417/SA2418) and both the BraD/BraE and VraD/VraE ABC transporters. *Mol. Microbiol.* **81,** 602–622 (2011).
23. Ammam, F. *et al.* The functional vanGCd cluster of Clostridium difficile does not confer vancomycin resistance. *Mol. Microbiol.* **89,** 612–625 (2013).
24. Boneca, I. G. & Chiosis, G. Vancomycin resistance: occurrence, mechanisms and strategies to combat it. *Expert Opin. Ther. Targets* **7,** 311–328 (2003).
25. Deaths involving MRSA. *Office for National Statistics* (2013).
26. Goldberg, E. J. *et al.* Clostridium difficile infection: A brief update on emerging therapies. *Am. J. Health. Syst. Pharm.* **72,** 1007–1012 (2015).
27. Ma, P. *et al.* Expression, purification and activities of the entire family of intact membrane sensor kinases from Enterococcus faecalis. *Mol. Membr. Biol.* **25,** 449–473 (2008).
28. Morel, F. M. M. & Price, N. M. The Biogeochemical Cycles of Trace Metals in the Oceans. *Science* **300,** 944–947 (2003).
29. Urich, T. *et al.* The sulphur oxygenase reductase from Acidianus ambivalens is a multimeric protein containing a low-potential mononuclear non-haem iron centre. *Biochem. J.* **381,** 137–146 (2004).
30. Grabarczyk, D. B. *et al.* Mechanism of thiosulfate oxidation in the SoxA family of cysteine-ligated cytochromes. *J. Biol. Chem.* **290,** 9209–9221 (2015).
31. Denkmann, K. *et al.* Thiosulfate dehydrogenase: a widespread unusual acidophilic c-type cytochrome. *Environ. Microbiol.* **14,** 2673–2688 (2012).
32. Kabsch, W. *XDS*. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132 (2010).
33. Terwilliger, T. C. *et al.* Decision-making in structure solution using Bayesian estimates of map quality: the *PHENIX AutoSol* wizard. *Acta Crystallogr. D Biol. Crystallogr.* **65,** 582–601 (2009).
34. Chen, V. B. *et al. MolProbity* : all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 12–21 (2010).
35. Blanc, E. *et al.* Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2210–2221 (2004).
36. Chen, F. PyMOL. (2014).
37. Brito, J. A., Denkmann, K., Pereira, I. A. C., Archer, M. & Dahl, C. Thiosulfate dehydrogenase (TsdA) from Allochromatium vinosum: structural and functional insights into thiosulfate oxidation. *J. Biol. Chem.* **290,** 9222–9238 (2015).
38. Matthews, B. W. Solvent content of protein crystals. *J. Mol. Biol.* **33,** 491–497 (1968).
39. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38,** W545–W549 (2010).

Appendix

Appendix A

Table A.1 Target gene information

Uniprot	Target	*Organism*	Gene size (b)	Amino acid number	Protein size (kDa)
Q5HG05	ArlS	MRSA *(COL)*	1356	451	52,4
Q5HHW5	SaeS	MRSA *(COL)*	1056	351	39,713
Q5HCS4	BceS	MRSA *(COL)*	888	295	34,019
Q5HFT1	SrrB	MRSA *(COL)*	1752	583	66,075
Q5H3G3	AgrC	MRSA *(COL)*	1293	430	49,628
Q5HJX6	KinB	MRSA *(COL)*	1827	608	69,923
Q5HF81	PhoR	MRSA *(COL)*	1665	554	63,904
Q5HI08	GraS	MRSA *(COL)*	1041	346	41,078
Q5HJF6	Unk	MRSA *(COL)*	1557	518	61,046
Q5HDG4	NreB	MRSA *(COL)*	1035	344	39,12
Q186I0	VanS	*C.difficile (630)*	1143	380	44,721
Q188S3	SpaK	*C.difficile (630)*	1425	474	54,321
Q182X0	PhoR	*C.difficile (630)*	1290	429	49,012
Q18BY5	SrrB	*C.difficile (630)*	1254	417	48,199
Q186P7	Unk	*C.difficile (630)*	1233	410	47,584
Q186D3	GraS	*C.difficile (630)*	1857	618	71,996
Q186Z1	KinB	*C.difficile (630)*	1434	477	54,509
Q181K0	SaeS	*C.difficile (630)*	1194	397	45,927
Q18C51	BceS	*C.difficile (630)*	1005	334	39,606
Q189F2	AgrC	*C.difficile (630)*	1338	445	51,563

Table A.2 Primers for MRSA target genes

Target	Organism	Forward	Reverse
ArlS	MRSA *(COL)*	TCCCCCCGGGATGACAAAACGTAAATTGCGCAATAACTGG	TATGAGCTCAAATATGATTTTAAACGTTGTTCCTTTG
SaeS	MRSA *(COL)*	TCCCCCCGGGATGGTGTTATCAATTAGAAGTCAA	TATGAGCTCTGACGTAATGTCTAATTTGTGTAATGT
BceS	MRSA *(COL)*	TCCCCCCGGGATGACCTTTCTTAAAAGTATTACTCAG	TATGAGCTCTTCATCTGGAAATTGAATCGTAAATGTTG
SrrB	MRSA *(COL)*	TCCCCCCGGGATGATGAGCCGGCTAAATAGTGTCGTAATTAAACTGTG	TATGAGCTCTTCTGGTTTTGGTAGTTTAATAATAAATGTTGTGCC
AgrC	MRSA *(COL)*	TCCCCCCGGGATGGAATTATTAAATAGTTATAATTTTG	TATGAGCTCGTTGTTAATAATTTCAACTTTTT
KinB	MRSA *(COL)*	TCCCCCCGGGATGAAGTGGCTAAAACAACTACAAT	TATGAGCTCTTCATCCCAATCACCGTCTTCAAT
PhoR	MRSA *(COL)*	TCCCCCCGGGATGATGAAGTTTCACCACCGCTTAATGTTAT	TATGAGCTCTTCTTTATAATCTTTTAGAATAACTTTGAAC
GraS	MRSA *(COL)*	TCCCCCCGGGATGAATAATTTGAAATGGGTAGCTT	TATGAGCTCAAATGACAAATTTGTCACTTCCGACATGC
Unk	MRSA *(COL)*	TCCCCCCGGGATGACGGCATACAAACCTTATAGACATCAATTAAGACGA	ATAAGAATGCGGCCGCAACATCTACATTCCCCCTTGAAAGTGGT
NreB	MRSA *(COL)*	TCCCCCCGGGATGATTAATGAGGACAGTATACAGTTAGATACTTTATTGA	TATGAGCTCAATTGGAATGTTCAATGTAACATTGGTACCCTCAC

Table A.3 Primers for *C. difficile* targets

Target	Organism	Forward	Reverse
VanS	*C.difficile (630)*	TCCCCCCGGGATGAAAGGTAAAAATATTAAAGTCAATCCTGACTTTACTC	TATGAGCTCCAAATTTCCTACAATCGGAATAATAACCTT
SpaK	*C.difficile (630)*	TCCCCCCGGGATGAAAAATCAATCCCTTATTACCCAATTTAGACATAC	TATGAGCTCAATGCTATGCCAAAATTTCACACATGCTCCACCTGAT
ZPhoR	*C.difficile (630)*	TCCCCCCGGGATGGAAGATTTAAATAGTATATATTTTTTTATTGTT	TATGAGCTCTTTAAGGAAATTTTTTTGTAATATTGTAACTATAAATTCCGCTACCA
SrrB	*C.difficile (630)*	TCCCCCCGGGATGAAAATAGTGTTTCTATATAATCCAGAAGTGAAGAAG	TATGAGCTCGTCACTATCATACATTTTATGAAATATTATATGAAATTCTG
Unk	*C.difficile (630)*	TCCCCCCGGGATGATAAATAAGAATGTATTTACCAGTACTAAAA	TATGAGCTCTTTTAACTTTTTTGGTATTGACAATTCAAATGTAG
GraS	*C.difficile (630)*	TCCCCCCGGGATGGATACCCATAATAAATATGT	TATGAGCTCTAAATCCGAGAACTCTATTTCTAAT
KinB	*C.difficile (630)*	TCCCCCCGGGATGTTACATTTTGAAGGTCTGAGAAAAAAAGT	TATGAGCTCTGCTTGTTTAGCAAATTTATTTTTTATATTAAAAGTT
SaeS	*C.difficile (630)*	TCCCCCCGGGATGGATATAAAATTGAAAAATAAAAGTTTATTAAATC	TATGAGCTCAGCTGATTGTATTATTGGCAGTTGTA
BceS	*C.difficile (630)*	TCCCCCCGGGATGAGCTTTAAAGATTTTCTCAAGG	TATGAGCTCTAAATCACATTTGAATGTAATAGTA
AgrC	*C.difficile (630)*	TCCCCCCGGGATGGGCACTTATTCTTTAATCTCTAC	TATGAGCTCAAATAATCTAATATAAAGAAATAAAG

Note: CCCGGG – Xma I restriction site; GAGCTC – Sac I-HF restriction site; GCGGCCGC – Not I-HF restriction site.

Appendix B

Protocol nr 1: DNA gel extraction (Zymoclean™ Gel DNA Recovery Kit)

DNA band was cut from agarose gel and weighed. Three volumes of Agarose Dissolving Buffer were added to each volume of gel (for 100 mg gel slice add 300 µl of buffer). Sample was incubated at 45ºC until the gel was dissolved. Mixture was loaded into a Zymo-Spin column, placed into a 2 ml Collection Tube, centrifuged at full speed (10 000 xg) for 30 sec and flow through was discarded. 200 µl of Wash Buffer were added to the column and centrifuged for 30 sec twice. Zymo-Spin Column was placed into a new 1.5 ml eppendorf tube and 10 µl of DNA Elution Buffer were added directly to the column matrix and centrifuged to elute DNA (10 000 xg during 15 sec).

Analysis of results

0.8% Agarose gel (100 ml):

0.8 g of agarose was dissolved in 100 ml 1x TAE in the microwave. 1x TAE is obtained by dissolution of 50x TAE 50x [1 L: Tris-Acetate (242 g), 20 mM EDTA (18 g), Acetic acid (57.1 ml) pH 7,5]. After cooling, 1 µl of SyBr Safe was added. Gel runs at 100 V for 25 min.

Protocol nr 2: SDS-PAGE

SDS-PAGE gel:

Running gel was prepared according to table 5.4 and was transferred to specific support. While gel polymerises, stacking gel is prepared (according to table 5.4) and transferred to the support under the running gel.

Table B.1 SDS-PAGE gel preparation

	10%	12%
	Running gel	
H_2O	3.4 ml	4.1 ml
1,5 M Tris-HCl pH 8.8	2.5 ml	2.5 ml
20% (w/v) SDS	0.05 ml	0.05 ml
Acrylamide/Bis-acrylamide (30%/0.8% w/v)	3.3 ml	4.0 ml
10% (w/v) ammonium persulfate (APS)	0.05 ml	0.05 ml
TEMED	0.005 ml	0.005 ml
Total	10.005 ml	
	Stacking gel	
H_2O	3,075 ml	
1,5 M Tris-HCl pH 6.8	1.25 ml	
20% (w/v) SDS	0.025 ml	
Acrylamide/Bis-acrylamide (30%/0.8% w/v)	0.67 ml	
10% (w/v) ammonium persulfate (APS)	0.025 ml	
TEMED	0.005 ml	
Total	5.05 ml	

Gel is submerged in 1x Running buffer (10xRunning buffer (1 L): 30.3 g Tris base, 144 g Glycine, 10 g SDS). 5 µl of loading buffer (2x Loading Buffer: 35.5 ml MilliQ water, 12.5 ml 0.5 Tris-HCl pH 6.8, 25 ml Glycerol, 20 ml 10% (w/v) SDS, 2 ml 0.5% (w/v) bromophenol blue, 5% β-mercaptoethanol) were added to 15 µl of each sample and incubated at 37ºC for 30 min. Samples were loaded in the gel. Gels runs at 300 mA and 180 W during 1 hour. Gel was revealed by incubation with Bio-Safe Coomassie (Bio-Rad).

Protocol nr 3: Western Blot

An SDS-PAGE gel is incubated for 5 min in transfer buffer and a nitrocellulose membrane is activated in methanol during 30 sec. Both are placed in the transfer cassette according to figure B.1. Transfer protocol was selected (Medium Weight Program Trans-blot Turbo Transfer System- BioRAD) during 7 min. After transfer, membrane is washed with TBST or TBSTT for 5-10 min and blocked with 3% low fat dry milk in TBST/TBSTT at room temperature for 40-60 min. Membrane is washed twice with TBST/TBSTT for 5-10 min and incubated with probe diluted 1:10000 in TBST/TBSTT for 1 hour. Washing step is repeated three times with 10 ml TBST/TBSTT for 5 min and BCIP/NBT Liquid Substrate System is used to reveal western blot results.

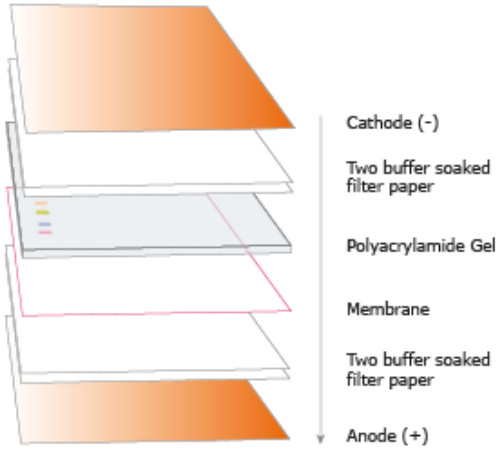

Figure B.1 Western blot placing sample scheme

10xTransfer buffer (5 L): 145 g Tris base, 725 g Glycine

1xTransfer buffer (3 L): 700 ml EtOH, 300 ml 10x Transfer buffer

TBST: 50 mM Tris pH= 7,6, 150 mM NaCl, 0,05% Tween 20

TBSTT: 50 mM Tris pH= 7,6, 150 mM NaCl, 0,05% (v/v) Tween 20, 0,2% (v/v) Triton X-100

Protocol nr 4: InVision His-tag In-gel Stain

After running, an SDS-PAGE gel is fixed with 200 ml of fixing solution (40% methanol and 20% acetic acid) for 1 hour and washed twice with ultrapure water during 10 min. Then, gel was incubated in 25 ml ready-to-use solution of InVision His-tag In-gel Stain for 1 hour and was washed twice with 200 ml of 20 mM phosphate buffer pH 7.8 for 10 min. Gel was visualized in the Fuji scanner.

Media:

- LB agar (1 L): 10 g NaCl, 10 g Tryptone, 5 g Yeast extract, 20 g agar
- LB broth (1 L): 10 g NaCl, 10 g Tryptone, 5 g Yeast extract
- 2xYT (1 L): 16 g Tryptone, 10 g Yeast extract, 5 g NaCl
- TB (1 L): 12 g Tryptone, 24 g Yeast extract, 4 ml Glycerol, 100 ml of sterilized solution of 0,17 M KH_2PO_4 and 0,72 M K_2HPO_4
- Autoinduction: 1 L of LB media, 1x 5052 solution, 1xNPS solution, 1 mM $MgSO_4$

  50x5052 solution (100 ml): 2,5% glucose (*w/v*) – 2,5 g, 25% glycerol (*w/v*) 25 ml, 10% lactose (*w/v*) 10 g

  20xNPS solution (250 ml): 6,6% $(NH)_4SO_4$ (*w/v*) 16,5 g, 13,6% K_2HPO_4 (*w/v*) 34 g, 14,2% Na_2HPO_4 (*w/v*) 35,5 g
- M9 (1 L): 200 ml of M9 salts, 2 ml of 1 M $MgSO_4$, 20 ml of 20% glycose, 100 µl of 1 M $CaCl_2$.
- M9 salts (1 L): 64 g $NaHPO_4$-7 H_2O, 15 g KH_2PO_4, 2,5 g NaCl, 5 g HN_4Cl.

Appendix C

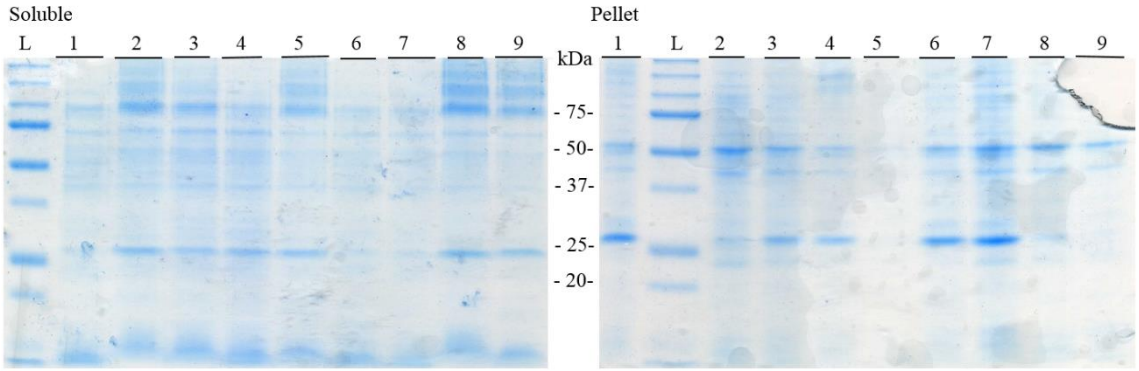

Figure C.1 - 12% SDS-PAGE to analyse ArlS solubilization during 2 hours at 4ºC. 1-Hega 8, 2 - DDM, 3 - Cymal-4, 4 - LDAO, 5 - n-Octyl-β-D-Thioglucopyranoside, 6 - Fos-choline, 7 - Mega-8, 8 - DM and 9 - n-Octyl-β-D-glucopyranoside, L - Precision Plus Protein Standards all blue ladder (BioRAD)

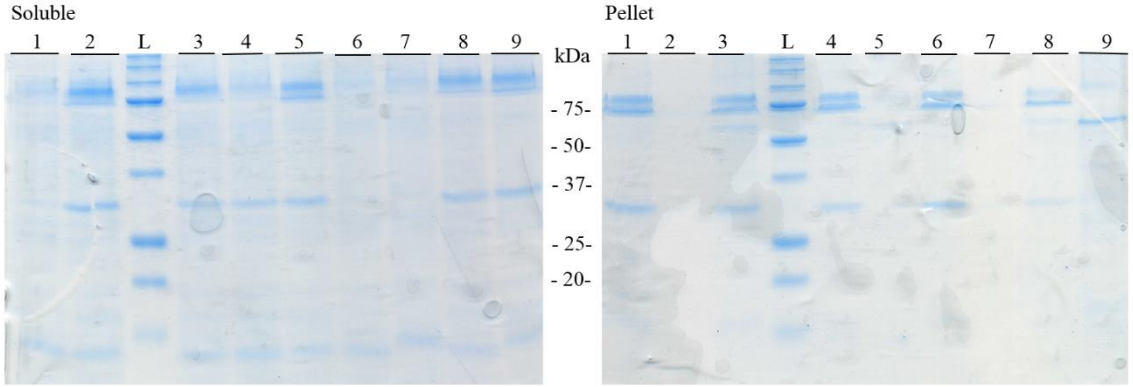

Figure C.2 - 12% SDS-PAGE to analyse ArlS solubilization during 4 hours at 4ºC. 1-Hega 8, 2 - DDM, 3 - Cymal-4, 4 - LDAO, 5 - n-Octyl-β-D-Thioglucopyranoside, 6 - Fos-choline, 7 - Mega-8, 8 - DM and 9 - n-Octyl-β-D-glucopyranoside, L - Precision Plus Protein Standards all blue ladder (BioRAD)

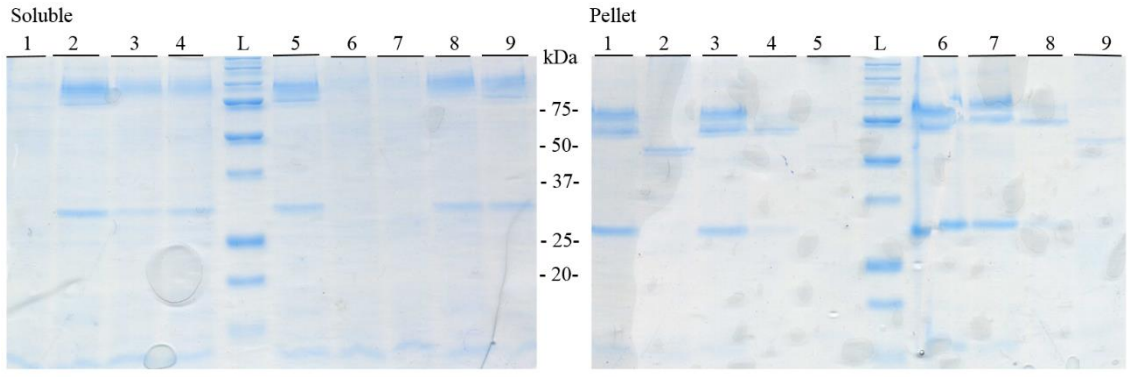

Figure C.3 - 12% SDS-PAGE to analyse ArlS solubilization overnight at 4ºC. 1-Hega 8, 2 - DDM, 3 - Cymal-4, 4 - LDAO, 5 - n-Octyl-β-D-Thioglucopyranoside, 6 - Fos-choline, 7 - Mega-8, 8 - DM and 9 - n-Octyl-β-D-glucopyranoside, L - Precision Plus Protein Standards all blue ladder (BioRAD)

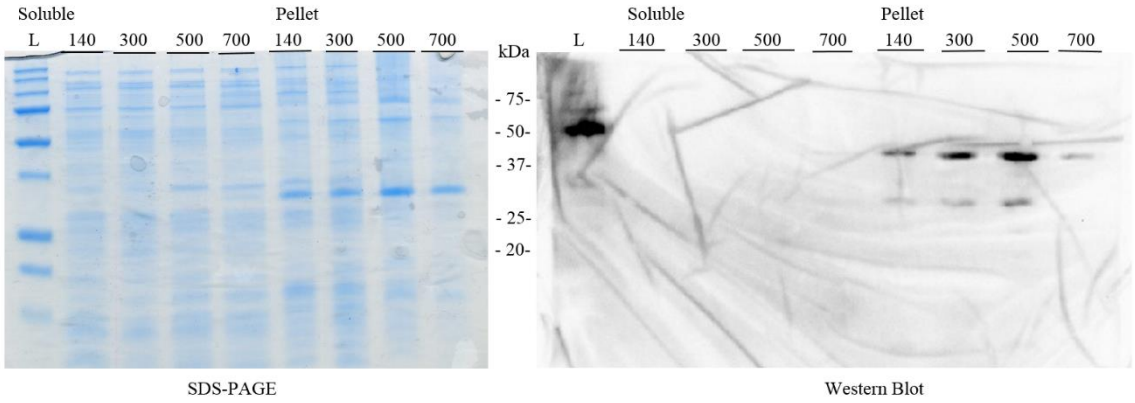

Figure C.4 - 12% SDS-PAGE to analyse ArlS solubilization with LDAO at 4ºC. NaCl concentrations (mM): 250, 500 and 750; L - Precision Plus Protein Standards all blue ladder (BioRAD)

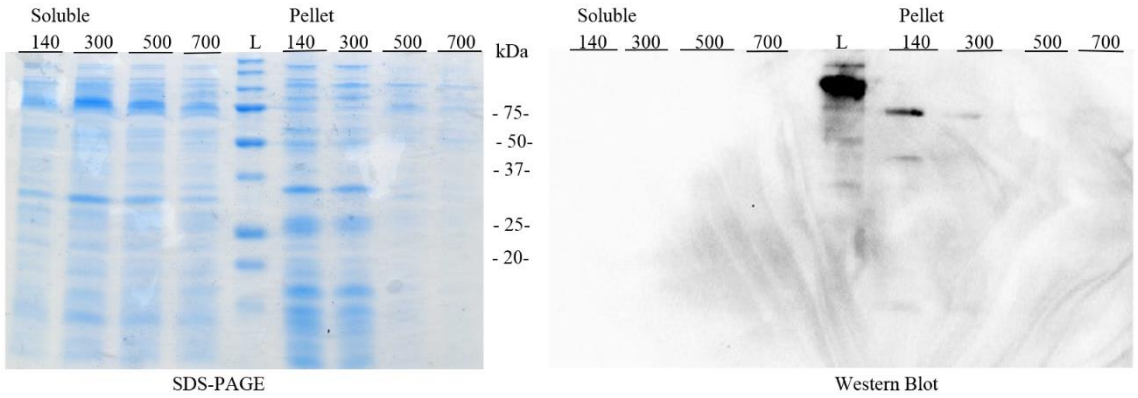

Figure C.5 - 12% SDS-PAGE to analyse ArlS solubilization with Fos-choline at 4ºC. NaCl concentrations (mM): 250, 500 and 750; L - Precision Plus Protein Standards all blue ladder (BioRAD)

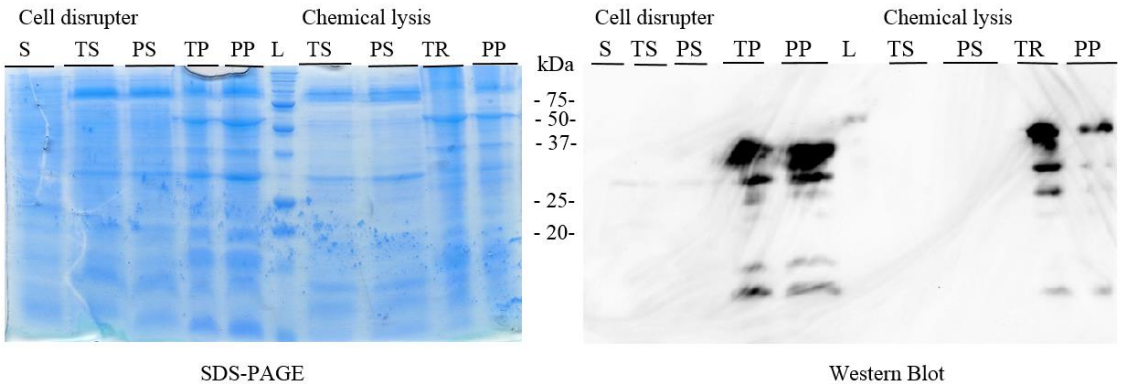

Figure C.6 – 12% SDS-PAGE and Western blot to analyse solubilization of ArlS obtained from cell disrupter and chemical lysis with 1% DDM and 700 mM NaCl at 4ºC. Cell disrupter – cell lysis by cell disruption, Chemical lysis – cell lysis with chemicals, S – soluble fraction; 20 mM Tris pH 7: TS – Tris soluble fraction, TP - Tris insoluble fraction; 10 mM PBS pH 7.4: PS – PBS soluble fraction, PP - PBS insoluble fraction. L- Precision Plus Protein Standards all blue ladder (BioRAD)

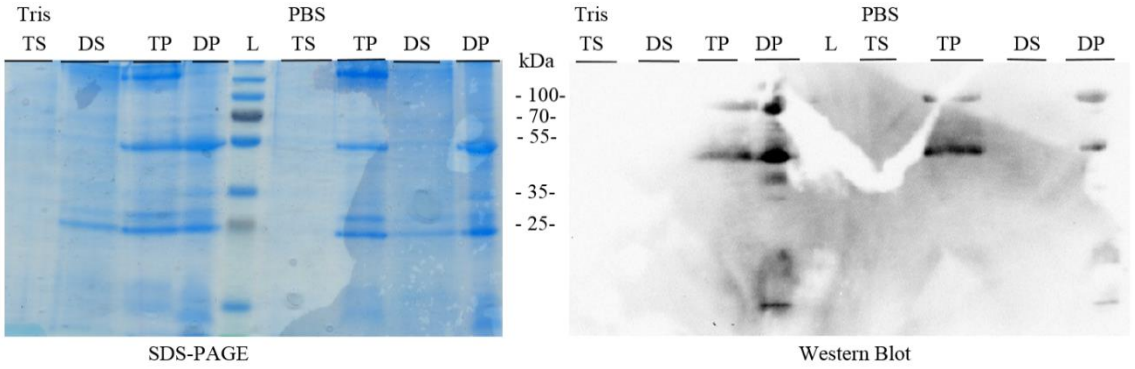

Figure C.7 - 12% SDS-PAGE and Western blot to analyse solubilization of ArlS at room temperature. 20 mM Tris pH 7 and 10 mM PBS pH 7.4; 1% Triton X-100: TS – triton soluble fraction, TP – triton insoluble fraction; 1% DM: DS – DM soluble fraction, DP – DM insoluble fraction. L - Page Ruller™ Plus Prestained Protein ladder

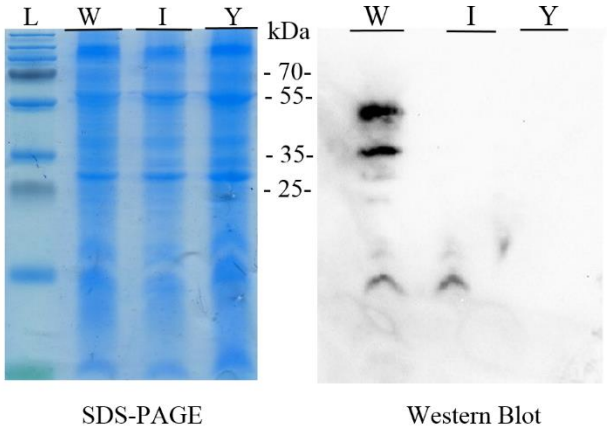

Figure C.8 – 12% SDS-PAGE and Western blot to analyse ArlS outer and inner membrane separation with sucrose gradient technique. W-outer membrane, I – intermediate, Y – inner membrane fraction. L - Page Ruller™ Plus Prestained Protein ladder

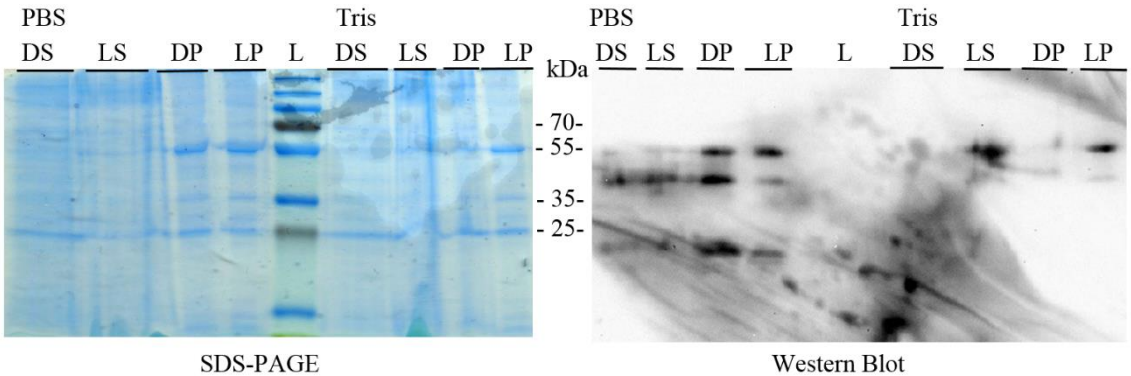

Figure C.9 – 12% SDS-PAGE and Western blot to analyse outer membrane solubilization at 4ºC. Buffer: 20 mM Tris pH 7, 10 mM PBS pH 7.4; Detergents: DDM and LDAO. DS – DDM soluble fraction, DP – DDM insoluble fraction; LS – LDAO soluble fraction; LP – LDAO insoluble fraction. L - Page Ruller™ Plus Prestained Protein ladder

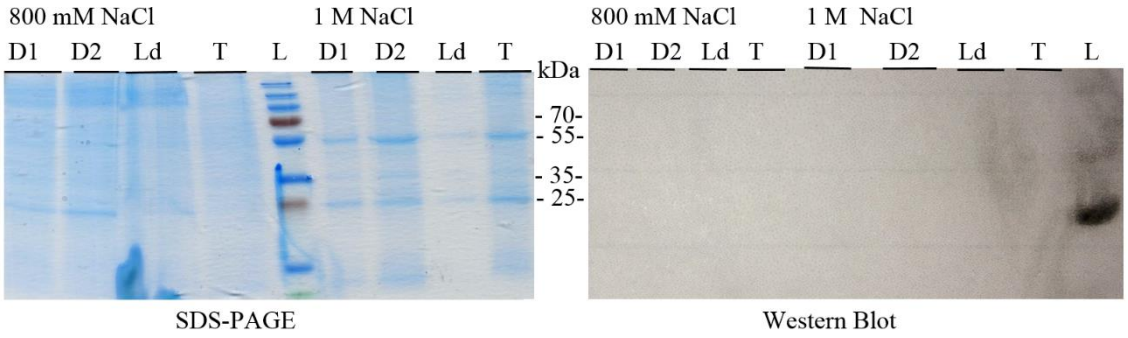

Figure C.10 – 12% SDS-PAGE and western blot to analyse ArlS outer membrane soluble fraction at 4ºC. Buffer: 10 mM PBS pH 7.4; Detergents: Triton X-100, DM, DDM and LDAO; D1 – DM, D2 – DDM, Ld – LDAO, T – Triton x-100; NaCl concentration: 800 mM and 1 M; L - Page Ruller^{TM} Plus Prestained Protein ladder

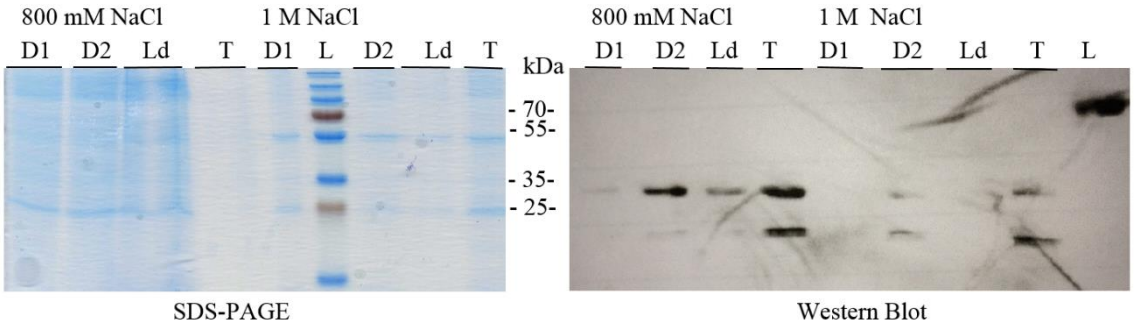

Figure C.11 – 12% SDS-PAGE and western blot to analyse ArlS outer membrane insoluble fraction at 4ºC. Buffer: 10 mM PBS pH 7.4; Detergents: Triton X-100, DM, DDM and LDAO; D1 – DM, D2 – DDM, Ld – LDAO, T – Triton x-100; NaCl concentration: 800 mM and 1 M; L - Page Ruller^{TM} Plus Prestained Protein ladder