TÉCNICO LISBOA

Studies with *Vibrio alginolyticus* OmpK towards the development of a vaccine for *Solea senegalensis*

João Bruno Fiúza Almeida Lopes

Thesis to obtain the Master of Science Degree in

Biotechnology

Supervisors:

Prof. Marília Clemente Velez Mateus

Prof. Jorge Humberto Gomes Leitão

Examination Committee

Chairperson: Prof. Leonilde de Fátima Morais Moreira

Supervisor: Prof. Jorge Humberto Gomes Leitão

Members of the Committee: Prof. Gabriel António Amaro Monteiro

October 2017

Agradecimentos

Gostaria de agradecer a algumas pessoas contribuíram directa e indirectamente para este projeto:

Gostaria de agradecer aos meus orientadores, Professor Jorge Leitão e Professora Marília Mateus pelo seu apoio e conselhos neste ano, pela a sua ajuda na escrita deste documento e por todas as correções e revisões feitas ao mesmo.

À Professora Isabel Sá-Correia por me ter recebido como aluno no Mestrado em Biotecnologia e no grupo do BSRG do Instituto Superior Técnico.

Gostaria de agradecer à Doutora Sílvia Sousa que me acompanhou desde a minha entrada no Grupo de Ciências Biológicas, supervisionando o meu trabalho laboratorial e aconselhando-me no mesmo. Obrigado pela paciência e disponibilidade.

Gostaria também de agradecer à Joana e ao Tiago por serem uns corações moles, pelo apoio, pelas conversas, e conselhos que me deram ao longo deste ano, apesar de fingirem que não se interessam.

Gostaria de agradecer a todos as pessoas do BSRG por todo o apoio na adaptação à vida laboratorial.

Gostava de agradecer aos meus colegas do Mestrado em Biotecnologia, pelo apoio, pelas conversas e por terem sempre uma palavra amiga para tornar algumas situações mais toleráveis.

Gostava de agradecer especialmente aos Gunas por serem as pessoas que são e nunca desistirem de me tentar animar.

Queria também agradecer à Maria Mendes, Iris Pereira, Margarida Venda e Tiago Reis porque apesar de cada um ter tomado o seu rumo depois da Licenciatura também a sua amizade ajudou-me a atravessar momentos mais difíceis. E as jantaradas eram sempre engraçadas.

Finalmente gostaria de agradecer à minha família porque foi graças ao seu apoio inesgotável, apesar de às vezes serem um pouco chatos, que consegui ultrapassar esta fase da minha vida. Obrigado por todas as vezes que me vieram buscar a Lisboa quando estava muito cansado para uma longa viagem de comboio. Obrigado por todas as vezes que esperaram que chegasse a casa e por me apoiarem em todas as decisões que tomei. E obrigado pela comida.

Abstract

Aquaculture is a growing industry due to the reduction of fish stocks in oceans. A major problem associated with high density fish cultures are infections such as *Vibrio*sis, affecting the productivity of aquaculture farms and forcing the use of antibiotics that can later contaminate the fish product.

The present study aims at the exploitation of the possibility of using *Vibrio alginolyticus* OmpK protein as putative immunogenic protein to be used as vaccine to protect aquaculture fish from *Vibrio* infections.

A bioinformatics analysis was initially performed to investigate the OmpK amino acid sequence conservation among *Vibrio* species and the occurrence of putative immunogenic segments within the proteins. The results obtained show that OmpK is a conserved protein potentially immunogenic.

The *Vibrio alginolyticus* gene encoding OmpK was amplified by PCR and cloned in pET23a+. The recombinant plasmid obtained was used to transform *E. coli* BL21 (DE3). The protein was overproduced in *E. coli* and purified by affinity chromatography. Initial attempts to purify recombinant OmpK by affinity chromatography were unsuccessful. Further experiments using recombinant OmpK denatured with urea 8 M allowed the one step purification of the protein, suggesting that the His-tag was hidden when the protein is overexpressed in *E. coli*. Successful recovery of the protein was achieved when it was denaturated prior purification to expose the histidine tail and to allow the interactions between the tail and the nickel during the purification process.

Batches of purified protein are being produced envisaging their use in immunization experiments to raise an antibody to be used in future work.

Keyword's: aquaculture, fish vaccine, recombinant proteins, OmpK, protein purification

Resumo

A industria da aquacultura é uma industria em ascensão devido à redução das reservas de peixe no oceano. Um dos problemas na produção de peixe em grande escala é o aparecimento de infeções, como a vibrose, que afetam a produtividade dos tanques de aquacultura e forçam o uso em grande escala de antibióticos que podem depois afetar os consumidores.

O principal objetivo deste estudo é explorar a possibilidade de usar a proteína OmpK de *Vibrio alginolyticus* como proteína imunogénica para ser usada na vacinação dos peixes de aquacultura contra as infeções por *Vibrio*.

Inicialmente foram feitas análises bioinformáticas para investigar a conservação desta proteína e prever a ocorrência de segmentos imunogénicos na proteína. Os resultados obtidos indicam que a OmpK é uma proteína conservada e potencialmente imunogénica.

Procedeu-se à amplificação por PCR do gene que codifica OmpK e à clonagem no plasmídeo pET23a+. O plasmídeo recombinado obtido foi usado para transformar *E. coli* BL21e os transformantes foram usados para a sobreexpressão de OmpK. A purificação da proteína foi feita por cromatografia de afinidade. Inicialmente a purificação de OmpK recombinante não teve êxito, mas experiências posteriores usando proteína recombinante OmpK desnaturada com ureia 8M permitiram a purificação da proteína, sugerindo que a His-tag se encontra escondida quando sobreexpressa em *E. coli*. A proteína foi obtida com sucesso quando esta é desnaturada antes da purificação para expor a cauda de histidinas e permitir a ligação entre a cauda e o níquel durante o processo de purificação.

Estão a ser produzidos batches de proteína purificada de modo a serem usadas em experiências de imunização para produzir um anticorpo a ser usado em futuras experiências.

Palavras chave: aquacultura, vacinação de peixes, proteínas recombinantes, OmpK, purificação de proteínas

Table of contents

List of tables

List of figures

List of abbreviations

A - absorbance
bp – base pair
BSA - Bovine serum albumin
Cov - coverage
DNA - deoxyribonucleic acid
DTT - Ditiotreitol
EDTA - Ethylenediamine tetraacetic acid
ELISA - Enzyme linked immune-sorbent assay
EMBOSS - European Molecular Biology Open Software Suite
HMM - Hidden Markov Model
HPLC - High-performance liquid chromatography
IDEN - Identity
IMAC - immobilized metal ion affinity chromatography
IPTG - Isopropyl β-D-1-thiogalactopyranoside
I-TASSER - Iterative Threading ASSEmbly Refinement
LB - Lennox broth
LOMETS - Local Meta-Threading-Server
MWCO - Molecular weight cut-off
NCBI - National Center for Biotechnology Information
NOAA - National Oceanic and Atmospheric Administration
ºC – degrees Celsius
OD – Optical density
OM - outer membrane
Omps – Outer membrane proteins
PBS - Phosphate-buffered saline
PCR - Polymerase chain reaction
PDB - Protein Data Bank
rev. - revolutions
rpm - revolutions per minute
SB - Super Broth
SDS - Sodium dodecyl sulfate
SDS-PAGE - Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TAE - Tris-Acetate-EDTA
TCM - Tris-Calcium-Magnesium
Tris - tris(hidroximetil)aminometano

1 – Introduction

1.1 - Aquaculture

As defined by NOAA (National Oceanic and Atmospheric Administration), aquaculture (also known as fish or shellfish farming) refers to the breeding, rearing and harvesting of plants and animals in all types of water environments, including ponds, rivers, lakes, and the ocean. Fish and shellfish can be grown in aquaculture to market size in ponds, tanks, cages, or raceways. Stock restoration is a form of aquaculture in which the species are released into the wild environment to rebuild natural populations or coastal habitats. Growth of plant species and fish in aquaculture can also be performed in the framework of the pharmaceutical, nutritional and biotechnology industries.

1.1.1 - Benefits and risks of aquaculture

Although aquaculture is apparently highly beneficial with no drawbacks. The aquaculture industry poses several risks regarding occupational hazards and might cause several deleterious effects to the environment and to humans. When considering occupational hazards, safety concerns and risks to human health due to aquaculture industry is dependent of the type of operation, scale of production, and the cultured species. These risks can be categorized into physical work hazards, chemical and toxic exposure, and risks associated with infectious diseases[1].

Physical work hazards are mainly associated with the use of machinery and tools, electrocution and high-voltage accidents, as well as hazards associated with the industry like drowning, musculoskeletal injuries, and decompression sickness[2].

Toxic exposures include the exposure of workers to unique and toxic reagents, such as hydrogen sulphide that can cause anoxic brain injury and unconsciousness at high levels[2]. The toxicity of products is usually focused on the toxicity of a single chemical reagent but, in the aquacultures, multiple chemicals might be used producing unexpected toxic by-products and harmful gases[1].

Aquaculture workers or even consumers can be exposed to uncommon microbes including bacteria, viruses, algae, and parasites, potentially causing zoonotic diseases or emerging infections. Various fish pathogens are known to be contagious to humans, including several species of the genera *Mycobacterium* and *Vibrio*[1].

Aquaculture practice also causes environmental degradation. The negative effects include organic pollution and eutrophication due to a build-up of excess of nutrients and wastes that can cause algal blooms, reduction in water quality and natural habitat destruction[1,3]. These effects are usually located in the surrounding regions and can be transitory. Therefore, a good management protocol can minimize these effects. Chemicals used in aquaculture can also cause pollution in the environment. These chemicals include antibiotics, pesticides, herbicides,

hormones, pigments, and vitamins for example, with antibiotics being the most used. The main concern about the use of antibiotics is the selection of antibiotic resistances among bacteria in natural niches.

Aquaculture farms modify habitats to increase the productivity. These modifications include habitat conversion and predator-control programs[1]. Examples of these modifications include changes in biological composition of rivers, due to the biological processes of non-native species; farmed species may feed on native species and cause local extinction of aquatic species (plant or animal).

Aquaculture also presents benefits, by reducing the dependence of natural stocks, allowing the conservation of endangered species. Aquaculture can indirectly benefit the environment by providing methods to convert agriculture wastes in high-quality fish protein, by enriching pond mud for use as fertilizer and improving the quality of soil used as crop land[4].

1.1.2 - Effects of *Vibrio* contaminations in Portuguese aquacultures

Many bacterial diseases affect fishes in the aquaculture environment, with the risk of transmission to consumers. Bacteria of the *Vibrio* genus represent a major group in fish pathogens. *Vibrio* members are gram-negative bacteria, mesophilic and chemoorganotrophic, presenting a facultative fermentative metabolism[5]. These bacteria are found abundantly in aquatic habitats and in association with eukaryotes.

Vibrio are responsible for the contamination of fish in aquaculture. In Portugal these bacteria contaminate the sole *Solea senegalensis* (Kaup)[6] cultures. The sole is a flatfish, well adapted to temperate waters with a high economic value, that is farmed mostly in Spain and Portugal.

Diseases of bacterial origin are the most significant causes for economic losses affecting fish cultures, and cultured sole. The economic losses caused by these microorganisms indicate the urging of the development of prophylactic strategies, such as vaccination, to prevent disease[7].

Vibrio species that can cause disease in sole include *Vibrio harveryi* and *V. parahaemolyticcus* involved in an outbreak of farmed sole in south Spain causing a moderate level of mortality[8]. Main external signs of the bacterial disease were skin ulcers and haemorrhagic areas near the fins and mouth[8] (Figure 1).

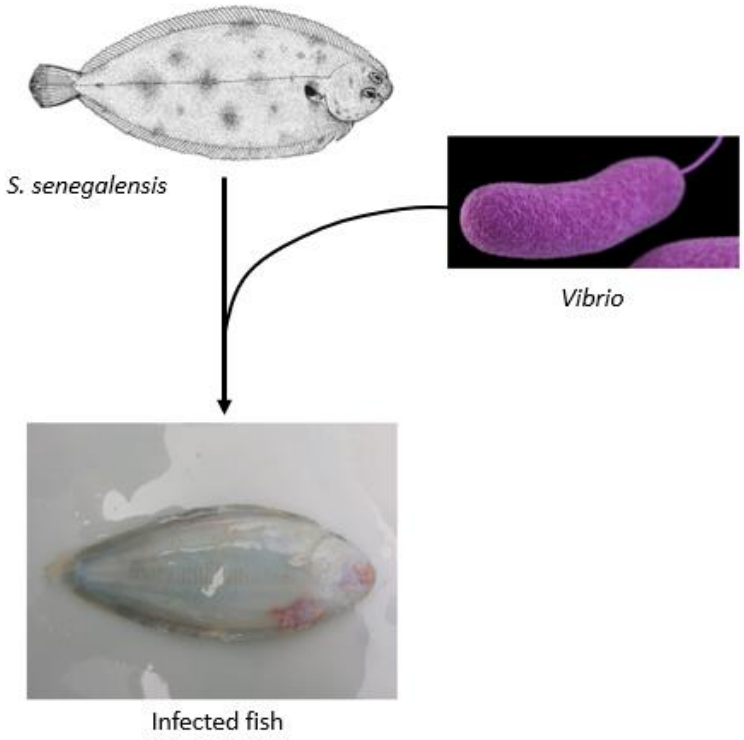


Figure 1 - *Vibrio*sis effect on soles. Adapted from Food and Agriculture Organization of the United Nations, Fisheries and Aquaculture Department; Centers for Disease Control and Prevention and Tapia-Paniagua et al.

1.1.3 - Advantage of fish immunization

Outbreaks of infectious disease or infections results in a reduced production of fish in aquaculture. Infection control represents additional costs associated with labour and antibiotic therapy[9]. Furthermore, extensive use of antibiotics potentiates the risk of contamination of fish products and selection of resistant strains that would damage the aquatic environment. For these reasons, aquaculture industry has given priority to vaccination as a mean to control endemic diseases. For the fish industry, vaccination is a matter of economy, because it has a relatively low cost and the potential to increase fish production[10].

Several vaccination methods of fish can be used in aquaculture: injection, immersion, and oral delivery. Some parameters have to be considered to choose the best method, like protective immunity after vaccination, vaccine delivery, labour requirements, and risk of side-effects[9].

1.1.3.1 - Injection

This method requires the use of an anaesthetic, because fish need to be immobilized. In this case, vaccines are injected intraperitoneally, most often to the cranial pelvic fins. The injection method should not be used when the fish is smaller than 10g due to difficulties in handling such

small fish and the risk posed by injecting an excessive volume. In addition, besides the large volume of non-physiological suspension, the vaccine may contain potentially toxic components[11].

The major advantage of injection is the superior protective immunity afforded after vaccination[12,13], with the relative percentage survival surpassing 95% in some cases. Another advantage is a lower dosage of vaccines that in this case remains the same for all sizes bigger than 10g, while in other methods it will depend on the fish biomass[9].

Risks of undesirable side-effects are greater with the injection methods[14] because of the anaesthetics and handling of fish, and the possibility of contamination of vaccine equipment and solutions. An undesirable side-effect of vaccination is the possibility of long-term persistence of local reactions in the fish abdomen[15]. Vaccination may also pose at risk humans that in the process of vaccinating the fishes, the operator can "self-vaccinate" by accident. This might cause local pain and swelling and, in extreme cases, severe anaphylactic shock.

1.1.3.2 - Immersion

There are several methods to vaccine the external surface of a fish. The most common is the dip method, in which the fish is immersed for 20-30 seconds in a 1:10 vaccine solution. Dipping provides a relatively high efficacy with minimal stress and side effects and only moderate labour requirement[14]. However, the amount of vaccine required is higher and the protection is poorer compared to methods like injection. The dip method is mostly used for small fish or fish that are particularly susceptible to handling[9].

1.1.3.3 - Oral vaccination

A vaccine-containing feed is usually given to fish over one or more prolonged periods. The main objection to oral vaccination is the poor protection achieved in comparison to the injection and dip methods[16]. Furthermore, the large consumption of vaccine makes the method relatively expensive.

1.1.4 - Immunogenicity of the proteins to prevent infections of sole in aquaculture

With the aim of producing a vaccine to prevent *Vibrio* infections of sole, an immunogenic molecule needs to be identified. The outer membrane of gram-negative pathogenic bacteria plays an important role in this interaction with the hosts namely in adherence to surfaces, uptake of nutrients from the host and subverting host defence mechanisms[17]. Examples of potentially immunogenic outer membrane proteins are OmpK and OmpW. OmpK is a receptor for the broad-host range *Vibrio* phage KVP40, and is widely distributed among species of the *Vibrionaceae* family[18,19]. . OmpK is an outer membrane protein found in pathogenic bacteria. The OMPs of Vibrio species are highly immunogenic components due their exposed epitopes on the cell surface[20] . OmpK, as one of the most important OMPs, is widely distributed among Vibrio and was shown to be the receptor of KVP40, a broad-host-range vibriophage in V. parahaemolyticus[21]

. By polymerase chain reaction (PCR) and sequencing, open reading frame (ORF) of ompK gene obtained from prepared genome DNA of V. alginolyticus strain ATCC 17749 presented a full length of 811 bp. The deduced amino acid sequence of ompK gene consists of 271 amino acid residues with predicted molecular weight of 30.17 kDa.

1.1.5 - Development of new ELISA methods for early detection of the infection

In addition to their exploitation as vaccines, immunogenic proteins have the potential of being used in infection detection methods. One of these methods is the Enzyme linked immune-sorbent assay (ELISA). This technology is based on the ability of a specific antibody to bind to a given antigen present in solution[22].

Some advantages of ELISA are the fact that it can be easily used in high throughput assay with low sample volume requirements and often less sample extract clean-up procedure when compared with other methods such as HPLC. The method is rapid, simple, sensitive, and portable for use in the field.

Epidemiological studies are very important since it is possible to discover new immunogenic markers that can be used to produce more sensitive and specific antibodies that allow for a rapid and more sensitive detection of infectious agents based on ELISA methods[22].

1.2 – Production of immunogenic proteins

Proteins are synthesized by all living forms as part of their natural metabolism. Some proteins play a significant role in immune responses and are commercially produced in industries with the aid of genetic engineering and protein engineering. The production of recombinant proteins is an important sector in biopharmaceutical and agriculture industry[23].

1.2.1 - *Escherichia coli* as a factory of proteins: advantages and disadvantages

Escherichia coli has been the pioneering host for recombinant protein production since the recombinant DNA procedures were developed. Because of the simplicity of the organism and the accumulated knowledge on its genetics and physiology, there is a large array of tools available for genetic engineering adapted to the bacterium. For these reasons, *E. coli* is the preferred host when trying to produce a new protein[24]. Other advantages of *E. coli* as a recombinant protein producer include fast growth rate and simple culture procedures[24].

Although *E. coli* offers several advantages, is not an ideal protein producer as it is difficult to find the optimal conditions for protein production and downstream processing. Different problems like plasmid loss and antibiotic-based maintenance, undesired chemical inducers of gene expression, plasmid/ protein-mediated metabolic burden and stress responses, lack of post-translational modifications, none or poor secretion, protein aggregation and proteolytic digestion,

endotoxin contamination and complex downstream processing are among the main obstacles for protein production in *E. coli*[24].

In industries like the pharmaceutical industry, the proper glycosylation of proteins is often required and considered of extreme importance. Cost-effective bioproduction is required to decrease the costs in the downstream process. Due to these limitations other bacteria, yeast and even mammalian cells are being used for the production of specific recombinant proteins[25]. Mammalian cells are used because of the ability to properly glycosylate and give the correct conformation to the recombinant protein, which greatly reduces the cost of the downstream processing[24]. Despite all the potential and value added by using non-*E. coli* as cell factories, adapting the biological processes for large scale production is often unaffordable[24].

1.2.2 - Molecular vectors of heterologous gene expression

An adequate recombinant protein production system requires the host's efficient expression of the gene of interest, it is required the use of a molecular vehicle to insert the gene of interest. However, due to structural changes or random mutation, the recombinant gene can be changed or lost in the host cell. There are three alternative locations for the gene of interest: plasmids, integrated in the host's chromosome, or a virus for delivery. In *E. coli* a carrier plasmid is the most used method.

1.2.2.1 - Plasmid-based expression systems

Plasmids are extrachromosomal self-replicating cytoplasmic DNA elements that can be found both in eukaryotes and prokaryotes. Plasmids are used as molecular vehicles for recombinant genes and are the most popular choice when using prokaryotes as hosts, due to the easiness of genetic manipulation. Furthermore, gene dose (number of copies of a gene) is higher than if the recombinant gene is integrated in the chromosome. Plasmid copy number is particular for each expression system and depends on the plasmid, the host and the culture conditions, and can vary from a few up to two hundred[26]. Plasmids impose a metabolic burden on the host, as cellular resources need to be utilized for their replication as well as for the expression of the plasmid-encoded genes and production of the recombinant protein. This metabolic burden can increase with an increase in the size of the insert, expression level, recombinant protein yield and toxicity of the protein to the host[27]. One of the consequences of the metabolic load is a decrease of the growth-rate of the cells hosting the plasmid.

Cells harbouring recombinant plasmids often exhibit a lower growth-rate than plasmid-free cells. If in a culture there is a high number of plasmid-free cells they will overtake the culture and lead to a loss of the plasmid, which is the main cause of reduced recombinant protein productivity in plasmid-based systems. The two main reasons for plasmid lost are an unequal plasmid distribution upon cell division, also called plasmid segregational instability[26] and plasmid multimerization. Other parameters like plasmid size (larger plasmids are less stable), presence of foreign DNA, nutrient availability, cell growth rate, temperature and mode of culture can also lead

to plasmid loss. Plasmid segregational instability happens when the plasmid copy number is not high enough when compared to the number of daughter cells; several plasmid-free cells can be generated, which exhibit higher specific growth rate, leading to an overtake of the culture and loss of the plasmid of interest[27]. Plasmid multimerization occurs because plasmids have the same sequence so they can recombine and form a single dimeric circle with two origins of replication. The main consequences are fewer independent units to be segregated, leading to an increase of the plasmid loss, together with an increase of the size of the plasmid uptaken by the host cell. This increase in plasmid size will reduce the growth-rate of the harbouring cells, increasing the chance of plasmid loss[27,28].

Several mechanisms can be used to ensure plasmid maintenance in cell populations. One of the most used strategies is to provide plasmid-bearing cells with a competitive advantage over plasmid-free cells by means of conferring a competitive advantage to plasmid bearing cells[29]. The introduction of a selective pressure will prevent plasmid instability and the overtake of the culture by plasmid-free cells. The selective pressure most commonly used given to the cells by introducing into the plasmid a gene or genes that provide resistance to a certain antibiotic to be supplemented in the culture medium. This procedure ensures that only the cells containing the plasmid and the genes for the resistance will be able to survive. However this approach can be ineffective if the antibiotics are in low concentration or degraded, or if detoxifying enzymes leak from plasmid-containing cell, therefore the amount of time for cell culture is limited if no more antibiotic is added to compensate deactivation or degradation[30]. Major disadvantages of this approach include the high costs of antibiotics and the problems associated with downstream processing and wastewater treatment.

Despite antibiotics usage is still the most used strategy, other strategies have been explored. These include deletion of an essential gene from the bacterial chromosome and its inclusion in the plasmid, and the introduction of a growth repressor in the bacterial genome and its antidote in the plasmid[31].

Plasmids can also be lost by structural instability. In this case, a reorganization of the genetic material occurs by recombination, leading to a non-productive vector[28]. Although structural instability is less common than segregational instability, it cannot be prevented by selective pressure. Structural instability can result in a complete elimination of the recombinant protein production or in the production of aberrant proteins with minor changes (deletions, additions or substitutions). The presence of proteins with minor changes is very difficult to detect and can result in major functional differences, depending on specific amino acid residues affected. The complete amino acid sequencing of the recombinant protein or the DNA sequencing of the gene of interest is necessary to avoid erroneous results[32].

Plasmid copy number is an important factor in plasmid-based systems for heterologous protein production. Although high plasmid copy numbers are usually required to improve protein yield, it may also result in protein degradation and deficient post-translational modifications[31]. Cells with high plasmid copy numbers can also have low recombinant protein yield due to a

reduction of the translation efficiency[33]. The plasmid copy number should be optimized for each process[34].

1.2.3 - Processes leading to non-functional proteins

In addition to plasmid segregational and/or structural instability post-translational processing can also contribute to the production of non-functional recombinant proteins. Post-translational processing of a protein comprises several steps like folding, aggregation, solubility of the protein, proteolytic processing, glycosylation and other modifications.

1.2.3.1 - Folding, aggregation, and solubility

Folding is a complex process assisted by foldases, that accelerate protein folding, and chaperones, that prevent the formation of non-native insoluble folding intermediates[35]. Misfolded proteins might accumulate as intracellular particles called inclusion bodies. Protein misfolding is mainly caused by cell stress, but can also be caused by heat shock, nutrient depletion, or other stimuli[36,37]. Incorrect protein folding has adverse consequences. For example, human pathologies such as Alzheimer's disease, Parkinson disease and Huntington disease, are characterized by intracellular protein aggregation and accumulation[36].

The formation of inclusion bodies is usually the result of overexpression of heterologous proteins. This phenomenon is not yet fully understood although several hypotheses have been proposed. It is possible that heterologous proteins often reach non-physiological concentrations that may promote their aggregation[38]. Another explanation for inclusion bodies formation is that aggregation results from lack of disulphide bond formation due to the reducing environment of bacterial cytosol and the lack of dedicated cellular machinery[30]. Rapid intracellular protein accumulation[31] and expression of large proteins[37] increase the probability of aggregation.

In some circumstances, protein aggregation and the formation of inclusion bodies can have a positive effect on the cell and for the production process since aggregation protects the protein from proteolysis and facilitate protein recovery by simply breaking the cells and recovering the inclusion bodies by centrifugation[31]. In addition, when the expressed protein is toxic to the host, toxicity can be prevented when the protein forms inclusion bodies[39]. In these situations, the production of inclusion bodies is not only desirable, but is also promoted though the use of molecular biology strategies, such as protease-deficient strains or cultivation at high temperatures[38]. Nevertheless, refolding step is still an empirical process, often inefficient, with yields usually lower than 10%[40]. Therefore, inclusion body formation likely results in an expensive and difficult downstream processing. For example, if instead of bacterial cells, mammalian cells are used, the folding would be correct and it not so many steps would be required in the downstream process.

No methods are presently available to predict if a specific protein will or not aggregate in a particular expression system. In addition, solubilisation and renaturation is also empirical[31,39].

Nevertheless, several strategies to avoid protein aggregation have been tested, including the use of foldases and chaperones to facilitate the process, but this strategy is not always successful because it's impossible to predict which chaperones will facilitate the folding of a particular protein[32].

1.2.3.2 - Glycosylation

Glycosylation is one of the various post-translational modifications that proteins undergo. This modification is very complex, involving various cellular processes and requiring several enzymes. Glycosylation usually occurs in the endoplasmic reticulum and Golgi apparatus of eukaryotic cells, although N-glycosylation can be performed by some bacteria[41]. Three types of glycosylation are known: N-(glycans linked to an Asn of a AsnXaaSer/Thr consensus sequence, where Xaa is any amino acid), O-(glycans linked to a Ser or a Thr), and C-(glycans linked to a tryptophan [Trp]). Little is known about C-linked glycosylation and biological significance[42]. N-linked glycosylation is considered the most relevant for recombinant protein production.

Glycosylation can determine the protein stability, solubility, antigenicity, folding, localization, biological activity and circulation half-life. Authentic glycosylation is especially important when recombinant proteins are used as drugs because glycosylation profiles are protein-, tissue- and animal-specific; non-authentic glycosylation can trigger adverse immunogenic responses[43]. Several factors can affect glycosylation. One of these factors is that different glycosylation sites are often glycosylated in different ways[44].

Culture conditions can also affect glycosylation, as lower growth rate may result in a reduced protein production rate and an increase of the extent of glycosylation. For O-linked glycosylation, as in N-glycosylation, non-authentic profiles can also elicit an immune response towards the recombinant protein[45]. The strategies used for N-glycosylation also improve the amount of O-glycosylation.

The present study aims for the production and purification of recombinant OmpK protein from *E. coli* to produce a vaccine that allow aquacultured fish immunization against vibrio infections (Figure 2).

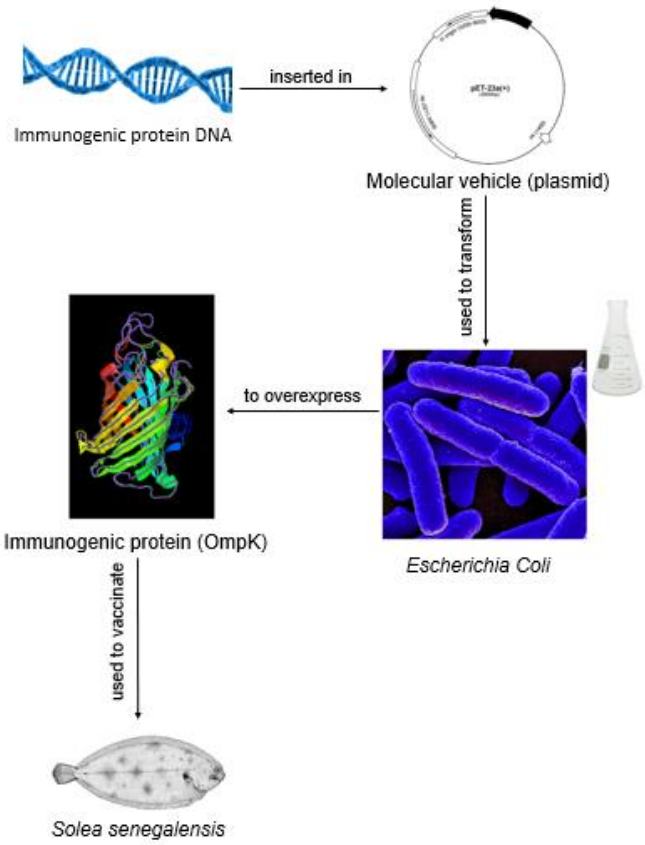


Figure 2 - Summary of a possible strategy for vaccination of fishes in aquaculture against *Vibrio* infections. Adapted from Novagen; National Institute of Allergy and Infectious Diseases; I-TASSER software and Food and Agriculture Organization of the United Nations, Fisheries and Aquaculture Department.

2 - Materials and Methods

2.1 - Bacterial strains, plasmids and culture conditions

Strains and plasmids used in this study are listed in table 1. To preserve cultures for longer periods of time, bacterial strains were kept at -80ºC in 40% (v/v) glycerol. When in use, *Vibrio alginolyticus* and Escherichia coli strains were maintained, respectively, in Lennox broth (LB, Pronadisa) plates with 3% NaCl and Lennox broth (LB, Pronadisa) agar plates (2% agar). Cultivation of *V. alginolyticus* and *E. coli* strains on solid media was carried out by incubation at 37ºC for 1 day. When required, growth media were supplemented with ampicillin to maintain selective pressure, at the final concentrations of 150 µg/ml.

Table 1 - Strains and plasmids used in this study

Strain/ Plasmids	Description	Source
Vibrio strains		
V. alginolyticus ATCC 17749	Strain type; Isolated from spoiled horse mackerel which caused food poisoning, Japan	Myamoto et al.[46]
E. coli strains		
DH5α	supE44 (ϕ80 lacZΔM15) hsdR17(rK - mK +) recA1 endA1 gyrA96 thi-1 relA1 deoR Δ(lacZYA-argF)U169	Invitrogen
BL21 (DE3)	FompT hsdSB (rB - mB -) dcm gal λ(DE3).	Stratagene
Plasmids		
pET23a+	Cloning vector, Ap	Novagen
pJBF2	pET23a+ containing the *ompK* gene, Ap	this work

2.2 – Molecular biology methods

2.2.1 – DNA manipulation techniques

2.2.1.1 – Extraction and purification of genomic and plasmid DNA

The genomic DNA was extracted from overnight cultures of *V. alginolyticus* using the High Pure PCR Template kit from Roche. The concentration of genomic DNA solutions was estimated by measuring the absorbance at 260 nm in a ND-1000 spectrophotometer (nanodrop) from alfagene, assuming that 1 unit of absorbance at 260 nm corresponds to 50 µg/ml of dsDNA. The ratios between absorbance at 260 nm (A_{260}) and 280 nm (A_{280}), that informs about contamination by proteins and between A_{260} and 230 nm (A_{230}) (that informs about contaminations by reagents) were also recorded. Samples with ratios A_{260}/A_{280} and A_{260}/A_{230} below 1.8 and 1.7, respectively, were discarded[47].

2.2.1.2 - Polymerase chain reaction (PCR) conditions.

The oligonucleotides primers were designed based of the gene encoding for OmpK sequence (UniProtKB - B8Y1F2_VIBAL) of *Vibrio alginolyticus* ATCC 17749 and were synthetized by Eurofins MWG Operon (Germany) (Table 2).

Table 2 - Oligonucleotides used for OmpK gene amplification by PCR. The recognition sites for restriction enzymes are in bold. Abbreviations used: A, Adenine; C, Cytosine; G, Guanidine; T, Thymine

Primer	Sequence	Size (bp)	Temperature of melting (ºC)
Up-OmpK	5'-AA**GGATCC**ATGCGTAAATCAC-3'	21	60
Lw-OmpK	5'-AA**CTCGAG**GAACTTGTAAGTT A-3'	22	60

Genomic DNA from *Vibrio alginolyticus* ATCC 17749 were used as template for PCR amplification of OmpK with an expected size of 811 bp. PCR amplification was performed in a 2720 thermal cycler (Applied Biosystems). The amplification mixture composition is shown in table 3. The sample was subjected to an initial denaturation at 98ºC for 3 minutes, followed by 30 cycles of: denaturation (98ºC for 10 seconds), annealing (59ºC for 30 seconds) and elongation (72ºC for 27 seconds). After the final elongation at 72ºC for 7 minutes, the sample was stored at 4ºC until further use. The Phusion DNA polymerase from Thermo was used.

Table 3 - Summary of the reaction components used to amplify V. alginolyticus OmpK by PCR.

Proteins	Primers		T HF 5x (µl)	dNTP's (µM)	DNA (ng)	Phusion (U)	H_2Od (µl)	Total volume (µl)
	Upper (µM)	Lower (µM)						
OmpK	0.4	0.4	4.0	200	100	0.4	required for the final volume	20

2.2.1.3 - Agarose gel electrophoresis

Agarose gels were prepared as described by Maniatis et al.[48], using agarose (Lonza) at 1% (w/v) in buffer Tris-Acetate-EDTA (TAE) 1X (TAE 50X was composed of 242 g/l Tris.base, 57.1 ml/l acetic acid, 37.2 g/l $Na_2EDTA.2H_2O$, pH 8.0). Dilution of 1:50 of this buffer in distilled water was performed to obtain TAE 1X. Samples were prepared by adding 1 µl of gel loading dye (Thermo) to each 10 µl of DNA sample solution. The molecular weight marker 1kb Plus DNA ladder (Thermo) (Annex 1) was used. DNA fragments were separated by electrophoresis in TAE 1X buffer at 4.0 V/cm, at 100V and 200 mA for more or less 30 minutes. The gels were stained with Gel Red (Biotium) for 20 minutes in a concentration of 100 µl per litre of Gel Red commercial solution 10,000X in TAE 1X buffer and the DNA was visualized in a transilluminator (Bio-Rad).

After electrophoresis, gel slices containing the DNA of fragments of interest were purified from agarose using the NZYGelpure kit from NZYtech (Portugal), and following the

manufacturer's instructions. The concentration of the extracted DNA solutions was estimated by measuring the A_{260} in a ND-1000 spectrophotometer (nanodrop).

2.2.1.4 – Construction of pJBF2 plasmid

After extraction for the agarose gel, the DNA fragment obtained by PCR was re-dissolved in molecular biology grade water at a final volume of 30 µl. The DNA fragment was then restricted with adequate enzymes [BamHI (NZYtech) and XhoI (Thermo)] in a total volume of 20 µl, in a reaction mixture with a composition described in Table 4. Plasmid pET23a+ was also digested with BamHI/ XhoI (Tables 4 and 5)

Table 4 - Summary of the volumes required for the digestion step mixture of OmpK

OmpK					
DNA (ng)	BamHI (U)	XhoI (U)	10X K Buffer (Takara) (µl)	H_2Od (µl)	Final volume (µl)
300	10	10	3.0	10	30

Table 5 - Summary of the volumes required for the digestion step mixture of pET23a+

pET23a$^+$					
DNA (ng)	BamHI (U)	XhoI (U)	10X K Buffer (Takara) (µl)	H_2Od (µl)	Final volume (µl)
100	10	10	2.0	required for the final volume	20

After the digestion step, the DNA was precipitated by adding sodium acetate 3 M (1/10 volume) and ethanol 100% 2.5 times the DNA volume used, respectively. The mixture was kept for at least one hour at -80ºC. After this step, the sample was centrifuged for 30 minutes at 4ºC, in a (12145 rotor, 15000 rpm) centrifuge (J2-21, Beckman). The supernatant was removed with a syringe, and 500 µl of ethanol 70% was added to the precipitated DNA and centrifuged again for 15 minutes at room temperature. The supernatant was removed with a syringe, and the sample was dried in a speed vacuum (Eppendorf AG) for 15 minutes at 45ºC with the program set for alcohol dry (V-AL). The dry sediments were resuspended in 16 µl of H_2O and stored at -20ºC until further use.

In order to overexpress the gene encoding for OmpK as a His6-tagged protein, the DNA fragments were ligated directionally into the expression vector pET23a+. The composition of the ligation mixture is summarized in Table 6.

Table 6 - Summary of the reaction mixture used for the insertion of the gene of interest in pET23a+

plasmid DNA (ng)	Insert DNA (ng)	10 x T4 DNA ligase buffer (Thermo) (µl)	T4 DNA ligase (Thermo) (U)	H_2O (µl)	Final volume (µl)
20-100	60-300	2.0	2.5	required for the final volume	20

The reaction mixture was incubated for one hour at 22ºC. Afterwards the enzymes were heat-inactivated by placing the sample in a dry bath at 70ºC for 5 minutes and stored at -20ºC until further use. The plasmid obtained, was named pJBF2 and was used for the overproduction of OmpK in *E. coli.*

2.2.2 – Methods for insertion of foreign DNA in bacterial cells

2.2.2.1 – Preparation of competent *E. coli* DH5α cells

E. coli DH5α were made competent as follow: a colony of *E. coli* DH5α was inoculated in 6 ml of LB liquid medium and incubated overnight with orbital agitation (250 rev./min) at 37ºC. Aliquots from overnight cultures were used to inoculate 100 ml of LB liquid medium at a OD_{640nm} of 0.1 and then incubated at 37ºC, with orbital agitation (250 rev./min), until the bacterial culture reached the mid-log phase ($OD_{640nm} \approx 0.5$). The cells were harvested by centrifugation at 5600 rev./min. (12145 rotor, J2-21, Beckman) for 5 minutes at 4ºC, in sterile centrifuge bottles. To eliminate salts, the cell pellets were washed twice with $MgCl_2$ 0.1 M and $CaCl_2$ 0.1 M. The first washing step with a volume of 100 ml of $MgCl_2$ 0.1 M. After resuspension, the sample was centrifuged at 3400 rev./min. for 5 minutes at 4ºC. The second washing step was performed with 100 ml of $CaCl_2$ 0.1 M and the sample was resuspended and placed on ice for 30 minutes. Afterwards the sample was centrifuged at 3400 rev./min. for 5 minutes at 4ºC. The pellets from the second washing step were then resuspended in 22 ml of $CaCl_2$ 0.1 M and 3.5 ml of glycerol 86% (w/v) was added. The resulting suspension was distributed as 500 µl aliquots that were frozen in 1.5 ml cryovials and kept at -80ºC, until further use.

2.2.2.2 – Classical Transformation

The recombinant plasmids were transformed into competent *E. coli* DH5α. Classic transformation was performed by adding into a sterile Eppendorf tube 50 µl of buffer Tris-Calcium-Magnesium (TCM), 150 µl of competent cells suspension and 15 µl of recombinant plasmid DNA. TCM buffer contained $CaCl_2$ 10 mM, $MgCl_2$ 10mM and Tris 10 mM.

The transformation mixture was incubated on ice for 30 minutes, heated to 42ºC for 3 minutes in a dry bath. After the heat shock, samples were placed on ice for 5 minutes and after this time, 800 µl of LB medium with 0.5 g/L of NaCl was added and the sample was incubated for one hour at 37ºC with agitation at 250 rev./min to allow the recovery of the cells. Aliquots of 100 µl were then spread onto the surface of solid LB medium with 10 g/L of NaCl supplemented with

ampicillin (150 µg/ml) and incubated for 24 hours at 37ºC. The plates are then stored at 4ºC until further use.

Since the constructed plasmid only contains one marker of resistance (Annex 3) it is impossible to ensure that the colony has the desired plasmid inserted. Therefore, plasmids hosted by the transformed cells have to be extracted and analysed with the restriction enzymes XhoI and BamHI to confirm that the correct gene was inserted. For this step the alkaline lysis method was used to extract the plasmids. For this purpose, a mass of cells of one isolated colony from the plate was resuspended in 150 µl of solution I (Table 7), previously prepared and stored at 4ºC, and is incubated at room temperature for 5 minutes to lyse the cells.

Table 7 - Summary of the composition of Solution I

Solution I - pH 8.0	
glucose (mM)	50
EDTA (mM)	10
Tris-base (mM)	25
Lysozyme (mg/ml)	5.0
Rnase (µg/ml)	1.0

After incubation, 200 µl of solution II was added freshly and containing 0.2 M NaOH and 1% SDS. Afterwards, the sample was mixed by inverting the tube repeatedly and placed on ice for 5 minutes.

150 µl of solution III was then added (Table 8). The sample was mixed by vigorously inverting the tubes and was placed on ice for 10 minutes.

Table 8 - Summary of the composition of Solution III

Solution III - pH 4.5 – 5.5	
Potassium acetate 5M (ml/ml $_{\text{of solution III}}$)	0.600
Glacial acetic acid (ml/ml $_{\text{of solution III}}$)	0.115
H_2O (ml/ml $_{\text{of solution III}}$)	0.285

The sample was centrifuged for 10 minutes at 13500 rpm (Microstar12), at room temperature. After centrifugation, the supernatant was transferred to a new eppendorf and 1 ml of ethanol 100% (v/v) was added. The sample was then placed on ice for at least 10 minutes to precipitate the DNA.

After precipitation, the DNA was centrifuged at 13500 rpm (Microstar12), at room temperature. The supernatant was discarded with a syringe and 500 µl of ethanol 70% were added. The sample is vortexed and centrifuged again at 13500 rpm (Microstar12), at room temperature. Afterwards the supernatant was discarded and the sample precipitate was dried in a speed vacuum for 15 minutes at 45ºC with the program set for alcohol dry (V-AL), to remove the remaining ethanol. The pellet was then resuspended with 30 µl of H_2Od.

To reduce the amount of restriction enzyme used, 8.5 µl of plasmid DNA samples were mixed with 1.5 µl of loading buffer 6X (Thermo), the samples were applied in an agarose gel 1%, electrophoresed at 100 V and 200 mA for 30 minutes and then stained with Gel Red for 20 minutes. The images obtained were analysed to access the sizes of gene fragments obtained in order to choose the transformed *E. coli* colonies with higher probability of containing the inserted gene. The extracted plasmid DNA resulting from these colonies should be longer than that from the others and the difference varies with the size of the gene inserted, being approximately 811 bp for OmpK in addition to the cloning vector.

The plasmids DNA from chosen colonies were digested with BamHI and XhoI for 3 hours at 37ºC, with a reaction mixture as summarized in Table 9. 2 µl of loading buffer 6X were then added to each mixtures that were then placed in an agarose gel 1%, side by side with a ladder composed by 1 µl of gene ruler 1kb Plus (Thermo), plus 1.5 µl loading buffer 6X and 7.5 µl of H_2Od was used as molecular mass marker.

Table 9 - Summary of the volumes used in the digestion of the plasmid DNA.

OmpK				
DNA (ng)	BamHI (U)	XhoI (U)	10X K Buffer (Takara) (µl)	Final volume (µl)
200	10	10	2.0	25

The colonies with the plasmid of interest (811 bp for OmpK) were kept, plasmids were extracted from their fresh cultures and sent for sequencing by Eurofins MWG Operon (Ebersberg). Plasmids for sequencing were extracted using the kit NZYMiniprep from nzytech, according to the manufacturer's instructions.

2.2.2.3 – Preparation of electrocompetent *E. coli* BL21(DE3) cells

Cells of *E. coli* BL21 were made electrocompetent as follows: a colony of *E. coli* BL21 was inoculated in 30 ml of LB liquid medium and incubated overnight with orbital agitation (250 rev./min) at 37ºC. Aliquots from overnight cultures were used to inoculate 100 ml of LB liquid medium at a final OD_{640nm} of 0.1 that were incubated at 37ºC, with orbital agitation (250 rev./min), until the bacterial culture reached the mid-log phase ($OD_{640nm} \approx 0.8$). The cells were harvested by centrifugation at 12000xg (centrifuge J2-21, Beckman) for 15 minutes at 4ºC, in sterile centrifuge bottles. To eliminate salts, the cell pellets were washed three times with sterile ice-cold distilled sterile water and centrifuged as described above. Volumes of 100 ml, 60 ml and 20 ml of ice-cold distilled water were used in the sequential washes. The pellets were then resuspended in 4 ml of 10% (w/v) glycerol and centrifuged at 12000xg (J2-21, Beckman) for 15 minutes at 4ºC. The resulting cell pellets were resuspended in 2 ml of 10% (w/v) glycerol and 110 µl aliquots were frozen in 1.5 ml cryovials and stored at -80ºC, until used.

2.2.2.4 – Transformation by electroporation

In order to transform *E. coli* BL21 cells, a cryovial of frozen bacterial cells, prepared as described in 2.2.2.3, was thawed in ice. Then 5-10 µl of a solution of the plasmid of interest were added and gently stirred. The mixture was pipetted into an ice-cold cuvette with electrodes, without leaving air bubbles in the droplet. The cuvette was placed in a slot in the chamber rack of a Gene Pulser™ apparatus (Bio-Rad). A short electric pulse was applied, after setting the device at the following conditions resistance 4kΩ; capacitance 25 µF, and voltage 2.5 kV. After the electric-pulse, cells were incubated in LB liquid medium for 1 hour or overnight. After incubation, the bacterial cultures were plated on LB medium with 150 µg/ml of ampicillin.

2.3 - Protein manipulation techniques

2.3.1 – Overproduction of His_6-tagged proteins and preparation of cell extracts

A colony of *E. coli* BL21 containing the recombinant pJBF2 plasmid for the production of the OmpK protein was inoculated in 30 ml of LB liquid medium and incubated for approximately three hours with orbital agitation (250 rev./min) at 37ºC until an optical density of at least 1 was reached. An aliquot of the culture was used to inoculate 100 ml of Super Broth (SB) medium (32g/L tryptone, 20g/L yeast extract, 5g/L NaCl) with 150 µg/ml of ampicillin with an optical density of 0.1. The inoculated SB medium was incubated at 30ºC with orbital agitation of 250 rev./min until an optical density between 0.60 and 1 was reached (approximately two hours). Before induction of expression with IPTG 0.4 mM a sample was retrieved from the inoculum with a volume in millilitres following the equation 0.6/optical density, the volume is centrifuged at maximum speed for five minutes at room temperature and resuspended in 40 µl of Gel loading buffer (20% (v/v) glycerol, 4% (wt/v) SDS, 100 mM Tris.Cl (pH 6.8), 0,2% (wt/v) bromophenol blue, 200 mM DTT). After resuspension, the samples were heated for 5 minutes at 95ºC to stop the activity of the enzymes. Afterwards the samples were frozen at -20 ºC. The cells were induced with 0.4mM of IPTG at an OD_{640nm} of 0.6 and grown at 30 ºC. Samples are taken each hour, following the method previously used for the sample taken before inducing, to control the overexpression of the protein of interest. In the ideal point of overexpression, the totality of the culture was centrifuged for 5 minutes, at 7000xg and 4ºC (J2-21, Beckman). The cells were then resuspended in 8 ml of buffer containing 500 mM NaCl and 20 mM Tris-HCl and stored at -80 ºC.

Aliquots from cell suspension prepared from cultures carried out for 16 hours were slowly defrosted on ice for approximately 2 hours. In order to disrupt the cell wall the aliquots were sonicated with a Branson sonifier 250 (Branson) using eight cycles of 30 seconds, 50% of duty cycle, output of 60 with two minutes break between each cycle. After sonication aliquots were centrifuged 60 minutes at 12 000 rpm and 4ºC (J2-21, Beckman). The supernatants were transferred to a clean falcon and kept in ice until further processing. Samples of the pellet and supernatant were collected to quantify the amount of protein.

2.3.2 - Purification of C-terminal His_6-tagged proteins by affinity chromatography

The purification of OmpK in its native state was carried out by affinity chromatography using a HisTrap FF column (GE Healthcare) with a volume of 1 ml. The column was equilibrated with 10 ml of buffer 1 (Table 10). After equilibrating the column, the protein sample was applied to the column. The column was then washed with 10 ml of buffer 1, to wash proteins weakly bound to the column. Afterwards successive buffers with the same composition as buffer 1 and crescent concentrations of imidazole (60, 100, 200, 300 and 500 mM) were used to elute the contaminant proteins that do not have a strong affinity to the nickel in the column. The protein of interest, due to the histidine tail should be tightly bound to the nickel only being eluted at higher imidazole concentrations. The eluates from the washes with different concentrations of imidazole were collected in 1 ml fractions. To regenerate the column, 10 ml of buffer 1 was used; followed by buffer 2 containing EDTA to release all the nickel and proteins bound to the column. Finally, a solution containing 500 mM of nickel prepared from a 2M solution of nickel sulfate was added to the column. The column was then stored in 4ºC until further use.

Table 10 - Summary of the buffers and solutions used in the purification in HisTrap FF columns

Buffer	Buffer 1	Buffer 2
Components	10 mM imidazole 500 mM NaCl 20 mM Tris-HCl	Phosphate buffer 1x 0.05 M EDTA

2.3.3 – Purification of OmpK proteins from inclusion bodies

2.3.3.1 – Column-based purification

2.3.3.1.1 – Denaturation of OmpK proteins from inclusion bodies for the column-based purification

The recombinant His6-tagged OmpK protein was denatured using urea as the denaturation agent. After the sonication cell suspension were centrifuged (30 min at 12500xg and 4ºC) and resuspended with 10 ml of buffer 3, and centrifuged again (30 min at 12500xg and 4ºC), and resuspended in 10 ml of buffer 4 to denature the proteins. The suspensions were stirred at room temperature for 30 minutes. A solution of 1M of DTT was added to meet 2.5 mM and stirring was maintained for 45 minutes. The suspension was then centrifuged for 30 minutes at 12500xg and 15ºC, resulting in a supernatant containing the denatured protein. In order to be used in the His-tag FF columns (GE Healthcare), the concentration of urea (Sigma) was decreased from 8M to 6M by adding 3.3 ml of buffer 5 due to the proximity to the maximum concentration of urea allowed for the column[49].

Table 11 - Summary of the buffers and solutions used in the column based denaturation of OmpK

Buffer	Buffer 3	Buffer 4	Buffer 5
Composition	10 mM imidazole 500 mM NaCl 20 mM Tris-HCl 1% Triton X-100	10 mM of imidazole 500 mM NaCl 20 mM of Tris-HCl 8 M of Urea	10 mM imidazole 500 mM NaCl 20 mM of Tris-HCl

2.3.3.1.2 - Renaturation of denatured OmpK proteins for column-based purification

The column was initially equilibrated with 10 ml of buffer 6. After equilibrating the column, 13.3 ml of the protein sample was applied to the column. The protein of interest has a histidine tail that will bind to the nickel on the column. The column was then washed with 10 ml of buffer 6, to wash proteins weakly bound to the column. Afterwards, 5 ml of buffer 7 was added to the column in order to elute all the contaminant proteins present. After elution, the renaturation process started with the reduction of the concentration of urea in the sample by adding 20 aliquots of 2 ml of buffer 8 with decreasing concentrations of urea, 0.25 M decrease between each buffer from 6M to 1M concentration. After adding each aliquot, a 5 minutes break was taken before adding the following aliquot to ensure the drop in urea concentration. Afterwards 5 ml of buffer 9 was added to collect the protein of interest for dialysis. 5 ml of buffer 10 was then used to discard the high concentration of imidazole solution present in the column and 10 ml of buffer 5 was added to clean the column which was then stored at 4ºC.

Table 12 - Summary of the buffers and solutions used in the purification of OmpK using the column based protocol

Buffer	Buffer 6	Buffer 7	Buffer 8	Buffer 9	Buffer 10
Composition	10 mM imidazole 500 mM NaCl 20 mM Tris-HCl 6 M Urea	50 mM imidazole 500 mM NaCl 20 mM Tris-HCl 6 M Urea 10% Glycerol	10 mM Tris-HCl 0.5M NaCl 10% Glycerol	300 mM imidazole 500 mM NaCl 20 mM Tris-HCl 10% Glycerol	500 mM imidazole 500 mM NaCl 20 mM Tris-HCl 10% Glycerol

The volumes collected from the column eluates at 300 mM of imidazole were dialyzed using a Slide-A-Lyzer dialysis cassette (Pierce) and placed in buffer 8 overnight at 4ºC with gentle agitation. After dialysis and to conclude de renaturation process the sample was added to a Hist-tag column previously equilibrated with buffer 10. The fractions of protein that do not bind to the column were collected and should contain the renatured protein of interest. In order to collect the non renatured and clean the column, 10 ml of buffers 10 and 5, respectively were used. Afterwards the column was stored at 4ºC.

2.3.3.2 – Batch-based purification

2.3.3.2.1 – Denaturation of OmpK proteins from inclusion bodies for the batch-based purification

The denaturation of OmpK proteins from inclusion bodies for the batch-based purification was described in 2.3.3.1.1 with the exception of the suspension being left overnight at room temperature in buffer 7 to finish the denaturation process.

2.3.3.2.2 - Renaturation of denatured OmpK proteins for batch-based purification

For the renaturation process 2 ml of the denaturated protein sample was added to an amount of 8ml of slurry containing 4 ml of resin of Ni sephase6FF (GE Healthcare) was used accordingly to the manufacturer's instructions centrifuged at room temperature for 5 minutes at 1000xg. The supernatant was discarded and 5 ml of H_2O_d were added. After a 3 minutes rest at room temperature, the resin was centrifugated at room temperature for 5 minutes at 1000xg. The supernatant was discarded again and 5 ml of buffer 6 (Table 12) was added. After a 3 minutes rest at room temperature, the resin was centrifuged at room temperature for 5 minutes at 1000xg. Resuspended in 20 ml the protein sample and incubated for one hour at room temperature and then centrifuged at room temperature for 5 minutes at 1000xg. 5 ml of buffer 6 was added and incubated for 5 minutes at room temperature and centrifuged at room temperature for 5 minutes at 1000xg. 7,5 ml of buffer 7 was added followed by incubation for 10 minutes at room temperature. The resin was centrifuged at room temperature for 5 minutes at 1000xg and a sample was collected for SDS-PAGE analysis. The renaturation process started by the addition of 10 aliquots of 2 ml of buffer 8 with decreasing concentration of urea, 0.5 decrease between each buffer addition from 6M to 1M. After the addition of each the slurry rested for 10 minutes and was then centrifuged at room temperature for 5 minutes at 1000xg before the addition of following aliquot, to ensure the drop in urea concentration. The process was continued for buffers 9 and 10. The regeneration of the resin was performed as described in the manufacturer's instructions. The protein of interest pool was that from supernatant after incubation with buffer 9, which was dialysed as described in 2.3.3.1.2.

2.3.4 - Extraction of His_6-tagged periplasmatic proteins from *E. coli*

A 1-litre of *E. coli* BL21 transformed with the recombinant plasmid harbouring the *ompK* gene was carried out as described in 2.3.1. The cells were harvested by centrifugation (4000xg for 20 minutes at 4ºC). The pellet was resuspended in buffer 11 (Table 13) at 80 ml per gram wet weight of cells. The cell suspension was kept on ice and 500 mM EDTA were added dropwise to a final concentration of 1mM, and left on ice for 5 to 10 minutes with gentle agitation. Afterwards the cell suspensions were centrifuged (8000xg for 20 minutes at 4ºC) to remove all supernatant,

and resuspended in the same volume of ice-cold buffer 12. The cell suspension were gently stirred for 10 minutes in an ice bath and subjected to another centrifugation (8000xg for 20 minutes at 4ºC). The resulting supernatant, i. e., the supernatant is the low osmolarity fluid used in the osmotic shock procedure containing the periplasmatic proteins, was kept at 4ºC until further processing.

Table 13 - Summary of the buffers and solutions used for periplasmatic protein extraction from E. coli BL21

Buffer	Buffer 11	Buffer 12
Components	30mM Tris 20% sucrose pH 8.0	5mM $MgSO_4$

2.3.5 – Analysis of proteins by electrophoresis in polyacrylamide gels

2.3.5.1 – Denaturating polyacrylamide gel electrophoresis (SDS-PAGE)

Gels for SDS-PAGE were prepared as described by Maniatis et al.[48], using 15% acrylamide. Samples were prepared by adding 10 µl of gel loading buffer to each 40 µl of protein sample solution. The molecular weight marker used was precision plus protein dual xtra standards (Bio Rad) (Annex 2). The gel caster was assembled and water was added to detect leaks in the system. The system was filled with running gel until 1.5 – 2 cm of the top and a fin layer of isopropanol was added. After polymerization the isopropanol was removed and the combs were inserted. After placing the combs the stacking gel solution was poured and the system was left until the polymerisation of the stacking gel. Protein fragments were separated by electrophoresis in running buffer 1X buffer with a voltage of 150V and 200 mA for more or less 50 minutes and then stained with Coomassie blue staining solution for 30 minutes with agitation and destained with a destaining solution twice 30 minutes each with agitation. The protein bands and the marker were directly visualized after the destaining step.

Table 14 - Summary of the buffers and solutions used in the SDS-PAGE gels preparation

Solution	Gel loading buffer	Running buffer 10x	Coomassie blue staining solution	Destaining solution
Composition	100 mM of Tris-base pH 6.8 4% SDS (p/v) 20% glycerol 0.2% of bromophenol blue 200 mM of DTT	0.25 M Tris-base ph 6.8 1.92 M glycine 1% SDS (p/v)	0.2% (p/v) coomassie blue R-250 10% (v/v) acetic acid 47.5% (v/v) ethanol	10% (v/v) acetic acid 26.25% (v/v) ethanol

2.3.5.2 – Western Blot analysis of His_6-tagged proteins

A SDS-PAGE gel after electrophoresis was incubated for 15 minutes in Bjerrun-Schaffer solution (48 mM of Tris-HCl, 39 mM of glycine, 20% methanol v/v, and 0.04% SDS p/v, pH 9.2). A system was made so that the bands present in the gel were transfered to a nitrocellulose membrane. The system was composed of a nitrocellulose membrane with the acrylamide gel on top in the middle of 2 sheets of 3 MM paper. The nitrocellulose membrane and the sheets of paper were incubated in Bjerrun Shaffer buffer for 5 and 15 minutes, respectively. The system was transferred to the Trans-Blot SD, semi-dry transfer cell and left transferring for 50 minutes at 20 V, 120 mA transferring. After transfer the membrane was blocked with 4 ml of a solution with PBS 1x with 5% non-fat milk. The incubation was carried out overnight at 4ºC or for 1 hour at room temperature. The milk proteins block the membrane preventing non-specific binding to the membrane. After blocking the membrane was washed three times 5 minutes each with PBS solution with 0.05% Tween 20 and then incubated with diluted Anti-polyhistidine peroxidase conjugate with 0.05% Tween 20, 1x BSA and PBS 1x for 2 hours. The dilution of the antibody used was 1:2000. The membrane was then washed again 3 times with a solution of PBS 1x with 0.05% Tween 20. To visualize the results the membrane was incubated for 2 to 3 minutes with a mixture of two buffers (buffer 13 and 14) 1:1 to allow the visualization in the FUSION Solo apparatus.

Table 15 - Summary of the buffers used in the western blot protocol

Buffer	Buffer 13	Buffer 14
Composition	Luminol 2.5 mM p-Coumanic acid 400 µL Tris-HCl pH 8.6 100 mM	H_2O_2 0,15% Tris-HCl pH 8,6 100mM

2.3.5.3 - Bradford test protocol

This test was performed to determine the concentration of protein in solution after the renaturation process. In this test 50 µl of sample were added to 2.5 ml of Bradford reagent[50] followed by incubation at room temperature for 15 minutes. After 15 minutes absorbance of samples were measured in a spectrophotometer at $OD_{595\ nm}$. A calibration curve was built for each buffer used, in this case buffer 12. The calibration curve was obtained using BSA concentrations 100,200,350 and 500 µg/ml, diluted in the same buffer used for the sample. The blank contains only the buffer used in the samples and the Bradford solution. The calibration line was used to estimate the concentration of protein in the solution tested.

2.4 - Bioinformatics analysis of the *Vibrio alginolyticus* OmpK protein

The National Center for Biotechnology Information (NCBI) databases was used to retrieve the amino acid sequence of OmpK from *Vibrio alginolyticus* ATCC 17749. The protein amino acid sequence was aligned with sequences of the OmpK proteins from other *Vibrio* species known to infect fishes[51], also retrieved from NCBI. The program used for this alignment was Protein Basic Local Alignment Tool (BLASTP; https://blast.ncbi.nlm.nih.gov/) from the NCBI. The *Vibrio* species chosen to be analysed were *Vibrio alginolyticus*[52], *Vibrio parahaemolyticus*[53], *Vibrio salmonicida*[54], *Vibrio anguillarum*[55], *Vibrio ordalli*[56], *Vibrio proteolyticus*[57] and *Vibrio campbelli*[58] due to being able to infect fish.

To analyse regions with high a probability of being immunogenic, the free-to-use softwares used were: NetSurfP (Center for biological sequence analysis; http://www.cbs.dtu.dk/services/NetSurfP/), that predicts the surface accessibility and secondary structure of amino acids in an amino acid sequence; Pred-TMBB (Biophysics and Bioinformatic Laboratory; http://bioinformatics.biol.uoa.gr/PRED-TMBB/), a method based on the Hidden Markov Model (HMM), that predicts the transmembrane beta-strands of gram-negative bacteria outer membrane proteins, and discriminating such proteins from other water-soluble proteins when screening large datasets; BepiPred (Center for biological sequence analysis; http://www.cbs.dtu.dk/services/BepiPred/), that predicts the location of linear B-cell epitopes using a combination of a hidden Markov model (HMM) and a propensity scale method; and Antigenic (EMBOSS; http://www.bioinformatics.nl/cgi-bin/emboss/antigenic) that predicts potentially antigenic regions of a protein sequence using a semi-empirical process based on physicochemical properties of amino acid residues and their frequencies of occurrence in experimentally known segmental epitopes.

To verify if the antigenic regions predicted in the above tools are conserved between all the bacterial species known to infect fish, an alignment was made using Clustal Omega (EMBL-EBI; https://www.ebi.ac.uk/Tools/msa/clustalo/). Clustal Omega is a multiple sequence alignment program that uses seeded guide trees and HMM profile-profile techniques to generate alignments between three or more sequences. The information generated by Clustal Omega was then used in a program for multiple sequence alignment editing, visualisation and analysis (Jalview; http://www.jalview.org) to analyse the conservative regions and compare with the immunogenic regions predicted.

Using the Pred-TMBB software, a 2D representation of the protein was produced to analyse the regions that are predicted to interact with the extracellular medium. Analysis of OmpK structure was also conducted using the program I-TASSER (Iterative Threading ASSEmbly Refinement; https://zhanglab.ccmb.med.umich.edu/I-TASSER/), from the University of Michigan. This program starts by identifying structural templates from the PDB (Protein Data Bank; https://www.rcsb.org/pdb/home/home.do) by the multiple threading approach LOMETS (Local Meta-Threading-Server;https://zhanglab.ccmb.med.umich.edu/LOMETS/), with full-length atomic models constructed by iterative template fragment assembly simulations. Function insights of the

target were then derived by threading the 3D models through the protein function database BioLiP (Ligand-protein Binding Database; https://zhanglab.ccmb.med.umich.edu/BioLiP/).

3 - Results and Discussion

3.1 – Bioinformatic analysis

Previous work has identified *Vibrio* strains associated with fish infection. The *Vibrio* species identified were: *Vibrio alginolyticus*, *Vibrio parahaemolyticus*, *Vibrio salmonicida*, *Vibrio anguillarum*, *Vibrio ordalli*, *Vibrio proteolyticus*, and *Vibrio campbelli*. Outer membrane proteins are recognized as important for bacterial protection against harsh environmental conditions[59]. Among Omps, OmpK was chosen as a model protein to develop vaccines against *Vibrio* infections. In order to analyse the prospective efficacy of the protein chosen as a vaccine candidate, a BLASTP analysis of OmpK protein sequences of these species was performed. The published amino acid sequences of proteins from all the strains of the target species (*Vibrio alginolyticus* ATCC 17749) were retrieved and the results of the BLASTP analysis are summarized in Table 16 obtained from the alignment of OmpK proteins from *Vibrio* species.

The results for *ompK* gene, show that the protein has a high level of identity among all the strains of the *Vibrio* species know to infect fish, with values ranging from 65% to 100% with the exception of *Vibrio Salmonicida*, in which no identity or similarity was found. The similarity for most of the species analysed ranging between 76% and 100% (Table 16) showing that some of the altered codons did encode the same amino acids. Identity is the extent to which two amino acid sequences have the same residues at the same positions in an alignment, while similarity is the extent to which protein sequences are related some amino acids can be substituted by other with the same characteristic. The results indicate that OmpK has a conserved sequence among the analysed *Vibrio* species

Table 16 - Results of the BLASTP analysis for the alignment of OmpK sequences with *Vibrio alginolyticus*, *Vibrio parahaemolyticus*, *Vibrio salmonicida*, *Vibrio anguillarum*, *Vibrio ordalli*, *Vibrio proteolyticus* and *Vibrio campbelli* strains. The (-) means a lack of correspondence in the alignment

Results from BlastP for OmpK protein							
Species	strain	Identity	Similarity	Species	strain	Identity	Similarity
Vibrio alginolyticus	ATCC 17749	100	100	*Vibrio anguillarum*	O1	86	91
	NBRC 15630	100	100		O2	88	93
	12G01	99	99		775	86	91
	E0666	78	85		M3	86	91
	40B	99	100		RV22	88	93
	chemovar iophagus	-	-		96F	86	91
	subs. Iophagus	-	-	*Vibrio ordalli*	phage PV05	-	-
Vibrio parahaemolyticus	AQ3810	86	91		ATCC 33509	-	-
	V-223/04	86	91		LMG 13544	-	-
	RIMD 2210633	85	89		NCIMB 2167	-	-
	O1:K33	86	91		12B09	88	93
	O1:KUK	86	89		FF-93	88	93
	str. 16	65	76		FS-144	88	93
	VPCR-2010	86	91		FS-238	88	93
	EKP-028	85	89	*Vibrio proteolyticus*	NBRC 13287	75	82
	EKP-021	85	89	*Vibrio campbelli*	ATCC BAA-1116	86	89
	VPTS-2010	86	91		ATCC 25920	85	89
	VP 2007-007	85	92		CAIM 519/NBRC 15631	85	89
	VPTS-2010-2	86	91		LMG 11216	85	89
	EKP-008	86	91		HY01	81	87
	IDH 02640	85	89		CAIM 333	-	-
	10296	86	91		LMG 21362	-	-
	861	85	89		M1	-	-
	3259	86	91		DS40M4	81	87
Vibrio salmonicida	LFI 1238	-	-				

An *in silico* analysis of the potential antigenic regions of each protein amino acid was performed. For this approach, the four softwares NetSurfP, Pred-TMBB, BepiPred and Antigen were used (Table 17). NetSurfP and Pred-TMBB predict regions with surface accessibility, while BepiPred and Antigen predict localization of potential epitopes.

Table 17 - NetSurfP and Pred-TMBB predicted regions of *Vibrio alginolitycus* OmpK with higher surface accessibility and BepiPred and Antigenic predicted regions of *Vibrio alginolitycus* OmpK epitopes

OmpK											
Surface Accessibility						Epitopes					
NetSurfP			Pred-TMBB			BepiPred			Antigenic		
Start	End	Size (aa)	Start	End	Size (aa)	Start	End	Size (aa)	Start	End	Size (aa)
79	91	13	76	99	24	79	89	11			
128	131	4	127	146	20	127	134	8	133	138	6
200	207	8	176	216	41	199	217	19	187	196	10
245	249	5	238	254	17	242	253	12	238	244	7

The objective of this analysis was the identification, for the proteins under study, of amino acid segments of at least 10 residues. The results in Table 17 shows that at least 4 regions composed of more than 10 amino acids residues were predicted for OmpK by the program Pred-TMBB and only one segment was predicted by NetSurfP. Results from the immunogenic activity predicted by softwares BepiPred and Antigenic are also presented in Table 17. These results show some similarity with the results in the other two programs but low intersection.

We have then analysed the conservation degree of the predicted amino acid segments among *Vibrio* species with available genome sequences, through an alignment using Clustal Omega. The Jalview software was used to plot the results as shown in Figure 3.

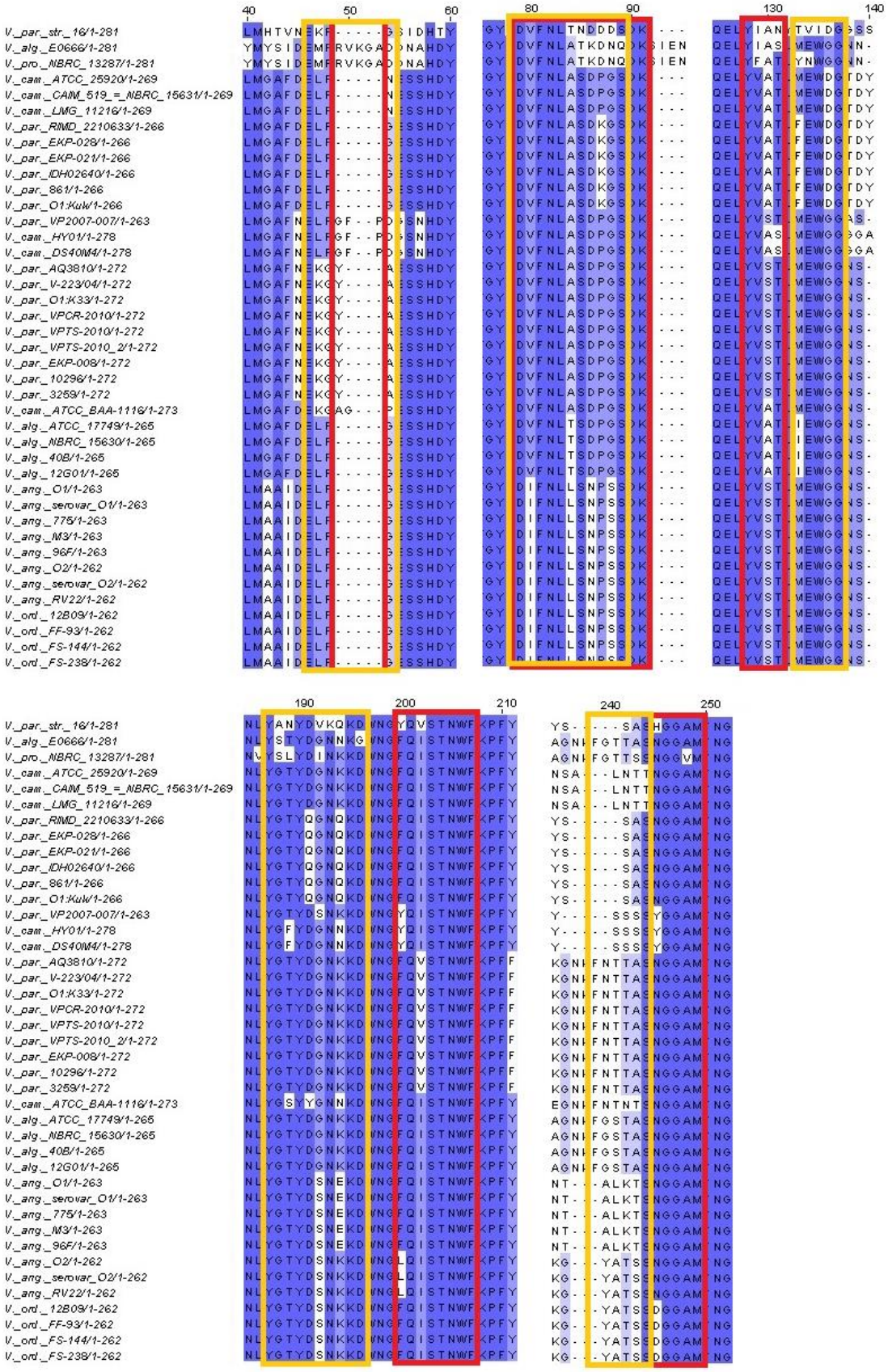


Figure 3 – Regions of interest from the results of the multiple alignment performed by Clustal Omega analysed with Jalview software. Regions predicted by NetSurfP and Pred-TMBB are marked in red and regions predicted by BepiPred and Antigenic are marked in yellow.

Figure 3 shows overlap of regions predicted by softwares NetSurfP and Pred-TMBB and softwares BepiPred and Antigen. The intersection of the two surface accessibility programs, shown in red revealed one region in the non conserved amino acid 49 to 53, while the other four regions are mostly conserved in all the species. Results for the softwares BepiPred and Antigen show that two of the regions predicted are not conserved and correspond to amino acid 49 to 53 and 238 to 244. A higher confidence on the occurrence of immunogenic segments would result if there is an intersection between the programs that predict surface accessibility and programs that predict epitopes in a conserved section, which is the case of the amino acid 79 to 89 segment in Figure 3. Another difference in the outputs from the softwares is that NetSurfP, Pred-TMBB and BepiPred predicted regions at the beginning and at the end of the protein, while the Antigen software only predicted the occurrence of these regions at the end of the protein.

Pred-TMBB software was used to obtain 2D representation of the protein that allow for the visualization of the protein regions that interact with the extracellular space (Figure 4).

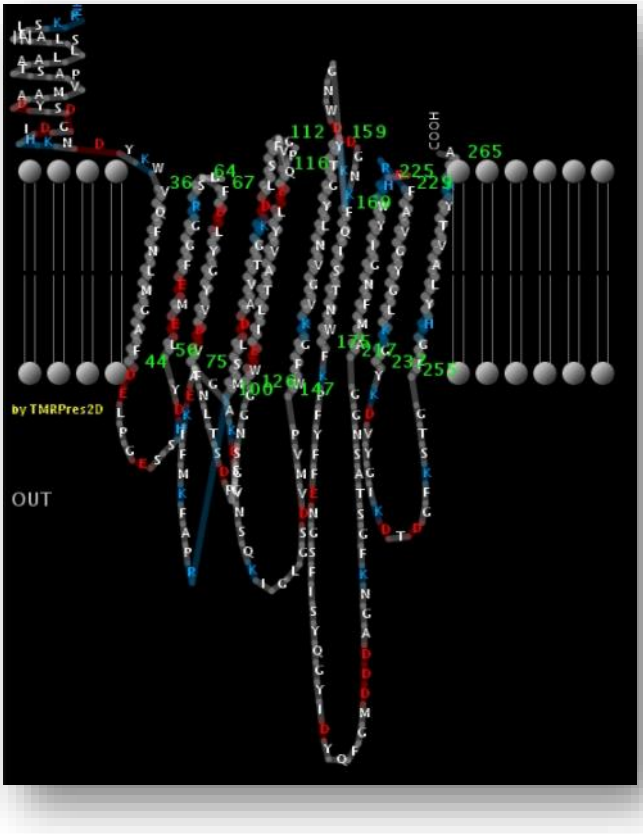


Figure 4 – 2D image of the protein structure showing the surface regions as predicted by Pred-TMBB

The results show that the predicted 2D image of OmpK in the membrane exhibits to the extracellular space the amino acid residues that correspond to the predicted immunogenic regions shown using the software Pred-TMBB.

The prediction of the *V. alginolyticus* OmpK structure was also conducted using the program I-TASSER (Iterative Threading ASSEmbly Refinement) resident at the University of Michigan web server. The results are shown in Figure 5.

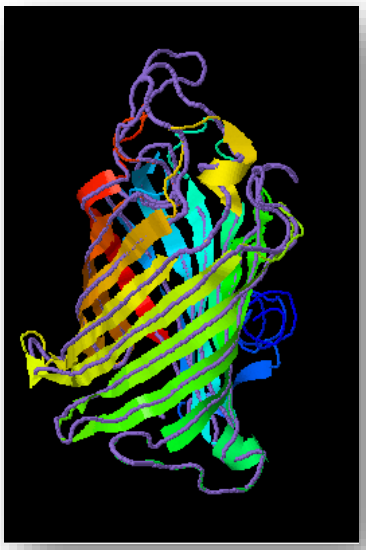


Figure 5 – 3D structure predicted by the program I-TASSER

The I-TASSER program also identify the top 10 analog proteins in the PDB database which are shown in Table 18.

Table 18 - Top 10 analogs, of V. alginolyticus OmpK protein present in the PDB database predicted by the I-TASSER program.

Rank	Organism	Protein	PDB Hit	TM-score	IDEN	Cov
1	*E. coli*	Tsx	1tlyA	0.825	0.275	0.838
2	*E. coli*	NanC	2wjqA	0.598	0.056	0.728
3	*E. coli*	Phospholipase A	1ilzA	0.581	0.059	0.766
4	*Pseudomonas aeruginosa*	EstA	3kvnX	0.569	0.077	0.747
5	*Pseudomonas putida F1*	COG4313	4rl8A	0.563	0.049	0.777
6	*Klebsiella oxytoca*	CymA	4d51A	0.559	0.064	0.834
7	*Pseudomonas aeruginosa*	OprP	2o4vC	0.557	0.075	0.774
8	*Neisseria meningitidis*	NalP	1uynX	0.554	0.071	0.728
9	*Providencia stuartii*	Omp-Pst1	4d64A	0.550	0.088	0.743
10	*Pseudomonas sp. UK4*	FapF	5o65A	0.549	0.093	0.774

The top 10 analogs present in Table 18 were analysed by the software and were ranked based on the TM-score of the structural alignment between the query structure and known structures in PDB library. Two other parameters were used to analyse the analogs found (IDEN and Cov). IDEN is the percentage sequence of identity in the structurally aligned region and Cov

represents the coverage of the alignment by TM-align and is equal to the number of structurally aligned residues divided by length of the query protein. The PDB hit with better score was Tsx structure in *E. coli*. Tsx is a nucleoside-specific outer membrane (OM) transporter of Gram-negative bacteria[60].

3.2 – Construction of transformants of *E. coli* BL21 with the OmpK gene

Our bioinformatics analyses suggest that OmpK contains immunogenic segments and therefore, the protein might be used as a potential vaccine. In order to produce this *Vibrio* protein in *E. coli* a strategy was designed including the amplification of the gene by PCR, and its cloning in a controlled expression vector that would lead to the production of a protein with a his-tag at its C-terminus to facilitate the purification.

Therefore, cultures of *Vibrio alginolyticus* ATCC 17749 were performed, cells were harvested by centrifugation, and total DNA was extracted. The quantity and quality of extracted DNA was assessed by spectrophotometry (nanodrop) (Table 19) as described in 2.2.1.1.

Table 19 - Concentration and purity indicators of genomic DNA extracted from *V. alginolyticus* cells.

Genomic DNA			
Name of the sample	Quantity (ng/µl)	A(260/280)	A(260/230)
Vibrio	73.9	1.97	2.46

The extracted DNA was then used as a template in reaction mixtures to amplify by PCR the genes encoding for OmpK, with adequate primer pairs. The results obtained are shown in Figure 6.

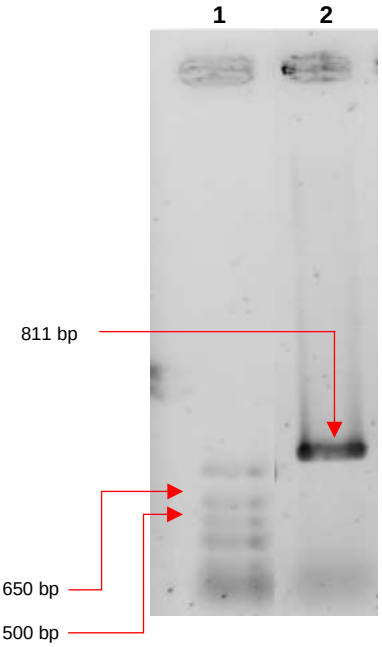


Figure 6 - Photograph of agarose gel prepared to analyse the PCR products obtained using *Vibrio alginolyticus* total DNA as template and primers adequate to amplify the OmpK encoding genes. Lane 1 contains the DNA molecular mass standard, while lane 2 contains the amplification product with the expected size.

The analysis of the amplified gene encoding OmpK agarose by gel electrophoresis indicates that the amplified DNA fragments had the expected size. The band was then extracted from the agarose gel in order to be cloned in the pET23a+ expression vector.

Plasmid pET23a+ was extracted from *E. coli* BL21 using the kit NZYMiniprep from nzytech, according to the manufacturer's instructions. Plasmid pET23a+, as well as the DNA fragment corresponding to OmpK amplified by PCR, were digested with the restriction enzymes XhoI and BamHI. After recovering the digested plasmid and DNA fragments, ligation mixtures were prepared to ligate each fragment to the linearized plasmid pET23a+. The resulting ligation mixtures were used to transform *E. coli* DH5α. The transformants were plated on LB medium containing 150 µg/ml of ampicillin. After incubation at 37ºC, a few colonies were selected and plasmid DNA was extracted. Results presented in Figure 7 show the restriction analysis of recombinant plasmid harbouring the *ompK* gene.

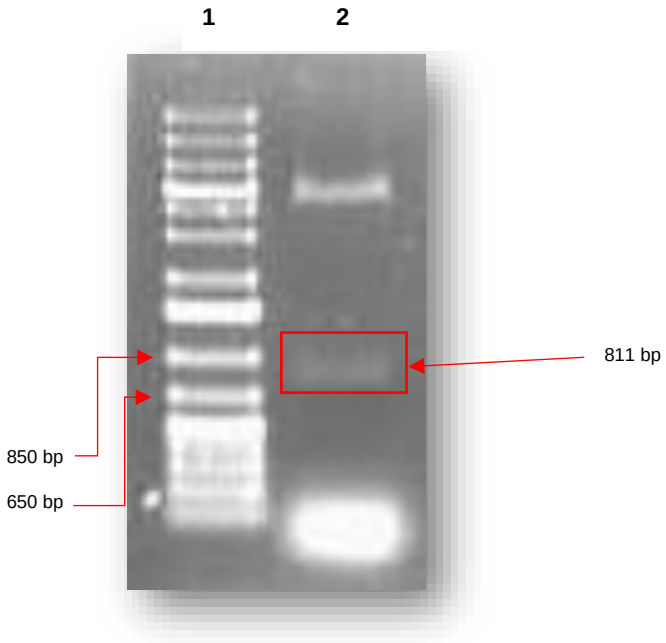


Figure 7 - Photograph of the agarose gel used to confirm that the fragment corresponding to the *ompK* gene was correctly ligated to the expression vector pET23a+. Lanes 1, molecular mass standard; 2, restriction analysis of recombinant plasmid pET23a+ with the DNA fragment corresponding to the *Vibrio alginolyticus ompK* gene.

The recombinant plasmid was purified from *E. coli* DH5α cultures using the commercial kit NZYMiniprep from NZYtech. The plasmid was sent for sequencing by a commercial company (Eurofins MWG Operon) to confirm that the correct sequence was cloned. A region from the results obtained are shown below, where the *ompK* gene sequence is in bold.

GATAATTTTGTTTACTTTAAGAAGGAGATATACATATGGCTAGCATGACTGGTGGAC
AGCAAATGGGTCGCGGATCC**ATGCGTAAATCACTTTTAGCTCTTAGCCTTCTAGCGGCTA**
CTTCAGCTCCAGTTATGGCGGCAGATTACTCTGACGGCGATATTCACAAAAACGATTACA
AGTGGATGCAATTTAACCTAATGGGTGCATTTGACGAACTTCCAGGTGAGTCTTCACACG
ACTACCTAGAAATGGAATTCGGCGGCCGCTCTGGTATCTTCGACCTTTACGGCTACGTTG
ATGTATTCAACCTAACTTCTGATCCAGGTAGCGACAAAGCTGGCGCAGAGAAGATCTTTA
TGAAATTCGCTCCACGTATGTCTCTAGATGCGGTAACTGGTAAAGACCTATCTTTCGGTC
CAGTTCAAGAGCTATACGTTGCAACTCTAATCGAGTGGGGTGGTAACTCTGGTGTTAACT
CTCAAAAAATCGGTCTAGGTTCTGACGTAATGGTTCCTTGGTTTGGTAAAGTTGGTGTTAA
CCTATACGGCACTTACGACGGCAACAAAAAGGATTGGAACGGTTTCCAAATCTCGACTA
ACTGGTTCAAACCATTCTACTTCTTCGAGAACGGTTCATTCATTTCTTACCAAGGCTACAT
CGATTACCAATTCGGTATGGATGATGATGCTGGTAACAAGTTTGGTTCAACCGCGTCAAA
CGGTGGTGCAATGTTCAACGGTATCTACTGGCACTCAGACCGCTTCGCAGTGGGTTACG
GTCTAAAAGGTTACAAAGACGTGTACGGTATCAAAGATACTGATGGGTTCAAGTCAACTG
GCTTTGGCCACTACCTTGCTGTAACTTACAAGTTCCTCGAGCACCACCA

Results of nucleotide sequencing show that the DNA insert was correct and therefore the recombinant plasmid can be used to transform *E. coli* BL21.

3.3 – Overexpression of OmpK protein by the *E. coli* BL21 (DE3) transformants

After obtaining *E. coli* BL21 colonies with the recombinant plasmid bearing *ompK* experiments were carried out to find the best conditions for the overexpression of OmpK. Several culture media, temperature and time of incubation were tested (Table 20).

Table 20 - Summary of the culture conditions tested to optimize OmpK overexpression, the respective OD values and a qualitative analysis of the amount of OmpK protein present.

Medium	Temperatures (ºC)	Time (hours)	OD_{640nm}	Quantity of OmpK
LB	37	1	0.85	+
		2	1.73	++
		3	2.96	++
		16	5.35	+++
SB	20	16	8.78	+
	25	16	8.83	+
	30	16	8.93	++++
	37	1	1.20	++
		2	2.01	+++
		3	2.60	+++
		16	8.84	++++

The initial objective was to find the time of cultivation and medium composition that led to a higher amount of protein overexpressed. The results in Table 20 show that 16 hour incubation with SB medium at 30 and 37ºC were the optimal conditions found for overexpression.

3.4 - Purification of the *V. alginolyticus* OmpK protein

After optimization of the conditions for a high production of OmpK, the purification of the protein was started. Initially the protocol described in 2.3.1 included the sonication of each sample with 6 cycles. The number of sonications was changed due to the high amount of cells in the 16 hour cultures (2,89 g) that required more sonication cycles to efficiently disrupt all the cells in the suspension and release the OmpK produced. The next step was to optimize conditions for protein purification by affinity chromatography namely to determine the necessary concentrations of imidazole necessary so that firstly all the impurities leave the column and afterwards the OmpK protein elutes from the column. Figure 8 shows the SDS-PAGE image in which fractions from a column affinity chromatography are compared.

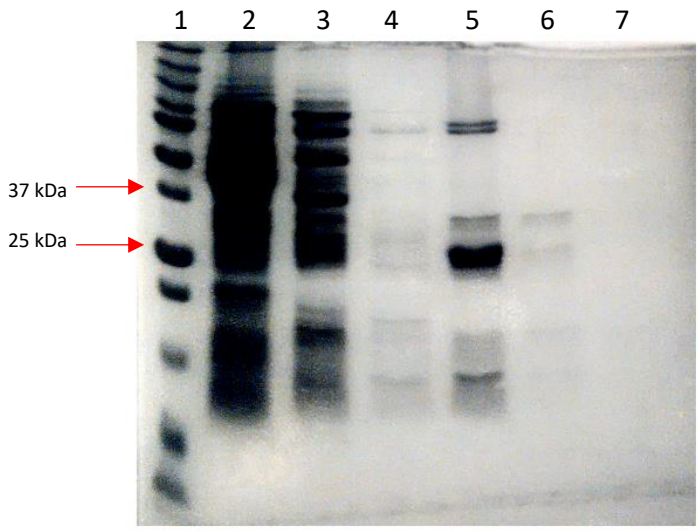


Figure 8 - SDS-PAGE gel showing profiles of chromatographic fractions from steps performed at different imidazole concentrations on a His-tag FF column. Lanes: 1 – molecular mass marker; 2 – chromatographic feed; 3 - wash fraction with 10 mM of imidazole; 4, 5, 6, 7 – elution fraction containing 60, 100, 200, 300 mM of imidazole.

The results presented in Figure 8 show that using the standard conditions of imidazole concentrations of 60, 100, 200, 300 and 500 mM, impurities were eluted at the same concentration of imidazole required to elute OmpK, not allowing its purification. In order to decrease the amount of impurities several other conditions were tested: decreasing the concentration of imidazole and increasing the volume of buffers used (Figure 9); using filters with a specific cut-off that would allow concentration of the protein and elimination of part of the impurities (Figure 10) and use of glycerol (Figure 11). The rational using of glycerol was to

stabilize the protein in the renaturation process preventing precipitation of OmpK during the dialysis step[61].

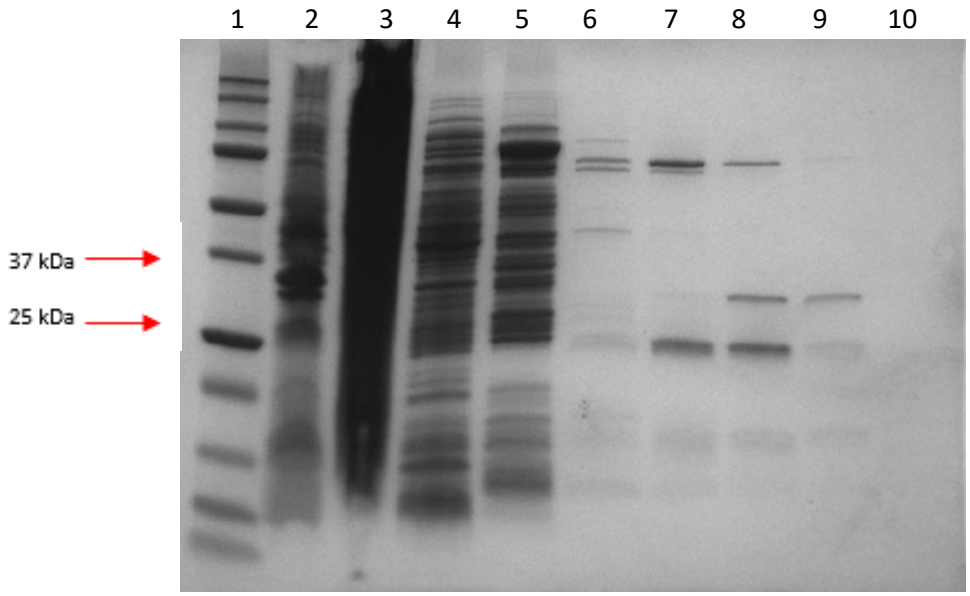


Figure 9 – SDS-PAGE gel showing profiles of chromatographic fractions from steps performed at different imidazole concentrations and increased volume from 5 ml to 10 ml of the fraction containing 70 mM of imidazole on a His-tag FF column. Lanes: 1 – molecular mass marker; 2 – sample of the pellet collected after the first centrifugation; 3 – chromatographic feed; 4 – wash fraction with 10 mM of imidazole; 5, 6, 7, 8, 9, 10 - elution fraction containing 40, 50, 70, 100, 200, 500 mM of imidazole.

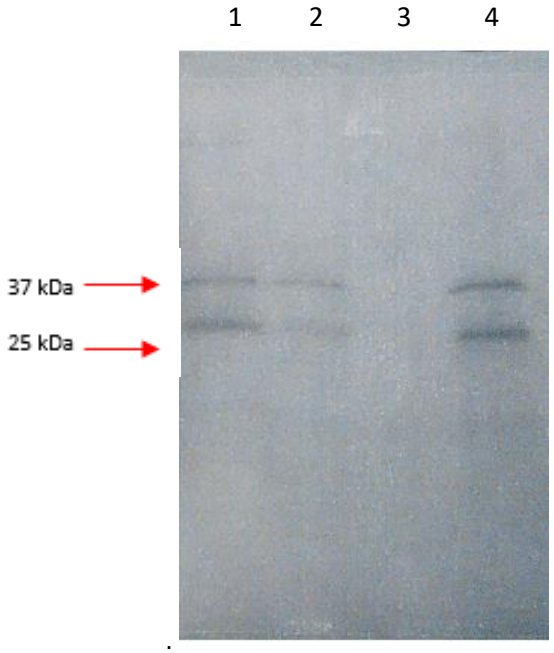


Figure 10 - SDS-page gel showing the results obtained when cut-off filters of 30 kDa were used. Lanes: 1 – eluate obtained with buffer containing 100 mM of imidazole, 2 – eluate obtained with buffer containing 200 mM of imidazole, 3 – eluate obtained with the filtrate of the filter with a cut-off of 30 kDa, 4 - sample of the retentate of the filter with a cut-off of 30 kDa.

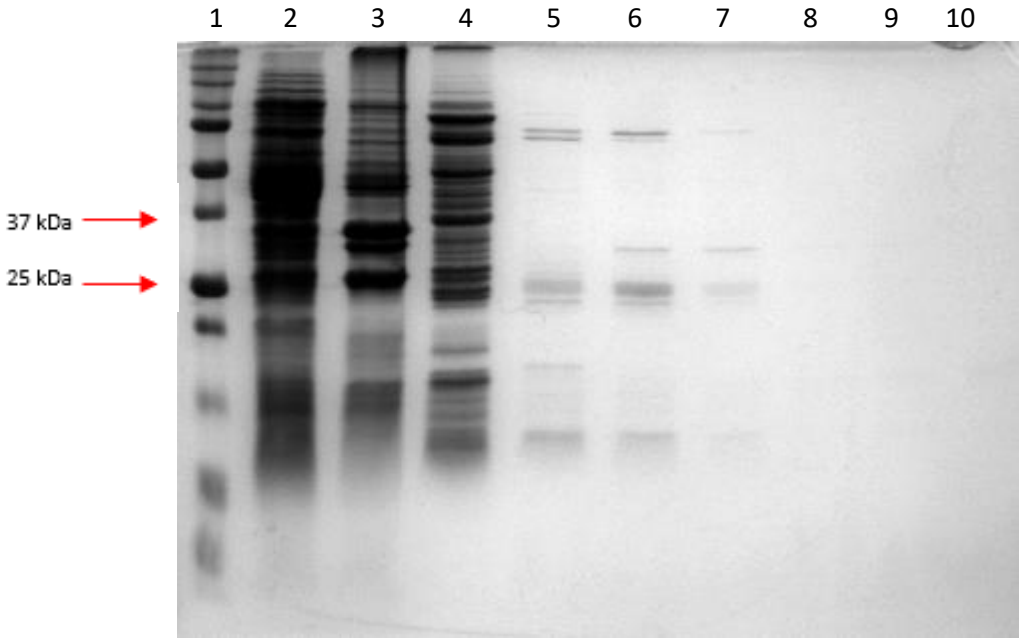


Figure 11 - SDS-PAGE gel showing profiles of chromatographic fractions from steps performed at different imidazole concentrations and using 10% glycerol in the buffers on a His-tag FF column. Lanes: 1 – molecular mass marker, 2 – sample of the culture after 16 hours of cultivation, 3 – sample of the pellet collected after the first centrifugation, 4, 5, 6, 7, 8, 9, 10 – elution fraction containing 40, 50, 70, 85, 100, 200, 500 mM of imidazole.

As shown in Figures 8 to 11, with the exception of the use of concentration of glycerol of 10% all the other conditions did not improve the purification process.

The use of filters with specific cut-off had the objective of concentrating the sample and try to separate some of the impurities from OmpK. OmpK as a molecular mass of 30.11 kDa so to separate the impurities with lower molecular weight a cut-off off 30 kDa was used (Figure 10). The results showed that both the protein of interest and the lower molecular weight impurities were retained by the filter.

Although some of these modifications led to a decrease of impurities, they did not allow the purification of OmpK. Due to the contamination present, a protocol for extracting the periplasmatic protein was tested (Figure 12) relying on the fact that the recombinant OmpK produced contains a signalling peptide sequence that allows its transportation for the periplasm. The temperature used in this protocol (20ºC) was different than in the normal growth (30ºC) to increase the amount of OmpK being transported to the periplasm[62,63].

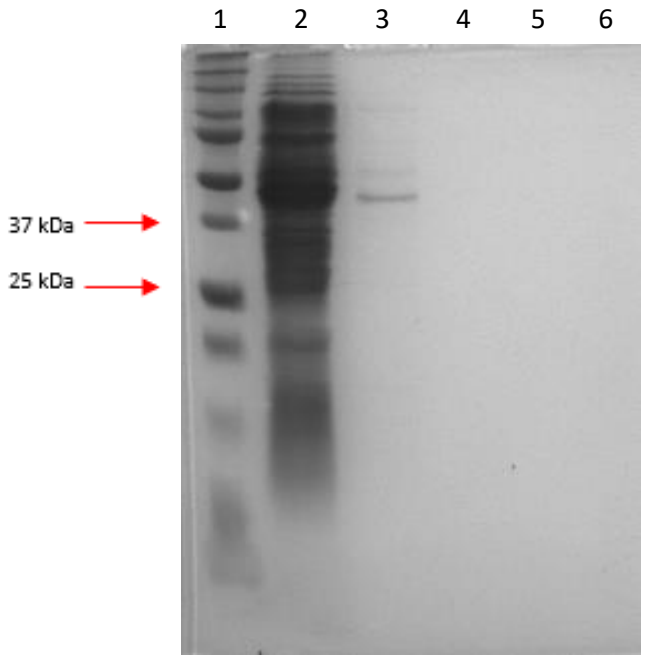


Figure 12 - SDS-PAGE gel showing profiles of chromatographic fractions from steps performed at different imidazole concentrations from the purification of periplasmatic proteins on a His-tag FF column. Lanes: 1 – molecular mass marker, 2 – sample of the culture after 16 hours of cultivation, 3 – sample of the pellet collected after the first centrifugation, 4, 5, 6 – elution fraction containing 40, 50, 200 mM of imidazole.

The results presented in Figure 12 show no bands with the correct weight due to the high volume of the final solution (200 ml). The final solution was concentrated using a 9K MWCO concentrator with 7 ml (Pierce) from 200 ml to 40 ml. However, SDS-PAGE gel analysis of the purified protein factions still show no bands with the correct weight. After testing these conditions, a western blot was performed to verify the authenticity of the protein present in the SDS-PAGE gel (Figure 13).

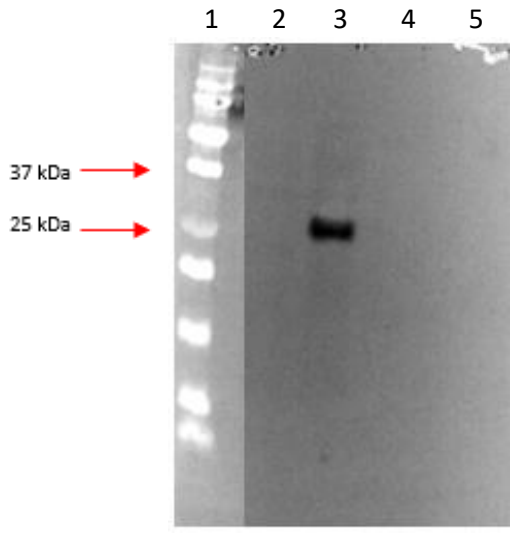


Figure 13 – western blot image of periplasmatic proteins with an antibody anti-His. Lanes: 1 – molecular mass marker, 2 – sample of the culture after inoculation, 3 – sample of the culture after 16 hours of cultivation, 4 – eluate obtained with buffer containing 40 mM of imidazole showed in figure 9 lane 4, 5 - eluate obtained with buffer containing 100 mM of imidazole showed in figure 9 lane 8.

The western results show that the protein present in the samples was not OmpK due to the appearance of a signal only in the sample corresponding to the 16 hour culture. These results prompted us to hypothesize that the histidine tail of OmpK was hidden in the protein native structure and therefore not allowing its binding to the nickel ions in the column.

3.5 – Purification of denatured OmpK

To test the hypothesis that the histidine tail was hidden, the protein recovered as inclusion bodies was denatured prior to purification according to the protocol described in 2.3.3.1.1 with some minor differences. The purification protocol was carried out with imidazole concentrations of 10, 60, 100, 200, 300 and 500 mM; glycerol was not added to the buffers and DTT was added only 15 minutes after resuspending in buffer with 8M of urea. The SDS-PAGE gel obtained for samples using denatured protein showed a strong band with the correct weight being eluted in a buffer with high concentration of imidazole (200mM) (Figure 14), thus supporting our hypothesis.

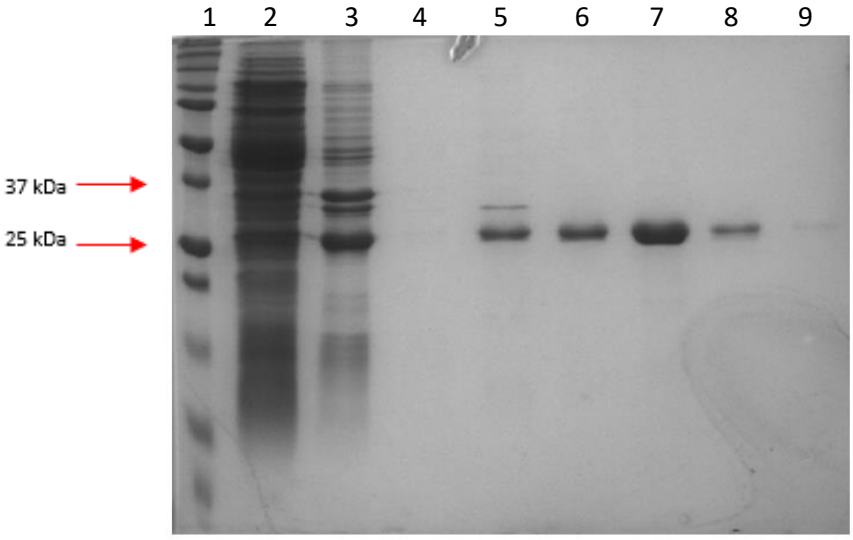


Figure 14 - SDS-PAGE gel showing profiles of chromatographic fractions from steps performed at different imidazole concentrations with OmpK denatured prior the purification on a His-tag FF column. Lanes: 1 – molecular mass marker, 2 – sample of the culture after 16 hours of cultivation, 3 – sample of the supernatant collected after the first centrifugation, 4 – wash fraction with 10 mM of imidazole, 5, 6, 7, 8, 9 - elution fraction containing 60, 100, 200, 300, 500 mM of imidazole.

Results in Figure 14 also show that a protein with the corresponding weight to OmpK is being eluted with all the buffers used, although in higher concentrations when using buffers with 100, 200 or 300 mM of imidazole. The presence of the protein in eluate buffers might be explained by the existence of a fraction of protein that did not completely denaturated, and thus led to a lower affinity to the nickel in the column.

To confirm the success of this protocol a western blot was performed (Figure 15).

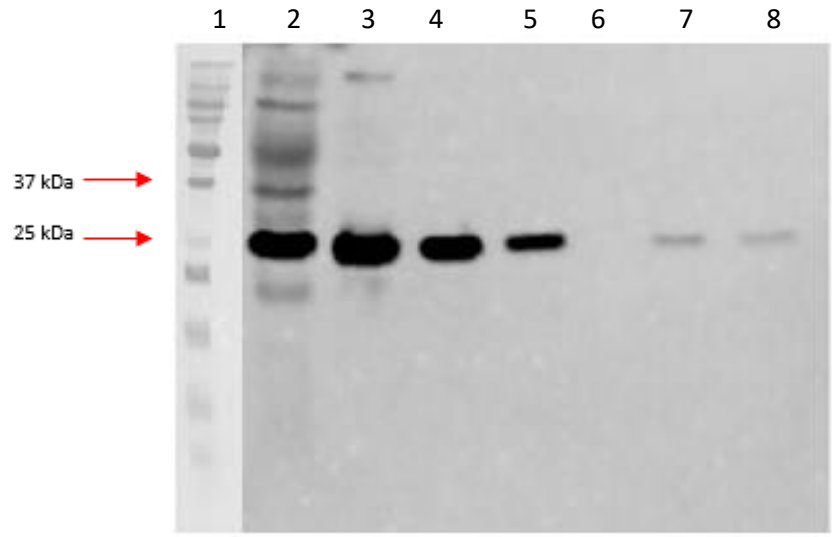


Figure 15 – Photograph of the western blot obtained using an antibody anti-His. Lanes: 1 – molecular mass marker, 2 – sample of the culture after 16 hours of cultivation, 3 – eluate obtained with buffer containing 60 mM of imidazole, 4 – eluate obtained with buffer containing 200 mM of imidazole, 5 - eluate obtained with buffer containing 300 mM of imidazole, 6 - empty, 7 - eluate obtained with from the protein solution post-dialysis, 8 - eluate obtained with buffer containing 200 mM of imidazole post-dialysis.

The results in Figure 15 show a strong signal corresponding to the bands with the correct weight and that could also be seen in the respective SDS-PAGE gel. These results allow us to conclude that the histidine tail of OmpK is hidden when the protein is overexpressed by *E. coli*. Although the protein denaturation allows the purification of OmpK without any impurities, the protein is required in the native state when envisaging its future use. To finalize the renaturation process and eliminate urea from the protein solution, a final step of dialysis was carried out (figure 16). The fractions collected with the buffer with 100, 200 and 300 mM of imidazole were pooled and dialyzed.

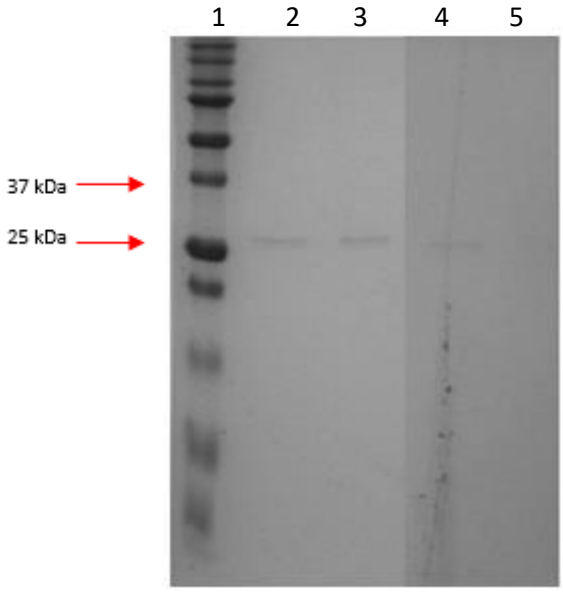


Figure 16 – Photograph of a SDS-PAGE gel analysis of samples of denatured OmpK purification post-dialysis. Lanes: 1 – molecular mass marker, 2 – sample of the solution post-dialysis before passing though the column, 3 – sample of the post-dialysis solution that passed though the column, 4 – eluate obtained with buffer containing 200 mM of imidazole.

Figure 16 shows a significant decrease of protein in the affinity eluate when compared with the pre-dialysis sample. In order to reduce the losses of purified protein, the protocol described in 2.3.3.1.1 was used. In this protocol, DTT was added 30 minutes after resuspension in the denaturation buffer to increase the amount of denaturated proteins. The increase of time without DTT benefits the denaturation although DTT is used as an unfolding agents in concentrations between 1 – 10 mM[64]. As described in Alliegro[65] DTT function is not always connected with the presence of cysteine. The concentration used cannot be higher than 5 mM due to limits of concentration of this reagent that can be used in the affinity column (figure 17).

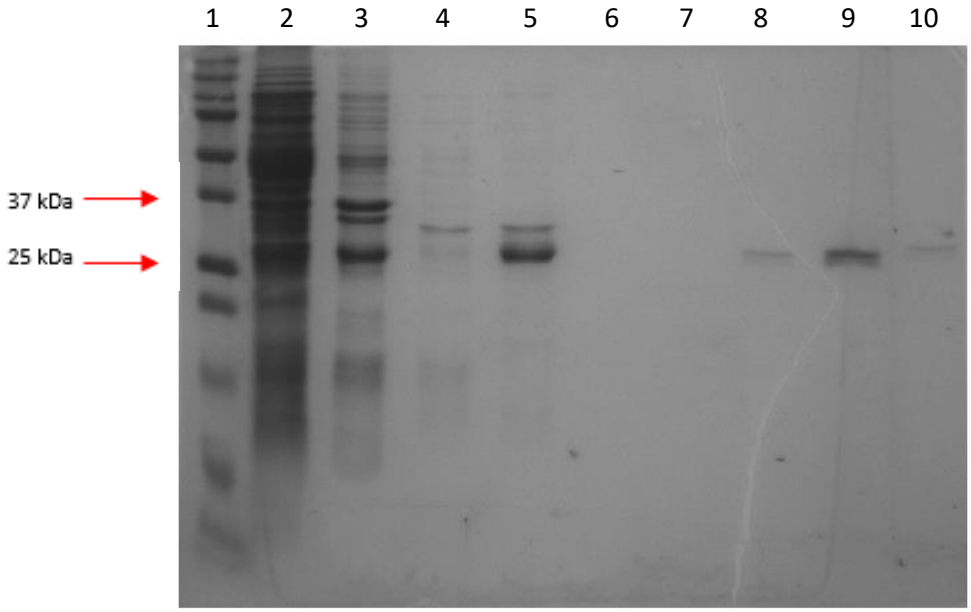


Figure 17 – SDS-PAGE gel showing profiles of chromatographic fractions from steps performed at different imidazole concentrations of denatured OmpK purification using DTT 30 minutes after resuspension on a His-tag FF column. Lanes: 1 – molecular mass marker, 2 – sample of the culture after 16 hours of cultivation, 3 – sample of the supernatant collected after the first centrifugation, 4, 5 – wash fraction with 10, 60 mM of imidazole, 6, 7, 8, 9 - eluate fraction with 100, 200, 300, 500 mM of imidazole.

As shown in Figure 17 the use of DTT led to a decrease of the protein loss in the buffer with 60 mM of imidazole[66].

3.6 - Column and batch protocols

To increase the amount of protein purified and renatured two methods were tested: renaturation in column and in batch as described in 2.3.3.1 and 2.3.3.2. Results presented in figures 18, 19 and 20, 21 show the results pre- and post-dialysis results for the column-based protocol and batch-based protocol, respectively.

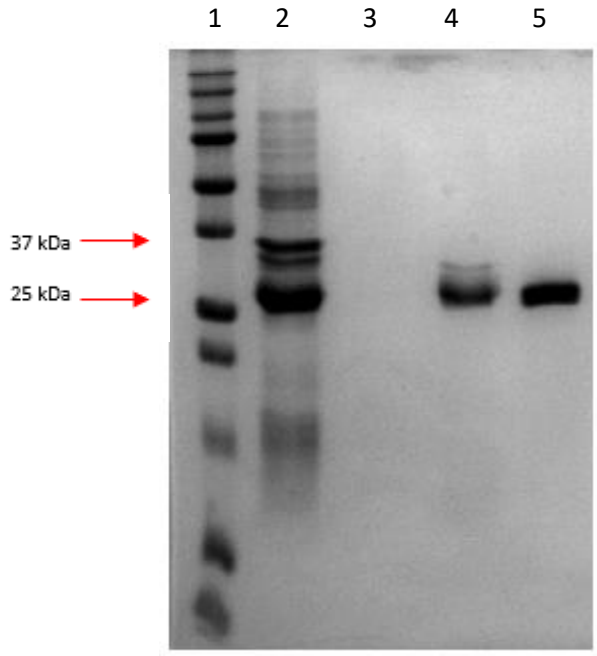


Figure 18 – SDS-PAGE gel showing profiles of chromatographic fractions from steps performed at different imidazole concentrations of denatured OmpK purification samples using the column-based protocol pre-dialysis. Lanes: 1 – molecular mass marker, 2 – sample of the protein solution before passed in the column, 3, 4 – wash fraction with 10, 50 mM of imidazole, 5 – eluate fraction with 300 mM of imidazole.

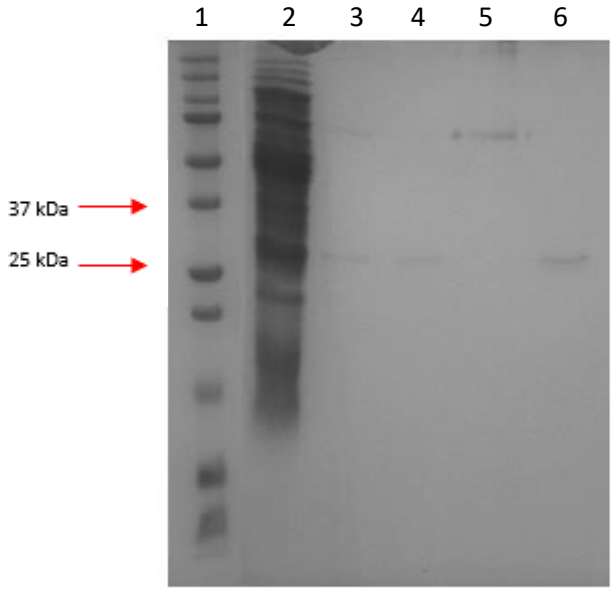


Figure 19 – SDS-PAGE gel showing profiles of chromatographic fractions from steps performed at different imidazole concentrations of denatured OmpK purification samples obtained with the column-based protocol post-dialysis. Lanes: 1 – marker, 2 - sample of the culture after 16 hours of cultivation, 3 – sample of the post-dialysis protein solution before passed in the column, 4 – sample of the post-dialysis protein solution after passed in the column, 5 – eluate obtained with buffer containing 10 mM of imidazole, 6 - eluate obtained with buffer containing 500 mM of imidazole.

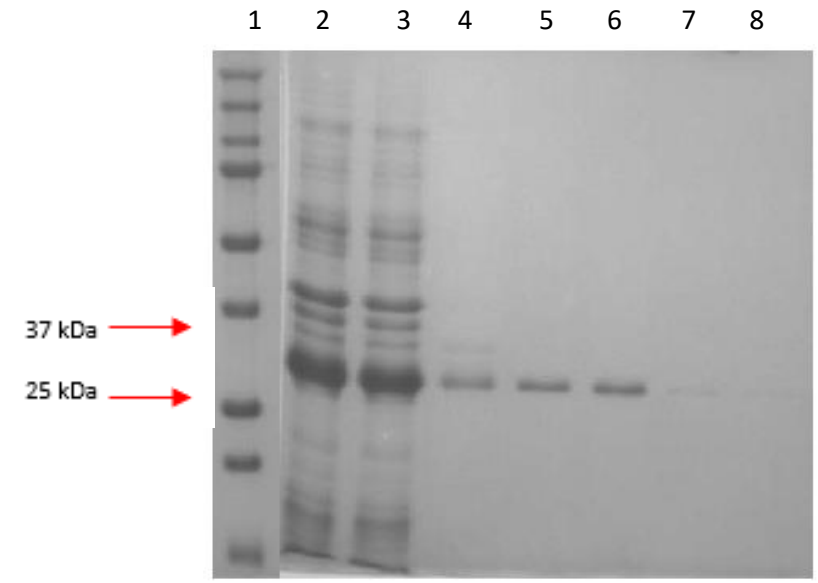


Figure 20 – SDS-PAGE gel showing profiles of chromatographic fractions from steps performed at different imidazole concentrations of denatured OmpK purification samples obtained with the batch-based protocol pre-dialysis. Lanes: 1 – molecular mass marker, 2 – sample of the protein solution before passed in the column, 3, 4 – wash fraction with 10, 50 mM of imidazole, 5, 6 – eluate fraction 300 mM of imidazole; 7, 8 -. eluate fraction 500 mM of imidazole.

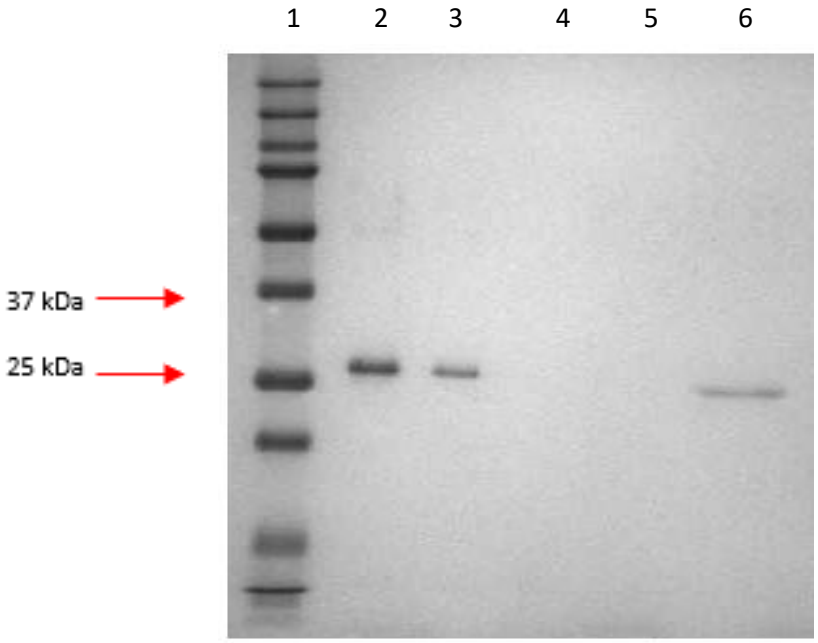


Figure 21 – SDS-PAGE gel showing profiles of chromatographic fractions from steps performed at different imidazole concentrations of denatured OmpK purification samples obtained with batch-based protocol post-dialysis. Lanes: 1 – molecular mass marker, 2 – sample of the protein solution before passed in the column and post-dialysis, 3 – sample of the protein solution after passed in the column and post-dialysis, 4 – empty, 5 – eluate obtained with buffer containing 10 mM of imidazole, 6 - eluate obtained with buffer containing 500 mM of imidazole.

The column-based and batch-based protocols for the IMAC affinity chromatography are similar, but the final concentration of OmpK obtained in the batch protocol is higher. The difference might be explained by the higher time of renaturation, increased incubation time with the different buffers and lower final concentration of urea in the buffer before dialysis might have favoured the OmpK yield. In the batch-based protocol the proteins were denaturated with 6M of urea overnight,

instead of 1 hour and 15 minutes with urea 8M as in the case of the column-based protocol. This might allow for a more complete denaturation of the proteins avoiding the losses obtained for column-based protocol in the 60 mM buffer. The batch protocol also allows a longer contact time of the buffers and resin beads.

The test used to quantify the protein present in the final solution was Bradford test. This test is an analytical procedure to measure the concentration of protein in a solution. The calibration curve prepared to estimate protein concentration is presented in Annex 4. The results of quantification of OmpK using the two protocols are shown in Table 21

Table 21 - maximum A_{595nm} and concentration values obtained for the column based and batch based protocols

Column protocol		Batch protocol	
A_{595nm}	Concentration (µg/ml)	A_{595nm}	Concentration (µg/ml)
0.026	54	0.052	92

Analysis of the renaturated protein was made by spectrophotometry in the UV in collaboration with a colleague (Mariana Jerónimo) to confirm the renaturation of the purified protein. With the purpose of evaluating the heat-induced unfolding of the renatured protein, the sample (and controls) were subjected to progressive heating steps and immediately its absorbance at 280 nm was evaluated. Samples of renatured protein collected after passing the dialysis suspension in an His-tag FF columns, were compared with two samples of denatured proteins (controls). One of the controls was the eluate obtained with buffer containing 50 mM of imidazole in the column-based protocol and the second control was the protein that remains on the column after adding the post-dialysis protein sample and that is eluted with 500 mM of imidazole. The samples were heated to different temperatures 4, 30, 40, 50, 60, 70, 80, 90 and 100ºC. The absorbances of the sample and controls were measured at 280 nm and plotted against the incubation temperature. Hydrophobic residues like tryptophan and tyrosine when exposed increase the A_{280nm}. If the protein is renaturated these residues due to their hydrophobic nature remain hidden. When the melting temperature of the protein is reached, the unfolding of proteins starts, becomes irreversible and hydrophobic residues are progressively exposed increasing the absorbance values.

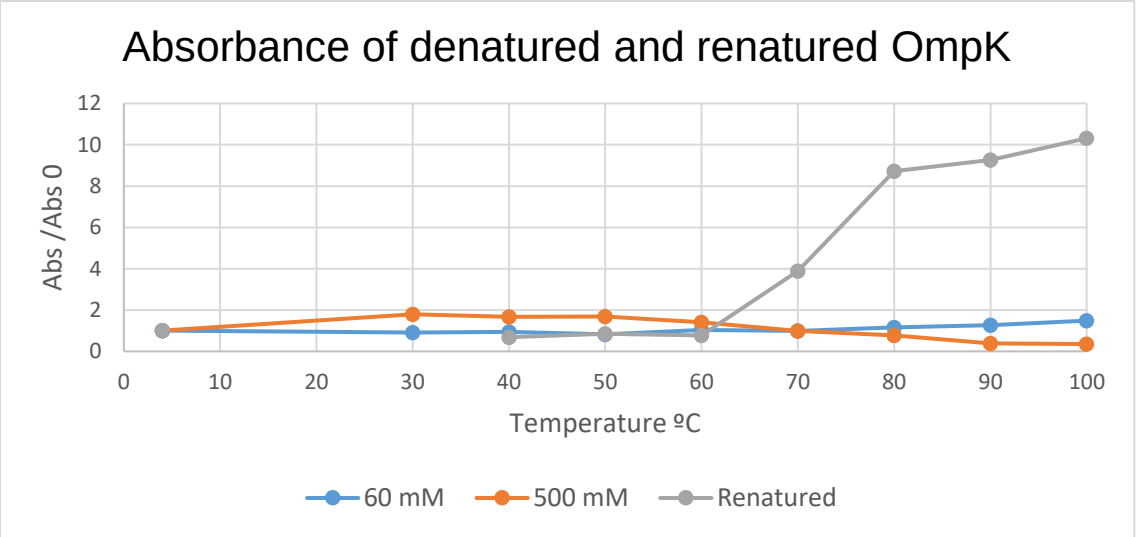


Figure 22 – Temperature profile of normalized protein solution absorbance at 280 nm.

The results presented in Figure 22 show that the normalized absorbance of the denaturated OmpK protein samples remained the same with a variation between 0 and 2. Conversely, the renatured OmpK sample showed an increase in his absorbance ratio with temperatures higher than 60ºC. The increase of the absorbance indicates that hydrophobic residues of the protein that were previously hidden inside the protein conformation became exposed. This result supports the conclusion that recombinant OmpK has been renaturated by the protocol described in 2.3.3.2.

4 – Conclusions and Future Perspectives

The results obtained in our bioinformatics analysis indicate that the *V. alginolyticus* OmpK is a protein highly conserved among all the *Vibrio* species known to infect fish. The bioinformatics predictions indicate that OmpK has the potential of being immunogenic, since at least two programs predicted the occurrence of at least two immunogenic segments composed of more than 10 amino acid residues. Therefore, we have cloned the gene in the pET23a+ and transformed *E. coli* BL21 with it, in order to overproduce the protein envisaging its future use in vaccine experiments. Obtained results indicate that recombinant *E. coli* BL21 harboured the constructed plasmid. However, several unexpected difficulties were faced when purifying the protein that was mainly aggregated in inclusion bodies. Once solubilized and denatured the protein instead of recovering the maximal amounts from a IMAC affinity chromatography, with the concentrations of imidazole of 200 - 300 mM, most of it eluted with impurities with low concentrations of imidazole. We therefore hypothesized that the His-tag was occluded in the protein structure, and thus not accessible to bind to the nickel ions in the resin. Two different methods were used in the purification of OmpK and although more tests are required to optimize the purification in batch higher concentrations of pure OmpK, were obtained. The concentration of OmpK obtained by the batch purification is 92 µg/ml in 16 ml from a 600 ml culture in SB medium at a final OD_{640nm} of 8.93. Pure OmpK (3 mg) is being prepared to send to a commercial company (SicGen) to produce antibodies that will allow future localization of possible immunogenic segments. To produce antibodies for OmpK, the samples need to be absent of impurities inclusive endotoxins. However, due to absorption of OmpK to the membrane in the kit used to eliminate endotoxins, this last requirement is turning out to be difficult to achieve.

To allow an easier recovery of the protein of interest, a leaky strain of *E. coli* BL21 with the *ompK* gene is being constructed by cloning an antisense of the *lpp* gene encoding for the Braum's lipoprotein in a controlled expression plasmid. The recombinant plasmid will then be inserted into E. coli, leading the expression of the lpp antisense to the weakening of the bacterial cell wall. In order to be efficiently secreted by a leaky E. coli prepared as described above, recombinant OmpK should be expressed in conditions that do not promote its aggregation and/or the formation of inclusion bodies, since translocation of the protein to the periplasmatic space involved a signal peptide included in the recombinant OmpK, and the protein has to be in a soluble form. To be used in large scale, recombinant OmpK should be produced in bioreactors. Batch and fed-batch cultivation strategies will be implemented in 1-2L fermenters. If we succeed in the construction of a leaky *E. coli* able to secrete OmpK, integration of protein expression and excretion with *in-situ* product recovery is recommended. This can be achieved by coupling batch fermentation with chromatographic protein capture.

The immunogenicity of the OmpK protein produced will be assessed in soles from aquaculture to test its efficacy as a protective vaccine. The protein can also be used in ELISA methods to detect early infections caused by *Vibrio sp*. in aquaculture fish. In this case, the

immunogenic protein can be immobilized in a solid support to bind to the antibodies from sera of infected fish.

5 – References

1. Cole, D. W. et al. Aquaculture: Environmental, toxicological, and health issues. *Int. Jl Hyg. Env. Heal.* **212,** 369–377 (2009).

2. Durborow, R. M. Health and safety concerns in fisheries and aquaculture. *Occup. Med.* **14,** 373–406 (1998).

3. Aubin, J., Papatryphon, E., Van der Werf, H. M. G., Petit, J. & Morvan, Y. M. Characterisation of the environmental impact of a turbot (*Scophthalmus maximus*) re-circulating production system using Life Cycle Assessment. *Aquaculture* **261,** 1259–1268 (2006).

4. Muir, J. F., Brugere, C., Young, J. A. & Stewart, J. A. The solution to pollution? The value and limitations of environmental economics in guiding aquaculture development. *Aquac. Econ. Manag.* **3,** 43–57 (1999).

5. Thompson, F. L. et al. Phylogeny and molecular identification of vibrios on the basis of multilocus sequence analysis. *Appl. Environ. Microbiol.* **71,** 5107–15 (2005).

6. Dinis, M. T., Ribeiro, L., Soares, F. & Sarasquete, C. A review on the cultivation potential of *Solea senegalensis* in Spain and in Portugal. *Aquaculture* **176,** 27–38 (1999).

7. Arijo, S. et al. Effectiveness of a divalent vaccine for sole, *Solea senegalensis* (Kaup), against *Vibrio harveyi* and *Photobacterium damselae* subsp. *piscicida*. *J. Fish Dis.* **28,** 33–38 (2005).

8. Zorrilla, I., Balebona, M. C., Mori.igo, M. A., Sarasquete, C. & Borrego, J. J. Isolation and characterization of the causative agent of pasteurellosis, *Photobacterium damsela* ssp. piscicida, from sole, *Solea senegalensis* (Kaup). *J. Fish Dis.* **22,** 167–172 (1999).

9. Press, C. M. & Lillehaug, A. Vaccination in European salmonid aquaculture: A review of practices and prospects. *Br. Vet. J.* **151,** 45–69 (1995).

10. Horne, M. T. & Robertson, D. A. Economics of vaccination against enteric redmouth disease of salmonids. *Aquac. Res.* **18,** 131–137 (1987).

11. Nilsem, H., Lillehaug, A., Taksdal, T. & Nordmo, R. Toxicity of intraperitoneally injected formalin in Atlantic salmon, Salmo salar L. *J. Fish Dis.* **15,** 323–329 (1992).

12. Hastein, T. & Refsti, T. Vaccination of rainbow trout against vibriosis by injection, dip and bath. *Bull. Eur. Ass. Fish pathol* **645,** (1986).

13. Johnson, K. A. & Amend, D. F. Comparison of efficacy of several delivery methods using *Yersinia ruckeri* bacterin on rainbow trout, *Salmo gairdneri* Richardson. *J. Fish Dis.* **6,** 331–336 (1983).

14. Lillehaug, A. A survey on different procedures used for vaccinating salmonids against vibriosis in Norwegian fish-farming. *Aquaculture* **83,** 217–226 (1989).

15. Lillehaug, A., Lunder, T. & Poppe, T. T. Field testing of adjuvanted furunculosis vaccines in Atlantic salmon, *Salmo salar* L. *J. Fish Dis.* **15,** 485–496 (1992).

16. Hart, S., Wrathmell, A. B., Harris, J. E. & Grayson, T. H. Gut immunology in fish: A review. *Dev. Comp. Immunol.* **12,** 453–480 (1988).

17. Ningqiu, L. et al. An outer membrane protein, OmpK, is an effective vaccine candidate for *Vibrio harveyi* in Orange-spotted grouper (*Epinephelus coioides*). *Fish Shellfish Immunol.*

25, 829–833 (2008).

18. Inoue, T., Matsuzaki, S. & Tanaka, S. A 26-kDa outer membrane protein, OmpK, common to *Vibrio* species is the receptor for a broad-host-range vibriophage, KVP40. *FEMS Microbiol. Lett.* **125,** 101-5 (1995).

19. Inoue, T., Matsuzaki, S. & Tanaka, S. Cloning and sequence analysis of *Vibrio parahaemolyticus ompK* gene encoding a 26-kDa outer membrane protein, OmpK, that serves as receptor for a broad-host-range vibriophage, KVP40. *FEMS Microbiol. Lett.* **134,** (1995).

20. Li, Y.-D. *et al.* Cloning, Expression, and Genus-Specificity Analysis of 28-kDa OmpK from Vibrio alginolyticus. *J. Food Sci.* **75,** M198–M203 (2010).

21. Qian, R.-H. *et al.* A conserved outer membrane protein as an effective vaccine candidate from Vibrio alginolyticus. *Aquaculture* **278,** 5–9 (2008).

22. Zheng, M. Z., Richard, J. L. & Binder, J. A Review of Rapid Methods for the Analysis of Mycotoxins. Mycopathologia 161, 261–273 (2006).

23. Demain, A. L. & Vaishnav, P. Production of recombinant proteins by microbes and higher organisms. *Biotechnol. Adv.* **27,** 297–306 (2009).

24. Ferrer-Miralles, N. & Villaverde, A. Bacterial cell factories for recombinant protein production; expanding the catalogue. *Microb. Cell Fact.* **12,** 113 (2013).

25. Sørensen, H. Towards universal systems for recombinant gene expression. *Microb. Cell Fact.* **9,** 27 (2010).

26. Paulsson, J. & Ehrenberg, M. Noise in a minimal regulatory network: plasmid copy number control. *Q. Rev. Biophys.* **34,** 1–59 (2001).

27. Summers, D. Timing, self-control and a sense of direction are the secrets of multicopy plasmid stability. *Mol. Microbiol.* **29,** 1137–1145 (1998).

28. Corchero, J. L. & Villaverde, A. Plasmid maintenance in *Escherichia coli* recombinant cultures is dramatically, steadily, and specifically influenced by features of the encoded proteins. *Biotechnol. Bioeng.* **58,** 625–632 (1998).

29. Lewin, B. *Genes VI. 1997*. **29,** (Oxford University Press, 1994).

30. Baneyx, F. Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotechnol.* **10,** 411–421 (1999).

31. Swartz, J. R. Advances in *Escherichia coli* production of therapeutic proteins. *Curr. Opin. Biotechnol.* **12,** 195–201 (2001).

32. Palomares, L. A., Estrada-Mondaca, S. & Ramírez, O. T. Production of recombinant proteins: challenges and solutions. *Methods Mol. Biol.* **267,** 15–52 (2004).

33. Sayadi, S., Nasri, M., Berry, F., Barbotin, J. N. & Thomas, D. Effect of temperature on the stability of plasmid pTG201 and productivity of xylE gene product in recombinant *Escherichia coli*: development of a two-stage chemostat with free and immobilized cells. *J. Gen. Microbiol.* **133,** 1901–1908 (1987).

34. Prazeres, D. M. F., Ferreira, G. N. M., Monteiro, G. A., Cooney, C. L. & Cabral, J. M. S. Large-scale production of pharmaceutical-grade plasmid DNA for gene therapy: Problems and bottlenecks. *Trends Biotechnol.* **17,** 169–174 (1999).

35. Ailor, E. & Betenbaugh, M. J. Modifying secretion and post-translational processing in

insect cells. *Curr. Opin. Biotechnol.* **10,** 142–145 (1999).

36. Kopito, R. R. Aggresomes, inclusion bodies and protein aggregation. *Trends Cell Biol.* **10,** 524–530 (2000).

37. Schlieker, C., Bukau, B. & Mogk, A. Prevention and reversion of protein aggregation by molecular chaperones in the *E. coli* cytosol: implications for their applicability in biotechnology. *J. Biotechnol.* **96,** 13–21 (2002).

38. Carrió, M. M. & Villaverde, A. Construction and deconstruction of bacterial inclusion bodies. in *J. Biotechnol.* **96,** 3–12 (2002).

39. Clark, E. D. Protein refolding for industrial processes. *Curr. Opin. Biotechnol.* **12,** 202–207 (2001).

40. Datar, R. V, Cartwright, T. & Rosen, C. G. Process economics of animal cell and bacterial fermentations: a case study analysis of tissue plasminogen activator. *Biotechnology.* **11,** 349–357 (1993).

41. Benz, I. & Schmidt, M. A. Never say never again: Protein glycosylation in pathogenic bacteria. *Molecular Microbiology* **45,** 267–276 (2002).

42. Furmanek, R., Furmanek, R. & Hofsteenge, Review Protein C-mannosylation: Facts and questions?. in *J. Quarterly.* (2000).

43. Lisowska, E. The role of glycosylation in protein antigenic properties. *Cell. Mol. Life Sci*. **59,** 445–455 (2002).

44. Gu, X., Harmon, B. J. & Wang, D. I. C. in *Advances in Bioprocess Engineering.* **2**, 1–24 (Springer Netherlands, 1998).

45. Lopez, M. et al. O-Glycosylation potential of lepidopteran insect cell lines. *Biochim. Biophys. Acta - Gen. Subj.* **1427,** 49–61 (1999).

46. Miyamoto, Y., Nakamuma, K. & Takizawa, K. Pathogenic halophiles. Proposals of a new genus ' *Oceanomonas* ' and of the amended species names. *Jpn. J. Microbiol.* **5,** 477–486 (1961).

47. Technologies, N. ND-1000 Spectrophotometer V3.5 User's Manual. (2007).

48. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual*. (USA:Cold Spring Harbor Laboratory Press, 1989, 1989).

49. GE Healthcare. *Instructions 11-0008-88 AH*.

50. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72,** 248–54 (1976).

51. Wiik, R. et al. Classification of fish-pathogenic vibrios based on comparative 16S rRNA analysis. *Int. J. Syst. Bacteriol.* **45,** 421–8 (1995).

52. Mustapha, S., Mustapha, E. M. & Nozha, C. *Vibrio Alginolyticus* : An Emerging Pathogen of Foodborne Diseases. International Journal of Science and Technology. **2,** 302–309 (2013).

53. Defoirdt, T. et al. Quorum sensing-disrupting brominated furanones protect the gnotobiotic brine shrimp *Artemia franciscana* from pathogenic *Vibrio harveyi*, *Vibrio campbellii*, and *Vibrio parahaemolyticus* isolates. *Appl. Environ. Microbiol.* **72,** 6419–6423 (2006).

54. Egidius, E., Wiik, R., Andersen, K., Hoff, K. A. & Hjeltnes, B. *Vibrio salmonicida* sp . nov ., a New Fish Pathogen. *Int. J. o Syst. Bacterioloy* **36,** 518–520 (1986).

55. Frans, I. et al*. Vibrio anguillarum* as a fish pathogen: Virulence factors, diagnosis and prevention. *J. Fish Dis.* **34,** 643–661 (2011).., Heang, H., Criel, G., Sorgeloos, P. & Verstraete, W. Selected bacterial strains protect Artemia spp. from the Pathogenic effects of Vibrio proteolyticus CW8T2. *Appl. Environ. Microbiol.* **66,** 1139–1146 (2000).

56. Schiewe, M. H., Trust, T. J. & Crosa, J. H. *Vibrio ordalii* sp. nov.: A causative agent of vibriosis in fish. *Curr. Microbiol.* **6,** 343–348 (1981).

57. Verschuere, L., Heang, H., Criel, G., Sorgeloos, P. & Verstraete, W. Selected bacterial strains protect *Artemia* spp. from the pathogenic effects of *Vibrio proteolyticus* CW8T2. *Appl. Environ. Microbiol.* **66,** 1139–1146 (2000).

58. Haldar, S. et al. Identification of *Vibrio campbellii* isolated from diseased farm-shrimps from south India and establishment of its pathogenic potential in an *Artemia* model. *Microbiology* **157,** 179–188 (2011).

59. Koebnik, R., Locher, K. P. & Van Gelder, P. Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol. Microbiol.* **37,** 239–53 (2000).

60. Ye, J. & van den Berg, B. Crystal structure of the bacterial nucleoside transporter Tsx. *EMBO J.* **23,** 3187–95 (2004).

61. Arnold, T. & Linke, D. Phase separation in the isolation and purification of membrane proteins. *Biotechniques* **43,** 427–30, 432, 434 (2007).

62. pET System Manual.

63. San-Miguel, T., Pérez-Bermúdez, P. & Gavidia, I. Production of soluble eukaryotic recombinant proteins in *E. coli* is favoured in early log-phase cultures induced at low temperature. *Springerplus* **2,** 89 (2013).

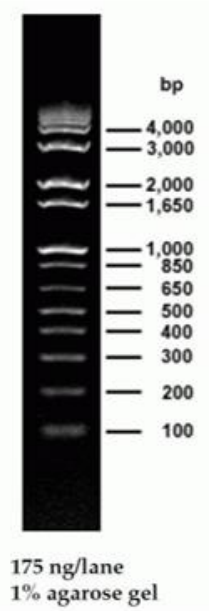
64. Lin, J.-L., Ruaan, R.-C. & Hsieh, H.-J. Refolding of partially and fully denatured lysozymes. *Biotechnol. Lett.* **29,** 723–729 (2007).

65. Alliegro, M. C. Effects of dithiothreitol on protein activity unrelated to thiol–disulfide exchange: For consideration in the analysis of protein function with Cleland's reagent. *Anal. Biochem.* **282,** 102–106 (2000).

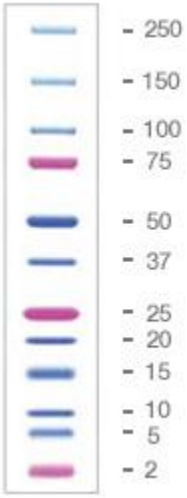
66. Mollania, N. et al. An efficient *in vitro* refolding of recombinant bacterial laccase in *Escherichia coli*. *Enzyme Microb. Technol.* **52,** 325–330 (2013).

6 - Annexes

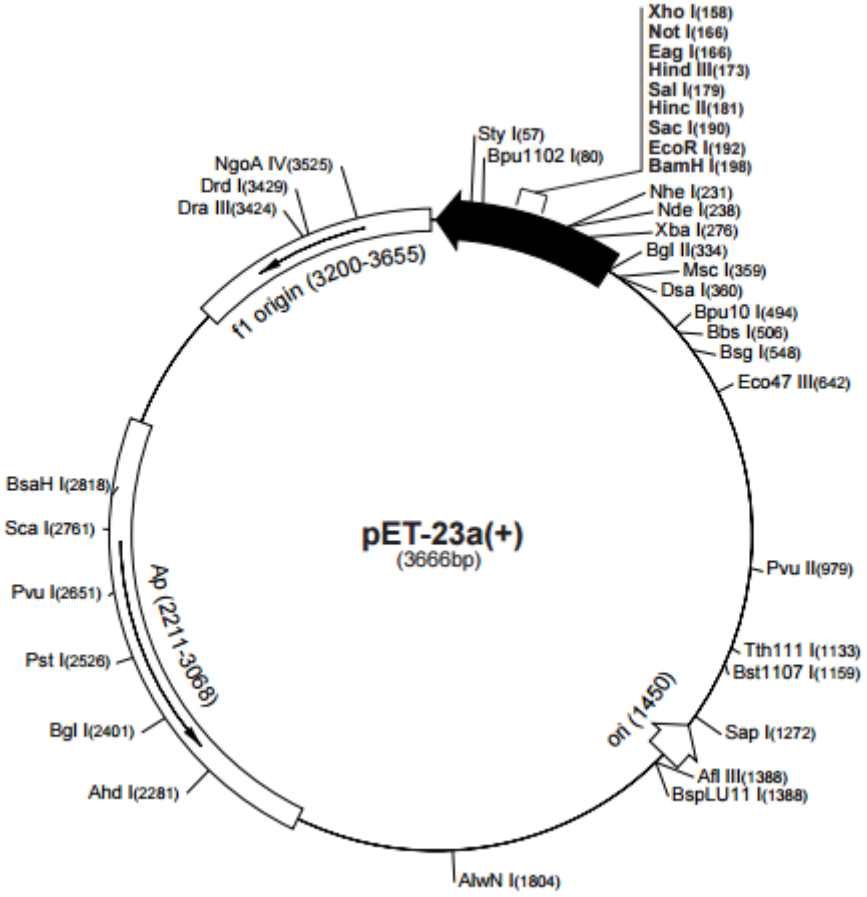
Annex 1 – Ladder 1 Kb Plus Compatibility used for the confirmation of the relative sizes of the nucleic acid fragments in the agarose gels (adapted from Thermo Fisher Scientific)



Annex 2 – Precision plus protein Dual xtra standards used for the confirmation of the relative sizes of the proteins in the SDS-PAGE gels (adapted from Bio Rad)



Annex 3 – restriction sites of the plasmid pET23a+ (adapted from Novagen, url http://www.chemistry.mcmaster.ca/berti/safety/Manuals/pET23a.pdf), The ampiclin resisntace is refered as Ap in the figure bellow



Annex 4 – Construction of the calibration curve for the Bradford test

Table 1 – A_{595nm} concentration for the buffers with different concentration of BSA

BSA concentration (µg/ml)	A_{595nm}
100	0.066
200	0.118
350	0.238
500	0.342

With the ODs measured it is possible to build a calibration curve to allow the calculation of the concentration of OmpK in the solution (figure 23 and equation 1).

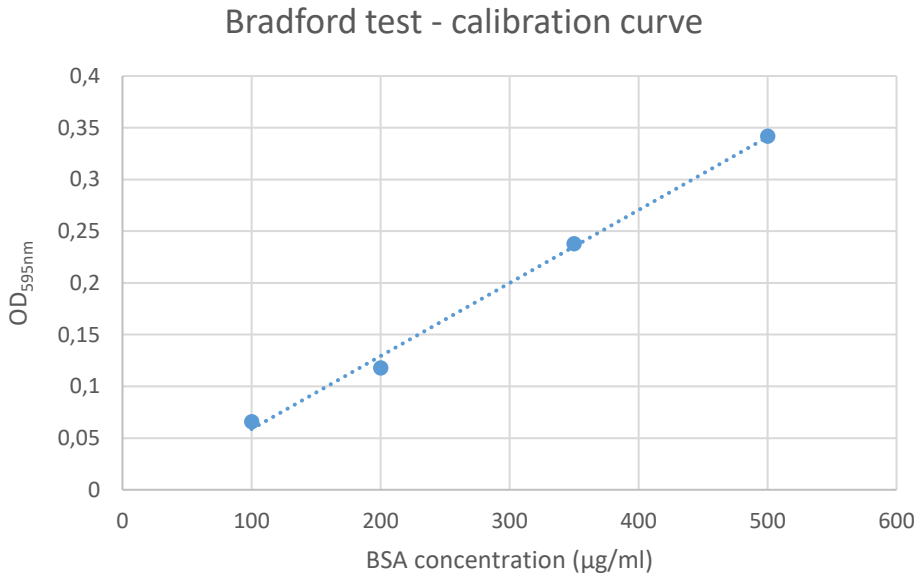


Figure 1 – Graphic plot of Absorbance of BSA standard solutions at 280 nm and trend line obtained from the points

Equation 1 – equation of the trend line correlating the points in the graphic

$$OD_{595} = 0.0007 * protein\ concentration\ (\frac{\mu g}{ml}) - 0.0119$$

$$R^2 = 0.9959$$

The R value is high meaning the trend line fits with the data reasonably well in the range of 100 – 500 mg/L. The equation obtained was used to calculate the concentration of the protein samples in the final OmpK solution.