Development of an oral pDNA vaccine against *Vibrio* species

Carolina Alfaiate Fonseca

Thesis to obtain the Master of Science Degree in

Biotechnology

Supervisors

Prof. Gabriel António Amaro Monteiro

Prof. Marília Clemente Velez Mateus

Examination Committee

Chairperson: Prof. Arsénio do Carmo Sales Mendes Fialho

Supervisor: Prof. Marília Clemente Velez Mateus

Member of the Committee: Prof. Sílvia Andreia Bento da Silva Sousa Barbosa

September 2020

Preface

The work presented in this thesis was performed at BioEngineering Research Group (BERG) of the institute for Bioengineering and Biosciences (iBB) at Instituto Superior Técnico (Lisbon, Portugal), from September 2019 to September 2020, under the supervision of Professor Gabriel Monteiro and Professor Marília Mateus.

I declare that this document is an original work of my own authorship and that it fulfils all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Agradecimentos

Para a concretização deste projeto, contei com a disponibilidade e o incentivo de muitos, que de várias formas contribuíram para chegar até aqui. Grata por todo o apoio e inspiração que me deram, grata por ter feito um percurso de crescimento e de conhecimento, para a realização do projeto que agora entrego.

Aos meus orientadores, Professor Gabriel Monteiro e Professora Marília Mateus, gostaria de agradecer, em especial, por todo o apoio prestado, pela disponibilidade, pelo saber transmitido e por toda a orientação dada durante o projeto.

À Ana Rita Santos, Sara Sousa Rosa e Cristiana Ulpiano, que foram grandes pilares nestes meses e que se mostraram sempre disponíveis. Queria agradecer pela paciência e pela ajuda que sempre me deram com alguns procedimentos laboratoriais.

Aos meus amigos, que me acompanham desde o primeiro dia, esta caminhada não tinha sido a mesma sem vocês. À Margarida P., João, Luís, Érica e Margarida B, por todos os cafés, almoços e jantares, e por todas as conversas ao longo destes dois anos.

Ao Ricardo, por todo o incentivo, carinho e companheirismo ao longo deste percurso e por ser um dos meus maiores suportes desde o início, sempre fundamental tanto a nível laboratorial como emocional.

Um enorme obrigada à minha família, pelo apoio incondicional durante este percurso académico e pelo espírito vencedor que sempre me incutiram. Aos meus pais e ao meu irmão, um obrigada especial por serem os meus pilares, fonte de apoio e inspiração. Às minhas tias e grandes amigas, Helena Alfaiate e Sandra Alfaiate, por estarem sempre presentes e pelas palavras de incentivo nas horas certas.

Abstract

Nowadays, vaccines are recognized worldwide as one of the most important tools to fight human and animal infectious diseases. Fish consumption has been growing over the years, and so is aquaculture production to meet the current demand. Increased fish stock densities are responsible for high levels of stress, enhancing the vulnerability to infections like vibriosis. Thus, the main goal consists in creating and developing an oral plasmid DNA vaccine against *Vibrio* species.

Pure plasmids were produced in *Escherichia coli* at the following yields and respective final concentrations: 67.4µg of pVAX eGFP, 7.31µg of pVAX OmpK, and 23.52µg of pVAX OmpK-frag per gram of cell dry weight and 526.5ng/µL, 56ng/µL, and 175ng/µL.

Chitosans and tripolyphosphate were used to synthesize different nanoparticles and later complexation with pDNA. Produced nanoparticles were characterized, regarding average hydrodynamic diameter, zeta potential, and polydispersity index. A long-chain chitosan formulation (LC2, 5mL chitosan (3mg/mL), 2mL TPP (0,75mg/mL)) showed fairly small diameters, low Pdi value, and significant positive zeta potential and was chosen for pDNA complexation. Short-chain chitosan resulted in greater diameters and Pdi values, meaning that some aggregation may have occurred.

Bioinformatics was used to predict new immunogenic epitopes among outer membrane proteins of *Vibrio* bacteria. Analyses suggest that OmpK should be targeted since, according to BLAST analysis, this protein is surface-exposed in the majority of *Vibrio* strains. Within the OmpK sequence, the peptide between positions 200 and 228 is the most hydrophilic and most flexible, having just a small region of five amino acids more prone to aggregation.

Keywords: Aquaculture, DNA Vaccination, Oral Vaccination, Nanoparticles, Chitosan, Bioinformatic Tools

Resumo

As vacinas são uma das principais ferramentas contra doenças infeciosas em humanos e animais. O acréscimo na produção em aquacultura originou um aumento de peixes em reservatórios, causando-lhes elevados níveis de stress e tornando-os mais vulneráveis a infeções, como a Vibriose. O objetivo deste projeto consiste em criar e desenvolver uma vacina de administração oral de ADN plasmídico contra espécies *Vibrio*.

Plasmídeos puros foram produzidos em *Escherichia coli* com os seguintes rendimentos e respetivas concentrações finais: 67,4µg de pVAX eGFP, 7,31µg de pVAX OmpK e 23,52µg de fragmento pVAX OmpK por grama de peso seco celular e 526,5ng/µL, 56ng/µL e 175ng/µL.

Chitosano e tripolifosfato foram utilizados para sintetizar diferentes nanopartículas, para efetuar uma posterior complexação com ADN. Estas foram caracterizadas pelo seu diâmetro hidrodinâmico médio, potencial zeta e índice de polidispersidade. Uma formulação de chitosano de cadeia longa (LC2, 5mL chitosano (3mg/mL), 2mL TPP (0,75mg/mL)) apresentou diâmetros pequenos, baixos valores do índice de polidispersidade e um potencial zeta positivo, selecionando-se para a complexação com o ADN. O chitosano de cadeia curta apresentou diâmetros e valores de índice de polidispersidade maiores, significando que alguma agregação ocorreu.

A bioinformática previu novos epítopos imunogénicos entre as proteínas da membrana externa das bactérias *Vibrio*. As análises BLAST sugerem que a OmpK deve ser selecionada, prevendo que esta se encontre nas superfícies das membranas de muitas estirpes *Vibrio*. Na sequência OmpK, o péptido entre as posições 200 e 228 é o mais hidrofílico e flexível, contendo apenas uma região de cinco aminoácidos mais propensos à agregação.

Palavras-chave: Aquacultura, Vacinação de ADN, Vacinação Oral, Nanopartículas, Chitosano, Ferramentas Bioinformáticas

Table of Contents

List of Figures

List of Tables

Abbreviations

APCs - Antigen presenting cells
a^3v - Amino Acid Aggregation-propensity Value
a^4v - Amino Acid Aggregation-propensity Value window Average
a^4vAHS - Amino Acid Aggregation-propensity Value window Average (a^4v) Average in the Hotspot
***bgh*-PolyA -** Bovine growth hormones polyadenylation
BLAST- Basic Local Alignment Search Tool
bp - Base pair
Da - Dalton
DLS - Dynamic light scattering
DNA - Deoxyribonucleic acid
eGFP - Enhanced green fluorescent protein
EMA - European medicine agency
FDA - Food and drug administration (U.S.A)
GALT - Gut-associated lymphoid tissue
GFP - Green Fluorescence Protein
GI - Gastrointestinal
IEDB - Immune Epitope Database
Ig - Immunoglobulins
ISG - Invariant Surface Glycoprotein
MHC - Major histocompatibility complex
MW - Molecular weight
NCBI - National Center for Biotechnology Information
OMPs - Outer membrane proteins
OmpK - Outer membrane protein K
OmpW - Outer membrane protein W
OmpV - Outer membrane protein V
OmpU - Outer membrane protein U
PdI - Polydispersion index
pDNA - Plasmid DNA
RNA - Ribonucleic acid
RFU - Relative Fluorescence Units
TPP - Sodium tripolyphosphate

1. Introduction

1.1. *Vibrio* Infections

Vibrios are gram-negative, usually motile rods, mesophilic and chemoorganotrophic organisms. They also have a facultatively fermentative metabolism and belong to the *Gammaproteobacteria*, according to 16S rRNA gene sequence analysis.[1]

These species are found in aquatic environments throughout the world, more commonly in warmer waters (17ºC to 20ºC). Depending on the species, they can also tolerate a wide range of salinities.[2] They belong to the *Vibrionaceae* family and most of the species are widely distributed in the environment. Those species have also been extensively used in physiological, biochemical, molecular biology, and pathogenicity studies.[3]

Therefore, Vibriosis, caused by *Vibrio* species, is one of the most common bacterial diseases among marine fish and shellfish, affecting all fish stages of growth, leading to mortality for up to 50%. *V. parahaemolyticus, V. alginolyticus, V. harveyi, V. owensii and V. campbelli* are the most frequent species infecting farmed aquatic animals. [4] V. *anguillarum*, *V. ordalii*, *V. salmonicida*, *V. vulnificus, V. ponticus,* and *Photobacterium damselae* subsp. *damselae* are also relevant species that can harm cultured fish. [5]

Vibrio can be divided into two categories: nonepidemic strains and epidemic strains (for example, *V. cholerae,* carrying a suite of specific virulence genes, causes the disease cholera). The latter is associated with several diseases such as gastroenteritis, wound infections, and septicemia in susceptible hosts. Some species are responsible for the serious enteric disease after ingestion of raw or undercooked seafood carrying these bacteria and they can even be fatal.[6]

Nonetheless, most of those species are beneficial for many aquatic animals, having symbiosis relationships with them. Furthermore, these species are also able to contribute to the cycling of carbon and other nutrients in the aquatic environment. The ones that are associated with diseases in animals and humans have virulence-associated genes. These genes are frequently not present or available in the environmental *Vibrio* species. [5]

Some studies were done to try to understand the virulence characteristics of some *Vibrio* species. *V. harveyi* was isolated from diseased Japanese flounder and compared with the environmental *V. harveyi* isolates.[7] The same thing was done with *V. alginolyticus*, where bacteria from diseased fish and shrimp and from a healthy shoal were isolated. For this latter study, they concluded that the number of extracellular enzymes is much higher in diseased fish and shrimp when compared with healthy organisms. Therefore, several extracellular products, such as chitinases, hemolysins, alkaline proteases, cysteine proteases, alkaline metal-chelator-sensitive proteases, serine proteases, and metalloproteases have been proposed as virulence factors for fish and other marine organisms.[8]

It´s well known that the primary mode of transmission of *Vibrio* species in fish is through the water. Bacteria can penetrate through the fish skin, being the principal route of infection. Gills and gastrointestinal tract can also work as the portal of entry, even though physical and chemical barriers

(such as mucus and epidermis) will act as the first barrier against invasion by the pathogen. So, if the mucosal layer is damaged, the possibility of infection is much higher. [4,9]

Isolating fish through quarantine can avoid some serious problems, since healthy and sick fish are physically separated, avoiding the contagion. Other practices that can also avoid the spread of the disease in aquaculture are optimizing water quality, providing high-quality feed, breeding of disease-resistant broodstocks and also very importantly, the establishment of a vaccination program. [4]

The vaccine development for *Vibrio* strains has been challenging and slow mainly because the specific-specific vaccines only protect against a specific bacterial strain. Essentially, the antigenic diversity is very high, and their serotypes have made the vaccines unable to protect against several *Vibrio* infections. Thus, the elaboration and improvement of these vaccines is fundamental because they cannot only protect the fish against *Vibrio* infections but they can also avoid the use of antibiotics in the aquaculture industry.[10]

The use of antibiotics is a good option since *Vibrio* species are known to be extremely vulnerable to the majority of clinically used antibiotics. But the extensive use of antibiotics in aquaculture is being discouraged due to the appearance of antimicrobial resistant pathogenic bacteria in shellfish.[11] Accumulation in the flesh of animals and contamination of the aquatic environment are other undesirable side effects of chemicals and antibiotics. [12]

The multidrug-resistance in pathogens may have been acquired through horizontal gene transfer, using mobile genetic elements such as plasmids, transposons, integrons, and integrative conjugative elements. [13]

Regarding humans, the risk is greater for people who are immunocompromised, have cirrhosis or have conditions predisposing them to increased saturation of transferrin with iron. People in these risk groups should avoid eating raw oysters, particularly during the summer and early fall, when water temperatures may exceed 20^{o}C and should try to minimize exposure of wounds to warmer estuarine or marine waters.[2]

So, due to all these factors mentioned above, vaccination is a much safer option than others that can be used to treat or prevent vibriosis.

1.2. Principles of Innate and Adaptative Immunity in Fish

Since fishes are a representative population of lower vertebrates, they are important models in comparative immunology as they can provide a link to the early evolution of vertebrates.[14]

Fish immune system can be divided into two components: innate and adaptive **(Figure 1)**. The innate defense mechanism in fish is activated quickly upon infection. It includes surface barriers (such as mucus, gills, gastrointestinal tract), growth inhibitors (transferrin, interferon), enzyme inhibitors, lysins (complement, antimicrobial, peptides, lysozyme), precipitins and agglutinins (pentraxins, lectins), nonspecific cellular factors like phagocytes (macrophages, neutrophils), phagocyte activating molecules (opsonins, cytokines), natural cytotoxic cells, eosinophils, basophils, mast cells and inflammation.[15,16]

The strength of innate defense against pathogens is impressive, since the recognition machinery is very limited. The very efficient immune defense of invertebrates, which exclusively rely on innate parameters for coping with a large variety of pathogens in diverse environmental conditions, is also of primary importance in combating infections in fish. [17]

In opposition, adaptative immune response takes several days to become effective, although it provides specific memory. Specific memory is essential to complete the elimination of the pathogen.[16] The lymphocytes mediate three aspects of adaptative immune system: humoral immunity, cell-mediated immunity and immunological memory. Humoral immunity is responsible for the production of immunoglobulins (Ig) by B-cells. Until now, three types of Igs are known in fish.[15, 18]

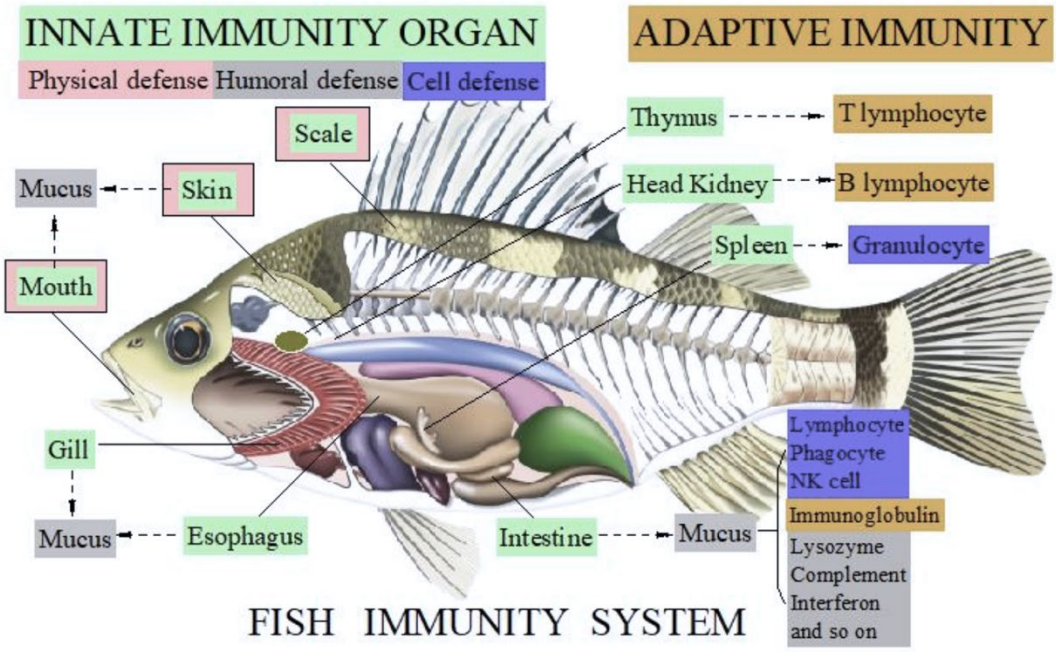

Figure 1 - *Innate and adaptive immunity in fish. Obtained from: Fan G, Chen J, Jin T, et al. Report on Marine Life Genomics.; 2018. doi:10.20944/preprints201812.0156.v1*

The immunological memory is a significant aspect in specific immune response, which comprises the adaptive change in lymphoid cells. When fish are exposed to a pathogen more than once, the immune system will recognize it immediately and will destroy it before it can do serious damages. This is the basis for successful vaccination approach. Triggering the innate and adaptive immune system is the main goal when vaccines are being developed. [15,16]

To enhance and prolong the immune response, it's possible to use adjuvants and delivery systems. Due to their unique properties (**Section 6**), nanoparticles-based vaccines can induce the up-regulation of several inflammatory, innate and specific immune responsive genes. [15,19]

1.3. Vaccination in Aquaculture

Vaccines are recognized worldwide as one of the most important tools for fighting infectious diseases.[20] Although vaccines have already long proven their efficacy in shoal protection, in the aquaculture industry, they are in an early phase of development. For instance, the shoal immunity threshold in Atlantic salmon farms has been estimated to be 52% for infectious salmon anemia and 66% for pancreas disease.[21] The commercially available vaccines for fishes are made of inactivated bacteria and are applied by emersion or injection with an oil adjuvant.[12]

Aquaculture plays an important role in economic development and social stability worldwide since it improves the nutritional standards and reduces poverty in some poor countries.[4] It has been the fastest-growing food-producing sector for years, and the yield of aquaculture has overgrown the yield of capture fisheries. This causes an increase in stocking densities that give rise to high-stress levels, making fish more vulnerable to infections.[22] The expected demand for fish meal in fed-aquaculture diets is present in **Figure 2** [23]. To avoid serious economic losses, it's important to prevent disease outbreaks. This can be achieved by developing or refining vaccines for aquaculture species.[22]

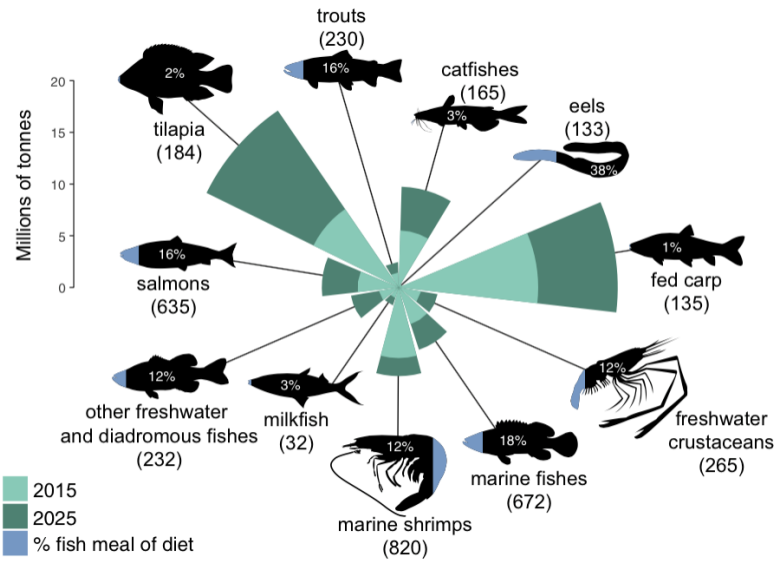

Figure 2 - Expected demand for fish meal in fed-aquaculture diets. *The percentage inside each fish represents the estimated fish meal inclusion on 2015 and the values in brackets are the estimated volume of fish meal included in the diets also in 2015 (thousands of tons). Obtained from: Hua K, Cobcroft JM, Cole A, et al. Review the future of aquatic protein: Implications for protein sources in aquaculture diets. One Earth. 2019;1(3):316-329. doi:10.1016/j.oneear.2019.10.018*

Focusing on the nutritional value, fish is one of the most nutrient-dense animal-source food, containing not only animal protein but also unique long-chain polyunsaturated fatty acids (LC-PUFAS) and highly bioavailable essential micronutrients (vitamins D and B and minerals such as calcium, phosphorus, iodine, zinc, iron, and selenium). Because of these benefits, it's fundamental to include fish in the meals, since all of those compounds are often not present elsewhere in the diet, having beneficial effects for adult and child cognitive development. [24] Additionally, one good advantage of aquaculture is that farmed aquatic food products are capable of reducing the impact on the wild fish stock since overfishing may end up in ecological collapse. [25]

From a health perspective, the presence of by-products and toxic chemicals that are released into the environment by industries can harm humans who consume contaminated fish. Some examples are heavy metals, that can accumulate in fatty tissues due to their lipophilic properties, methylmercury, and polychlorinated biphenyls (PCB). The latter has an exceptionally long half-life, making aquatic organisms such as fish, vulnerable to PCB exposure. In farmed fish, the methylmercury is rarely found, but PCBs can be found in higher concentrations due to polluted feed. [24,26, 27] Nonetheless, those toxic components can be easily controlled in farmed fish if raised under appropriate conditions, and they can be as beneficial as the wild fish. [28]

To maintain a healthy fish population, it's very important to have good environmental conditions. Nevertheless, for species reared in an open aquatic environment, it's impossible to avoid exposure to pathogens. [29]

Pathogens can spread quickly within a population of cultured fish not only because of the high density of animals used but also because of the effectiveness of pathogen transportation in water. This happens in all forms of intensive culture and is very common in aquaculture. For example, carps can be very resistant and are much more robust than, for instance, Atlantic salmon. The Atlantic salmon is adapted to grow and thrive in clean, running fresh water. [29]

There are vaccines available for most aquaculture fish species. Most of them are targeting bacterial pathogens, and only a few are targeting viruses.[22]

There are four vaccine concepts: attenuated, inactivated, purified subunits of the pathogens (proteins or glycoproteins), and DNA vaccines. [19]

The live vaccines, although they provide acceptable or even good protection under experimental conditions, the safety considerations stopped further work. This happens because some vaccines revealed residual virulence to groups of vaccinated fish at an unacceptable level. So, to overcome this problem, most vaccines commercially available for aquaculture are inactivated, non-replicating vaccines. Nevertheless, these vaccines have low protective immunity when compared with the protection achieved by the live vaccines.[19] Commercial vaccines can be administered orally (by mixing with the feed), by immersion (dip or bath), or by injection (intraperitoneal or intramuscular route) **(Figure 3)**. The route of administration depends on the age and the size of the fish.[19]

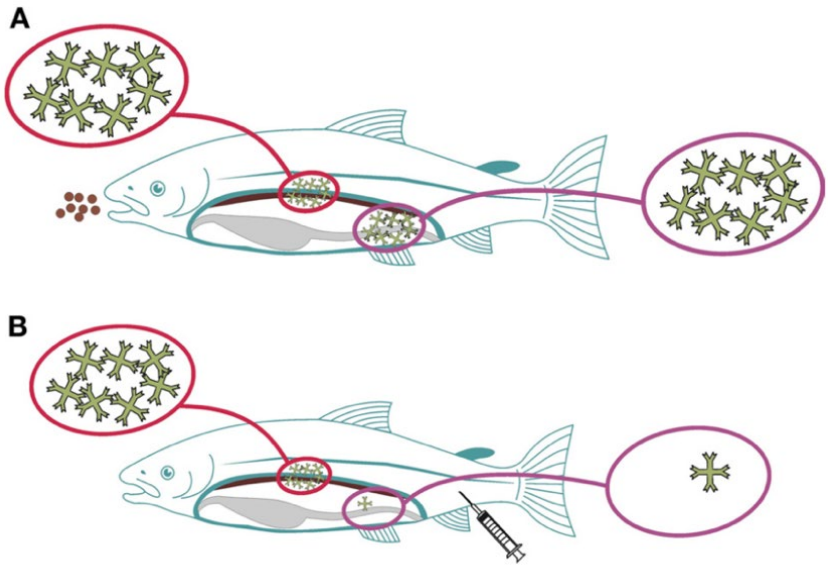

Figure 3 - *Immune responses induced via mucosal (gut) versus parenteral routes in fish.* ***A*** *– Antigens are delivered via the gut. Local and systemic immune responses will be elicited.* ***B*** *– Antigens are delivered parenterally. Systemic responses will be strong and local responses (gut) will be almost absent. Obtained from: Mutoloki, S., Munang'andu, H. M., Evensen Ø. Oral vaccination of fish – Antigen preparations, uptake and immune induction. Front Immunol. 2015;6:1-10. doi:10.3389/fimmu.2015.00519*

The largest body area of living organisms is constituted by mucosal surfaces. They are in constant contact with the external environment, and they play an important role in the maintenance of immunological homeostasis. [30]

Intraperitoneal injection is conventionally used to deliver water-in-oil (w/o)-based injectable vaccines while the intramuscular injection is most often used to deliver DNA plasmids.[22]

Protection is usually higher with injection-vaccination but is also linked with intensive handling and stress for fish. Moreover, the type of adjuvant used in w/o-based injectable vaccines can also be associated with local side effects such as tissue inflammation, adhesion, and necrosis.[22]

Oral and immersion vaccination are usually used for fishes that are too small to be injected, although these routes have low efficacy and short protection. Immersion vaccines are suitable for several bacterial pathogens and are very cheap and easy to administer to small fish. Conversely, injection is labor-intensive and, more importantly, requires the fish to be over a certain size.

The advantages of injected vaccines are that the volume required is small and that every fish is vaccinated with the correct dose, stimulating effectively the body to produce antibodies. [29,31]

Because of this, to ensure protection throughout the entire production cycle, vaccination regimes have been developed for various species in which a combination of immersion, oral, and injection vaccination is used (multicomponent vaccines). It is only after injection with w/o-based vaccines that strong and long-lasting protection is achieved. [22,29]

Assessing vaccines involves determining their effectiveness in stimulating their immune response by evaluating antibodies in the blood of vaccinated fish. The increased expression or the synthesis of some molecules associated with stimulating the innate immune response in fish, such as some inflammatory cytokines, interferon (IFN), and associated molecules are also used as a measure of the efficacy of vaccines. [32]

From date to date, many fish aquaculture vaccines have been developed and approved against various pathogens. Numerous studies have been conducted to increase their efficiency by testing new adjuvants and performing the rational identification of the antigen formulations and pathogen contaminations. [20] However, before such vaccines can be successfully commercialized, several hurdles have to be overcome regarding the production of cheap but effective antigens and adjuvants, while bearing in mind environmental and associated regulatory concerns. [29]

In the 1970s, the effectiveness of fish immersion vaccines based on formalin-inactivated broth cultures was proven. Since then, similar vaccines against the salmonid *Vibrio* diseases were developed. The excellent efficacy of these vaccines resulted in their extensive use and an immediate and permanent reduction in the use of antibiotics. To prevent those diseases, it's very important to have proper fish management with good hygiene and limited stress. [29]

1.3.1 Oral Vaccination in Aquaculture

Theoretically, the ideal method of vaccine delivery in fish is oral vaccination. The oral vaccine with antigens can be included in the feed and much effort has been put into the elaboration and improvement of those vaccines. For antigen delivery in fish, the gut is an attractive route because it offers an easy way of administering antigens and is less stressful. [30]

However, due to the harsh conditions found in the gastric environment (which is also a drawback for other types of oral vaccination) and the degradation of DNA by endogenous nucleases, this procedure has its limitations. Moreover, since the mucosal epithelium has a mucus layer that is highly viscous and has specialized enzymatic processes, only some biomaterials can resist to that. [33] As a result of this, poor and inconsistent responses have been reported by conventional oral vaccines. To overcome this critical problem, different approaches were created to protect the antigen from degradation. Some examples with promising results are: entrapping in liposomes or alginate beads, neutralization of gastric secretions, and application of biofilm vaccines.[29]

Although fish's immune system is based on the fundamental receptors, pathways, and cell types found in all groups of vertebrates, the current knowledge on adaptive immunity in fish is limited compared to the mammalian vertebrates. [34,35] The lack of relevant information about the immune system restrains the development of high-quality oral vaccines, making this complex and difficult.[31] Understanding what constitutes good immunological induction is the main goal to develop good oral vaccines. [30]

Currently, it is still necessary to have a large quantity of antigen and the protection achieved is weak and of short duration. Because of this, immersion in a diluted suspension of the vaccine and injections into the body cavity are still the two main methods of vaccine delivery. [36]

The above-described information is summarized in the following **Table 1.**

Table 1 *- Advantages and challenges of oral vaccination.* [29,30]

Advantages	Challenges
Safety	Destruction in the gut
Simplicity	Fail to induce strong immune responses
Versatility	Autoimmunity
Stability	-
Nontoxicity	-
Long cytotoxic T lymphocyte responses	-

1.3.2 Immune Induction regarding Oral Vaccination

Comparing with mammals, instead of having lymph nodes or Peyer's patches, fish have a less organized, diffuse gut-associated lymphoid tissue (GALT). Although GALT is functionally different from lymph nodes or Peyer's patches, it is also capable of local immune responses. [30]

Oral administration of antigens will up-regulate the genes related to the recruitment of immune cells and local antibody production. However, protection is variable. The nature of antigens, formulation, and dosage can also influence the efficacy of oral vaccines. Since the dosage is the key to any vaccine regime, it is extremely important to determine the dose at an individual level. To accomplish this, it is important to examine feed residues and weight gain at the population level following feeding. This gives an idea of the average intake in the population. [30]

1.3.3 Oral Tolerance

Oral tolerance is described as the hypo-responsiveness to a fed antigen.[37] It is a very well-known phenomenon in fish, and it is a result of the suppression of the cellular and/humoral immune response. It has been recognized as the suppression of antibodies and is easily induced. [30]

Some factors can be responsible for tolerance. Repeated administration of small amounts of antigens, vaccination of too young (immunocompetent) fish, low temperatures (lower end of the permissive limit), type of antigens, administration regime, and genetics have been associated with the induction of tolerance in fish. [30]

1.3.4 DNA Immunization

DNA vaccination is based on plasmid DNA containing antigen-encoding sequences and appropriate promoter/enhancer control sequences. This will lead to the expression of the antigen and its immunogenic presentation by *in vivo* transfected cells. This technique can efficiently stimulate humoral (antibody) and cellular (T-cells) immune responses to the produced protein antigens. In this type of vaccination, immunogenic proteins are expressed *in vivo* by the transfected cells of the vaccine recipients in their native conformation with correct posttranslational modifications from antigen-encoding expression plasmid DNA. Plasmid DNA vaccination will ensure the integrity of antibody-defined epitopes and supports the generation of protective (neutralizing) antibody titers. [38]

This approach is unique, technically simple, and safe since it is a non-live vaccine. Additionally, DNA vaccines induce killer cytotoxic T lymphocytes (CTLs), suggesting that an important shift had occurred in non-live vaccine platforms. [39] Moreover, since DNA is a very stable molecule, they are easy to transport to the site of use, to store and are highly specific. [40]

Several technical improvements occurred in this field. Gene optimization strategies, improved RNA structural design, novel formulations, immune adjuvants, and more delivery approaches made this possible. These methods, when combined together, cause increased levels of immune responses. [39]

Due to its large cellular surface area, highly vascularized nature, and ability to generate mucosal immunity, the intestinal epithelium happens to be a unique target. The intestinal mucosa is composed of lamina propria, which is rich in antigen-presenting cells (APCs). These cells can sample antigens not only directly from the intestinal lumen but also antigens that have been transported across the intestinal epithelium.[33] Professional APCs, which include B cells, macrophages, and dendritic cells, are broadly distributed in tissues and are the major cells of the immune system. The macrophages and dendritic cells will capture the nanoparticles that encapsulate DNA at the beginning of the immune response. [41] Regarding the dendritic cells, they are the most potent immune-initiator cells and exhibit antigen presentation because they can express high levels of Major Histocompatibility Complex (MHC) I and II and co-stimulatory molecules (B7), triggering the activation of naive CD8 and CD4 T cells. [41,42]

While endogenous antigens, like tumor antigens, are processed in the cytosol, degraded by proteasomes, transported to the endoplasmic reticulum, and then associated with the newly formed complexed of class I MHC, the exogenous antigens, which can be inhaled, ingested, or injected like particulate antigen carriers, are going to be processed by MHC II molecule on APCs.

Fundamentally, antigens are taken up by the cells by either MHC class I or II. MHC I is present in all nucleated cells and, MHC II is only present in professional APCs. [43,41]

With MHC II, the antigen is encapsulated into endosome and is exposed to a series of intracellular compartments of increasing acidity. There are 3 different stages: early endosome (pH 6.5–6.0), late endosome (pH 6.0–5.0), and phagolysosome (pH <5.0). The endosome will fuse with lysosomes, forming phagolysosomes (which contain lysozymes). The acidic environment and the action of degradable enzymes of phagolysosomes will promote the degradation. [41]

The step of T cell receptors recognizing peptides presented by MHC molecules on an APCs is fundamental to initiate an adaptive response, which requires the binding of MHC molecules to a CD4 or CD8 receptor. These membrane proteins CD4 and CD8, which are expressed on T helper cells and cytotoxic lymphocytes respectively, are capable of augmenting the sensitivity and response of T cells, allowing the binding to MHC ligands.[44,45]

Fundamentally, using the host cellular machinery, the plasmid enters the nucleus of transfected local cells, such as myocytes or keratinocytes, including APCs. Then, the expression of plasmid-encoded genes is followed by the production of foreign antigens. These antigens can become the subject of immune surveillance in the context of both MHC class I and class II molecules of APCs in the vaccinated host. [39] APCs can provide complete immune protection since they are able to manage both the innate and the adaptive responses of the immune system. This happens because APCs interact with T cells, linking innate and adaptive immune responses. They are essential for successful immune protection, but the efficient delivery of DNA vaccines must also occur in the correct target tissue and to the proper cell type. The professional APCs and other immune cells are going to induce the adaptative immune system, and B and T lymphocytes will mediate the responses. [33,] The above-described information is summarized in the following **Figure 4**.

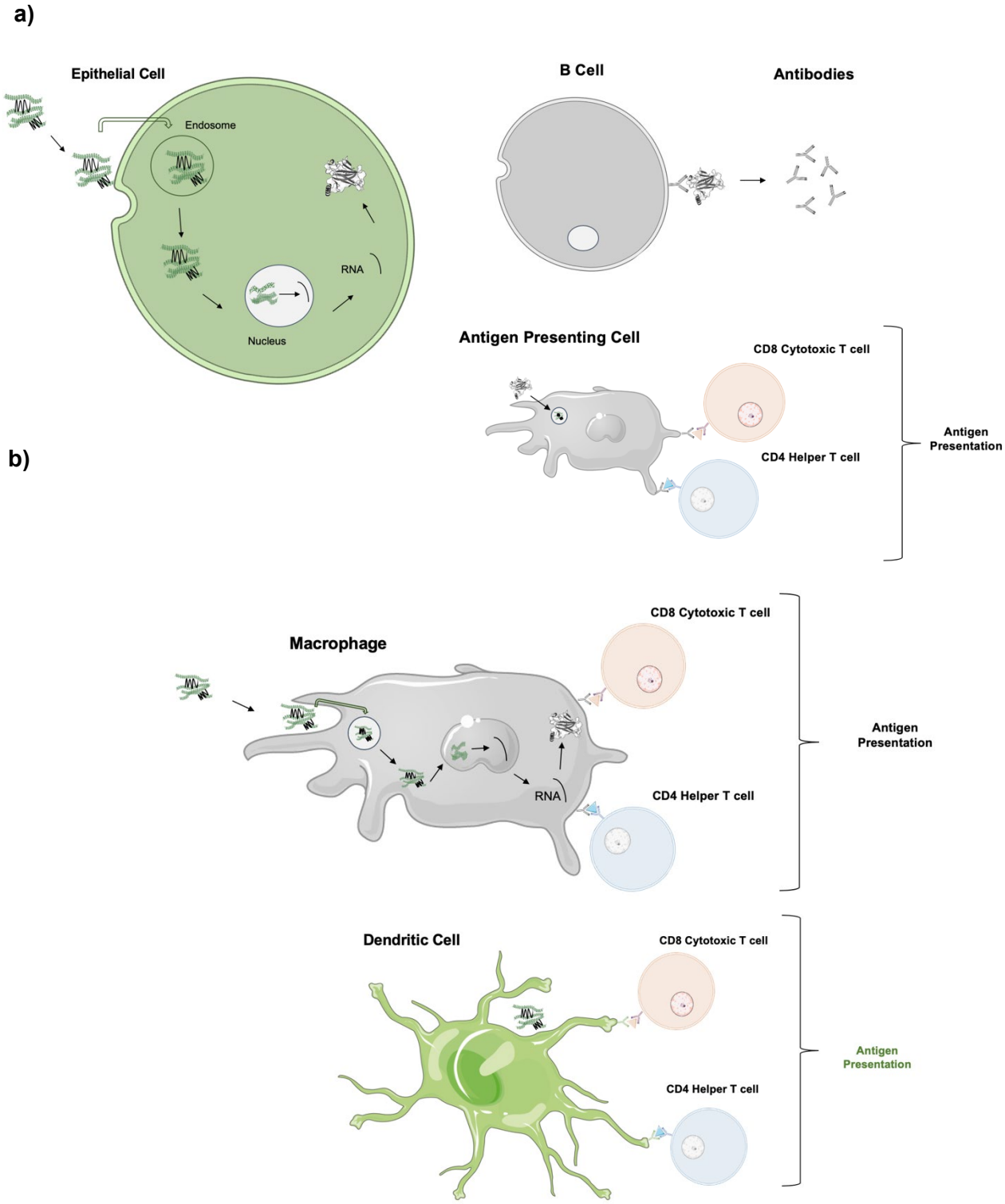

Figure 4 - a) *Nanoparticles encapsulating plasmid DNA can be uptaken by epithelial cells. The pathogen-derived antigens are transcribed and translated from plasmid DNA and secreted into extracellular space. They can be taken up by B-cell receptor-mediated endocytosis oy by professional APCs.*
b) *APCs can also be directly transfected by the uptake of nanoparticles encapsulating plasmid DNA. Professional APCs such as macrophages are more efficient for larger microparticles by phagocytosis and dendritic cells are more effective at the uptake of nanoparticles by micropinocytosis. Adapted from: Farris E, Brown DM, Ramer-tait AE, et al. Chitosan-zein nano-in-microparticles capable of mediating in vivo transgene expression following oral delivery. J Control Release. 2017;249:150-161. doi:10.1016/j.jconrel.2017.01.035*

1.3.5 Challenges

DNA vaccines have some disadvantages, mainly because of health and safety issues. Three main problems are related to this: integration into cellular DNA, development of autoimmunity, and antibiotic resistance. [39]

Most of the safety issues are related to the activation of oncogenes as a result of genomic incorporation of immunizing DNA, as well as producing anti-DNA antibodies. Inactivation of tumor suppressor genes can also occur. Essentially, foreign DNA can be integrated into cellular DNA, causing several complications. [39,40] However, DNA vaccines were tested and didn't indicate relevant levels of integration into host cellular genomic DNA. The integration that is detected occurs at rates that are orders of magnitude below the spontaneous mutation frequency. [39] Even though these issues need to be solved based on both scientific and clinical research studies. [40]

Another undesirable consequence is autoimmunity, although preclinical studies in non-human primates and early studies in humans did not detect increases in anti-nuclear or anti-DNA antibodies. [39]

The presence of antibiotic resistance genes in the delivered plasmids is a major drawback of modern gene therapy and DNA vaccine application. So, to improve those techniques, it's essential to avoid antibiotic resistance genes or other sequences used for the selection and production of the therapeutic plasmid. Better techniques are creating new challenges, and it is likely that in the near future, the regulatory status for antibiotic-free selection will progressively move from preferred to highly recommended and mandatory. Antibiotic-free selection is a field of investigation that will benefit from accumulating knowledge of the biological path.[46]

Also, the separation of the plasmid DNA (pDNA) into minicircles and miniplasmids after the pDNA production can avoid this issue. For example, minicircles are plasmids-based vectors for gene delivery containing only the expression cassette, not having any bacterial sequences. [47]

Nevertheless, genetic immunization is limited to protein antigens, not being useful for non-proteic antigens such as bacterial polysaccharides. [48] The above-described information is summarized in the following **Table 2.**

Table 2 *- Advantages and challenges of DNA immunization.* [39,48]

Advantages	Challenges
Design	Integration
Time to manufacture	Autoimmunity
Safety	Antibiotic Resistance
Stability	Low immunogenicity
Mobility	Limited to protein antigens
Immunogenicity	-

1.4 Plasmid DNA vectors

Plasmids are circular or linear extrachromosomal replicons, found in many microorganisms in the domains *Bacteria*, *Archaea*, and *Eukaryota*. [49] Those double-stranded DNA molecules range in size from 2-3 genes (2-3kb) to elements that can contain 400 genes or more, and they do not contain any of the set of core genes needed by the cell for the basic growth and multiplication. Instead, they carry genes that may be useful periodically to enable the cell to exploit particular environmental situations, such as genes that confer antibiotic resistance. Basically, the plasmid is essential for adaptation by means of genetic diversity. Genetic diversity in bacteria is due to the active transport of genes which are incorporated with accessory elements like plasmids. [50]

Plasmid vectors are DNA platforms, deoxyribonucleic acid vectors, that are comprised of a viral gene promotor, a gene encoding resistance to an antibiotic, a bacterial origin of replication, and a multiple cloning site (MCS) for a transgenic region. Herein, one or several antigenic genes of interest can be inserted. [51]

DNA vaccine combines a eukaryotic region that directs expression of the transgene in the target organism and a bacterial region that provides selection and propagation in the *Escherichia coli* host.[52]

For plasmid DNA production, several genetic elements are needed **(Figure 5)**. A replication region and a selection marker are mandatory to allow the maintenance of multiple copies of the plasmid per host cell and stable inheritance of the plasmid during the bacterial growth, respectively.[53]

The eukaryotic expression unit comprises an enhancer/promoter region, intron, signal sequence, vaccine gene, and a transcriptional terminator. Immune stimulatory sequences (ISS) will enhance the potency of the DNA vaccine and also add adjuvanticity. [53] The polyadenylation signal will mediate mRNA cleavage and polyadenylation, leading to efficient mRNA export to the cytoplasm. [52]

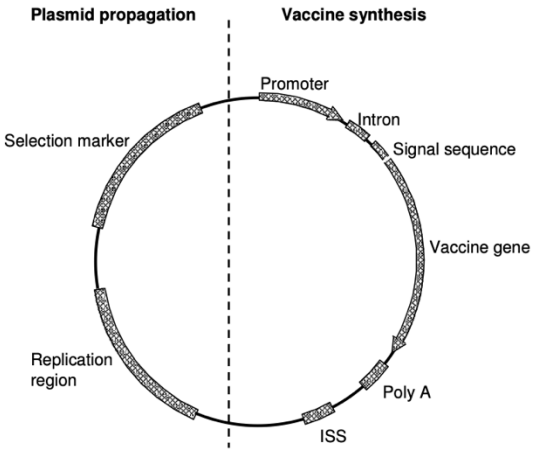

Figure 5 - *Genetic elements of a plasmid DNA vaccine. Plasmid DNA vaccines consists of a unit for propagation in the microbial host and a unit that drives vaccine synthesis in the eukaryotic cells. Replication region and a selection marker are employed. The eukaryotic expression unit comprises an enhancer/promoter region, intron, signal sequence, vaccine gene and a poly A sequence. Immune stimulatory sequences (ISS) add adjuvanticity and may be localized in both units. Obtained from: Glenting J, Wessels S. Ensuring safety of DNA vaccines. Microb Cell Fact. 2005;4:1-5. doi:10.1186/1475-2859-4-26*

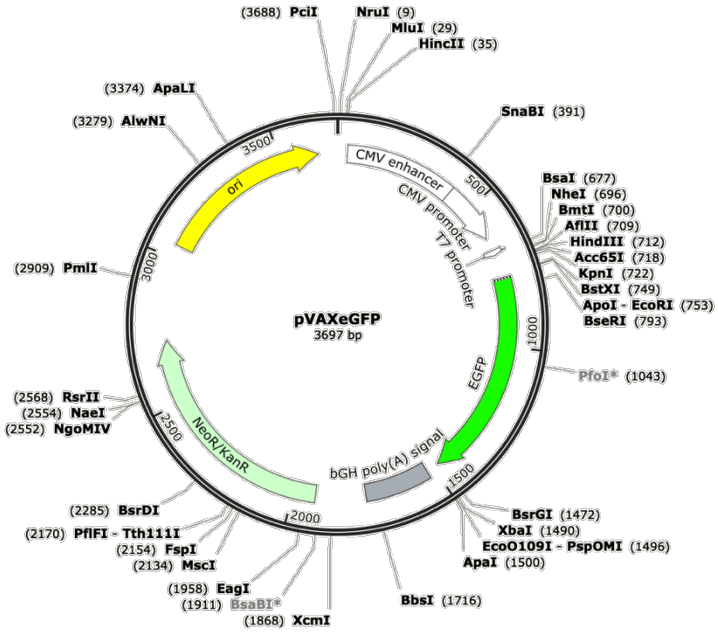

Figure 6 - *pVAXeGFP map and features. This image was obtained using SnapGene.*

In the work of *Azzori et al.*, pVAX eGFP (**Figure 6**) was created by modification of the commercial plasmid pVAX1lacZ, with 6050 bp, where the β-galactosidase reporter gene was replaced by the enhanced Green Fluorescent Protein (eGFP) gene. This vector also contains the human cytomegalovirus (CMV) immediate-early promoter, allowing the expression of the desired recombinant protein in eukaryotic cells; the bovine growth hormones polyadenylation (*bgh*-PolyA) sequence, allowing the nuclear export and consequent translation of the mRNA; a kanamycin resistance gene (kan) for bacterial selection, preventing the growth of plasmid-free bacteria during fermentation and a pUC origin of replication, that permits a high-copy number replication in bacteria. [54]

1.5 Outer membrane proteins (OMPs)

Outer membrane proteins (OMPs) have been revealed as good vaccine candidates against bacterial infections since they are highly antigenic proteins. There are already some studies regarding this topic.[10] One example is the recombinant OmpK, which is a good protective antigen against *V. harveyi* in Orange-spotted grouper, meaning that this protein should be considered as a vaccine candidate for vaccine development.[55] Likewise, OmpU can also be used as an efficient vaccine candidate for the disease caused by *V. harveyi* [56] and OmpW against *V. alginolyticus*.[57]

OMPs have been proven to play an essential role in the survival and multiplication of bacteria in a host system [58] and are responsible for the interactions between bacteria with hosts in adherence, uptake of nutrients, and destabilizing host-defense mechanisms. [55]

OMPs are exclusively present in Gram-negative bacteria, being one of the constituents of the outer membranes (with sugars and lipids) and are also important adhesion factors [10,57,59] Since they can be easily recognized as foreign substances by the host's immunological defense system, they are suitable and convenient components to use for diagnosis purposes.[59]

OMPs are highly hydrophobic, suggesting that they can preserve their binding interactions. These proteins are also resistant to bile, meaning that bacteria can colonize the intestine of the fish and also stimulate the biofilm production, providing an adaptative and survival advantage for bacteria in an aquatic environment. [59,60]

1.5.1 OmpK and OmpK-frag

Particularly, the recombinant OmpK, which is widely distributed among species of the family *Vibrionaceae*, is a good protective antigen against vibriosis. Because of this, the OmpK should be considered as a vaccine candidate for vaccine development. [55]

The two pDNA encoding OmpK and immunogenic portions of OmpK were designed by an IBB (Institute for Bioengineering and Biosciences) collaborator who had previously worked with them, in order to be used for immunization of fish in aquaculture. The backbone used for the two plasmids was the pVAX GFP expression vector, with 3697 bp [54] **(Figure 6, Annex 1)**.

Both OmpK and immunogenic portions of OmpK (**Annexes 2 and 3)**, respectively were fused to the reporter gene *gfp,* in order to detect fluorescence after the expression in the host cells. At the end of this process, two different plasmids: pVAX ompK and pVAX ompK-frag were created. This process was similar to the one developed by Rauta *et al.* [61, 62]

1.6 Nanoparticles for pDNA Delivery

Nanoparticles are solid, colloidal particles, varying from 10nm to 1,000nm. In fact, for nanomedical application, is important that the size is less than 200nm.[63] They are classified into different classes, based on their properties, shapes, and sizes. They have unique physical and chemical properties due to their high surface area and nanoscale size. The reactivity, toughness, and other properties are also dependent on their unique size, shape, and structure. [64]

Due to these characteristics, nanoparticles are suitable candidates for various applications.[64] Nanoparticles in vaccine development have opened up incredible opportunities in the field of biomedicine. They can be classified according to their action, which means that they can be grouped either as an efficient mode of delivery system or as an adjuvant. The first ones can deliver the antigen to targeted immune cells while protecting it, and the adjuvant nanoparticles will activate specific pathways, to provide an efficient antigen uptake and processing. [15]

To improve vaccine efficiency, these nanocarriers should protect the antigens from premature proteolytic degradation, facilitate antigen uptake and processing by antigen-presenting cells, control release, and should be safe. [65]

Although nanoparticles provide a promising approach for improving the biodistribution of drugs, they still display several limitations, such as clearance by the immune system and impaired diffusion in the tissue microenvironment. [66]

1.7 Chitosan Nanoparticles

Chitosan is the most important derivative of chitin (**Figure 7**). Chitin, β-(1→4)-linked polymer of 2-acetamido-2-deoxy-d-glucose (N-acetyl-d-glucosamine), is present in the exoskeletons of insects, crustaceans and in the cell walls of fungi and algae. [67, 68]

Chitosan is a naturally occurring polysaccharide, cationic, highly basic, mucoadhesive, and biocompatible polymer that is produced by removing the acetate moiety from chitin. [67]

Chitosan-derived biomaterials have received substantial attention as anti-microbial, functional, renewable, non-toxic, biocompatible, bioabsorbable, and biodegradable biopolymer agents. Is insoluble in water and organic solvents but soluble in acetic, nitric, hydrochloric, perchloric, or phosphoric acids[69] and widely used in many areas such as biomedical engineering, fiber industry, environmental protection, and cosmetics. [68]

The solubility range can be changed upon depolymerization and chemical modification of the primary and secondary hydroxyl groups. [69]

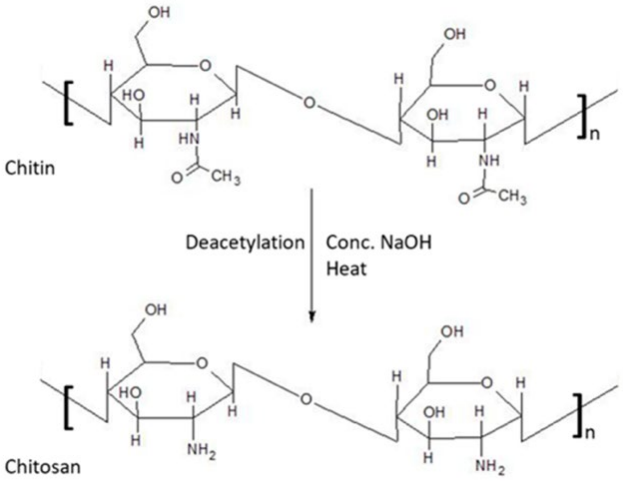

Figure 7 - *Deacetylation of chitin to chitosan. Obtained from: Mohammed MA, Syeda JTM, Wasan KM, et al. An overview of chitosan nanoparticles and its application in non-parenteral drug delivery. Pharmaceutics. 2017;9(4). doi:10.3390/pharmaceutics9040053*

To produce chitosan, chitin is first extracted/separated from its natural sources through processes of acidification and alkalization. Purified chitin is then N-deacetylated to chitosan **(Figure 7)**.[67] Basically, the acetyl groups in chitin can be removed by deacetylation to convert insoluble chitin into a more soluble compound, chitosan. This name is given when there is at least 50% degree of deacetylation. [70]

The quality and chemical properties of both chitin and chitosan are closely related to the degree of deacetylation, which is defined as the molar fraction of deacetylated units in the polymer chain. [68] This step is fundamental to control the end product properties, such as molecular weight (MW) and pKa (6–7.5). Moreover, solubility, viscosity, ion-exchange capacity, the ability of flocculation, and reaction with the amino group are also related to the degree of deacetylation. [67]

It's also known that chitosan can interact with mucus since the latter is negatively charged. They will form a complex by ionic or hydrogen bonding as well as though hydrophobic interactions. [67] At acidic pH, the amino groups situated on chitosan become protonated. This will transmit a positive charge to the chitosan chains, and the protonated chitosan in an acidic environment condense negatively charged DNA into nanoscale polyplexes through electrostatic interaction. Then, those polyplexes can provide efficient protection of DNA against nuclease degradation and deliver DNA into target the cell via endocytosis.[69,71]

Since most biological cell surfaces are anionic, chitosan will strongly adhere to the tissues at the site of a wound via electrostatic interactions due to its cationic characteristics. [69]

Chitosan nanoparticles are a drug carrier and have the advantage of slow/controlled drug release, improving drug solubility and stability, enhancing efficacy, and reducing toxicity.

Because of their small size, they can pass through biological barriers *in vivo*, such as the blood-brain barrier, to deliver drugs to the lesion site. [72]

Chitosan opens the tight junctions of the epithelium, facilitating paracellular and transcellular transport of drugs **(Figure 8)**. [67]

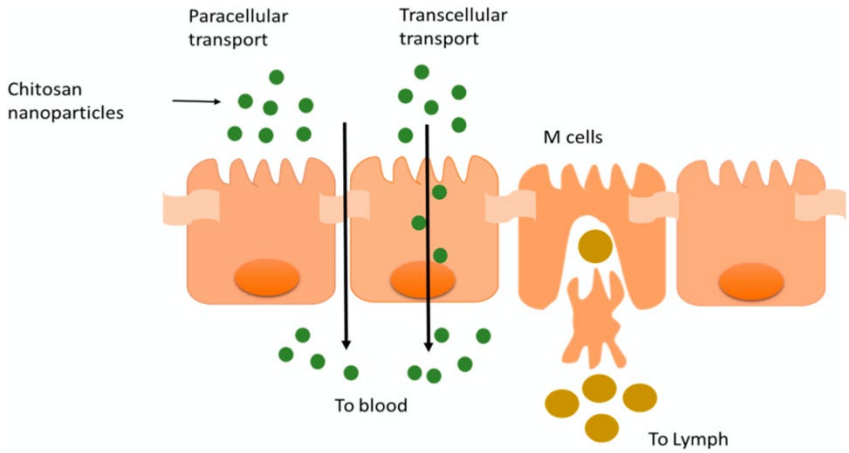

Figure 8 - *Presumed mechanism of transcellular and paracellular transport of chitosan NP across the epithelium. Obtained from: Mohammed MA, Syeda JTM, Wasan KM, et al. An overview of chitosan nanoparticles and its application in non-parenteral drug delivery. Pharmaceutics. 2017;9(4). doi:10.3390/pharmaceutics9040053*

One of the most important factors that can influence gene expression is the molecular weight of the chitosan. Ranging from several thousand to millions of Daltons (Da), this factor is also very important regarding the biodegradability and cytotoxicity.[71]

Biodegradability and cytotoxicity are the most important advantages of low molecular weight chitosan over other gene carriers. It was shown that low molecular weight chitosan mediated higher transfection efficiency *in vitro* than did the high molecular weight chitosan. It was also stated that the high molecular weight chitosan degraded slowly *in vivo*, increasing the risk of accumulation in the tissues in a long period of administration. Shorter-chain low molecular weight chitosan was easily degraded to smaller oligo- and monosaccharides. [71] Under the action of endogenous enzymes *in vivo*, biodegradable nanoparticles can produce water and carbon dioxide without adverse effects.[72]

Nanoparticles produced by ionic gelation are one of the most studied nanosystems for drug delivery. Nevertheless, the lack of inter-laboratory reproducibility and poor physicochemical understanding of the process of the particle formation are some drawbacks that have been slowing their potential market application.[73]

One example of nanoparticles produced by ionic gelation is chitosan-tripolyphosphate (TPP) nanoparticles. For this chemical system, the average hydrodynamic diameter of the particles produced is strongly dependent on the initial chitosan concentration. The second most important factor to form particles is the degree of acetylation of the chitosan. [73]

Some advantages of ionic gelation over other methods are: the nanoparticles are obtained spontaneously under mild control conditions without involving high temperatures, organic solvents, or sonication and the TPP **(Figure 9)** is a multivalent polyanion, with low toxicity and cost, with no severe constrains of handling and storage, unlike other cross-linkers.

After the TPP is added, nanoparticles are form immediately through inter and intramolecular linkages created between TPP phosphates and chitosan amino groups. [74]

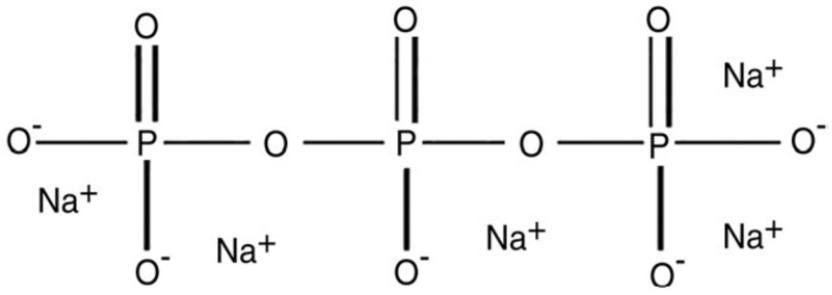

Figure 9 - *Chemical structure of sodium TPP. Obtained from: Sreekumar S, Goycoolea FM, Moerschbacher BM, et al. Parameters influencing the size of chitosan-TPP nano- and microparticles. Sci Rep. 2018;8(1):1-11. doi:10.1038/s41598-018-23064-4*

Afterwards, these nanoparticles are mixed with purified pDNA in order to encapsulate it. The N/P ratio [the ratio of positively-chargeable polymer amine (N=nitrogen) groups to negatively-charged nucleic acid phosphate (P) groups] is one of the most important physicochemical properties of polymer-based gene delivery vehicles. It can influence many other properties such as its net surface, size and stability. [75]

The ratio is expressed as the ratio of moles of the amine groups of chitosan to the phosphate ones of TPP and DNA.

$$N\ (amine\ moles) = \frac{m\ chitosan\ (g)}{\text{Molecular Weight glucosamine residue } (\frac{g}{mol})} \qquad \textbf{(Equation 1)}$$

$$P\ (phosphate\ moles) = \frac{m\ pDNA\ (g)}{\text{Molecular Weight nucleotide } (\frac{g}{mol})} + \frac{3 \times m\ TPP\ (g)}{\text{Molecular Weight TPP}(\frac{g}{mol})} \qquad \textbf{(Equation 2)}$$

The number of moles of the amine groups of chitosan can be obtained using **Equation 1**. Herein, m represents the mass, in grams, of chitosan in the complexation. The molecular weight of chitosan glucosamine residues is 231 g/mol.[76]

The number of the phosphate moles from both pDNA and TPP can be calculated using **Equation 2**, where m also represents the mass, in grams, of pDNA, and TPP.

The average molecular weight of a nucleotide is 330 g/mol, and the molecular weight of TPP is 367.86 g/mol. The TPP mass must be multiplied by three since it will contribute with three phosphates per molecule.

1.8 Nanoparticles Characterization

To characterize nanoparticles, important properties such as Z-average, zeta potential and polydispersity index (PdI) must be determined. Z-average can be obtained with Dynamic Light Scattering (DLS) and zeta potential with electrophoretic mobility (EM), using an equipment like the Zetasizer Nano.

1.8.1 Dynamic Light Scattering

One important basic information about nanoparticles is their size, as it is one of the main determinants of biodistribution and retention of the nanoparticles in the target tissues. The particle size is defined as the size of a hypothetical hard sphere that diffuses in the same fashion as that of the nanoparticles being measured. [77]

After a DLS measurement, the Z-average (or cumulants mean) is the first and more accurate size value. However, this value is only meaningful if the provided sample has monomodal, spherical, and monodisperse particles. For bimodal distribution, where nanoparticles have clear size differences, the Z-average value has no meaning. It is also the primary parameter produced by the method but very sensitive to small changes in the sample. [78]

For particle size distribution, DLS is commonly used. This technique, also known as photon correlation spectroscopy, allows the calculation of diffusion coefficients associated with the Brownian movement of particles dispersed in liquid media. Thus, by irradiating a sample with a laser beam, it is possible to analyze the intensity fluctuations of the light scattered by the particles. [79] Additionally, the nanoparticles are always in a random motion due to their kinetic energy. Because of this, the variation of the intensity with time will contain information on that random motion. This will be very useful to measure the diffusion coefficient of the particles. [80]

Brownian movement is defined as the movement of particles due to random collisions with molecules of liquid in their surroundings. Small particles move faster than larger particles. This means that the translational diffusion coefficient (D) of the particles is inversely related to their size. This can be shown by the Stokes-Einstein equation (**Equation 3**), where k is the Boltzmann constant, T is the temperature in Kelvin, η is the viscosity of the dispersing medium, and $\mathrm{R_H}$ is the hydrodynamic radius of the particle. [79]

$$D = \frac{\mathrm{kT}}{6\pi\eta R_{\mathrm{H}}} \qquad \textbf{(Equation 3)}$$

The obtained results are reported as a mean particle size and homogeneity of size distribution. Polydispersity index (PdI) is a dimensionless parameter that can be obtained from a cumulant analysis of the DLS-measured intensity autocorrelation function. [79]

PdI is used to estimate the average uniformity of a particle colloidal solution and the larger the PdI values, the larger the size distribution in the particles in the sample. A higher PdI value can also suggest particle aggregation. A monodisperse distribution of particles in solution shows PdI lower than 0.1.

PdI is defined as the squared of ratio between the standard deviation (σ) of the particle diameter distribution and the mean particle diameter (**Equation 4**). [81]

$$PdI = \left(\frac{\sigma}{2R_{\mathrm{H}}}\right)^2 \qquad \textbf{(Equation 4)}$$

The measuring time for DLS is short and almost automated, meaning that the entire process is not labor intensive. Additionally, DLS is a non-invasive technique and the sample can be used for other purposes after the measurements. [80]

To reduce the diameter, sonication or vigorous vortexing can be performed. It was detected that sonication can reduce agglomeration and has minimal effect on particle surface charge.[82] These techniques can help by homogenizing the suspensions.

1.8.2 Zeta Potential

Regarding the surface charge, the Zetasizer can measure the zeta potential, which is indicative of the particle surface charge, being a characterization method of nanometer-sized objects in liquids, such as pharmaceuticals [83]

Since most colloidal dispersions in aqueous media carry an electric charge, the development of this charge at the particle surface affects the distribution of ions in the surrounding interfacial region. [84]

Furthermore, the zeta potential is a measure of the effective electric charge in the nanoparticle surface, quantifying the charges. If a nanoparticle has a net surface charge, the charge is screened by the concentration of ions of opposite charge near the nanoparticle surface. [85]

Two liquid layers are surrounding the nanoparticles: the Stern layer (strongly bound inner part) and a weakly bound outer layer. One significant limitation is that in bimodal samples, the zeta potential value of larger particles dominates the scattering signal of smaller particles. [77]

In zeta potential measurements, an electric field is applied across the sample. The movement of nanoparticles, electrophoretic mobility, is measured by Doppler velocimetry (LDV). Laser Doppler electrophoresis is an accepted method for the measurement of particle electrophoretic mobility and hence zeta potential of dispersions of colloidal size materials. [84] To calculate the zeta potential (ζ), Henry's Equation is used:

$$U_{\mathrm{e}} = \frac{2}{3}\,\frac{\varepsilon\zeta}{\eta}\,f(ka) \qquad \textbf{(Equation 5)}$$

where U_{e} is the electrophoretic mobility, ε is the dielectric constant or permittivity of the medium, η is the absolute zero-shear viscosity of the medium, *f(ka)* is the Henry's function, and *ka* is a measure of the ratio of the particle radius to the Debye length.

The Henry's function varies monotonically between 1.0 (at k=0) and 1.5 (at k=∞) and k, the Debye-Hückel parameter, corresponds to inverse of Debye length [86,87]

The zeta potential measurements depend on the strength and valency of ions that are present in the nanoparticle's suspension. High ionic strength and high valency ions compress the electric double layer, reducing the zeta potential. Moreover, the pH can greatly influence the zeta potential. If the suspension is acidic, the nanoparticles will have a positive charge and vice versa. This means that a zeta potential value without the indication of pH is meaningless. [77]

Moreover, the magnitude of the zeta potential provides information about particle stability. The higher the magnitude of the potential, the higher the electrostatic repulsion. Thus, the electrostatic repulsion will promote the stability of the nanoparticle. [85]

Furthermore, one important aspect is also the sample concentration. The upper limit of concentration depends on the particle size, PdI, and other optical properties of the particles. If the sample concentration is too high, the laser beam does not penetrate through the sample, which will decrease the scattered light that is going to be detected. So, there is a maximum turbidity of the sample that can be measured. Some samples may require a solution, however, it is not desirable to dilute the samples, in order to minimize any changes in the zeta potential. [84]

1.9 Encapsulation of pDNA

1.9.1 Fluorescence DNA Intercalators

DNA intercalators are widely used to visualize DNA and DNA transactions as fluorescent probes *in vitro* and *in vivo*. They can perturb DNA structure and stability, influencing DNA-processing by proteins. They are small molecules that can reversibly bind between adjacent base pairs of double-stranded DNA.[88]

The ability of nanoparticles to form complexes with pDNA can be assessed by using the fluorescent intercalating dye (for instance, OliGreen) to label-free plasmid DNA. Fluorescence from the OliGreen nucleic acid staining will be an indication of free or incompletely complex DNA. The sample that only contains free plasmid in water will exhibit the maximum fluorescence. [61]

In complexed pDNA-nanoparticles samples, a decrease in the fluorescence intensity is expected, when compared with naked pDNA. [61] Such a decrease in fluorescence indicates that the pDNA is complexed to the point that DNA dye association and fluorescence are prevented. [89]

1.9.2 Gel Electrophoresis

The lack of electrophoretic mobility of the pDNA in the complexed form while subjected to gel electrophoresis conditions, confirms complete encapsulation. This lack of mobility was seen since no free plasmid bands were visible in complexes with pDNA/nanoparticles in the respective electrophoresis gel. [89]

Moreover, after the treatment with DNase, the protection from enzymatic digestion that nanoparticles complexes confer to pDNA can also be confirmed by gel electrophoresis (**Figure 10**). If free pDNA is present, typical bands of open circular and supercoiled pDNA will no longer be observed after incubation with DNase, while degraded fragments can be seen. However, such a pattern will not occur after DNase treatment over the encapsulated samples. [61]

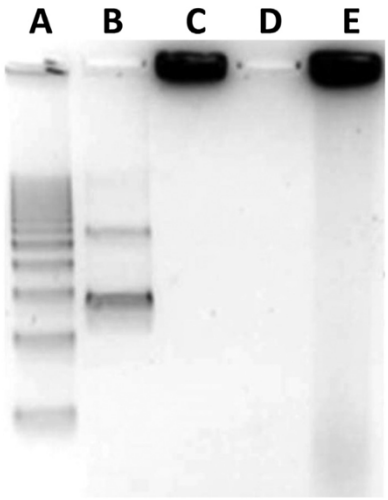

Figure 10 - *Protection of plasmids from nuclease digestion by biopolymer complexation. Lane A − 1 kb molecular weight marker. The 100 ng of pDNA, naked (Lane B) or complexed in nanoparticles (Lane C), become fully degraded when naked and incubated with DNase I (Lane D) or stay intact if complexed in nanoparticles before incubation with DNAse I (Lane E). Obtained from: Rauta PR, Nayak B, Monteiro GA, et al. Design and characterization of plasmids encoding antigenic peptides of Aha1 from Aeromonas hydrophila as prospective fish vaccines. J Biotechnol. 2017;241:116-126. doi:10.1016/j.jbiotec.2016.11.019*

2. Objectives and Motivation

High stocking densities, poor water quality, low oxygen levels, and high pathogen loads can expose fish to high-stress levels. The aquaculture environment, such as tanks and cages, will increase the susceptibility of infectious diseases, affecting economic development, social stability worldwide, and health problems associated with the consumption of diseased fish. Therefore, one approach to avoid infectious diseases is vaccination.

Nowadays, the commercially available vaccines for fish are made of inactivated bacteria and are applied by emersion or injection with an oil adjuvant. The oil can increase the efficiency and the duration of the protection, but it can cause local side effects such as tissue inflammation, adhesion, and necrosis. These previous complications can be avoided if oral vaccination is used.

Oral vaccines can be mixed with the feed, even though it is necessary to examine feed residues and the population weight gain, to understand the dose at the individual level. However, to improve vaccine efficacy, it should be coated (it can be coated with several materials such as chitosan or PLGA), even though, in most cases, only limited protection can be obtained.

Premature DNA degradation can be solved with suitable carriers, such as nanoparticles. These carriers will increase the efficiency, by protecting the antigens from premature proteolytic degradation, facilitating antigen uptake, processing by antigen-presenting cells, and control release. This type of vaccination is easy to administer, safe (non-live vaccines), nontoxic, and versatile.

With this project, we aim to produce and characterize nanoparticles in order to find the best ones capable of delivering the pDNA to the cells. Chitosan-tripolyphosphate (TPP) nanoparticles were employed, and different concentrations of chitosan and TPP were tested. To characterize nanoparticles, the average hydrodynamic diameter, zeta potential, and polydispersity index were measured, using equipment like the Zetasizer Nano.

The backbone used for both plasmids (OmpK and OmpK-frag) was the pVAX eGFP expression vector. OmpK is an outer membrane protein present in *Vibrio* species. Vibriosis, caused by *Vibrio* species, is one of the most common bacterial diseases among marine fish and shellfish. OMPs are great vaccine candidates against bacterial infections since they are highly antigenic proteins. Moreover, some experiments were done using bioinformatics tools in order to predict other antigenic amino acid sequences. Besides OmpK, the selected proteins were OmpW, OmpV, and OmpU.

3. Workflow

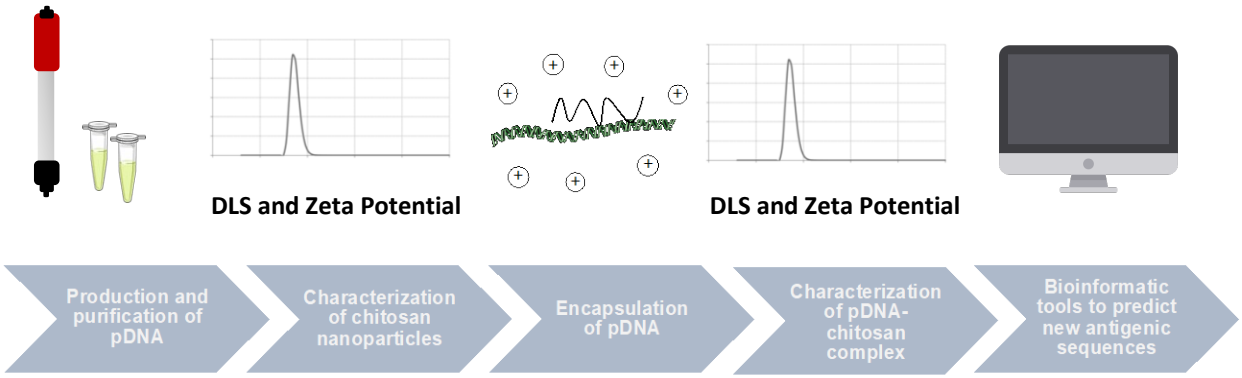

Figure 11 - Schematic workflow of the research work. *Upon the production and purification of pDNA, the nanoparticles were synthesized and both pDNA and nanoparticles were complexed together. To ensure the complexation, several methods were utilized, such as DLS and Zeta Potential. Bioinformatic tools were also used in order to predict new antigenic sequences.*

The first step into this project was to produce and purify the plasmids pVAX eGFP, pVAX ompK, and pVAX ompK-frag, in order to perform the encapsulation experiments. The two pDNA encoding OmpK and immunogenic portions of OmpK were already designed and ready to be utilized. Subsequently, the desired pDNA was transformed into chemically competent *Escherichia coli* cells (DH5 α strain). The plasmids were expressed, extracted, and purified, as described in **Sections 4.1** and **4.2**.

Alongside this, the nanoparticles were synthesized and characterized. The most desirable nanoparticles were then encapsulated with the previously purified plasmids and then were characterized again. Different chitosan nanoparticles formations were performed in order to find the most suitable ones for DNA vaccines (**Section 4.3** and **4.4**).

Moreover, with the intent of finding new antigenic sequences, other proteins besides the OmpK were analyzed. For this, we used several bioinformatic tools, so that we could predict the best antigenic regions in outer membrane proteins (**Section 4.5**).

4. Materials and Methods

4.1 *E. coli* expression and plasmid purification

4.1.1 Strains and plasmids

The *Escherichia coli* strain used for this experiment was DH5α. Regarding the plasmids, pVAX eGFP, pVAX ompK, and pVAX ompK-frag (**Figure 12** and **Annexes 1,2, and 3**) were obtained from a stock already prepared. The plasmids were used to transform chemically competent *E. coli* DH5α by thermal shock.

4.1.2 Medium and growth conditions

Luria-Bertani (LB) medium (NZYTech, 10 g/L of tryptone, 5 g/L yeast extract, 10 g/L NaCl), supplemented with kanamycin (30μg/mL), was used to grow *E.coli* DH5, at 37ºC, 250rpm in an orbital shaker.

4.1.3 Cell storage bank preparation

Firstly, the cells were grown overnight at 37ºC, 250rpm, in 5 mL LB medium in a 15mL centrifuge tube. In the next morning, the inoculum volume such that $OD_{660}\times V(mL)\approx 0.1$ was transferred to a new 15mL centrifuge tube with 5mL LB medium. Then, the cells were allowed to grow until the exponential phase and, after reaching $OD_{600}\approx 1$, aliquots containing 80μL of cell suspension and 20μL of 99% sterile glycerol (or 65μL of cell suspension and 35μL of 50% sterile glycerol) were prepared in 1.5mL Eppendorf's microtubes that were stored at -80ºC for further use.

4.1.4 Chemical competent cells

Regarding the chemical competent cells, after cells have grown overnight, the inoculum volume such that $OD_{660}\times V(mL)\approx 0.1$ was transferred to a 100mL Erlenmeyer flask containing 20mL LB medium. Again, cells were allowed to grow until $OD_{660}\approx 1$. The suspension was divided into two 15mL centrifuge tubes, and then both were centrifuged at 1000xg for 10 min, at 4ºC. The supernatants were discharged, and each pellet was resuspended in 1mL of cool transformation and storage solution (TSS) previously prepared.

TSS (20mL) was prepared with 0.2g of $MgCl_2.6H_2O$ (10g/L), 2g of PEG8000 (100g/L) and 0.4g of LB broth (20g/L), and then, resuspended in 19mL of MilliQ H_2O (95% (v/v)). After pH adjustment to 6.5, the solution was microfiltered (0.22μm filter). From this solution, 1900μL were mixed with 100μL of DMSO (5%(v/v)).

After the resuspension in TSS, 100μL aliquots were transferred to 1.5mL Eppendorf microtubes, whist rested on ice for 10 min, and were stored at -80ºC for further use.

4.1.5 Transformation by heat shock

The chemically competent cells contained in one microtube were mixed with 100ng of plasmid DNA and incubated on ice for 30 min. Then, the microtube was placed in a warm bath for 1 min at 42ºC and again on ice for 2 min. Next, 900μL of LB medium was added, and the suspension was left for 1 hour at 37ºC and 250 rpm.

Afterward, the suspension was plated in LB agar (NZYTech, 10 g/L of tryptone, 5 g/L yeast extract, 10 g/L NaCl and 15 g/L agar), supplemented with 30 μg/mL of kanamycin. This process was performed for the transformation of *E. coli* DH5α with pVAX eGFP, pVAX ompK, or pVAX ompK-frag.

The colonies that were grown in the LB agar with kanamycin were taken from the agar plate and inoculated into the liquid medium for a starter culture. The starter culture needs to be carefully monitored to obtain an optimum optical density at 600 nm (between 0.4 and 0.9, the log phase). The subsequent larger culture (100mL) also needs to be exactly observed.

4.1.6 pDNA production

To produce pDNA, the transformed *E. coli* DH5α was grown overnight in a 100mL Erlenmeyer containing 30mL of LB medium, supplemented with 30μg/mL of kanamycin in the conditions already described. In the next morning, the inoculum volume V(mL) such that $OD_{660} \times V(mL) \approx 0.1$ was transferred to a 250mL LB medium also with 30μg/mL of kanamycin in a 2L Erlenmeyer. The culture was allowed to grow before reaching the stationary growth phase when $OD_{600} \approx 4$.

4.1.7 Alkaline lysis and pDNA primary purification

According to the method previously described by Alves CPA *et al.*[90], cells were lysed following a modification of the alkaline lysis method. First, the pellet was resuspended in buffer I, which is composed of 50 mM glucose, 25 mM Tris–HCl pH 8, and 10 mM EDTA, pH 8. The cell suspension volume used for lysis corresponds to an OD_{600} of 60. Subsequently, the same volume of buffer II, which is composed of 0.2 M NaOH and 1% (w/v) SDS, was added to lyse the cells. The suspension was gently mixed, and the tubes were left resting for 10 min at room temperature. Buffer III, composed of 5M acetate/ 3M potassium (pH 5.0) solution prepared with potassium acetate, glacial acetic acid, and water, was added to stop cell lysis and neutralize the mixture. The volume used was the same as the one used for Buffer I and Buffer II. Then, the tubes were left resting on ice for 10 min. To remove cell debris, genomic DNA (gDNA), and proteins, the suspension was centrifuged at 13,000xg for 30 min at 4ºC in a Sorvall RC6 centrifuge with an SS-34 rotor. The supernatant was recovered to a clean tube and centrifuged again with the same settings to eliminate the remaining debris.

Afterward, the nucleic acids (pDNA, RNA, and traces of gDNA) were removed by precipitating them with 0.7 volumes of pure isopropanol for 2 hours at -20ºC. Then, the mixture was centrifuged again at 13,000xg for 30 min at 4ºC. The pellets were left to dry until the next morning at room temperature.

On the next morning, the pellets were resuspended in 1mL of 10 mM Tris–HCl, pH 8, and 2.5M of solid ammonium sulfate were added to the solution to precipitate the remaining proteins. The solution was gently mixed and then left to rest on ice for 15 min. Subsequently, the solution was centrifuged at 18,500xg for 30 min at 4ºC. The supernatants were kept at 4ºC until the purification step (Hydrophobic Interaction Chromatography).

4.1.8 Hydrophobic interaction chromatography

The solution obtained in the previous section (**4.1.7**) was purified by Hydrophobic Interaction Chromatography (HIC), using a membrane adsorber Sartobind® Phenyl Nano unit of 3mL bed volume. This membrane column was connected to an ÄKTA purifier 10 FPLC system (GE Healthcare). The membrane has phenyl moieties at its surface for binding the pDNA and hydrophobic impurities, which are separated with a step elution profile.

After equilibrating the membrane with the binding buffer, the pDNA-containing feed was injected into the column and allowed to flow through during 15 min and four elution steps were performed next (15, 25, 35, and 100% elution buffer-steps). The system was operated at a flow rate of 1mL/min as described previously.[91]

The binding buffer used was 1.8 M $(NH_4)_2SO_4$ in 10 mM Tris–HCl, pH=8.0, and the elution buffer was 10mM Tris–HCl, pH=8.0. The conductivity and UV absorbance of the eluate, at 260nm, were continuously monitored. Fractions were collected from the flowthrough and from the eluted peaks. More specifically, the eluate fractions of 1.5mL were collected during the run in 2mL Eppendorf microtubes with a fraction collector. The membrane was washed with 5mL of MilliQ water between runs. Lastly, the fractions were kept at 4ºC until the concentration, desalinization, and diafiltration steps in the next day.

4.1.9 Concentration, desalinization and diafiltration

Selected fractions obtained from procedure 4.1.8 were combined and then concentrated and desalted with 10mM Tris–HCl, pH=8.0 in a 15mL Vivaspin® Turbo 15, with a membrane cut-off of 30kDa. The combined fractions (pDNA-containing solutions) in the centrifugal ultrafiltration device were concentrated by centrifugation in a swing bucket rotor for 10 minutes at 3,000xg and the permeates were discharged.

Then, to desalt, a stepwise diafiltration was implemented: 15mL of 10mM Tris–HCl pH=8.0 of diafiltration buffer were added to the retentate solution and then centrifuged always using the same settings, always discharging the permeate. This step was repeated twice. The concentrated, desalted and purified pDNA solution was recovered from the retentate of the ultrafiltration unit with a micropipette and stored in an Eppendorf microtube. Subsequently, the pDNA concentration was measured in a NanoVue Plus (GE Healthcare) spectrophotometer and the quality of the purified pDNA was evaluated in an agarose gel electrophoresis. The purified pDNA was stored at 4ºC for further use (characterization and complexation in nanoparticles). The sample shouldn´t be stored at 4ºC for more than 12 months.

4.2 Agarose gel electrophoresis

A 1% (w/v) agarose gel (Fisher Scientific, Agarose (Low-EEO/Multi-Purpose/Molecular Biology Grade), Fisher BioReagents) was prepared, and it was placed in an electrophoretic chamber filled with Tris-acetate-EDTA (TEA). Then, all the nucleic acid samples were mixed with loading buffer and loaded in the gel as well as a molecular weight marker NZYDNA Ladder III (NZYTech). The electrophoretic conditions for this experiment were 100V, 400mA, and 1 hour runs.

Next, to visualize the bands using a UV gel documentation system, it was necessary to dip the gel to an ethidium bromide solution (0.4µg/mL) for 30 min. The gel was visualized on Eagle Eye II. If the gel is heavily stained, the excess can be removed with water.

4.3 Synthesis of chitosan-TPP nanoparticles

Long-chain (**LC**, MW = 100,000-300,000 from Acros Organics) and short-chain chitosan (**SC**, MW = 60,000-120,000 from Sigma Aldrich) were tested in different concentrations to find the concentration range where the best size distributions were formed. To achieve this, different chitosan and sodium tripolyphosphate (TPP) volumes and concentrations were tested.

Firstly, chitosan was dissolved overnight at 50ºC in a 0.5% acetic acid solution with constant stirring. The concentration used was 2, or 3mg/mL. Regarding the sodium tripolyphosphate solution (TPP, 85% pure from Acros Organics), it was prepared in MilliQ-H_2O at a concentration of 0.75, 1 or 2mg/mL. All the liquids used (the acetic acid solution and MilliQ-H_2O) were previously filtered with a 0.45µm filter.

At room temperature, the TPP solution was added dropwise under constant mild stirring over the chitosan solution. Then, the suspension was left stirring for 30 min.

To avoid aggregation and sedimentation phenomena, the particle characterization was assessed on the same day of particle formation. The particles were stored at 4ºC until particle characterization and subsequent experiments.

4.4 Characterization of nanoparticles

To characterize the nanoparticles, the Z-average hydrodynamic diameter (size), zeta-potential (surface charge), and polydispersity index (size distribution) are measured using Dynamic Light Scattering (DLS) and Zetasizer. All the measurements were performed at room temperature, using filtered solutions (to not interfere with the measurements), and all the formulations were replicated twice. The conditions were the same every time the process was executed. Moreover, it was not necessary to dilute the samples since they were not too concentrated, according to the instrument instructions.

4.5 Encapsulation of plasmid DNA

To encapsulate the nucleic acids, both suspensions of chitosan-TPP and pDNA were preheated separately at 55ºC. Then, an equal volume of both suspensions was quickly mixed together and vortexed at 2,500rpm for 30 seconds. Finally, the suspension was left to incubate for 30 min to stabilize the polyplexes.

To encapsulate the pDNA into the desired nanoparticles, the mixed and vortexed solutions were further homogenized by sonication while in an ice bath during 5min at 5W and using pulses of 5s followed by 10s between pulses with a Sonifier Sonoplus (Bandelin). This extra procedure is fundamental to guaranty small and narrow size distribution of these nanoparticles.

The speed-vac™ equipment was afterward used to concentrate the polyplexes by approximately 7-fold. Centrifugal vacuum concentration is a unique method used for removing solvents from samples to concentrate or dry biological and non-biological materials, residues, solutes, and analytes, for qualitative and quantitative analysis. It combines centrifugation, vacuum, and heat to efficiently evaporate a broad range of solvents. [92]

4.6 Fluorescence

The total pDNA amount was quantified using the Quant-iT™Picogreen™ ds-DNA Assay Kit (Invitrogen, USA). The assay was performed according to the manufacturer's instructions in a white 96-well plate (Grinder), and the samples were diluted in 1:2, 1:4, and 1:8 in PBS buffer. The fluorescence was measured at 480/520 nm using Varian Eclipse. The fluorophore was added before the encapsulation process.

The encapsulation efficacy refers to the amount of plasmid encapsulated into the nanoparticles when compared to the amount used in the encapsulation process. Fundamentally, the amount of plasmid that was not encapsulated and remained in the supernatant is measured.

4.7 Bioinformatic analysis

The amino acid sequences of OmpK, OmpW, OmpV and, OmpU from *Vibrio parahaemolyticus* were obtained from the National Center for Biotechnology Information (NCBI) databases. Those respective sequences were inserted in FASTA format, in a free-to-use portal, and the Immune Epitope Database (IEDB) Analysis Resource (http://tools.immuneepitope.org/bcell/), was used for this research, for which it was only required to choose a method (epitope, accessibility, flexibility, antigenicity, and hydrophilicity prediction).

All the required parameters, such as the threshold value, were predefined according to the software default settings. For the epitope prediction, two methods were used: BepiPred Linear Epitope Prediction and BepiPred Linear Epitope Prediction 2.0. Regarding the other analysis of prospective antigens, the chosen method was the Emini Surface Accessibility Prediction, the Karplus & Schulz Flexibility Prediction, the Kolaskar & Tongaonkar Antigenicity and, lastly, the Parker Hydrophilicity Prediction, for accessibility, flexibility, and hydrophilicity, respectively.

Additionally, we used Pred-TMBB software (http://bioinformatics.biol.uoa.gr/PRED-TMBB/input.jsp) to examine the 2D protein representation and evaluate the regions that are predicted to interact with the antigen-presenting cell exterior. It is expected that the obtained results with the IEDB software are in agreement with the ones obtained with the Pred-TMBB. The latter is based on the Hidden Markov Model, like the IEDB analysis, and predicts the transmembrane beta-strands of gram-negative bacteria outer membrane proteins. It can also discriminate these proteins from other water-soluble proteins.

4.7.1 Immune Epitope Database (IEDB) Analysis

There are different parameters such as hydrophilicity, flexibility, accessibility, turns, exposed surface, polarity and the antigenic propensity of polypeptides chains that can be associated with the position of continuous epitopes. These findings allow the search of positions of continuous epitopes from certain features of the protein sequence. The calculations are done based on propensity scales for each of the 20 amino acids. For each scale, there are 20 different values that are assigned to each of the amino acid residues. The values are assigned depending on their propensity to have the property described by the scale. In IEDB Analysis, the Y-axis represents, for each residue, the correspondent score (Bepipred score) and the X-axis represents the residue positions in the sequence. The larger the score, the higher the probability to be part of the epitope. Those residues are colored in yellow on the graphs. Nevertheless, these methods will not predict the epitopes *per se*. They only guide the researchers to search and investigate the protein regions on being valid epitopes. [123]

4.7.1.1 Bepipred Linear Epitope Prediction

Bepipred predicts the position of linear B-cell epitopes. B-cell epitopes are the sites of molecules that are recognized by antibodies of the immune system and understanding these epitopes can be fundamental to design and developing new vaccines. Bepipred uses the combination of the hidden Markov model and a propensity scale method because, according to *Erik J. et al*, the hidden Markov model is the best single method for predicting linear B-cell epitopes. Therefore, Bepipred was developed by joining this model with one of the best propensity scale methods. Each residue that is scored above the threshold, with a default value of 0.35, is predicted to be part of an epitope. Those residues are represented with the color yellow in the generated graphs where the Y-axes represent residue scores and X-axes represent residue positions in the sequence. There is also a relationship between the threshold and the sensitivity and specificity. For a score value of 0.35, the sensitivity value is 0.49, and specificity is 0.75. Those values are calculated based on the epitope/ non-epitope predictions and are based on a large benchmark calculation containing close to 85 B cell epitopes. [123,124]

4.7.1.2 Bepipred Linear Epitope Prediction 2.0

Bepipred 2.0 epitope prediction works very similar when comparing with Bepipred. Both will predict B-cell epitopes from a protein sequence, but Bepipred uses a Random Forest algorithm, trained on epitopes and non-epitope amino acids, obtained from crystal structures. B-cell epitopes can be divided into two groups: linear epitopes, formed by linear stretches of residues in the antigen protein sequence and discontinuous epitopes, that are formed by residues far apart in the antigen sequence that are brought together in space by its folding. The majority of epitopes are discontinuous but most of those sequences also contain few linear stretches. While Bepipred 1.0 used data from linear peptides to predict antibody recognition, Bepipred 2.0 used crystallography derived structural epitope data for training and evaluation. With Bepipred 2.0, the sequential prediction is also smoothed afterward. Likewise, the residues are assigned with scores and the threshold is 0.5. Again, if there is a probability of the residue belonging to the epitope, they are going to be colored in yellow on the graph. For a value of 0.5, the sensitivity value is 0.586 , and specificity is 0.572. Currently, these epitope prediction tools are essential only to filter and discard regions that are unlikely to have epitopes. Increasing accuracy and specificity possibly will allow for precise and targeted sequences. [123,125]

4.7.1.3 Emini Surface Accessibility Prediction

Emini surface accessibility prediction tool is based on Emini's surface accessibility scale. The accessibility profile was obtained using the formula:

$$S_n = (_{i-1}\ \pi^6\delta_n + 4 + i\ )\ (0.37)^{-6} \qquad \textbf{(Equation 6)}$$

This formula indicates that a surface probability (S) at sequence position n can be defined according to the above **Equation 6**. So, S_n is the surface probability, δ_n represents the fractional surface probability value for the amino acid at position n, and *i* can vary from 1 to 6. So, every hexapeptide sequence with a surface probability value of Sn=1 or higher is probably located on the cell surface.[123,126,127]

4.7.1.4 Kolaskar & Tongaonkar Antigenicity Prediction

Kolaskar and Tongaonkar antigenicity prediction is a semi-empirical method that uses the physicochemical properties of amino acid residues and also their frequencies of occurrence in experimentally known segmental epitopes. This method is capable of predicting antigenic elements on proteins. According to *Kolaskar et al.*, the hydrophobic residues cysteine, leucine, and valine are more prone to be a part of antigenic sites if they are on the surface of a protein. This was based on the analysis of data from experimentally determined antigenic sites on proteins. [123,128]

4.7.1.5 Karplus & Schulz Flexibility Prediction

This flexibility scale is based on the mobility of protein portions, using the known temperature B-factors of the α-carbons of 31 proteins with identified structures. Those 31 proteins that were selected had been refined with atomic temperature factors (B values), had more than 30 residues and their resolution been equal or greater than 0.3 nm. They were also 50% different in sequence when compared with all other included proteins.[123,129]

The B-factor is used in protein crystallography to explain the attenuation of X-ray or neutron scattering created by thermal motion. Greater B-factor values probably will indicate higher flexibility, when compared with the average flexibility. Thus, low B-factor values are found in more rigid regions.[130]

4.7.1.6 Parker Hydrophilicity Prediction

Regarding the hydrophilicity, this parameter was measured based on peptide retention times in High-Performance Liquid Chromatography (HPLC), using a reversed-phase column. The hydrophilicity scale was derived from the contribution of each amino acid side chain to the retention time of model synthetic peptides Ac-Gly-X-X-$(Leu)_3$-$(Lys)_2$-amide. Here, X corresponds to the 20 amino acids found in proteins. This scale can help us to predict the surface sites that correlate well with the antigenic regions for some proteins. To be more precise, this method uses 2 more different parameters: Janin´s scale accessibility and the flexibility of Karplus and Schultz.[123,131]

4.7.1.7 AGGRESCAN – Aggregation Prediction

AGGRESCAN is a web-based software that can predict the aggregation-prone segments in protein sequences. Fundamentally, AGGRESCAN is based on the aggregation-propensity scale for natural amino acids resulting from *in vivo* tests. Also, the algorithm considers that short and specific sequence stretches modulate protein aggregation and can identify the protein fragments that are involved in the aggregation, called hotspots.

The hotspot threshold is the average of the result of multiplying the amino acid aggregation-propensity value (a^3v, which stands for Amino Acid Aggregation-propensity Value) of each of the 20 natural amino acids by its frequency in the SwissProt database, which is currently -0.02. The hotspot regions have 5 or more residues on sequence with an a^4v (which stands for Amino Acid Aggregation-propensity Value (a^3v) window average) larger than the hotspot threshold and no proline, which is an aggregation breaker. [132]

4.7.1.8 Pred-TMBB software

The Pred-TMBB software was used to examine the 2D protein representation and evaluate the regions that are predicted to interact with the outside environment of the cell. Usually, the results obtained with the IEDB analysis are in agreement with the ones obtained with the Pred-TMBB. This method is based on the Hidden Markov Model, like the IEDB analysis, that predicts the transmembrane beta-strands of gram-negative bacteria outer membrane proteins. It can also discriminate these proteins from other water-soluble proteins. All the parameters are updated on a regular basis whenever new crystallographically solved structure become available.[133]

5. Results and Discussion

5.1. pDNA production and purification

After the transformation and the production of plasmids on a larger scale, the plasmids pVAX eGFP, pVAX ompK, and pVAX ompK-frag (**Figure 12** and **Annexes 1,2 and 3**) were purified from transformed *E. coli* DH5α cells by hydrophobic interaction phenyl membrane chromatography as described previously in materials and methods section. In order to choose the eluted fractions containing the supercoiled isoforms without traces of RNA, each fraction peak was evaluated through agarose gel electrophoresis.

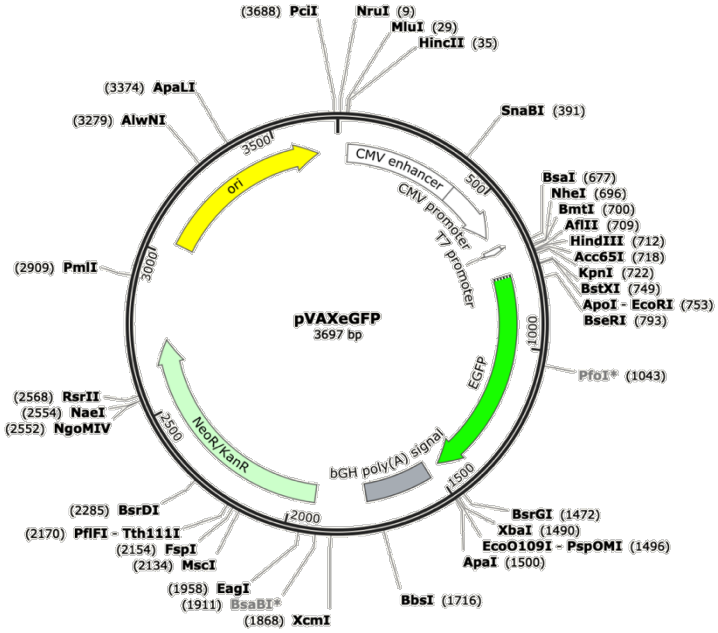

Figure 12 - *pVAXeGFP map and features. This image was obtained using SnapGene.*

Regarding the stability of the plasmids, the transformation of a plasmid vector containing exogenous genes into *E. coli* will bring a series of physiological burdens affecting plasmid stability. Plasmid instability is usually originated from either structural instability caused by changes in the plasmid itself, rearrangements in the plasmid DNA or segregational instability caused by defective partitioning of plasmids between the daughter cells during cell division. Therefore, plasmid instability is a key issue in plasmid DNA production processes, since it usually denotes a considerable decrease in plasmid DNA productivity, leading to poor economics of the whole process.[93]

Any plasmid preparation will contain different topologies of pDNA: supercoiled, open circular, relaxed and linear isoforms. The probability of chromosomal integration will increase if the introduced pDNA has been linearized, which can increase the risk of insertional mutagenesis or the spreading of antibiotic resistance genes. This is why regulatory agencies require a high percentage of supercoiled material in the plasmid preparation intended for vaccination or gene therapy.[47]

The gel electrophoreses were to characterize the size and, indirectly, the structure of the plasmids. The gel lanes with the samples from the peaks of flowthrough, from the eluted fractions and RNA, are shown in **Figure 13**. The chosen fractions, for each of the target plasmids, were pooled together to be desalted and concentrated as previously described above. The fractions number 5 present in Figure 13 (A, B, and C) were chosen to perform the encapsulation.

Regarding the production yield, 67.4µg of pVAX eGFP were produced per gram of cell dry weight (with a final concentration of 526.5ng/µL, a final volume of 200µL, and a total mass of 105.3µg). For pVAX ompK, 7.31µg of plasmid was produced per gram of cell dry weight (with a final concentration of 56ng/µL, a final volume of 200µL, and a total mass of 11.2µg) and for pVAX ompK-frag, 23.5µg of plasmid was produced per gram of cell dry weight (with a final concentration of 175ng/µL, a final volume of 200µL and a total mass of 35.0µg). For this calculation (*E. coli*), the value for the conversion between optical density (OD) and cell concentration was 0.4 (g/L)/OD_{600}.[94] The final OD_{600} for pVAX eGFP, pVAX ompK, and pVAX ompK-frag was, respectively, 3.91, 3.83, and 3.72.

Since the size is different between the three plasmids (3697bp for pVAX eGFP, 4442bp for pVAX ompK, and 3913bp for pVAX ompK-frag), it was expected to obtain different production yields since plasmid size can affect the performance of the *E. coli* host strain when attempting to obtain high-quality plasmid preparations.[95] Larger plasmids are always present in lower numbers in their bacterial host. Otherwise, the metabolic burden of maintaining and duplicating their genomes would be excessive.[96]

We were expecting these results, although other parameters, such as cell growth or DNA forms, could also affect the transfection.[97] There previous results are summarized in **Table 3**.

Table 3 - *Obtained results regarding the production of three different plasmids: pVAX eGFP, pVAX OmpK-frag and pVAX OmpK.*

Plasmids	Mass (ng)	Volume(µL)	Concentration (ng/µL)	Yield (µg/ gram of cell dry weight)	OD_{600}	Size (bp)
pVAX GFP	**105.3**	**200**	**526.5**	**67.4**	**3.91**	**3697**
pVAX OmpK-frag	**35.0**	**200**	**175**	**23.5**	**3.83**	**3913**
pVAX OmpK	**11.2**	**200**	**56**	**7.31**	**3.72**	**4442**

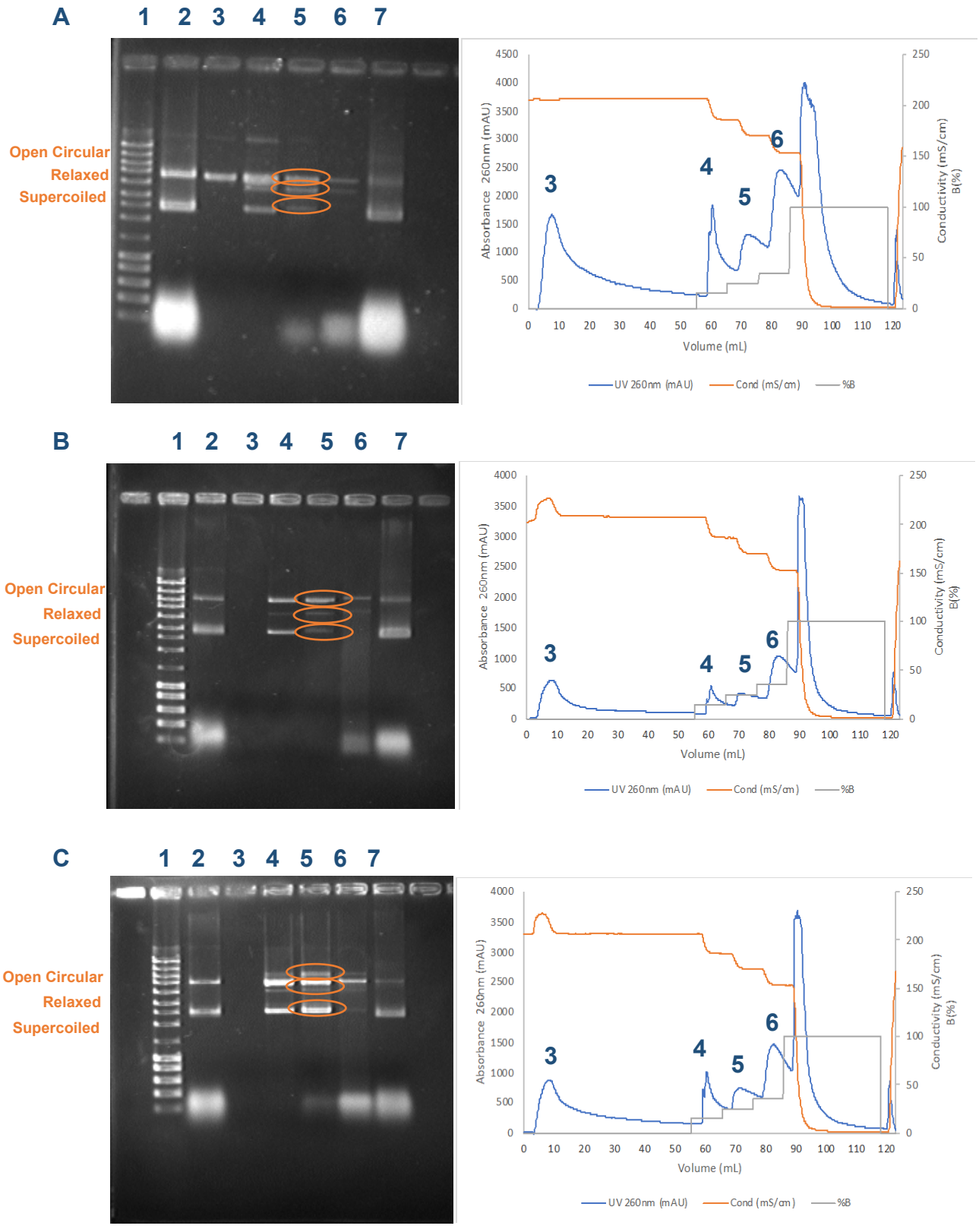

Figure 13 - *Agarose gel analysis of the separated fractions of pVAX eGFP (A), pVAX OmpK (B) and pVAX OmpK-frag (C). Lanes: 1 - Molecular weight marker NZYDNA Ladder III (NZYTech); 2 - Feed solution; 3 - Flowthrough; 4,5,6 - Fractions eluted in each chromatographic peak; 7 – RNA; and Fractionation of pVAX eGFP(A), pVAX OmpK(B) and pVAX OmpK-frag (C) solutions by stepwise elution on a Hydrophobic Interaction Chromatography. The chromatograms are showing peaks at 15, 25, 35 and 100% Elution Buffer-steps.*

5.2. Nanoparticles Characterization

5.2.1 Chitosan-TPP Nanoparticles

To study the stability of the nanoparticles in suspension before the complexation with pDNA is essential to analyze some parameters such as size (Z- average hydrodynamic diameter), size distribution (PdI, polydispersity index), and surface charge (zeta potential).

Different sodium tripolyphosphate (TPP) volumes and concentrations were tested to find the best conditions for the formation of stable nanoparticles (**Figure 14**). It's important to obtain not only stable nanoparticles but also with a unimodal and narrow size distribution, meaning that only one peak can be visualized in a size-distribution plot. Both long and short-chain chitosan (LC and SC) were used in this experiment. The chitosan stock solution was made with a concentration of 2 or 3 mg/mL and pH 3.5. The pH was measured only at the beginning of the process. The TPP stock solution was prepared with a concentration of 0.75, 1, or 2 mg/mL. The Z-average, polydispersity index (PdI), and zeta potential of nanoparticles from all the colloidal solutions are presented in **Table 4** and **Table 5**.

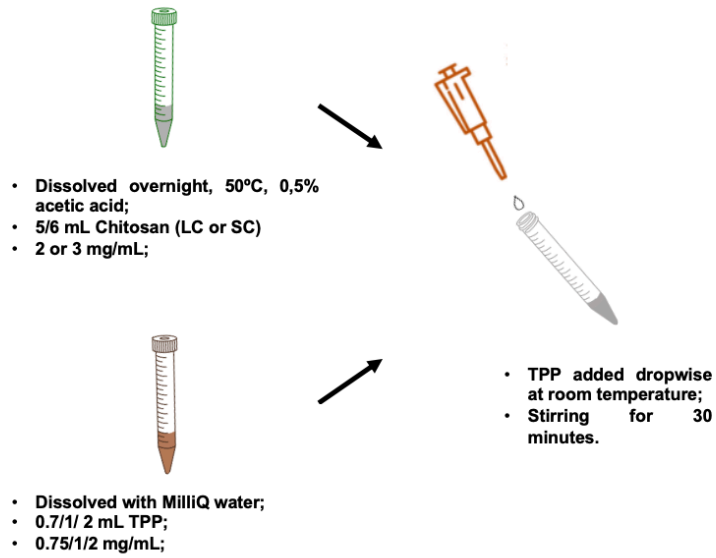

Figure 14 - *Illustration regarding the synthesis of chitosan-TPP nanoparticles.*

- **Long-Chain Chitosan**

In the first experience, with these conditions, it was possible to obtain nanoparticles with Z-average hydrodynamic diameters ranging from 174 and 610nm.

According to the **Table 4**, all of the formulations can be considered nanoparticles since nanoparticles are solid, colloidal particles with a size range from 10 nm to 1,000 nm, even though for nanomedical applications, the preferential size is less than 200 nm.[63]

The standard deviation values for long-chain chitosan were low, meaning that the measured values tend to be close to the expected value. Nevertheless, a perfect monomodal formulation

would have a zeta deviation of zero. The conductivity values were between 0.713 e 0.724 mS/cm and, the effective voltage was 148.8V.

Table 4 - *Z-average, PdI and zeta potential values of chitosan-TPP nanoparticles obtained using long-chain chitosan of MW= 100,000-300,000Da.*

Sample ID	Sample	Chitosan Concentration (mg/mL)	TPP (100,000-300,000Da) Concentration (mg/mL)	Z-average (d.nm)	PdI	Zeta Potential (mV)
LC1	5mL chitosan (3mg/mL, pH= 3.5) + 2mL TPP (1mg/mL)	2.14	0.286	244	0.321	36 ± 4
LC2	5mL chitosan (3mg/mL, pH= 3.5) + 2mL TPP (0,75mg/mL)	2.14	0.214	200	0.285	40 ± 4
LC3	5mL chitosan (3mg/mL, pH= 3.5) + 0,7 mL TPP (1mg/mL)	2.63	0.123	234	0.275	43 ± 4
LC4	5mL chitosan (3mg/mL, pH= 3.5) + 1mL TPP (1mg/mL)	2.50	0.167	210	0.252	39 ± 4
LC5	6mL chitosan (3mg/mL, pH= 3.5) + 2mL TPP (1mg/mL)	2.25	0.250	336	0.409	38 ± 4
LC6	5mL chitosan (2mg/mL, pH=3.5) + 2mL TPP (1mg/mL)	1.43	0.286	209	0.377	36 ± 3
LC7	5mL chitosan (2mg/mL, pH= 3.5) + 2mL TPP (0,75 mg/mL)	1.43	0.214	197	0.334	29 ± 3
LC8	5mL chitosan (2mg/mL, pH= 3.5) + 0,7 mL TPP (1mg/mL)	1.75	0.123	174	0.206	45 ± 4
LC9	5mL chitosan (2mg/mL, pH= 3.5) + 1mL TPP (1 mg/mL)	1.67	0.167	185	0.249	38 ± 4
LC10	6mL chitosan (2mg/mL, pH= 3.5) + 2mL TPP (1mg/mL)	1.50	0.250	267	0.544	47 ± 4
LC11	5mL chitosan (2mg/mL, pH= 3.5) + 2mL TPP (2 mg/mL)	1.43	0.571	610	0.830	39 ± 4

All the PdI values were quite good, ranging from 0.206 to 0.409, except for LC10 (0.544) and LC11 (0.830). This means that almost all suspensions have great homogeneous particle sizes. This index is dimensionless and such values smaller than 0.5 are mainly seen with a highly monodisperse standard. Samples showing PdI values greater than 0.7 have a very broad particle size distribution and are unsuitable for dynamic light scattering technique. Therefore, the formulations tested with long-chain chitosan (**Table 4**) led mainly to small-sized nanoparticles with narrow size distributions.

Regarding the zeta potential, all values measured were positive, meaning that the surface charge of the nanoparticles can interact not only with pDNA but also with the surface of the cells since both have a negative charge. This interaction will allow the loading of the pDNA into the cells.

With this, the eleven nanoparticle suspensions are good candidates, except for LC11, since they have good Z-average hydrodynamic diameters and PdI values. Even though, the LC2 formulation was one of the best chitosan-TPP nanoparticles because it showed one of the best

diameters (200nm), joining that to a good PdI value of 0.285 and a positive zeta potential of 40 mV. Because of this, LC2 was chosen to do the complexation assay. The LC7, LC8, and LC9 formulations were also an excellent choice, meaning that the polyplexes were very similar to the LC2 one.

LC1, LC2, LC3, LC4, LC5, and LC6 were made with a higher chitosan concentration (3mg/mL) while LC7, LC8, LC9, LC10, and LC11 were produced with a smaller concentration, 2 mg/mL. Additionally, the LC5 formulation had the same concentration as the first ones (3mg/mL), but instead of 5 mL, we used 6mL.

When comparing LC1 and LC6 (same formulating, only varying the chitosan concentration), we were expecting to obtain a higher diameter in LC1 because when the chitosan concentration is increased at a constant TPP concentration, the size of nanoparticles also increases.[98] This occurrence was observable not only with LC1 and LC6 but also with LC2 and LC7, LC3 and LC8 and LC4 and LC10, where the same condition was analyzed.

It should be noted that other factors that are not very well controlled, such as temperature or their interaction effects, can affect the average diameter of the particles. Despite that, the strongest factor determining the Z-average hydrodynamic diameter of the particles is, in fact, chitosan concentration. [73] LC11 formulation was not adequate because the diameter is too large. This can be justified since the TPP concentration was too high. [99]

- **Short-Chain Chitosan**

A second experiment using short-chain chitosan polymer (MW= 60,000-120,000 Da) was also done. Again, different volumes and concentrations of stock solutions were tested, using chitosan solutions at concentrations of 2 or 3 mg/mL and pH 3.50 and TPP solutions at concentrations of 0.75, 1, or 2 mg/mL. The same conditions used for long-chain chitosan were used in this experiment.

Table 5 summarizes the mean diameter, PdI, and zeta potential for the measured samples of polyplexes prepared with the chitosan of short length polymer chains.

Regarding the size, these nanoparticles were all larger than the ones with the long-chain chitosan. The polydispersity index and zeta potential were also higher in these nanoparticles when compared with long-chain chitosan.

To reduce their diameter, after the formation of the complexes, one option is sonication. Sonication can reduce the average diameters, decreasing the aggregation, so all formulations were sonicated.

Here, all zeta potential values measured were also positive, but higher when compared with long-chain chitosan. The positive values indicate, again, that nanoparticles can interact with pDNA and with the surface of the cells (both negatively charged), allowing the loading of the pDNA into the cells.

Table 5 - *Z-average, PdI and zeta potential values of chitosan-TPP nanoparticles obtained using short-chain chitosan of MW= 60,000-120,000Da.*

Sample ID	Sample	Chitosan Concentration (mg/mL)	TPP (60,000-120,000Da) Concentration (mg/mL)	Z-average (d.nm)	PdI	Zeta Potential (mV)
SC1	5mL chitosan (3mg/mL, pH= 3.5) + 2mL TPP (1mg/mL)	2.14	0.286	416	0.559	50 ± 15
SC2	5mL chitosan (3mg/mL, pH= 3.5) + 2mL TPP (0.75mg/mL)	2.14	0.214	400	0.600	51 ± 7
SC3	5mL chitosan (3mg/mL, pH= 3.5) + 0,7 mL TPP (1mg/mL)	2.63	0.123	885	0.604	60 ± 9
SC4	5mL chitosan (3mg/mL, pH= 3.5) + 1mL TPP (1mg/mL)	2.50	0.167	719	0.586	52 ± 3
SC5	6mL chitosan (3mg/mL, pH= 3.5) + 2mL TPP (1mg/mL)	2.25	0.250	522	0.585	53 ± 4
SC6	5mL chitosan (2mg/mL, pH= 3.5) + 2mL TPP (1mg/mL)	1.43	0.286	347	0.487	50 ± 3
SC7	5mL chitosan (2mg/mL, pH= 3.5) + 2mL TPP (0,75 mg/mL)	1.43	0.214	231	0.523	47 ± 4
SC8	5mL chitosan (2mg/mL, pH= 3.5) + 0,7 mL TPP (1mg/mL)	1.75	0.123	338	0.520	50 ± 4
SC9	5mL chitosan (2mg/mL, pH= 3.5) + 1mL TPP (1 mg/mL)	1.67	0.167	425	0.530	50 ± 3
SC10	6mL chitosan (2mg/mL, pH= 3.5) + 2mL TPP (1mg/mL)	1.50	0.250	417	0.564	33 ± 3
SC11	5mL chitosan (2mg/mL, pH= 3.5) + 2mL TPP (2 mg/mL)	1.43	0.571	398	0.630	51± 4

It was found out that for given chitosan to TPP molar ratio, the average hydrodynamic diameter of the particles formed is strongly dependent on the initial chitosan concentration. The degree of acetylation of the chitosan was found to be the second most important factor involved in the system's ability to form particles.[73]

It's also believed that chitosans of higher molecular weight produce larger nanoparticles because this type of chitosan has a lower aqueous solubility. Also, it is acknowledged that the interaction with TPP is not so efficient when compared with lower molecular weight chitosan. Therefore, it would be expected that the low solubility of higher molecular weight chitosan would promote aggregation upon interaction with TPP, increasing the mean particle diameter of nanoparticles.

With this experiment, this behavior was not observed, perhaps because of the high aggregation with the short-chain chitosan (high PdI for short-chain chitosan nanoparticles).[100]

Regarding the short-chain chitosan, the best formulation is perhaps the SC7 with the smallest mean diameter of 231nm (and a PdI of 0.523). The one with the best polydispersity index is the SC6 (PdI=0.487), which led to nanoparticles with a mean diameter of 347nm.

Again, the first five formulations were made with a higher chitosan concentration (3mg/mL) while the other ones were produced with a smaller concentration, 2 mg/mL. The SC5 formulation had the same concentration as the first ones (3mg/mL), but instead of 5 mL, we used 6mL.

As expected, the SC1 formulation also had a higher diameter when compared with SC6 (same formulating, only varying the chitosan concentration). This occurrence was observable not only with SC1 and SC6 but also with SC2 and SC7, SC3 and SC8 and SC4 and SC10, where the same condition was analyzed.

5.3. pDNA Encapsulation

Based on the physicochemical characterization of chitosan-TPP nanoparticles tested previously, LC2 was chosen to encapsulate and complex with pVAX eGFP and pVAX OmpK-frag.

As above mentioned, this method comprises the preheating of the suspension of chitosan-TPP nanoparticles and the pDNA solution to 55°C, separately. Subsequently, an equal volume of heated chitosan-TPP and pDNA solutions was quickly mixed together, centrifuged at 2,500rpm for 30 seconds, and left 30 min at room temperature.

Regarding the plasmid pVAX OmpK-frag, for 1mL of suspension, 500μL of chitosan-TPP nanoparticles were mixed with 90μL of pDNA (175ng/μL) and 410μL of H_2O. Then, 200μL of the initial suspension was used to concentrate the sample, obtaining 3150ng of pDNA.

Considering the pVAX eGFP, for 1mL of suspension, 500μL of chitosan-TPP nanoparticles were mixed with 470μL of H_2O and 30μL of pDNA. Herein, a total of 15.8μg of pDNA (526 ng/μL), with a final concentration of 15.8ng/μL, was added to for the pDNA polyplexes. Only 200μL were concentrated in the speed-vac™, having 3.16μg of pDNA.

After the pDNA encapsulation, the Z-average, PdI, and zeta potential were also measured by DLS in a Zetasizer Nano. Those values for the complexed samples are present in **Table 6.**

Table 6 - *Z-average, PdI and zeta potential of chitosan-TPP nanoparticles complexed with pDNA (pVAX eGFP and pVAX Omp K-frag).*

Sample ID	Plasmid	Z-average (d.nm)	PdI	Zeta Potential (mV)
LC2	pVAX eGFP	171	0.172	41 ± 4
		141	0.132	39 ± 5
	pVAX OmpK-frag	188	0.269	40 ± 6

With the addition of pDNA, the mean particle size was reduced, suggesting that the presence of the additional negatively charged component in the mixture promotes the contraction of the positively charged polymer chain.

Also, it was expected that the neutralization degree of the charged amino groups is significantly improved originating a smaller net charge also responsible for the smaller particle size.[101] Regarding this, the zeta potential was almost the same with or without encapsulation.

5.3.1 Encapsulation Efficacy

5.3.1.1. Agarose Gel

Analyzing the agarose gel obtained after electrophoretic runs of the samples containing nude pDNA or its polyplexes (**Figure 15**), it is perceptible that in the lane where the nanoparticles were loaded, no migration of DNA was observed in the gel. In fact, in lane 3 of gel images A, B, and C, there is always a noticeable band in the gel wells that correspond to the complexes that exhibit no electrophoretic mobility due to the nature of chitosan, within the nanoparticles, which is a positively charged polymer. These positive complexes will not migrate in the gel, as won't the pDNA molecules firmly complexed with them. The complexes are truly encapsulating the pDNA.

In **Figure 15 A**, 6μL were loaded in lane number 2, having 3.16μg of pDNA. Moreover, the same amount of pDNA (3.16μg) was used in the encapsulation process in lane 3. Thus, we can be sure that all pDNA in lane 3 is in fact encapsulated.

For pVAX OmpK-frag (**Figure 15, B**), 18μL of the sample, with 3.15μg of nude pDNA, were loaded in lane number 2. Again, for lane number 3, the same amount of DNA was complexed with the nanoparticles and loaded in the gel.

Regarding **Figure 15 B**, it is possible to see that the results are smeared, mainly in lane 2. Smearing can occur either because those samples are degraded or denatured.

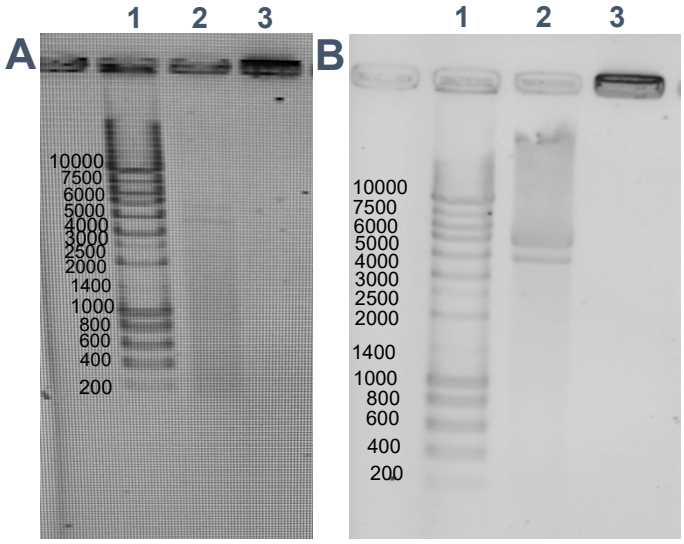

Figure 15 - *1% agarose gel after nucleic acids electrophoresis. Samples are:*
-(A) nude pVAX eGFP (Lane 2) and chitosan-TPP nanoparticles loaded with pVAX eGFP (Lane 3);
-(B) nude pVAX OmpK-frag (Lane 2) and chitosan-TPP nanoparticles loaded with OmpK-frag (Lane 3); LC2 formulation was used in these experiments.
1 – *Molecular weight marker NZYDNA Ladder III (NZYTech);* ***2 –*** *Free (nude) pDNA;* ***3 –*** *Nanoparticles loaded with pVAX eGFP or pVAX OmpK-frag.*

5.3.1.2. Fluorescence

The encapsulation efficacy was further investigated using fluorescence spectroscopy. The efficacy was evaluated by doing the difference between the amount of plasmid encapsulated into the nanoparticles and the one used in the encapsulation process. To achieve this, the amount of plasmid that was not encapsulated (in the supernatant) was measured.[102]

Essentially, the difference between free pDNA and encapsulated DNA is calculated.[103] Regarding the free pDNA and chitosan-TPP-pDNA plate wells, they were loaded with 875ng and 1.58μg of pDNA (**Figure 16**), respectively. Considering the measured fluorescence values in the wells, it was possible to conclude that the free pDNA fluorescence was 76-fold the fluorescence value obtained for chitosan-TPP-pDNA (quotient between the Relative Fluorescence Units (RFU) of those two parameters), despite the actual amount of free pDNA was almost twice the amount of pDNA in the nanoparticles. This means that nearly all pDNA was encapsulated in the nanoparticles since the fluorophore was not detected.

Those values are presented in Relative Fluorescence Units (RFU) because fluorescence intensity is not an absolute measurement. So, mixing the fluorophore with the sample of known concentration and measuring the RFUs will determine the relationship between those two.

We didn't see any difference in the fluorescence numbers when diluting them, meaning that all DNA was already encapsulated, and the fluorophore didn't bind with pDNA (there was not any free DNA). Moreover, the Chitosan-TPP was also measured in order to conclude that this formulation didn't have any fluorescence.

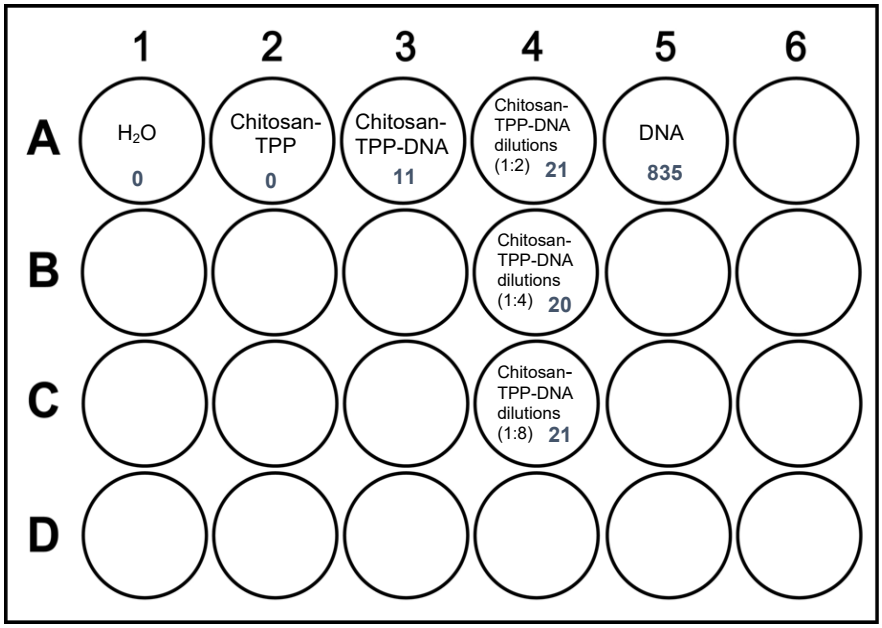

Figure 16 - *Illustration of a plate used in fluorescence-based assays to measure the encapsulation efficacy of LC2 chitosan-TPP nanoparticles. The measured values are presented in Relative Fluorescence Units (RFU).*

6. Bioinformatic Analysis

6.1. *Vibrio parahaemolyticus*

Vibrio parahaemolyticus was the chosen species to analyze because it is a model marine pathogen with a wide distribution in coastal marine waters and estuaries, infecting a broad spectrum of fish and mammals. Since genetically conserved OMPs are present in *Vibrio* pathogens, they are promising vaccine targets that could be used for developing effective subunit vaccines. [10] It is a versatile halophilic organism, adjusting to a wide variety of environments and causing infections to both humans and aquatic animals. The versatility regarding the host and the habitat is caused by the ability to acquire genes that improve the adaptation of the organism in different situations.[104]

V. parahaemolyticus and *V. vulnificus*-associated illnesses have been increasing worldwide recently. Besides human illness, these bacteria are destroying marine life in the coastal environment. [105]

6.2. BLASTN and BLASTP

The Basic Alignment Search Tool (BLAST) website at the National Center for Biotechnology (NCBI) is the main tool for searching and aligning sequences. BLAST uses nucleotides or proteins sequences as input and searches it against a database. Moreover, BLAST, which is one of the more popular bioinformatics tools, provides statistical information about each alignment, making the information more valuable.[106,107]

For this analysis, the strain *Vibrio parahaemolyticus* ATCC 17802 was chosen to perform the searching and aligning sequences.

Even though BLASTN and BLASP analyses were done (**Annex 4**), the results are not the same when comparing both outputs. This can be explained because there are only 20 amino acids available for protein synthesis, but 64 possible codons, creating a variable degree. Furthermore, since there are codon preferences in different organisms (codon usage bias), it is very likely to have a much lower level of sequence identity at the nucleotide level than at the protein level. The redundancy within the genetic code will lead to a variable number of codons to encode the same amino acid. Codon usage bias has been found in many genomes, where a clear preference for a particular codon exists.[108] Therefore, using a protein sequence can be more sensitive for coding regions.

It was desirable to obtain as many strains as possible so that the vaccine could work against the majority of them. Regarding the BLASTN, for OmpK, 10 strains were analyzed: *Vibrio parahaemolyticus*, *Vibrio fluvialis, Vibrio alginolyticus, Vibrio harveyi, Vibrio campbellii, Vibrio owensii, Vibrio vulnificus, Vibrio cholerae* and *Vibrio metoecus.* For OmpW, the strains (7) were *Vibrio parahaemolyticus, Vibrio diabolicus, Vibrio antiquarius, Vibrio neocaledonicus, Vibrio alginolyticus, Vibrio natriegens,* and *Vibrio campbellii.* For OmpV (7), the strains were *Vibrio*

parahaemolyticus, Vibrio diabolicus, Vibrio antiquarius, Vibrio neocaledonicus, Vibrio alginolyticus, Vibrio harveyi, and *Vibrio rotiferianus* and for OmpU (7), they were *Vibrio parahaemolyticus, Vibrio antiquarius, Vibrio diabolicus, Vibrio harveyi, Vibrio alginolyticus, Vibrio rotiferianus,* and *Vibrio jasicida.* With this outcome, the outer membrane protein K seems to be the broader one, since it is present in more *Vibrio* strains.

Regarding the BLASTP, other strains were also predicted such as *Vibrio breoganii* and *Vibrio coralliilyticus*. Several membrane proteins or hypothetical proteins were also predicted using this software.

Since the 4 proteins were predicted to be found in several strains, they seem to be good vaccine candidates against *Vibrio* infections, since they are highly antigenic proteins.

6.3. Peptide Sequence Selection

Predicting immunogenic epitopes using bioinformatics tools and designing them is no easy task. To obtain the right antigenicity, many parameters need to be taken into account, be optimized and analyzed, such as peptides purity levels, amino acid compositions, lengths, hydrophobicity, secondary structures, besides the inherent complexity of antigen recognition.

The most important element in a traditional peptide antibody production is the selection. For the immunization, peptides are usually 10 to 20 amino acids long because peptide sequences above those numbers may produce antibodies that do not recognize the protein with sufficient specificity and peptide sequences below those numbers may not recognize the native protein with sufficient affinity. Nevertheless, longer peptides can provide a good immune response and a relevant secondary structure, having a greater conformational similarity to the native protein. However, there are also short peptides that can also elicit a strong immune response. Thus, sequences of 10 to 20 amino acids are important since they can reduce synthesis problems and are more soluble in aqueous solutions, which is suitable for conjugation with the carriers and use in biological assays. [109,110]

For antibodies to bind to epitopes, epitopes must be found on the surface of proteins. They tend to have higher affinity when those regions are flexible enough to be into accessible positions and also when they are hydrophilic, surface orientated, and flexible. This can be justified because, in natural environments, hydrophilic regions are present on the surface while hydrophobic regions reside inside the protein structure. Nevertheless, hydrophilic regions can also contain hydrophobic regions (e.g. tryptophan, valine, leucine, isoleucine, and phenylalanine). Hydrophobic amino acids have aliphatic chains, hydrocarbons, aromatic side groups, making them insoluble or only slightly soluble in water. Therefore, it's essential to select a peptide sequence with as few of these residues as possible. It's also critical to avoid glutamine because not only may cause insolubility but also can form hydrogen bonds between peptide chains.[110,111]

Proline is a very much common residue in those antigen peptide sequences and there is a long list of known proline-containing epitopes. Usually, proline is present at solvent-exposed sites, such as loops, turns, N-terminal first turn of the helix, and random coils. Thus, the propensity of

proline to appear in these proteins regions can indicate that this residue has immunogenic properties. [110,112]

Choosing peptides that are located in the N- (capped with an acetyl group) or C- (capped with an amide group) terminal region of the protein are frequently more solvent-accessible and unstructured, meaning that antibodies can recognize the native protein easily. Therefore, the peptide will appear more like a native protein.[110]

Aggregation is also an essential factor to consider. In antibodies, the hydrophobic core of the protein tends to aggregate because this region will thermodynamically contribute to the stabilization of the tertiary interactions. [113]

Proteins tend to aggregate under a variety of environmental conditions. The extent of aggregation is dependent on many factors that are classified as intrinsic (primary, secondary, tertiary, or quaternary structure) or extrinsic (protein environment, processing conditions). Protein aggregates will have a reduced or no biological activity and this occurrence is still a major challenge in the development of biotechnology products. Those antibody aggregates can be reversible or irreversible, disturbing the therapeutic effect. [114,115]

Regarding the intrinsic factors, the aggregation propensity can be determined based on the primary sequence and structure. The three-dimensional structure of a native protein in its physiological conditions usually has the lowest Gibbs free energy, being the most stable conformation. Some amino acid sequence motifs are more prone to aggregation, such as sequences enriched hydrophobic or aromatic residues. Otherwise, charged residues such as lysine, arginine, and histidine are rare in those regions.[116,117]

Extrinsic factors refer to the environment that antibodies are exposed to or the process conditions that they go through. Some examples of extrinsic factors are temperature, pH, ionic strength, osmolality, dissolved oxygen, agitation, metals, and medium components. [118]

The above-described information is summarized in the following **Table 7**.

Table 7 - *Important selection factors regarding the antigen peptide sequences.*

Selecting Factors	Examples
Amino acid composition	Hydrophilic amino acids, proline, glutamine
Peptide length	10-20 amino acids
Peptide structure	Linear, flexible, helices
Protein target	Accessible epitope, N/C terminal,

6.4. Signal Peptides

Since eukaryotic cells have their genetic material enclosed within the nucleus, the transcriptional and translational events are physically separated. Thus, there must be a transport across the nuclear envelope to secrete the mRNA to the cytoplasm. Transport across the nuclear envelope occurs through the conserved nuclear pore complex, which is a proteinaceous structure creating an aqueous channel.[119] Subsequently, proteins must also be secreted to the extracellular space, in order to be captured by the immune system's cells.

For transportation to occur, genetic information must be secreted. The secretion will be mediated by these signal peptides, which are commonly between 15 and 30 amino acids in length but can have more than 50 residues. Essentially, the main function of signal peptides is to prompt the cell to translocate the protein to the desire location. The amino acid sequence is variable, but there are some common features such as an N-terminal polar region, a hydrophobic core region, and a C-terminal polar cleavage region. The polar region usually has 2 to 5 amino acids and a positive charge. The hydrophobic region contains between 6 to 15 amino acids and is the most essential part of the signal peptide for targeting.[120] These signal peptides are quickly degraded, but some will continue to have functions on their own.

Since the encoded immunogen expressed in the transfected cells may not enter in the MHCII antigen processing and presentation pathway, it is crucial to overcome this drawback. The secreted protein needs to be phagocyted by the APCs, and then gain entry into the MHCII exogenous pathway. If these proteins don't have a signal peptide, they will stay in the cytosol for the rest of the translation process. One way to solve this problem is to use the secretory signals to target the antigen to the extracellular medium. Thus, it is possible to increase antigen exposure to APCs, improving the effectiveness of DNA vaccines to elicit immune responses. Consequently, the plasmid will be able to gain access to both pathways (MHCI and MHCII). [121]

For instance, the lysosomes could be used to promote the protein sorting to the lysosomes, if a target signal is used. Those organelles are involved in the generation of epitopes that are going to be presented to the immune system when complexed with MHCII molecules.[122]

One example of a secretory sequence used for the development of a DNA vaccine against the Maedi-Visna virus was **MDAMKRGLCCVLLLCGAVFVSAR**.[122] Other antigen-targeting sequences used in DNA vaccine were: Lysosomal-associated membrane protein-1 (LAMP):

5´ -ISG GFP **SSLDPHCCGRCPGRAGPHRPHCLPHWQEEESRRLSDHLE**-3´

where the amino acid sequence that is represented above is going to produce the desired protein and Adenovirus e1a endoplasmic reticulum (e1a) [121]:

5'- **MRYMILGLLALAAVCSA** ISG GFP-3'

In both sequences, ISG denotes for Invariant Surface Glycoprotein, and GFP denotes for Green Fluorescence Protein. Since these regions are usually conserved in eukaryotes, it could be possible to attempt using these sequences in DNA vaccines for fish.

6.5. Outer Membrane Proteins

6.5.1. Outer Membrane Protein K

The obtained results for OmpK and its structure are presented below (**Figure 17**). In Bepipred Linear Epitope Prediction, there are 10 predicted peptides, as shown in **Figure 18** and **Table 8**. Theoretically, only five of them meet the requirements, having more than 10 amino acids and less than 20. Also, it may be possible to combine more than one antigen peptide sequence.

The same approach can be done with other programs. In Bepipred 2.0, Emini Surface Accessibility and Kolaskar & Tongaonkar Antigenicity, there are, respectively, 10, 7, and 6 predicted sequences (**Figure 19, 20 and 21** and **Table 9, 10 and 11**). Only 4 and 3 of those have more than 10 amino acids. In Kolaskar & Tongaonkar Antigenicity, all the predictions meet this requirement. In Parker Hydrophilicity and Karplus & Schulz Flexibility, there are 267 and 266 predicted residue scores, respectively.

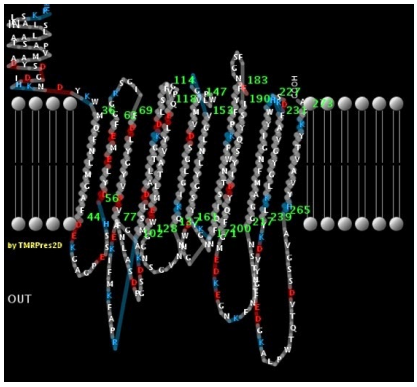

Figure 17 - *2D representation of OmpK Vibrio Parahaemolyticus, predicted by PRED-TMBB.*

- **Bepipred Linear Epitope Prediction**

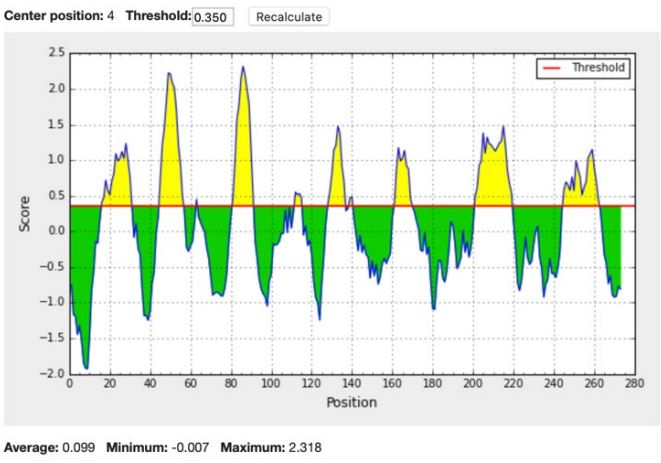

Figure 18 - *Predicted peptides for OmpK, using Bepipred Linear Epitope Prediction.*

Table 8 - *Predicted peptides for OmpK, using Bepipred Linear Epitope Prediction.*

No.	Start	End	Peptide	Length
1	16	30	APVMAADYSDGDIHK	15
2	45	56	DEKGAGPESSHD	12
3	63	63	G	1
4	81	91	ASDPGSDKAGA	11
5	112	115	LSFG	4
6	128	136	WGGNSGVNN	9
7	139	140	IG	2
8	161	169	YYGNNKDWN	9
9	201	219	GMEDKEGNKFNTNTSNGGA	19
10	245	262	GFEDGKALPWTQTVDSSG	18

- **Bepipred Linear Epitope Prediction 2.0**

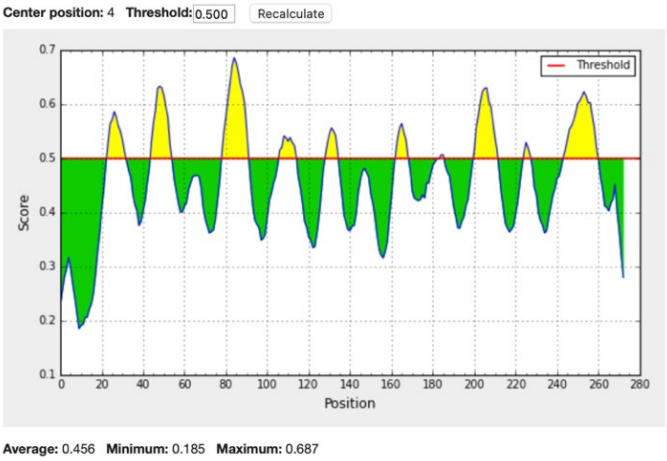

Figure 19 - *Predicted peptides for OmpK, using Bepipred Linear Epitope Prediction 2.0.*

Table 9 - *Predicted peptides for OmpK, using Bepipred Linear Epitope Prediction 2.0.*

No.	Start	End	Peptide	Length
1	24	32	SDGDIHKND	9
2	45	54	DEKGAGPESS	10
3	79	92	NLASDPGSDKAGAE	14
4	107	115	LTGKDLSFG	9
5	129	135	GGNSGVN	7
6	163	169	GNNKDWN	7
7	184	186	FEN	3
8	201	212	GMEDKEGNKFNT	12
9	225	228	YWHS	4
10	244	260	YGFEDGKALPWTQTVDS	17

- **Emini Surface Accessibility Prediction**

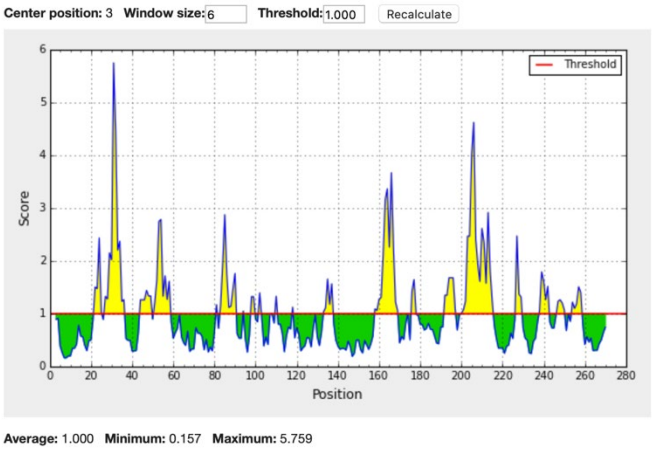

Figure 20 - *Predicted peptides for OmpK, using Emini Surface Accessibility Prediction.*

Table 10 - *Predicted peptides for OmpK, using Emini Surface Accessibility Prediction.*

No.	Start	End	Peptide	Length
1	27	36	DIHKNDYKWM	10
2	44	49	FDEKGA	6
3	51	58	PESSHDYL	8
4	83	90	DPGSDKAG	8
5	158	169	YGSYYGNNKDWN	12
6	192	197	YQSYID	6
7	200	215	FGMEDKEGNKFNTNTS	16

- **Kolaskar & Tongaonkar Antigenicity Prediction**

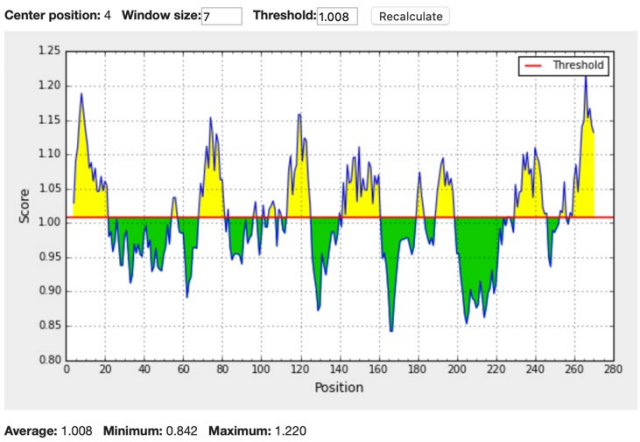

Figure 21 - *Predicted peptides for OmpK, using Kolaskar & Tongaonkar Antigenicity Prediction.*

Table 11 - *Predicted peptides for OmpK, using Kolaskar & Tongaonkar Antigenicity Prediction.*

No.	Start	End	Peptide	Length
1	4	21	SLLALSLLAATSAPVMAA	18
2	68	81	IFDLYGYVDVFNLA	14
3	113	125	SFGPVQELYVATL	13
4	142	160	GSDVMVPWLGKVGLNLYGS	19
5	189	198	FISYQSYIDY	10
6	230	246	RFAVGYGLKLYKDVYGF	17

- **Parker Hydrophilicity Prediction**

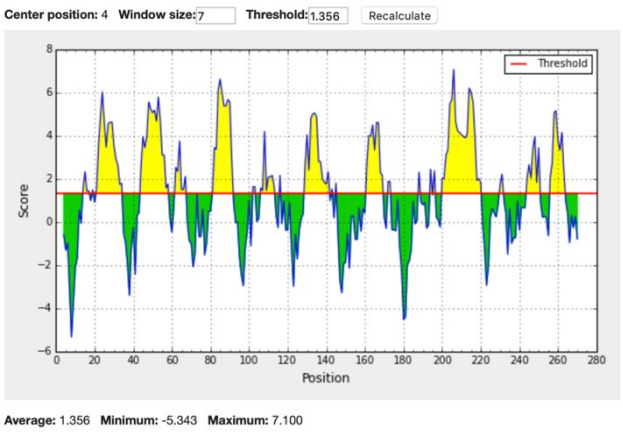

Figure 22 - *Predicted peptide for OmpK, using Parker Hydrophilicity Prediction.*

- **Karplus & Schulz Flexibility Prediction**

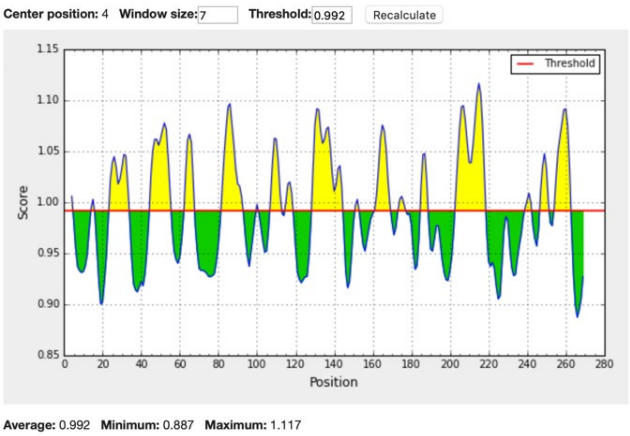

Figure 23 - *Predicted peptide for OmpK, using Karplus & Schulz Flexibility Prediction.*

- **AGGRESCAN – Aggregation Prediction**

Table 12 - *Predicted peptide sequences for OmpK, using AGGRESCAN, a server for the prediction and evaluation of "hot-spots" of aggregation in polypeptides. a^4vAHS stands for: Amino-acid aggregation-propensity value window average (a^4v) average in the Hotspot.*

No.	Start	End	Peptide	Length	a^4vAHS
1	1	16	MRKSLLALSLLAATSA	16	0.333
2	36	42	MQFNLMG	7	0.232
3	64	79	GRSGIFDLYGYVDVFN	16	0.324
4	92	99	EKIFMKFA	8	0.263
5	117	126	VQELYVATLM	10	0.379
6	142	147	GSDVMV	6	0.336
7	149	161	WLGKVGLNLYGSY	13	0.346
8	174	179	STNWFK	6	0.304
9	181	188	FYFFENGS	8	0.422
10	190	199	ISYQSYIDYQ	10	0.283
11	220	225	MFNGIY	6	0.283
12	232	243	AVGYGLKLYKDV	12	0.284
13	262	273	GVAHYVAVTYKF	12	0.420

After analyzing all the outputs from the algorithms, some plausible sequences of OmpK could have the potential to be applied in DNA immunization. From position 139 to 169, this sequence shows good scores regarding the epitope prediction, antigenicity, and accessibility, but has a high probability of aggregation (142 to 147 and 149 to 161, with a score of 0.336 and 0.346, respectively), making this sequence not suitable for the development of DNA vaccination.

Moreover, sequence 230-262 has the same drawback as the previous one, since the probability of aggregation is also high (score of 0.284, **Table 12**).

Finally, the best sequence seems to be from position 200 to 228, although there is a small region with a propensity of aggregation (220-225). Even though this region has a score of 0.283, this sequence has the smallest probability of aggregation, being the lowest value. Unfortunately, in this sequence, there is no proline, which is a very much common residue in those antigen peptide sequences. Usually, the propensity of proline to appear in these proteins' regions can indicate that this residue has immunogenic properties.

Regarding the hydrophilicity (**Figure 22**), the peptide sequence with the highest score (7.1) is EDKEGNK**,** from 203 to 209, and with a flexibility of 1.094. Regarding the Karplus & Schulz Flexibility Prediction **(Figure 23**), the best sequence is TNTSNGG (212 to 218)**,** with a score of 1.117 and the hydrophilicity has a score of 6.043.

Both regions are between positions 200 and 228, making this sequence a good selection. Therefore, the best sequence for OmpK is 200 to 228. In **Table 13**, we collected all the data already presented and summarized all the above information.

On the other hand, it is worth to mention that the OmpK sequence inserted in the plasmid (**Section 5.1**) differs from the one analyzed in here (**Section 6.6.1**), which can be explained by the fact that different model species and the software applied were different.

Table 13 - *Collection of data obtained from the previous software's already described. The sequence 200-228 appears to be the best option considering the Outer Membrane Protein K. Regarding the Parker Hydrophilicity Prediction and Karplus & Schulz Flexibility , we chose the best score individually, and the best one is indicated in bold.*

Peptides	Bepipred Linear Epitope	Bepipred Linear Epitope Prediction 2.0	Emini Surface Accessibility	Kolaskar & Tongaonkar Antigenicity	Parker Hydrophilicity Prediction	Karplus & Schulz Flexibility	AGGRESCAN – Aggregation
139-169	139-140 161-169	163-169	158-169	142-160	163-169(4.629)	162-168 (1.076)	142-147(0.336) 149-161(0.246)
200-228	**201-219**	**201-212** **225-228**	**200-215**	-	**203-209(7.1)**	**212-218(1.117)**	**220-225(0.283)**
230-262	245-262	225-228 244-260	-	230-246	256-262(1.091)	256-262 (1.091)	232-243(0.284)

6.5.2. Outer Membrane Protein W

Regarding the OmpW (**Figure 24**), in BepiPred, Bepipred 2.0, Emini Surface Accessibility and Kolaskar & Tongaonkar Antigenicity, there are, respectively, 12, 5, 6 and 7 predicted sequences (**Figure 25, 26, 27** and **28** and **Table 14, 15, 16, 17**). Only 3, 4, 2 and 3 of those sequences have more than 10 amino acids. In Parker Hydrophilicity and Karplus & Schulz Flexibility, there are 208 and 267 predicted residue scores, respectively.

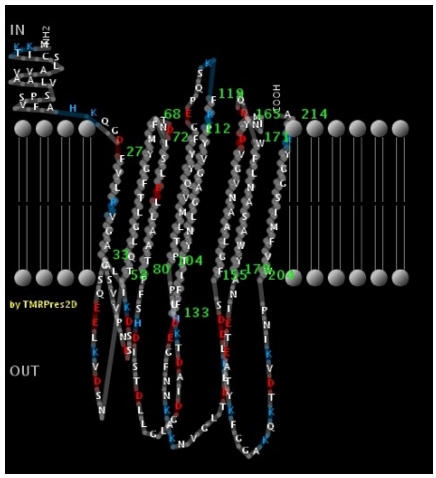

Figure 24 - *2D representation of OmpW Vibrio Parahaemolyticus, predicted by PRED-TMBB.*

- **Bepipred Linear Epitope Prediction**

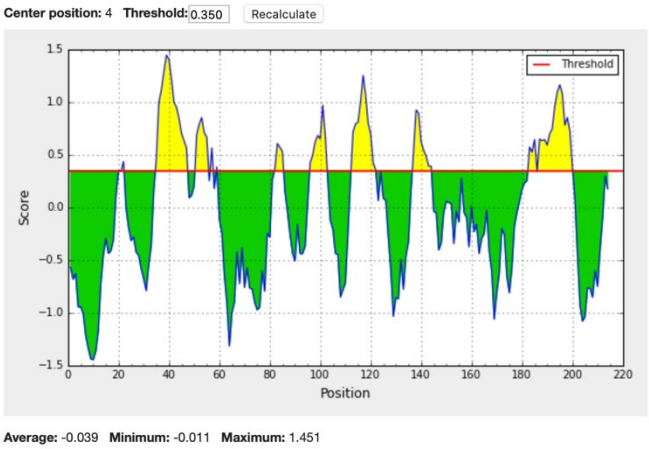

Figure 25 *- Predicted peptides for OmpW, using Bepipred Linear Epitope Prediction.*

Table 14 - *Predicted peptides for OmpK, using Bepipred Linear Epitope Prediction*

No.	Start	End	Peptide	Length
1	20	20	F	1
2	22	22	H	1
3	35	47	SVVPNDSSDKILG	13
4	51	55	ELKVD	5
5	57	57	N	1
6	59	59	Q	1
7	82	85	FSHD	4
8	96	102	IADTKHL	7
9	113	122	GEPQSKFRPY	10
10	137	144	FNNKAKNV	8
11	183	185	ETE	3
12	187	200	TYKFGGAKQKTDVK	14

- **Bepipred Linear Epitope Prediction 2.0**

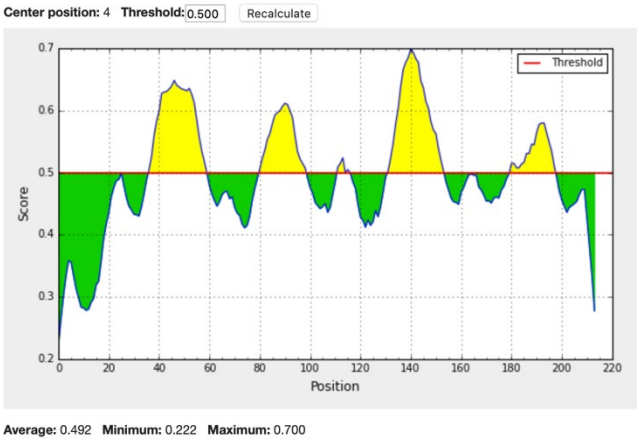

Figure 26 - *Predicted peptides for OmpW, using Bepipred Linear Epitope Prediction 2.0.*

Table 15 - *Predicted peptides for OmpW, using Bepipred Linear Epitope Prediction 2.0.*

No.	Start	End	Peptide	Length
1	37	59	VPNDSSDKILGSQEELKVDSNTQ	23
2	81	99	PFSHDISTDLLGLGDIADT	19
3	112	116	FGEPQ	5
4	132	154	FFDEGFNNKAKNVGLTDLKLDDS	23
5	181	198	NIETEATYKFGGAKQKTD	18

- **Emini Surface Accessibility Prediction**

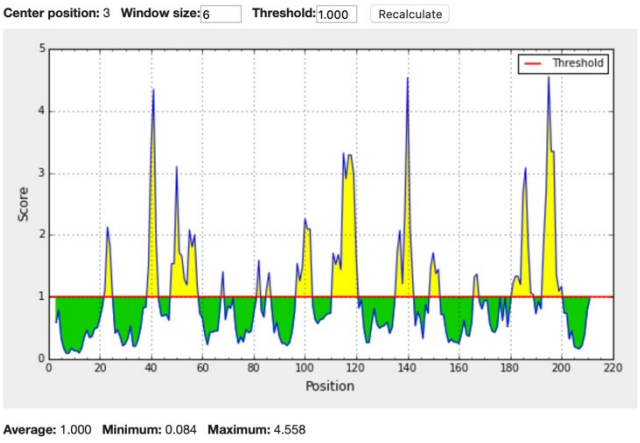

Figure 27 - *Predicted peptides for OmpW, using Emini Surface Accessibility Prediction.*

Table 16 - *Predicted peptides for OmpW, using Emini Surface Accessibility Prediction.*

No.	Start	End	Peptide	Length
1	48	58	SQEELKVDSNT	11
2	97	102	ADTKHL	6
3	111	120	YFGEPQSKFR	10
4	136	142	GFNNKAK	7
5	181	189	NIETEATYK	9
6	193	200	AKQKTDVK	8

- **Kolaskar & Tongaonkar Antigenicity**

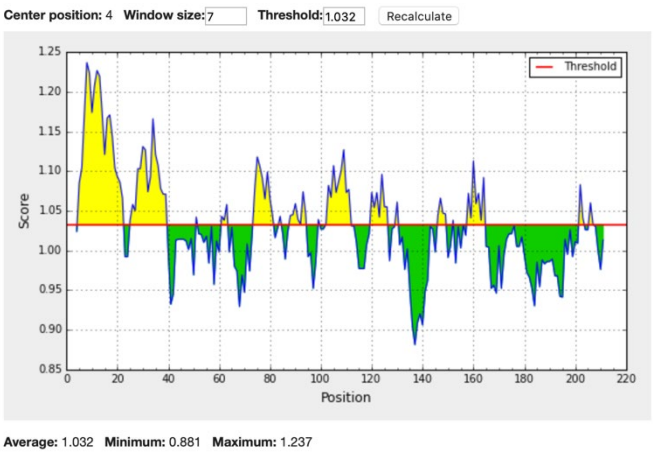

Figure 28 - *Predicted peptides for OmpW, using Kolaskar & Tongaonkar Antigenicity Prediction.*

Table 17 - *Predicted peptides for OmpW, using Kolaskar & Tongaonkar Antigenicity Prediction.*

No.	Start	End	Peptide	Length
1	5	22	ICSLAVVAALVSPSVFAH	18
2	25	39	GDFVLRVGAASVVPN	15
3	74	81	LELLAATP	8
4	88	94	TDLLGLG	7
5	102	111	LPPTLMVQYY	10
6	120	126	RPYVGAG	7
7	158	164	AANVGVD	7

- **Parker Hydrophilicity Prediction**

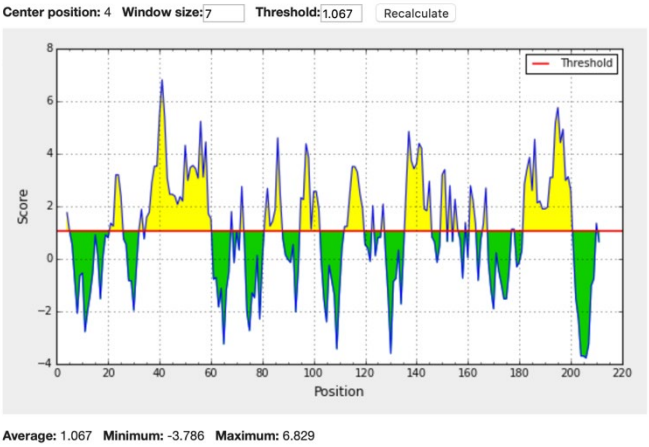

Figure 29 - *Predicted peptide for OmpW, using Parker Hydrophilicity Prediction.*

- **Karplus & Schulz Flexibility Prediction**

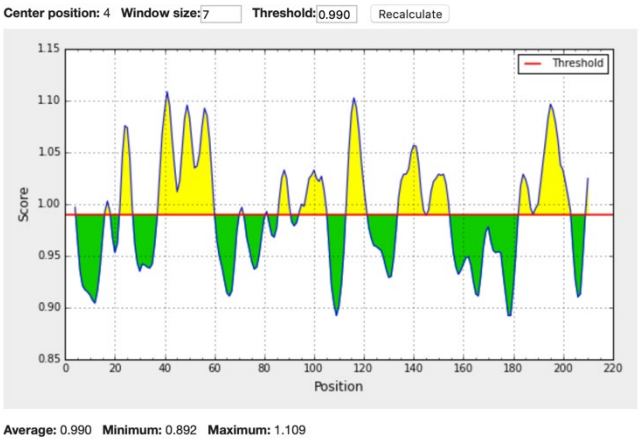

Figure 30 - *Predicted peptide for OmpW, using Karplus & Schulz Flexibility Prediction.*

- **AGGRESCAN – Aggregation Prediction**

Table 18 *- Predicted peptide sequences for OmpW, using AGGRESCAN, a server for the prediction and evaluation of "hotspots" of aggregation in polypeptides. . a^4vAHS stands for: Amino-acid aggregation-propensity value window average (a^4v) average in the Hotspot.*

No.	Start	End	Peptide	Length	a^4vAHS
1	1	16	MKKTICSLAVVAALVS	16	0.477
2	27	35	FVLRVGAAS	9	0.285
3	60	80	LGLTFGYMFTDNISLELLAAT	21	0.304
4	105	112	TLMVQYYF	8	0.440
5	122	135	YVGAGLNYTIFFDE	14	0.298
6	157	165	LAANVGVDY	9	0.215
7	167	171	INDQW	5	0.133
8	174	180	NASAWYA	7	0.221
9	204	214	WVFMISGGYKF	11	0.500

After analyzing the obtained results for OmpW, the best sequence was from amino acid 25 to 59. It is slightly bigger than 20 amino acids, but it presented promising results. There is only a small region that is more prone to aggregation (27-35, with a score of 0.285, **Table 18**) but has the best scores for hydrophilicity and flexibility (6.829 and 1.109, respectively, **Figure 29** and **Figure 30**). Regarding hydrophilicity, the best region is PNDSSDK (position 38 to 44)**,** with a score of 6.829, and regarding the flexibility, it is the same region, with a score of 1.109.

The sequence 74-122 is also suitable, having a higher probability of aggregation from 105 to 112 (score of 0.440). Moreover, the sequences 120-154 and 181-200 are also appropriate, according to the epitope prediction, and accessibility. The antigenicity feature was predicted for the sequence 120-154 (136-142), but not predicted for the 181-200 sequence. Despite that, the first one tends to aggregate from 122 to 135 (score of 0.298), and the last doesn't have any region prone to aggregation.

The selected regions are slightly bigger to attempt to have longer peptides, providing good immune response and a relevant secondary structure, allowing a greater conformational similarity to the native protein.

In **Table 19**, all the information is summarized, pointing out that sequence 25 to 59 is the best option regarding all the alternatives.

Table 19 - *Collection of data obtained from the previous software's already described. The sequence 25-59 appears to be the best option considering the Outer Membrane Protein W. Regarding the Parker Hydrophilicity Prediction and Karplus & Schulz Flexibility , we chose the best score individually, and the best one is indicated in bold.*

Peptides	Bepipred Linear Epitope	Bepipred Linear Epitope Prediction 2.0	Emini Surface Accessibility	Kolaskar & Tongaonkar Antigenicity	Parker Hydrophilicity Prediction	Karplus & Schulz Flexibility	AGGRESCAN – Aggregation
25-59	35-47 51-55 57-57 59-59	37-59	48-58	25-39	38-44(**6.829**)	38-44(**1.109**)	27-35(0.285)
74-122	82-85 96-102 113-122	81-99 112-116	97-102 111-120	74-81 88-94 102-111	83-89(4.614)	113-119(1.103)	105-112(0.440)
120-154	137-144	132-154	136-142	120-126	134-140(4.857)	137-143(1.057)	122-135(0.298)
181-200	187-200	181-198	181-189 193-200	-	191-197(5.157)	192-198(1.097)	-

6.5.3. Outer Membrane Protein V

For outer membrane protein V (**Figure 31**), in BepiPred, Bepipred 2.0, Emini Surface Accessibility, and Kolaskar & Tongaonkar Antigenicity web software's, there are 13, 9, 6 and 10 predicted sequences, respectively. Only 6, 5, 2, and 5 sequences have ten or more amino acids (**Figure 32**, **33**, **34** and, **35** and **Table 20**, **21**, **22** and, **23**). In Parker Hydrophilicity and Karplus & Schulz Flexibility, there are 252 and 251 predicted residue scores, respectively.

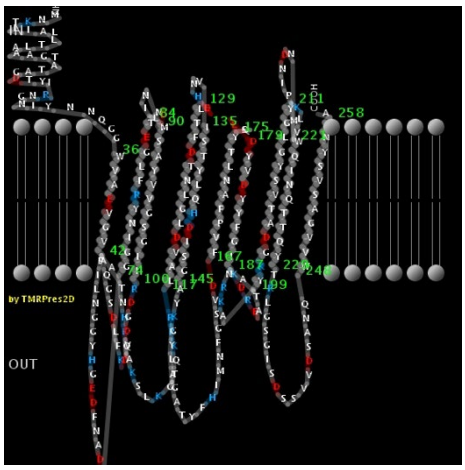

Figure 31 - *2D representation of OmpV Vibrio Parahaemolyticus, predicted by PRED-TMBB.*

- **Bepipred Linear Epitope Prediction**

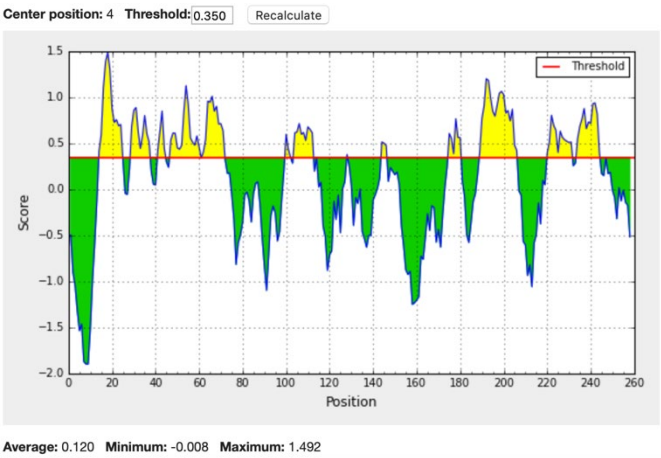

Figure 32 - *Predicted peptides for OmpV, using Bepipred Linear Epitope Prediction.*

Table 20 - *Predicted peptides for OmpV, using Bepipred Linear Epitope Prediction.*

No.	Start	End	Peptide	Length
1	14	24	AGTATAGDTYI	11
2	29	37	IYNNQGGWV	9
3	41	44	GVAQ	4
4	47	60	DLFKDQKHNTAPIL	14
5	62	71	GGYHGEDFNA	10
6	100	102	IRD	3
7	104	112	DVAKSLKGT	9
8	128	128	T	1
9	144	146	SGA	3
10	175	180	YQSEDY	6
11	189	206	DREATANRKAYKGDATVS	18
12	220	231	NWQINQTTQYTR	12
13	234	244	SGISDSSVVDS	11

- **Bepipred Linear Epitope Prediction 2.0**

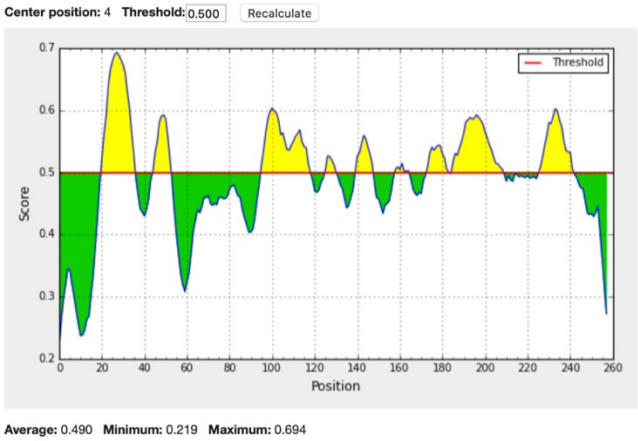

Figure 33 - *Predicted peptides for OmpV, using Bepipred Linear Epitope Prediction 2.0.*

Table 21 - *Predicted peptides for OmpV, using Bepipred Linear Epitope Prediction 2.0.*

No.	Start	End	Peptide	Length
1	21	36	DTYIRNGNIYNNQGGW	16
2	45	53	GSDLFKDQK	9
3	96	119	GSGIIRDGDVAKSLKGTQKRRLAV	24
4	126	131	DFTLDE	6
5	140	148	QHDISGAYK	9
6	159	165	IMNFGSV	7
7	174	184	TYQSEDYVDYY	11
8	186	209	GIKDREATANRKAYKGDATVSYGL	24
9	227	242	TQYTRLGSGISDSSVV	16

- **Emini Surface Accessibility Prediction**

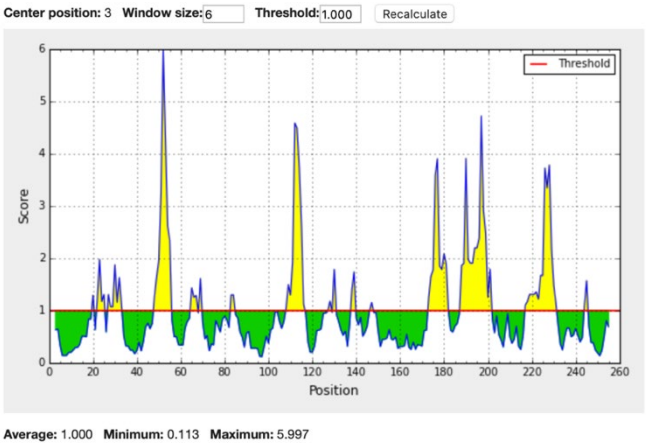

Figure 34 - *Predicted peptides for OmpV, using Emini Surface Accessibility Prediction.*

Table 22 - *Predicted peptides for OmpV, using Emini Surface Accessibility Prediction.*

No.	Start	End	Peptide	Length
1	27	33	GNIYNNQ	7
2	48	55	LFKDQKHN	8
3	108	116	SLKGTQKRR	9
4	173	181	LTYQSEDYV	9
5	188	201	KDREATANRKAYKG	14
6	217	231	INDNWQINQTTQYTR	15

- **Kolaskar & Tongaonkar Antigenicity**

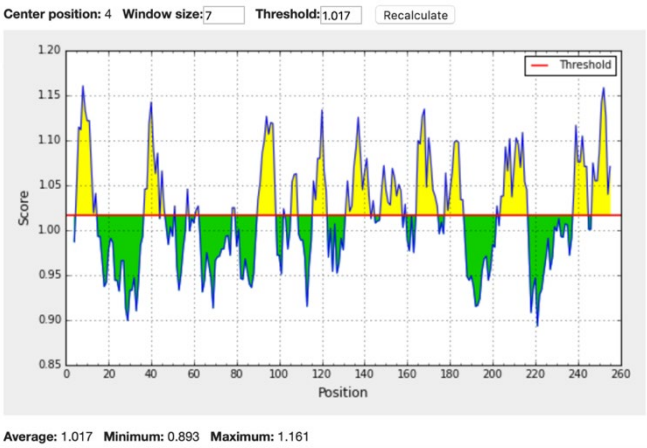

Figure 35 - *Predicted peptides for OmpV, using Kolaskar & Tongaonkar Antigenicity Prediction.*

Table 23 - *Predicted peptides for OmpV, using Kolaskar & Tongaonkar Antigenicity Prediction.*

No.	Start	End	Peptide	Length
1	5	14	LIALLTLAAA	10
2	37	43	VAEVGVA	7
3	90	98	MSAYVVGSG	9
4	116	122	RLAVDLG	7
5	131	142	EHNVISTYLQHD	12
6	148	157	KGYLAGATYF	10
7	165	174	VDFVPFANLT	10
8	178	186	EDYVDYYFG	9
9	204	216	TVSYGLGYKLVMP	13
10	238	244	DSSVVDS	7

- **Parker Hydrophilicity Prediction**

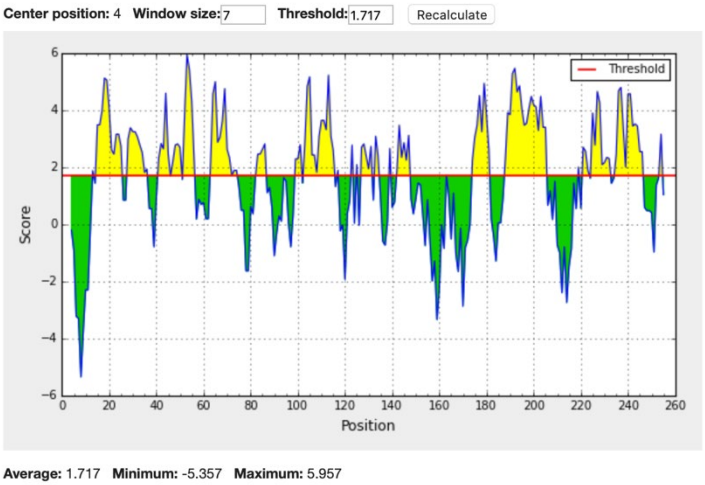

Figure 37 - *Predicted peptide for OmpV, using Parker Hydrophilicity Prediction.*

- **Karplus & Schulz Flexibility Prediction**

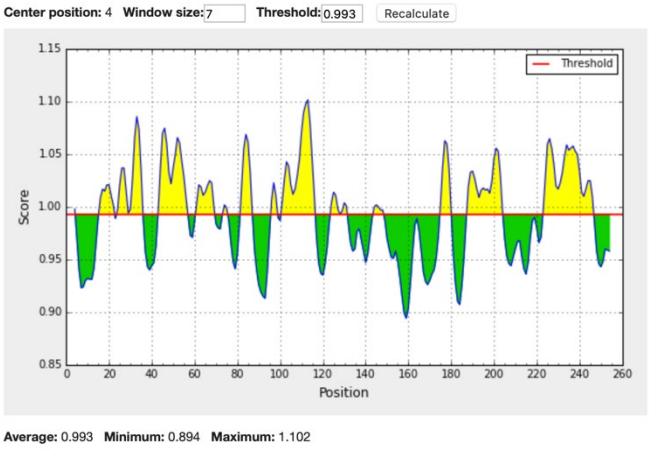

Figure 36 - Predicted peptide for OmpV, using Karplus & Schulz Flexibility Prediction.

- **AGGRESCAN – Aggregation Prediction**

Table 24 - *Predicted peptide sequences for OmpV, using AGGRESCAN, a server for the prediction and evaluation of "hotspots" of aggregation in polypeptides. a^4vAHS stands for: Amino-acid aggregation-propensity value window average (a^4v) average in the Hotspot.*

No.	Start	End	Peptide	Length	a^4vAHS
1	1	15	MNKTLIALLTLAAAG	15	0.357
2	37	41	VAEVG	5	0.212
3	76	80	INYRF	5	0.202
4	89	99	NMSAYVVGSGI	11	0.415
5	135	139	ISTYL	5	0.192
6	146	168	AYKGYLAGATYFHIMNFGSVDFV	23	0.337
7	170	174	FANLT	5	0.292
8	181	185	VDYYF	5	0.329
9	205	215	VSYGLGYKLVM	11	0.395
10	248	258	WVVGASVSYNF	11	0.245

For outer membrane protein V, there are several options: 14-37, 14-44, and 5-44 having 24, 31, and 40 residues, respectively. Since there is a region more prone to aggregation (37-41, score of 0.212), sequence 14 to 37 can be the best choice out of those three (**Table 24**).

The sequence 41 to 60, with 20 residues, also seems to be a good selection, not having any region more prone to aggregation and also containing the most hydrophilic region (50 to 56, KDQKHNT, with a score of 5.957, **Figure 36**).

Moreover, the sequences 90-122 and 96-122 are also a possibility, having 33 and 27 residues, respectively. Nonetheless, they tend to aggregate from position 89 to 99 (score of 0.415). The most flexible zone is also present in this interval, from 110 to 116 (KGTQKRR) with a score of 1.102 (**Figure 37**). All the sequences presented above are condensed in **Table 25**,

Table 25 - *Collection of data obtained from the previous software's already described. The sequence 41-60 appears to be the best option considering the Outer Membrane Protein V. Regarding the Parker Hydrophilicity Prediction and Karplus & Schulz Flexibility , we chose the best score individually, and the best one is indicated in bold.*

Peptides	Bepipred Linear Epitope	Bepipred Linear Epitope Prediction 2.0	Emini Surface Accessibility	Kolaskar & Tongaonkar Antigenicity	Parker Hydrophilicity Prediction	Karplus & Schulz Flexibility	AGGRESCAN – Aggregation
5-44	14-24 29-37 41-44	21-36	27-33	5-14 37-43	15-21(5.143)	30-36(1.086)	1-15(0.357) 37-41(0.212)
14-37	14-24 29-37	21-36	27-33	5-14	15-21(5.143)	30-36(1.086)	-
14-44	14-24 29-37 41-44	21-36	27-33	5-14 37-43	15-21(5.143)	30-36(1.086)	37-41(0.212)
41-60	41-44 47-60	45-53	48-55	-	50-56(**5.957**)	43-49(1.075)	-
90-122	100-102 104-112	96-119	108-116	90-98 116-122	110-116(5.246)	110-116(**1.102**)	89-99(0.415)
96-122	100-102 104-112	96-119	108-116	116-122	110-116(5.246)	110-116(**1.102**)	89-99(0.415)

6.5.4. Outer Membrane Protein U

For outer membrane protein U (**Figure 38)**, in BepiPred, Bepipred 2.0, Emini Surface Accessibility, and Kolaskar & Tongaonkar Antigenicity web software's, there are 14, 14, 10 and 12 predicted sequences, respectively (**Figure 39**, **40**, **41 and, 42** and **Table 26**, **27, 28** and **29**). Only 8, 8, 3, and 4 sequences have more than 10 amino acids. In Parker Hydrophilicity and Karplus & Schulz Flexibility, there are 317 and 316 predicted residue scores, respectively.

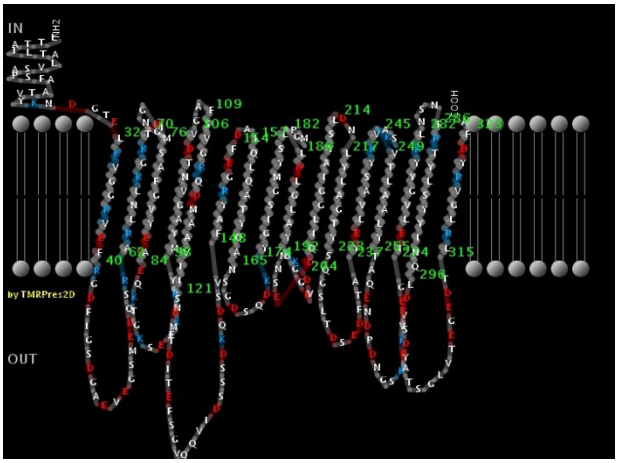

Figure 38 - *2D representation of OmpU Vibrio Parahaemolyticus, predicted by PRED-TMBB.*

- **Bepipred Linear Epitope Prediction**

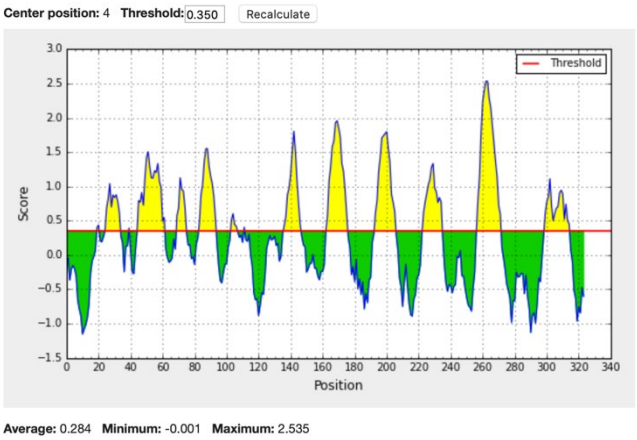

Figure 39 - *Predicted peptides for OmpU, using Bepipred Linear Epitope Prediction.*

Table 26 - *Predicted peptides for OmpU, using Bepipred Linear Epitope Prediction.*

No.	Start	End	Peptide	Length
1	19	20	SF	2
2	25	33	YKNDGTELK	9
3	39	39	E	1
4	45	61	IGSDGAEVEGSMEDQSR	17
5	69	74	KTDIGN	6
6	83	93	EAEQKTGKSEF	11
7	103	106	NTDV	4
8	111	111	V	1
9	136	146	VIDSSSDKQDS	11
10	162	175	TYQANSGDSQDKYG	14
11	193	205	SGGDVDKNNSEDQ	13
12	223	234	YSQGSLTDSEDF	12
13	257	271	AQENDPDNGSKYYDSV	15
14	299	314	EVKDATSGLVTEGEDT	16

- **Bepipred Linear Epitope Prediction 2.0**

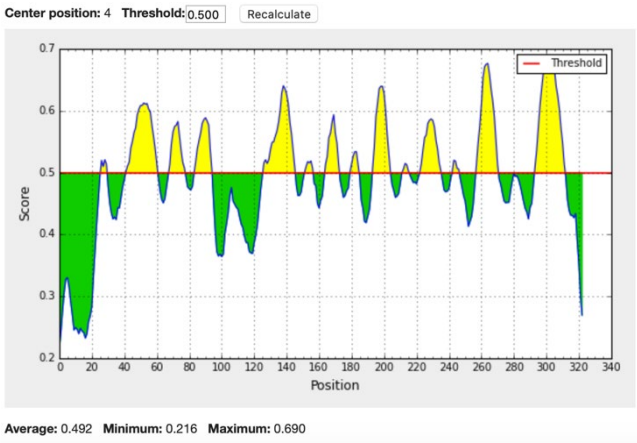

Figure 40 *- Predicted peptides for OmpU, using Bepipred Linear Epitope Prediction 2.0.*

Table 27 - *Predicted peptides for OmpU, using Bepipred Linear Epitope Prediction 2.0*

No.	Start	End	Peptide	Length
1	27	30	NDGT	4
2	42	61	GDFIGSDGAEVEGSMEDQSR	20
3	69	78	KTDIGNGMSA	10
4	85	95	EQKTGKSEFKN	11
5	127	146	ITEFSGVQQVIDSSSDKQDS	20
6	153	157	EFDAL	5
7	165	173	ANSGDSQDK	9
8	180	185	YSLPMG	6
9	194	204	GGDVDKNNSED	11
10	212	216	YSLDN	5
11	223	236	YSQGSLTDSEDFTA	14
12	243	247	YKVAS	5
13	258	271	QENDPDNGSKYDSV	14
14	295	312	INQLDEVKDATSGLVTEGE	19

- **Emini Surface Accessibility Prediction**

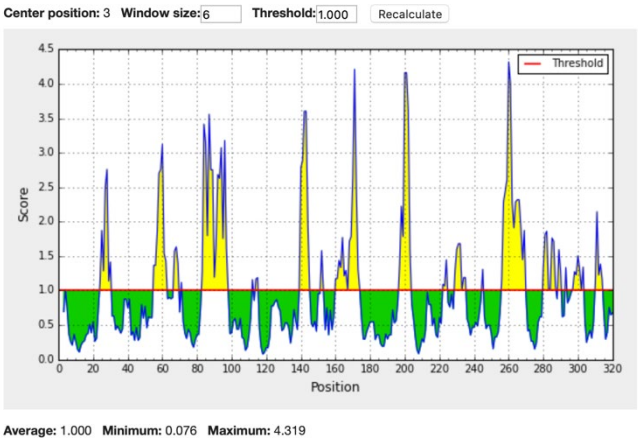

Figure 41 - *Predicted peptides for OmpU, using Emini Surface Accessibility Prediction.*

Table 28 - *Predicted peptides for OmpU, using Emini Surface Accessibility Prediction.*

No.	Start	End	Peptide	Length
1	24	30	VYKNDGT	7
2	55	62	SMEDQSRA	8
3	83	97	EAEQKTGKSEFKNRY	15
4	140	145	SSDKQD	6
5	160	174	QATYQANSGDSQDKY	15
6	197	203	VDKNNSE	7
7	229	235	TDSEDFT	7
8	256	269	TAQENDPDNGSKYD	14
9	280	287	YKLNSNFR	8
10	297	303	LDEVKDA	7

- **Kolaskar & Tongaonkar Antigenicity**

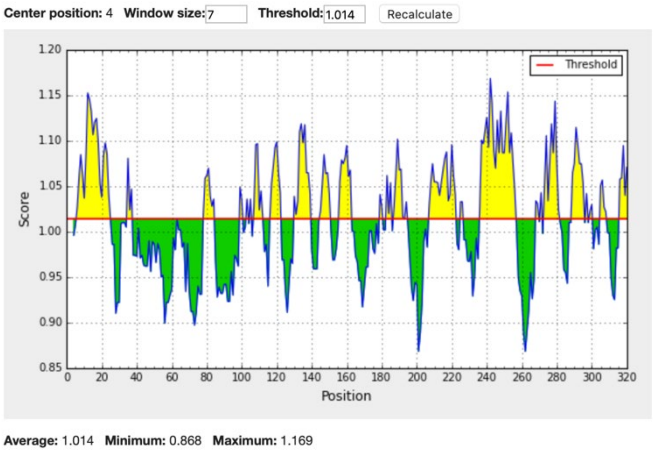

Figure 42 - *Predicted peptides for OmpU, using Kolaskar & Tongaonkar Antigenicity Prediction.*

Table 29 - *Predicted peptides for OmpU, using Kolaskar & Tongaonkar Antigenicity Prediction.*

No.	Start	End	Peptide	Length
1	6	24	LTTAILTALVSAPSFAATV	19
2	78	84	AFGVYEA	7
3	107	112	GAFSVG	6
4	117	122	AAVIIS	6
5	130	139	FSGVQQVIDS	10
6	145	150	DSVFAY	6
7	156	162	ALQLQAT	7
8	187	192	DLGLAY	6
9	207	222	LGGIAYSLDNLYLAGT	16
10	236	256	AYELVASYKVASKVTLAALYT	21
11	273	281	GIELVGYYK	9
12	289	295	YLSYYIN	7

- **Parker Hydrophilicity Prediction**

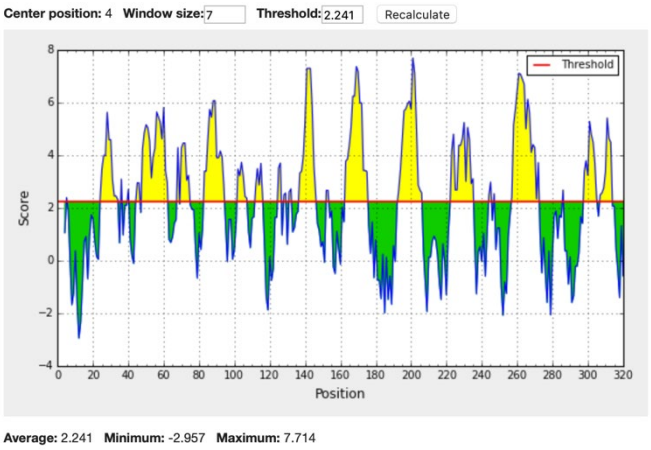

Figure 44 - *Predicted peptide for OmpU, using Parker Hydrophilicity Prediction.*

- **Karplus & Schulz Flexibility Prediction**

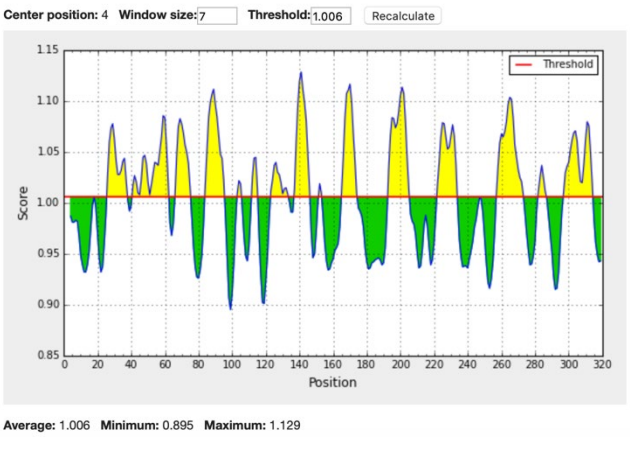

Figure 43 - *Predicted peptide for OmpU, using Karplus & Schulz Flexibility Prediction.*

- **AGGRESCAN – Aggregation Prediction**

Table 30 - *Predicted peptide sequences for OmpU, using AGGRESCAN, a server for the prediction and evaluation of "hotspots" of aggregation in polypeptides. a^4vAHS stands for: Amino-acid aggregation-propensity value window average (a^4v) average in the Hotspot.*

No.	Start	End	Peptide	Length	a^4vAHS
1	1	17	MKKAALTTAILTALVSA	17	0.316
2	76	81	MSAFGV	6	0.131
3	101	107	GVNTDVG	7	0.126
4	119	126	VIISDMTD	8	0.227
5	131	135	SGVQQ	5	0.088
6	177	182	SGMYSL	6	0.333
7	184	190	MGLDLGL	7	0.271
8	207	219	LGGIAYSLDNLYL	13	0.323
9	236	254	AYELVASYKVASKVTLAAL	19	0.298
10	272	282	EGIELVGYYKL	11	0.252
11	284	296	SNFRTYLSYYINQ	13	0.227
12	318	323	GVRYDF	6	0.095

When comparing outer membrane protein U with the previous proteins already analyzed, this protein is larger, having many more available sequences.

From position 19 to 33 (15 amino acids) and from 42 to 84 (43 amino acids), there is no region prone to aggregation (**Table 30**). From position 78 to 97 (20 amino acids), there is a possibility of aggregation from position 78 to 81 (score of 0.131). Also, the regions 140-157 and 160-175, with 18 and 16 amino acids, respectively, don't tend to aggregate, being a relevant detail because protein aggregates will have a reduced or no biological activity. Besides that, the sequence 138-157 has the most flexible region which comprises from 140 to 144, DSSSDKQ, with a score of 1.129 (**Figure 44**).

The sequence 187-205 (19 amino acids long) and 207-236 (29 amino acids) could also be a possibility but, again, tends to aggregate from 187 to 190 and 207 to 219, respectively (scores of 0.271 and 0.323). The sequence 187-205 has the most hydrophilic region (from position 198 to 204, DKNNSED, with a score of 7.714, **Figure 43**). The last two sequences, 236-271 and 289-314 (36 and 26 amino acids, respectively) are likewise prone to aggregation from 236 to 254 and from 289 to 296 (scores of 0.298 and 0.227, respectively).

Therefore, according to the collected data, we hypothesize that the sequence comprised between 138 to 157 is the best choice since it doesn't aggregate and, it has the most flexible region in the protein. All the information already examined is shown in **Table 31.**

***Table 31** - Collection of data obtained from the previous software's already described. The sequence 140-157appears to be the best option considering the Outer Membrane Protein U. Regarding the Parker Hydrophilicity Prediction and Karplus & Schulz Flexibility , we chose the best score individually, and the best one is indicated in bold.*

Peptides	Bepipred Linear Epitope	Bepipred Linear Epitope Prediction 2.0	Emini Surface Accessibility	Kolaskar & Tongaonkar Antigenicity	Parker Hydrophilicity Prediction	Karplus & Schulz Flexibility	AGGRESCAN – Aggregation
19-33	19-20 25-33	27-30	24-30	-	25-31(5.643)	26-32(1.078)	-
42-84	45-61 69-74	42-61 69-78	83-97	78-94	57-63(5.829)	56-62(1.086)	76-81(0.131)
78-97	83-93	85-95	83-97	78-84	85-91(6.086) 86-92(6.086)	86-92(1.112)	76-81(0.131)
138-157	140-146	153-157	140-145	145-150	140-146(7.314)	138-144(**1.129**)	-
160-175	162-175	165-173	160-174	-	166-172(7.386)	169-175(1.068)	-
187-205	193-205	194-204	197-203	187-193	198-204(**7.714**)	198-204(1.114)	184-190(0.271)
207-236	223-234	212-216 223-236	229-235	207-222	227-233(5.257)	222-228(1.079)	207-219(0.323)
236-271	257-271	243-247	256-260	236-256	258-264(7.129)	262-268(1.104)	236-254(0.298)
289-314	299-314	295-312	297-303	289-295	308-314(5.429)	308-314(1.080)	284-296(0.227) 318-323(0.095)

Summing up this analysis, if only one sequence could be selected to clone them in a DNA plasmid, we would suggest the OmpK. Since it is desirable to obtain as many strains as possible so that the vaccine could work against the majority of them, with BLASTN, ten strains were obtained and analyzed, being the highest number within the other OMPs (OmpW, OmpV, and OmpU). Within the OmpK sequence, the best option would be between positions 200 and 228.

7. Conclusions and Future Perspectives

The main goal of this project is to create and develop an oral pDNA vaccine against *Vibrio* species. To do so, the first step was to produce and purify two already constructed plasmid DNA molecules encoding for OmpK and immunogenic portion of OmpK, respectively. The backbone used for both plasmids was the pVAX eGFP expression vector, with 3697 bp. These two plasmids were prepared in order to be used for immunization of fish in aquaculture since OMPs are easily recognized as foreign substances by the host's immunological defense system, making them suitable and convenient for vaccination purposes.

Pure plasmids were produced in *Escherichia coli* at the following yields and respective final concentrations: 67.4µg of pVAX eGFP, 7.31µg of pVAX OmpK, and 23.52µg of pVAX OmpK-frag per gram of cell dry weight and 526.5ng/µL, 56ng/µL, and 175ng/µL. Since the size is different between the three plasmids (3697bp for pVAX eGFP, 4442bp for pVAX OmpK and 3913bp for pVAX OmpK-frag), it was expected to obtain different production yields. Usually, larger plasmids are present in lower numbers in their bacterial host due to the metabolic burden of also maintaining and duplicating their genomes.

Afterwards, chitosans and TPP were used to synthesize different nanoparticles and later complexation with pDNA. Produced nanoparticles were characterized regarding average hydrodynamic diameter, zeta potential, and polydispersity index (Pdi). A long-chain chitosan formulation (LC2, 5mL chitosan (3mg/mL, pH= 3.5) + 2mL TPP (0,75mg/mL)) showed fairly small diameters, low Pdi value and significant positive zeta potential and was chosen for pDNA complexation. The diameter, Pdi, and zeta potential obtained after complexation with pVAX eGFP were, respectively, 171nm, 0.172, and 41mV in the first experiment and 141nm, 0.132, and 39mV in the second trial. With pVAX OmpK-frag, the results were 188nm, 0.269, and 40mV, respectively. To corroborate the complexation of chitosan with DNA, fluorescence and agarose gel were also performed. When using fluorescence, the efficacy was evaluated by doing the difference between the amount of plasmid encapsulated into the nanoparticles and the one used in the encapsulation process. For agarose gel, a noticeable band is present in the gel wells that correspond to the complexes that exhibit no electrophoretic mobility due to the nature of chitosan, within the nanoparticles, which is a positively charged polymer. These positive complexes will not migrate in the gel, as won't the pDNA molecules firmly complexed with them.

In a different assay, short-chain chitosan resulted in greater diameters and Pdi values, meaning that some aggregation may have occurred. To avoid this problem, sonication can help by homogenizing the suspensions, decreasing the diameter, and consequently, the Pdi values.

Then, bioinformatics tools were used to predict new immunogenic epitopes. To obtain the right antigenicity, several parameters need to be taken into account, be optimized and analyzed, such as peptides purity levels, amino acid compositions, lengths, hydrophobicity, secondary structures, besides the inherent complexity of antigen recognition. Several software present in the Immune Epitope Database (IEDB) were useful in this project.

Afterwards, we concluded that within the analyzed OMPs (OmpK, OmpW, OmpV, and OmpU) the OmpK was the most suitable since it is desirable to obtain as many strains as possible in BLAST analysis. It is essential to observe this aspect so that the vaccine could work against the majority of Vibrio species. Thus, ten strains were obtained when using the BLASTN program for OmpK but only seven were found in each one of the remaining OMPs. Within the OmpK sequence, the best option would be between positions 200 and 228, since it has the most hydrophilic and most flexible regions, having just a small region of five amino acids more prone to aggregation.

To better understand the implications of these results, future studies could address the testing of new and different materials and applying them in the creation of new nanoparticles. One example is Poly lactic-co-glycolic acid (PLGA), which is a biodegradable synthetic copolymer of polylactic acid (PLA) and polyglycolic acid (PGA) and is highly biocompatible and non-toxic. PLGA will protect the DNA from nuclease degradation and will also control the DNA delivery system. This analysis will make us comprehend which materials are more appropriate for this specific investigation.

Also, it could be very interesting to develop new plasmids DNA (besides the pVAX eGFP OmpK-frag already tested), with different sequences and from different OMPs, to try to obtain the most efficient vaccine against vibriosis. To achieve an efficient vaccine, the *in vitro* studies would need to validate (or not) their ability to stimulate the immune response and also confirm the identity of the expressed antigen. Then, only if these results were encouraging, we could attempt to perform a vaccination trial in aquaculture fish (*in vivo*).

Besides the research proposals indicated above, it could also be attractive to use and try different software, such as NetSurfP for surface accessibility and Antigenic for epitopes, to obtain possible different outputs, making our results more meaningful and trustworthy.

8. References

1. Thompson FL, Gevers D, Thompson CC, *et al.* Phylogeny and molecular identification of vibrios on the basis of multilocus sequence analysis. *Appl Environ Microbiol*. 2005;71(9):5107-5115. doi:10.1128/AEM.71.9.5107
2. Morris JG, Acheson D. Cholera and other types of vibriosis: A story of human pandemics and oysters on the half shell. *Clin Infect Dis*. 2003;37(2):272-280. doi:10.1086/375600
3. Farmer J.J. The family vibrionaceae. *The Prokaryotes*. 2006;6:495-507. doi:10.1007/0-387-30746- X_17
4. Ina-Salwany MY, Al-Saari N, Mohamad A, *et al.* Vibriosis in fish: A review on disease development and prevention. *J Aquat Anim Health*. 2018;31(1):3-22. doi:10.1002/aah.10045
5. Mohamad N, Noor M, Amal A, *et al.* Vibriosis in cultured marine fishes: a review. *Aquaculture*. 2019;512:1-17. doi:10.1016/j.aquaculture.2019.734289
6. De Magny GC, Long W, Brown CW, *et al.* Predicting the distribution of *Vibrio spp.* in the Chesapeake bay: A *vibrio* cholerae case study. *Ecohealth*. 2009;6(3):378-389. doi:10.1007/s10393-009-0273-6
7. Li M, Wang C, Sun L. A pathogenic *Vibrio harveyi* lineage causes recurrent disease outbreaks in cultured Japanese fl ounder (Paralichthys olivaceus) and induces apoptosis in host cells. *Aquaculture*. 2011;319(1-2):30-36. doi:10.1016/j.aquaculture.2011.06.034
8. Bunpa S, Sermwittayawong N, Vuddhakul V. Extracellular enzymes produced by *Vibrio alginolyticus* isolated from environments and diseased aquatic animals. *Procedia Chem*. 2016;18:12-17. doi:10.1016/j.proche.2016.01.002
9. Frans I, Michiels CW, Bossier P, *et al.* Review *Vibrio anguillarum* as a fish pathogen: virulence factors, diagnosis and prevention. *J Fish Dis*. 2011;34(9):643-661. doi:10.1111/j.1365-2761.2011.01279.x
10. Li C, Ye Z, Wen L, *et al.* Identification of a novel vaccine candidate by immunogenic screening of *Vibrio parahaemolyticus* outer membrane proteins. *Vaccine*. 2014;32(46):6115-6121. doi:10.1016/j.vaccine.2014.08.077
11. Letchumanan V, Pusparajah P, Tan LTH, *et al.* Occurrence and antibiotic resistance of *Vibrio parahaemolyticus* from shellfish in Selangor, Malaysia. *Front Microbiol*. 2015;6:1-11. doi:10.3389/fmicb.2015.01417
12. Heppell J, Davis HL. Application of DNA vaccine technology to aquaculture. *Adv Drug Deliv Rev*. 2000;43(1):29-43. doi:10.1016/S0169-409X(00)00075-2
13. Microbiol A, Xu Y, Wang C, *et al.* IS CR 2 is associated with the dissemination of multiple resistance genes among *Vibrio para.* and *Pseudoalteromonas spp.* isolated from farmed fish. *Arch Microbiol*. 2017;199:891-896. doi:10.1007/s00203-017-1365-2
14. Fan G, Chen J, Jin T, *et al. Report on Marine Life Genomics*.; 2018. doi:10.20944/preprints201812.0156.v1
15. Vinay TN, Bhat S, Gon Choudhury T, *et al.* Recent advances in application of nanoparticles in fish vaccine delivery. *Rev Fish Sci Aquac*. 2017;26(1):29-41. doi:10.1080/23308249.2017.1334625
16. Secombes CJ, Ellis AE. The immunology of teleosts. In: *Fish Pathology*. ; 2012:144-166. doi:10.1002/9781118222942.ch4
17. Magnadóttir B. Innate immunity of fish (overview). *Fish Shellfish Immunol*. 2006;20(2):137-151. doi:10.1016/j.fsi.2004.09.006
18. Hansen JD, Landis ED, Phillips RB. Discovery of a unique Ig heavy-chain (IgT) in rainbow trout: Implications for a distinctive B cell developmental pathway in teleost fish. *Proc Natl Acad Sci U S A*. 2005;102(19):6919-6924. doi:10.1073/pnas.0500027102
19. Zhu B, Liu GL, Gong YX, *et al.* Protective immunity of grass carp immunized with DNA vaccine encoding the vp7 gene of grass carp reovirus using carbon nanotubes as a carrier molecule. *Fish Shellfish Immunol*. 2015;42(2):325-334. doi:10.1016/j.fsi.2014.11.026

20. Criscuolo E, Caputo V, Diotti RA, *et al.* Alternative methods of vaccine delivery: An overview of edible and intradermal vaccines. *J Immunol Res*. 2019;2019. doi:10.1155/2019/8303648
21. Rajan B, Løkka G, Koppang EO, *et al.* Passive immunization of farmed fish. *J Immunol*. 2017;198(11):4195-4202. doi:10.4049/jimmunol.1700154
22. Embregts CWE, Forlenza M. Oral vaccination of fish: Lessons from humans and veterinary species. *Dev Comp Immunol*. 2016;64:118-137. doi:10.1016/j.dci.2016.03.024
23. Hua K, Cobcroft JM, Cole A, *et al.* Review the future of aquatic protein: Implications for protein sources in aquaculture diets. *One Earth*. 2019;1(3):316-329. doi:10.1016/j.oneear.2019.10.018
24. Béné C; Arthur R; Norbury H *et al.* Contribution of fisheries and aquaculture to food security and poverty reduction : assessing the current evidence. *World Dev*. 2016;79:177-196. doi:10.1016/j.worlddev.2015.11.007
25. Gómez B, Munekata PES, Zhu Z, *et al.* Challenges and opportunities regarding the use of alternative protein sources: Aquaculture and insects. *Adv Food Nutr Res*. 2019;89:259-295. doi:10.1016/bs.afnr.2019.03.003
26. White SL, Demario DA, Iwanowicz LR, *et al.* Tissue distribution and immunomodulation in channel catfish (Ictalurus punctatus) following dietary exposure to polychlorinated biphenyl aroclors and food deprivation. *Int J Environ Res Public Heal*. 2020;17(4):1-17. doi:10.3390/ijerph17041228
27. Carubelli G, Fanelli R, Mariani G, *et al.* PCB contamination in farmed and wild sea bass (Dicentrarchus labrax L.) from a coastal wetland area in central Italy. *Chemosphere*. 2007;68(9):1630-1635. doi:10.1016/j.chemosphere.2007.04.004
28. Cahu C, Salen E, Lorgeril M De. Farmed and wild fish in the prevention of cardiovascular diseases: Assessing possible differences in lipid nutritional values. *Nutr Metab Cardiovasc Dis*. 2004;14(1):34-41. doi:10.1016/s0939-4753(04)80045-0.
29. Sommerset I, Krossøy B, Biering E, *et al.* Vaccines for fish in aquaculture. *Expert Rev Vaccines*. 2005;4(1):89-101. doi:10.1586/14760584.4.1.89
30. Mutoloki, S., Munang'andu, H. M., Evensen Ø. Oral vaccination of fish – Antigen preparations, uptake and immune induction. *Front Immunol*. 2015;6:1-10. doi:10.3389/fimmu.2015.00519
31. Kang M, Feng F, Wang Y, *et al.* Advances in research into oral vaccines for fish. *Int J Fish Aquac Sci*. 2018;8(1):19-40. doi:https://doi.org/10.1016/j.dci.2016.03.024
32. Rivas-Aravena A, Sandino AM, Spencer E. Nanoparticles and microparticles of polymers and polysaccharides to administer fish vaccines. *Biol Res*. 2013;46(4):407-419. doi:10.4067/S0716-97602013000400012
33. Farris E, Brown DM, Ramer-tait AE, *et al.* Micro- and nanoparticulates for DNA vaccine delivery. *Exp Biol Med*. 2016;241(9):1-11. doi:10.1177/1535370216643771
34. Petter M, Duffy MF. Antigenic Variation in Plasmodium falciparum. In: Hsu E, Pasquier L Du, eds. *Pathogen-Host Interactions: Antigenic Variation V. Somatic Adaptations*. Vol 57. Springer; 2015:47-90. doi:10.1007/978-3-319-20819-0
35. Salinas I. The mucosal immune system of teleost fish. *Biology (Basel)*. 2015;4(3):525-539. doi:10.3390/biology4030525
36. Sudheesh PS CK. Prospects and challenges of developing and commercializing immersion vaccines for aquaculture. *Int Biol Rev*. 2017;1(1):1-20. doi:https://doi.org/10.18103/ibr.v1i1.1313
37. Weiner HL, da Cunha AP, Quintana F, *et al.* Oral tolerance. *Immunol Rev*. 2011;241:241-259. doi:10.1111/j.1600-065X.2011.01017.x.
38. Schirmbeck R, Reimann J. Revealing the potential of DNA-based vaccination: Lessons learned from the hepatitis B virus surface antigen. *Biol Chem*. 2001;382(4):543-552. doi:10.1515/BC.2001.068
39. Kutzler MA, Weiner DB. DNA vaccines: Ready for prime time? *Nat Rev Genet*. 2008;9(10):776-788. doi:10.1038/nrg2432

40. Hasson SSAA, Al-Busaidi JKZ, Sallam TA. The past, current and future trends in DNA vaccine immunisations. *Asian Pac J Trop Biomed*. 2015;5(5):344-353. doi:10.1016/S2221-1691(15)30366-X
41. Sharma R, Mody N, Dubey S, *et al.* Nanoparticulate carrier(s): an emerging paradigm in new generation vaccine development. In: *Nanostructures for Drug Delivery*. Elsevier Inc.; 2017:523-550. doi:10.1016/B978-0-323-46143-6/00017-8
42. Lindenbergh MFS, Stoorvogel W. Antigen Presentation by Extracellular Vesicles from Professional Antigen- Presenting Cells. *Annu Rev Immunol*. 2018;36:435-459. doi:10.1146/annurev-immunol-041015-055700
43. Delamarre L, Holcombe H, Mellman I. Presentation of exogenous antigens on major histocompatibility complex (MHC) class I and MHC Class II molecules is differentially regulated during dendritic cell maturation. *J Exp Med*. 2003;198(1):111-122. doi:10.1084/jem.20021542
44. Li Y, Yin Y, Mariuzza RA. Structural and biophysical insights into the role of CD4 and CD8 inT cell activation. *Front Immunol*. 2013;4:1-12. doi:10.3389/fimmu.2013.00206
45. Artyomov MN, Lis M, Devadas S, *et al.* CD4 and CD8 binding to MHC molecules primarily acts to enhance Lck delivery. *Proc Natl Acad Sci*. 2010;107:16916-16921. doi:www.pnas.org/cgi/doi/10.1073/pnas.1010568107
46. Mignon C, Sodoyer R, Werle B. Antibiotic-free selection in biotherapeutics: Now and forever. *Pathogens*. 2015;4(2):157-181. doi:10.3390/pathogens4020157
47. Stenler S, Blomberg P, Smith CIE. Safety and efficacy of DNA vaccines: Plasmids vs. minicircles. *Hum Vaccines Immunother*. 2014;10(5):1306-1308. doi:10.4161/hv.28077
48. Robinson HL, Pertmer TM. DNA vaccines for viral infections: Basic studies and applications. *Adv Virus Res*. 2000;55:1-74. doi:10.1016/S0065-3527(00)55001-5
49. Shintani M, Sanchez ZK, Kimbara K. Genomics of microbial plasmids: Classification and identification based on replication and transfer systems and host taxonomy. *Front Microbiol*. 2015;6:242. doi:10.3389/fmicb.2015.00242
50. Bennett PM. Plasmid encoded antibiotic resistance: Acquisition and transfer of antibiotic resistance genes in bacteria. *Br J Pharmacol*. 2008;153(Suppl 1):347-357. doi:10.1038/sj.bjp.0707607
51. A. Gómez L, A. Oñate A. Plasmid-based DNA vaccines. In: *Plasmid*. ; 2019:1-14. doi:10.5772/intechopen.76754
52. Williams J. Vector design for improved DNA vaccine efficacy, safety and production. *Vaccines*. 2013;1(3):225-249. doi:10.3390/vaccines1030225
53. Glenting J, Wessels S. Ensuring safety of DNA vaccines. *Microb Cell Fact*. 2005;4:1-5. doi:10.1186/1475-2859-4-26
54. Azzoni AR, Ribeiro SC, Monteiro GA, *et al.* The impact of polyadenylation signals on plasmid nuclease-resistance and transgene expression. *J Gene Med*. 2007;(9):392-402. doi:10.1002/jgm.1031
55. Ningqiu L, Junjie B, Shuqin W, *et al.* An outer membrane protein, OmpK, is an effective vaccine candidate for *Vibrio harveyi* in Orange-spotted grouper (Epinephelus coioides). *Fish Shellfish Immunol*. 2008;25(6):829-833. doi:10.1016/j.fsi.2008.09.007
56. Wang Q, Chen J, Liu R, *et al.* Identification and evaluation of an outer membrane protein OmpU from a pathogenic *Vibrio harveyi* isolate as vaccine candidate in turbot (Scophthalmus maximus). *Lett Appl Microbiol*. 2011;53(1):22-29. doi:10.1111/j.1472-765X.2011.03062.x
57. Qian R, Chu W, Mao Z, *et al.* Expression, characterization and immunogenicity of a major outer membrane protein from *Vibrio alginolyticus*. *Acta Biochim Biophys Sin (Shanghai)*. 2007;39(3):194-200. doi:10.1111/j.1745-7270.2007.00268.x
58. Zhang W, Liao Z, Hu F, *et al.* Protective immune responses of recombinant outer membrane proteins OmpF and OmpK of Aeromonas hydrophila in European eel (Anguilla anguilla). *Aquac Res*. 2019;50(12):3559-3566. doi:10.1111/are.14311

59. Cai S, Yao S, Lu Y, *et al.* Immune response in Lutjanus erythropterus induced by the major outer membrane protein (OmpU) of *Vibrio alginolyticus*. *Dis Aquat Organ*. 2010;90:63-68. doi:10.3354/dao02206
60. Wang S, Lauritz J, Jass J, *et al.* Role for the major outer-membrane protein from *Vibrio anguillarum* in bile resistance and biofilm formation. *Microbiology*. 2003;149(Pt 4):1061-1071. doi:10.1099/mic.0.26032-0
61. Rauta PR, Nayak B, Monteiro GA, *et al.* Design and characterization of plasmids encoding antigenic peptides of Aha1 from Aeromonas hydrophila as prospective fish vaccines. *J Biotechnol*. 2017;241:116-126. doi:10.1016/j.jbiotec.2016.11.019
62. Cardeira M. Synthesis and characterization of nanoparticles as DNA vaccine carriers for fish immunization. Published online 2018.
63. Rizvi SAA, Saleh AM. Applications of nanoparticle systems in drug delivery technology. *Saudi Pharm J*. 2018;26(1):64-70. doi:10.1016/j.jsps.2017.10.012
64. Khan I, Saeed K, Khan I. Nanoparticles: Properties, applications and toxicities. *Arab J Chem*. 2019;12(7):908-931. doi:10.1016/j.arabjc.2017.05.011
65. Pati R, Shevtsov M, Sonawane A. Nanoparticle vaccines against infectious diseases. *Front Immunol*. 2018;9:1-16. doi:10.3389/fimmu.2018.02224
66. Busatto S, Pham A, Suh A, *et al.* Organotropic drug delivery: Synthetic nanoparticles and extracellular vesicles. *Biomed Microdevices*. 2019;21(2). doi:10.1007/s10544-019-0396-7
67. Mohammed MA, Syeda JTM, Wasan KM, *et al.* An overview of chitosan nanoparticles and its application in non-parenteral drug delivery. *Pharmaceutics*. 2017;9(4). doi:10.3390/pharmaceutics9040053
68. Zhang Y, Zhang X, Ding R, *et al.* Determination of the degree of deacetylation of chitosan by potentiometric titration preceded by enzymatic pretreatment. *Carbohydr Polym*. 2011;83(2):813-817. doi:10.1016/j.carbpol.2010.08.058
69. MH P, AS H, AZ S. Chitosan: A promising marine polysaccharide for biomedical research. *Pharmacogn Rev*. 2016;10(19):39-42. doi:10.4172/2155-952x.1000168
70. Roy JC, Salaün F, Giraud S, *et al.* Solubility of chitin: Solvents, solution behaviors and their related mechanisms. In: *Solubility of Polysaccharides*. ; 2017:109-127. doi:10.5772/intechopen.71385
71. Yang X, Yuan X, Cai D, *et al.* Low molecular weight chitosan in DNA vaccine delivery via mucosa. *Int J Pharm*. 2009;375(1-2):123-132. doi:10.1016/j.ijpharm.2009.03.032
72. Wang JJ, Zeng ZW, Xiao RZ, *et al.* Recent advances of chitosan nanoparticles as drug carriers. *Int J Nanomedicine*. 2011;6:765-774. doi:10.2147/ijn.s17296
73. Sreekumar S, Goycoolea FM, Moerschbacher BM, *et al.* Parameters influencing the size of chitosan-TPP nano- and microparticles. *Sci Rep*. 2018;8(1):1-11. doi:10.1038/s41598-018-23064-4
74. de Pinho Neves AL, Milioli CC, Müller L, *et al.* Factorial design as tool in chitosan nanoparticles development by ionic gelation technique. *Colloids Surfaces A Physicochem Eng Asp*. 2014;445(2014):34-39. doi:10.1016/j.colsurfa.2013.12.058
75. Gary DJ, Min J, Kim Y, *et al.* The effect of N/P ratio on the in vitro and in vivo interaction properties of pegylated poly[2-(dimethylamino)ethyl methacrylate]-based siRNA complexes. *Macromol Biosci*. 2013;13(8):1059-1071. doi:10.1002/mabi.201300046
76. Tammam SN, Azzazy HME, Breitinger HG, *et al.* Chitosan nanoparticles for nuclear targeting: The effect of nanoparticle size and nuclear localization sequence density. *Mol Pharm*. 2015;12(12):4277-4289. doi:10.1021/acs.molpharmaceut.5b00478
77. Cho EJ, Holback H, Liu KC, *et al.* Nanoparticle characterization: State of the art, challenges, and emerging technologies. *Mol Pharm*. 2013;10(6):2093-2110. doi:10.1021/mp300697h
78. Kumar CSSR. *Magnetic Characterization Techniques for Nanomaterials*. Springer; 2017. doi:10.1007/978-3-662-52780-1

79. Ramos AP. Dynamic Light Hamburan Terapan untuk Nano Partikel Karakterisasi. In: *Nanocharacterization Techniques*. Elsevier Inc.; 2017:99-110. doi:10.1016/B978-0-323-49778-7/00004-7
80. Lim J, Yeap SP, Che HX, *et al.* Characterization of magnetic nanoparticle by dynamic light scattering. *Nanoscale Res Lett*. 2013;8(1):1-14. doi:10.1186/1556-276X-8-381
81. Clayton KN, Salameh JW, Wereley ST, *et al.* Physical characterization of nanoparticle size and surface modification using particle scattering diffusometry. *Biomicrofluidics*. 2016;10(5). doi:10.1063/1.4962992
82. Murdock RC, Braydich-Stolle L, Schrand AM, *et al.* Characterization of nanomaterial dispersion in solution prior to in vitro exposure using dynamic light scattering technique. *Toxicol Sci*. 2008;101(2):239-253. doi:10.1093/toxsci/kfm240
83. Vogel R, Pal AK, Jambhrunkar S, *et al.* High-resolution single particle zeta potential characterisation of biological nanoparticles using tunable resistive pulse sensing. *Sci Rep*. 2017;7(1):1-13. doi:10.1038/s41598-017-14981-x
84. Kaszuba M, Corbett J, Watson FMN, *et al.* High-concentration zeta potential measurements using light-scattering techniques. *Philos Trans R Soc A Math Phys Eng Sci*. 2010;368(1927):4439-4451. doi:10.1098/rsta.2010.0175
85. Selvamani V. Stability studies on nanomaterials used in drugs. In: *Characterization and Biology of Nanomaterials for Drug Delivery*. Elsevier Inc.; 2019:425-444. doi:10.1016/b978-0-12-814031-4.00015-5
86. Clogston, J. D., & Patri AK. Zeta Potential measurement. In: *Characterization of Nanoparticles Intended for Drug Delivery*. Humana Press; 2010:63–70. doi:10.1007/978-1-60327-198-1
87. Hunter RJ. *Foundations of Colloid Science*. 5th ed. Oxford University Press Inc.; 2001.
88. Biebricher AS, Heller I, Roijmans RFH, *et al.* The impact of DNA intercalators on DNA and DNA-processing enzymes elucidated through force-dependent binding kinetics. *Nat Commun*. 2015;6:1-12. doi:10.1038/ncomms8304
89. Bordelon H, Biris AS, Sabliov CM, *et al.* Characterization of plasmid DNA location within chitosan/PLGA/pDNA nanoparticle complexes designed for gene delivery. *J Nanomater*. 2011;2011:9. doi:10.1155/2011/952060
90. Alves CPA, Šimčíková M, Brito L, *et al.* Development of a nicking endonuclease-assisted method for the purification of minicircles. *J Chromatogr A*. 2016;1443:136-144. doi:10.1016/j.chroma.2016.03.035
91. Raiado-Pereira L, Prazeres DMF, Mateus M. Impact of plasmid size on the purification of model plasmid DNA vaccines by phenyl membrane adsorbers. *J Chromatogr A*. 2013;1315:145-151. doi:10.1016/j.chroma.2013.09.076
92. Scientific TF. Choosing the best vacuum concentrator for your laboratory - Tips and best practices from industry pioneers The SpeedVac equation. Published online 2018:18.
93. Silva F, Queiroz JA, Domingues FC. Evaluating metabolic stress and plasmid stability in plasmid DNA production by Escherichia coli. *Biotechnol Adv*. 2012;30(3):691-708. doi:10.1016/j.biotechadv.2011.12.005
94. Orme N, Milo R, Phillips R. *Cell Biology by the Numbers*.; 2016. doi:10.2142/biophys.56.352
95. Yang J, Yang Y. Plasmid size can affect the ability of Escherichia coli to produce high-quality plasmids. *Biotechnol Lett*. 2012;34(11):2017-2022. doi:10.1007/s10529-012-0994-4
96. Sengupta M, Austin S. Prevalence and significance of plasmid maintenance functions in the virulence plasmids of pathogenic bacteria. *Infect Immun*. 2011;79(7):2502-2509. doi:10.1128/IAI.00127-11
97. Tang X, Nakata Y, Li H, *et al.* The optimization of preparations of competent cells for transformation of E . coli. 1994;22(14):2857-2858.

98. Naser MD, Eskandari R, Zolfagharian H, *et al.* Preparation and in vitro characterization of chitosan nanoparticles containing Mesobuthus eupeus scorpion venom as an antigen delivery system. 2012;18(1):44-52.
99. Deng Q, Zhou C, Luo B. Preparation and characterization of chitosan nanoparticles containing lysozyme. 2008;44(5):336-242. doi:10.1080/13880200600746246
100. Omar Zaki SS, Ibrahim MN, Katas H. Particle size affects concentration-dependent cytotoxicity of chitosan nanoparticles towards mouse hematopoietic stem cells. *J Nanotechnol*. 2015;1:1-5. doi:10.1155/2015/919658
101. Gaspar VM, Sousa F, Queiroz JA, *et al.* Formulation of chitosan-TPP-pDNA nanocapsules for gene therapy Applications. *Nanotechnology*. 2011;22(1). doi:10.1088/0957-4484/22/1/015101
102. Perez C, Sanchez A, Putnam D, *et al.* Poly (lactic acid) -poly (ethylene glycol) nanoparticles as new carriers for the delivery of plasmid DNA. 2001;75:211-224.
103. Farris E, Brown DM, Ramer-tait AE, *et al.* Chitosan-zein nano-in-microparticles capable of mediating in vivo transgene expression following oral delivery. *J Control Release*. 2017;249:150-161. doi:10.1016/j.jconrel.2017.01.035
104. Karunasagar I, Karunasagar I, Raghunath P. Editorial : Ecology , virulence , and detection of pathogenic and pandemic *Vibrio parahaemolyticus. Front Microbiol*. 2016;7(156):1-2. doi:10.3389/fmicb.2016.00156
105. Karunasagar I, Maiti B, Kumar BK. Molecular methods to study *vibrio parahaemolyticus* and *vibrio vulnificus* from atypical environments. In: *Methods in Microbiology*. 1st ed. Elsevier Ltd.; 2018:387-417. doi:10.1016/bs.mim.2018.09.001
106. Boratyn GM, Camacho C, Cooper PS, *et al.* BLAST : a more efficient report with usability improvements. 2013;41(Web Server issue):W29-W33. doi:10.1093/nar/gkt282
107. Camacho C, Coulouris G, Avagyan V, *et al.* BLAST + : architecture and applications. *BMC Bioinformatics*. 2009;10(421):1-9. doi:10.1186/1471-2105-10-421
108. Hilterbrand A, Saelens J, Putonti C. CBDB : The codon bias database. *BMC Bioinformatics*. 2012;13(62). doi:10.1186/1471-2105-13-62
109. Trier N, Hansen P. Peptides, antibodies, peptide antibodies and more. *Int J Mol Sci*. 2019;20(24):1-22. doi:10.3390/ijms20246289
110. Lee B, Huang J, Jayathilaka LP, *et al.* Antibody production with synthetic peptides. 2016;1474:25-47. doi:10.1007/978-1-4939-6352-2
111. Lodish H, Berk A, Zipursky SL *et al.* Hierarchical Structure of Proteins. In: *Molecular Cell Biology*. 4th ed. W. H. Freeman; 2000. https://www.ncbi.nlm.nih.gov/books/NBK21581/
112. Tchernychev B, Cabilly S, Wilchek M. The epitopes for natural polyreactive antibodies are rich in proline. 1997;94(12):6335-6339. doi:10.1073/pnas.94.12.6335
113. Kant R Van Der, Karow-zwick AR, Durme J Van, *et al.* Prediction and Reduction of the Aggregation of Monoclonal Antibodies. *J Mol Biol*. 2017;429(8):1244-1261. doi:10.1016/j.jmb.2017.03.014
114. Wang W, Nema S, Teagarden D. Protein aggregation — Pathways and influencing factors. 2010;390:89-99. doi:10.1016/j.ijpharm.2010.02.025
115. Nishi H, Miyajima M, Nakagami H, *et al.* Phase separation of an IgG1 antibody solution under a low ionic strength condition. *Pharm Res*. 2010;27(7):1348-1360. doi:10.1007/s11095-010-0125-7
116. Anfinsen CB. Principles that govern the folding of protein chains. 1973;181(4096):223-230. doi:10.1126/science.181.4096.223.
117. Wang X, Das T, Singh S, *et al.* Potential aggregation prone regions in biotherapeutics. 2009;1(3):254-267. doi:10.4161/mabs.1.3.8035
118. Paul A.J , Handrick R., Ebert S. HF. Identification of process conditions influencing protein aggregation in Chinese hamster ovary cell culture. *Biotechnol Bioeng*. 2017;115(5):1173-1185. doi:10.1002/bit.26534
119. Cour T, Kiemer L, Mølgaard A, *et al.* Analysis and prediction of leucine-rich nuclear export signals. 2004;17(6):527-536. doi:10.1093/protein/gzh062

120. Stern B, Olsen LC, Trösse C, *et al.* Improving mammalian cell factories : The selection of signal peptide has a major impact on recombinant protein synthesis and secretion in mammalian cells. *Trends Cell Mol Biol*. 2007;2:1-17.
121. Carvalho J, Azzoni A, Prazeres D, *et al.* Comparative analysis of antigen-targeting sequences used in DNA vaccines. *Mol Biotechnol*. 2009;44(3):204-212. doi:10.1007/s12033-009-9229-x
122. Henriques AM, Fevereiro M, Prazeres DMF, *et al.* Development of a candidate DNA vaccine against Maedi-Visna virus. 2007;119(3-4):222-232. doi:10.1016/j.vetimm.2007.05.004
123. Antibody Epitope Prediction. Accessed May 8, 2020. http://tools.immuneepitope.org/bcell/
124. Erik J, Larsen P, Lund O, *et al.* Improved method for predicting linear B-cell epitopes. *Immunome Res*. 2006;2(2):1-7. doi:10.1186/1745-7580-2-2
125. Jespersen MC, Peters B, Nielsen M, *et al.* BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res*. 2017;45(W1):W24-W29. doi:10.1093/nar/gkx346
126. Emini EA, Hughes J V, Perlow DS, *et al.* Induction of Hepatitis A Virus-Neutralizing Antibody by a Virus- Specific Synthetic Peptide. 1985;55(3):836-839.
127. Srinivasan P, Kumar SP, Karthikeyan M, *et al.* Epitope-based immunoinformatics and molecular docking studies of nucleocapsid protein and ovarian tumor domain of Crimean – Congo hemorrhagic fever virus. *Front Genet*. 2011;2(72):1-9. doi:10.3389/fgene.2011.00072
128. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic dete ~ inants protein antigens. *FEBS Lett*. 1990;276(1-2):172-174. doi:10.1016/0014-5793(90)80535-q
129. Karplus PA, Schulz GE. Prediction of chain flexibility in proteins synthesis of acetylcholine receptors in xenopus oocytes induced by poly (A) + -mRNA from locust nervous tissue. *Naturwissenschaften*. 1985;72:212-213.
130. Sun Z, Liu Q, Qu G, *et al.* Utility of B - factors in protein science : Interpreting rigidity , flexibility , and internal motion and engineering thermostability. *Chem Rev*. 2019;119(3):1626-1665. doi:10.1021/acs.chemrev.8b00290
131. Parker J, Guo D, Hodges R. New hydrophilicity scale derived from High-Performance Liquid Chromatography peptide retention data: Correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistr*. 1986;25(19):5425-5432.
132. Groot N De, Vendrell J, Ventura S. AGGRESCAN: a server for the prediction of “hot spots” of aggregation in polypeptides. *BMC Bioinformatics*. 2007;8(1):65. doi:10.1186/1471-2105-8-65
133. Bagos PG, Liakopoulos TD, Spyropoulos IC, *et al.* PRED-TMBB : a web server for predicting the topology of b-barrel outer membrane proteins. 2004;32:400-404. doi:10.1093/nar/gkh417

9. Annexes

- **Annex 1 – eGFP gene sequence, from SnapGene**

1 5’ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCT 56
57 GGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATG 112
113 CCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGC 169
170 CCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTAC 225
226 CCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTC 282
283 CAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGT 338
339 GAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCA 394
395 AGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAAC 450
451 GTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGC 507
508 CACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCC 563
564 CATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCG 619
620 CCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTG 675
676 ACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAA 3’720

- **Annex 2 – OmpK gene sequence** [62]

1 5’ATGATCGATGCTGACTACTCAGATGGCGATATCCACAAGAACGATTACAAGTGGATGC 58
59 AGTTCAACATCATGGGCGCAATAAATGAGAAAGGTCCTTACGAATCAACTCACGATTAC 117
118 CTAGAAATGGAATTTGGCGGTCGCTCTGGTATTTTCGACCTTTACGGCTACGTTGACG 175
176 TATTCAACCTAACTTCCGACAGCAGCAGTGATAAAGCTGATAAAGATGGTAAAATCTTT 234
235 ATGAAGTTCGCCCCTCGTATGTCTCTAGACGCTATTACTGGTAAAGACTTATCTTTCG 292
293 GTCCAGTACAAGAGCTTTACTTGTCTACTCTTTTTGAATGGGACGGTAACAATGGCGG 350
351 TGTTAACACTCAGAAAGTAGATTTTGGTTCTGACGTAATGGTTCCTTGGTTTGGTAAA 409
410 ATGGGTCTAAACCTTTACGGCACTTACGACTCAAACCAAAAAGATTGGAACGGTTTCC 467
468 AAATCTCGACTAACTGGTTCAAACCATTCTACTTCTTCGAGAACGGTTCATTCATCTCT 526
527 TACCAAGGTTACATCGATTACCAATTTGGTATGAAAGACGAGTATTCTCAAGTTAGTAA 586
587 GGCGGTGCTATGTTCAACGGTATCTACTGGCACTCTGACCGCTTCGCGGTAGGTTA 643
644 CGGTCTGAAGCTATACCATGACGTATATGGCTTCGAAGATGGTACTGGTCTACCATGG 702
703 GATTCTTCGACAAAATCTGAGTCTTCAGGCGTAGGTCACTACGTAGCTGTAACTTACA 760
761 AATTCTTA3’768

- **Annex 3 - OmpK immunogenic fragment gene sequence** [62]

1 5’ATGGGTTACATCGATTACCAATTTGGTATGAAAGACGAGTATTCTCAAGTTAGTAATGG 59
60 CGGTGCTATGTTCAACGGTATCTACTGGCACTCTGACCGCTTCGCGGTAGGTTACGGT 117
118 CTGAAGCTATACCATGACGTATATGGCTTCGAAGATGGTACTGGTCTACCATGGGATT 175
176 CTTCGACAAATCTGAGTCTTCAGGCGTAGGTCACTACGTAGCTGTAACTTACAAATTC 233
234 TTA3’236

- **Annex 4 – BLASTN and BLASTP Tables**

Table 1 - *Results of the BLASTN analysis for the alignment of OmpK sequences of Vibrio parahaemolyticus ATCC 17802.*

Species	Strain	E Value	Identity	Species	Strains	E Value	Identity
Vibrio parahaemolyticus	VPD14	0.0	100.00%	***Vibrio cholerae***	O1	0.0	83.54%
	FDAARGOS_191	0.0	100.00%		Isolate CTMA_1441	0.0	83.54%
	RIMD 2210633 DNA	0.0	100.00%		2011V-1043	0.0	83.54%
	20140829008-1	0.0	98.75%		F9993	0.0	83.54%
	20140722001-1	0.0	98.75%		C6706	0.0	83.54%
	20140624012-1	0.0	98.75%		2015V-1118	0.0	83.54%
	O1:Kuk str. FDA_R31	0.0	98.75%		3569-08	0.0	83.54%
	BB220P	0.0	98.63%		2010EL-1786	0.0	83.54%
	STO11	0.0	100.00%		0395	0.0	83.54%
	CPVP15		97.50%		O1 biovar El Tor str.N16961	0.0	83.54%
Vibrio fluvialis					E4	0.0	83.54%
	1.1533	0.0	99.38%		FORC_073	0.0	83.54%
	2015AW-0233	0.0	82.88%		V060002 DNA	0.0	83.54%
Vibrio alginolyticus	2012V-1235	0.0	82.88%		A1552	0.0	83.54%
					N16961	0.0	83.54%
Vibrio harveyi	CPVA2	0.0	99.00%		4295STDY6534248	0.0	83.54%
					4295STDY6534232	0.0	83.54%
	2011V-1164	0.0	97.50%		4295STDY6534216	0.0	83.54%
	WXL538	0.0	97.50%		A 1552	0.0	83.54%
	Isolate QT520	0.0	97.50%		FDAARGOS_223	0.0	83.54%
	TrH020803	0.0	97.50%		O1 biovar El Tor str. HC1037	0.0	83.54%
	FDAARGOS_107	0.0	97.38%		E7946	0.0	83.54%
	EsHS020801	0.0	97.38%		A19	0.0	83.54%
	SpGY020601	0.0	97.25%		NCTC9420	0.0	83.54%
	AnGS020805	0.0	97.25%		NCTC5395	0.0	83.54%
	E05003	0.0	97.12%		M2140	0.0	83.54%
	EpGS020805	0.0	97.00%		E9120	0.0	83.54%
	ATCC 14126	0.0	96.99%		E1320	0.0	83.54%
	C05011	0.0	96.99%		E1162	0.0	83.54%
	E05006	0.0	96.87%		E506	0.0	83.54%
	NBRC15634	0.0	96.75%		CRC1106	0.0	83.54%
	NB1017	0.0	96.64%		CRC711	0.0	83.54%
	E05008	0.0	96.74%		C5	0.0	83.54%
	Zj2008	0.0	97.30%		Env-390	0.0	83.54%
	EiGR021101	0.0	91.77%		2012Env-9	0.0	83.54%
Vibrio campbellii	FJXUW2	0.0	91.02%		2740-80	0.0	83.54%
	NB1014	0.0	90.29%		O1 str. KW3	0.0	83.54%
					TSY216	0.0	83.54%
Vibrio owensii	15112C	0.0	91.90%		O1 biovar El Tor str.FJ147	0.0	83.54%
	LA16-V1	0.0	91.77%		2012EL-2176	0.0	83.54%
Vibrio vulnificus	CAIM 519 = NBRC 15631	0.0	86.59%		MS6 DNA	0.0	83.54%
					2011EL-301	0.0	83.54%
	051011B		91.40%		IEC224	0.0	83.54%
Vibrio mimicus		0.0			O1 str. 2010EL-1786	0.0	83.54%
	FJ03-X2				MJ-1236	0.0	83.54%
	FORC_037	0.0	86.52%		O395	0.0	83.54%
Vibrio cholerae		0.0	86.27%		M66-2	0.0	83.54%
	ATCC 33653				O1 biovar El Tor str. N16961	0.0	83.54%
	1.1969	0.0	84.91%				
		0.0	84.91%				
	RFB16						
	3541-04	0.0	83.92%	***Vibrio metoecus***	2011V-1015	0.0	83.17%
	2015V-1126	0.0	83.67%				
	20000	0.0	83.67%				
	LMA3894-4	0.0	83.67%				

Table 2 - *Results of the BLASTN analysis for the alignment of OmpW sequences of Vibrio parahaemolyticus ATCC 17802.*

Species	Strain	E Value	Identity	Strain	E Value	Identity	Species	Strains	E Value	Identity
	AM43962	0.0	100.00%	2013V-1244	0.0	95.82%	***Vibrio diabolicus***	FDAARGOS_105	0.0	93.65%
	MPV1	0.0	100.00%	AM51552	0.0	95.82%		FDAARGOS_96	0.0	93.50%
	VPD14	0.0	100.00%	FORC_072	0.0	95.67%		LMG 3418	0.0	93.19%
	FDAARGOS	0.0	100.00%	VPL 4-90	0.0	95.67%				
	ATCC 17802	0.0	100.00%	FORC_018	0.0	94.74%	***Vibrio antiquarius***	EX25	0.0	93.50%
	HA2	0.0	100.00%	FORC_008	0.0	94.74%				
	1682	0.0	100.00%	UCM-V493	0.0	94.74%				
	FORC 014	0.0	100.00%	20160303005-1	0.0	94.59%	***Vibrio neocaledonicus***	CGJ02-2	0.0	93.19%
	RIMD 2210633 DNA	0.0	100.00%	Vb0624	0.0	94.59%				
	19-021-D1	0.0	99.84%	PB1937	0.0	94.59%				
	FORC 022	0.0	99.84%	160807	0.0	94.44%	***Vibrio alginolyticus***	2439-01	0.0	92.88%
	2012AW-0224	0.0	99.69%	S107-1	0.0	94.44%		K09K1	0.0	92.57%
	2013V-1136	0.0	99.69%	FDAARGOS_667	0.0	93.19%		K08M3	0.0	92.57%
	FORC071	0.0	99.69%					K10K4	0.0	92.57%
	D3112	0.0	99.69%					K06K5	0.0	92.57%
	CHN25	0.0	99.69%					K05K4	0.0	92.57%
	01:Kuk str. FDA	0.0	99.69%					K04M5	0.0	92.57%
	zj2003	0.0	99.69%					K04M3	0.0	92.57%
	2013V-1181	0.0	99.53%					K04M1	0.0	92.57%
	MAVP-R	0.0	99.53%					K01M1	0.0	92.57%
	MAVP-Q	0.0	99.53%					2010V-1102	0.0	92.11%
	ST631	0.0	99.53%					FDAARGOS_108	0.0	91.95%
	FORC 006	0.0	99.53%					HY9901	0.0	91.95%
	BB22OP	0.0	99.53%					ZJ-T	0.0	90.39%
	20151116002-3	0.0	99.38%					2013V-1302	0.0	88.87%
Vibrio parahaemolyticus	20130629002S01	0.0	99.38%					2014V-1011	0.0	88.72%
	O1:K33 str. CDC K4557	0.0	99.38%					2014V-1072	0.0	88.72%
	20140829008-1	0.0	99.22%					FDAARGOS_114	0.0	88.72%
	20140722001-1	0.0	99.22%					FDAARGOS_110	0.0	88.72%
	20140624012-1	0.0	99.22%					NBRC 15630	0.0	88.72%
	2012AW-0353	0.0	99.22%					ATCC 33787	0.0	88.72%
	2012AW-0154	0.0	99.22%					VA2	0.0	88.25%
	FORC_004	0.0	99.22%							
	2010V-1106	0.0	98.91%							
	2014V-1125	0.0	98.91%				***Vibrio natriegens***	CCUG 16371	0.0	87.38%
	2013V-1146	0.0	98.91%					CCUG 16373	0.0	87.37%
	2014V-1066	0.0	98.91%					CCUG 16374	0.0	87.23%
	2015AW-0174	0.0	98.91%					NBRC 15636	0.0	87.19%
	10329	0.0	98.91%							
	FDAARGOS_662	0.0	98.91%							
	Isolate R13	0.0	98.91%				***Vibrio sp.***	EJY3	0.0	87.21%
	R14	0.0	98.91%					dhg	0.0	87.06%
	FDAARGOS_51	0.0	98.91%							
	MAVP-26	0.0	98.91%							
	2013V-1174	0.0	98.45%				***Vibrio campbellii***	170502	2e-157	82.77%
	FORC_023	0.0	97.98%					BoB-90	2e-157	82.77%
	AM46865	0.0	96.13%							

Table 3 - *Results of the BLASTN analysis for the alignment of OmpV sequences of Vibrio parahaemolyticus ATCC 17802.*

Species	Strain	E Value	Identity	Species	Strains	E Value	Identity
Vibrio parahaemolyticus	VPDA14	0.0	100.00%	***Vibrio alginolyticus***	2439-01	0.0	82.24%
	FDAARGOS_191	0.0	100.00%		2014V-1011	0.0	82.24%
	RIMD 2210633 DNA	0.0	100.00%		2014V-1072	0.0	82.24%
	D3112	0.0	99.61%		2013V-1302	0.0	82.24%
	20140829008-1	0.0	99.49%		FDAARGOS_108	0.0	82.24%
	20140722001-1	0.0	99.49%		FDAARGOS_114	0.0	82.24%
	20140624012-1	0.0	99.49%		ATCC 33787	0.0	82.24%
	2012AW-0353	0.0	99.49%		NBRC 15630=ATCC 17749	0.0	82.24%
	19-021-D1	0.0	99.49%		FDAARGOS_110	0.0	82.11%
	2012AW-0154	0.0	99.49%		K09K1	0.0	82.11%
	Vb0624	0.0	99.49%		K08M3	0.0	82.11%
	160807	0.0	99.49%		K10K4	0.0	82.11%
	FORC_072	0.0	99.49%		K06K5	0.0	82.11%
	FORC_018	0.0	99.49%		K05K4	0.0	82.11%
	FORC_022	0.0	99.49%		K04M5	0.0	82.11%
	FORC_008	0.0	99.49%		K04M3	0.0	82.11%
	UCM-V493	0.0	99.49%		K04M1	0.0	82.11%
	AM43962	0.0	99.49%		K01M1	0.0	82.11%
	MVP1	0.0	99.36%		2010V-1102	0.0	81.98%
	ATCC 17802	0.0	99.36%		ZJ-T	0.0	81.98%
	CHN25	0.0	99.36%		Zj2004	0.0	82.92%
	FORC_023	0.0	99.23%		HY9901	0.0	81.47%
	20160303005-1	0.0	99.10%				
	2012AW-0224	0.0	99.10%				
	2013V-1244	0.0	99.10%				82.24%
	2013V-1174	0.0	99.10%	***Vibrio antiquarius***	EX25	0.0	
	AM51552	0.0	99.10%				82.11%
	AM46865	0.0	99.10%	***Vibrio neocaledonicus***	CGJ02-2	0.0	
	FORC_071PB1937	0.0	99.10%				77.86%
	PB1937	0.0	99.10%				77.86%
	20151116002-3	0.0	98.97%	***Vibrio rotiferianus***	AM7 DNA	0.0	
	2013V-1181	0.0	98.97%		B64D1	0.0	76.66%
	20130629002S01	0.0	98.97%	***Vibrio harveyi***			
	MAVP-R	0.0	98.97%		C05011	0.0	
	MAVP-Q	0.0	98.97%				
	ST631	0.0	98.97%				
	FORC_006	0.0	98.97%				
	O1:K33 str. CDC_K4557	0.0	98.97%				
	HA2	0.0	98.84%				
	1682	0.0	98.84%				
	O1:Kuk str. FDA_R31	0.0	98.84%				
	BB22OP	0.0	98.84%				
	2010V-1106	0.0	98.71%				
	2014V-1125	0.0	98.71%				
	2013V-1146	0.0	98.71%				
	2014V-1066	0.0	98.71%				
	10329	0.0	98.71%				
	FDAARGOS_662	0.0	98.71%				
	FDAARGOS_51	0.0	98.71%				
	MAVP-26	0.0	98.71%				
	FORC_014	0.0	98.71%				
	2013V-1136	0.0	98.58%				
	2015AW-0174	0.0	98.58%				
	FDAARGOS_667	0.0	98.58%				
	Isolate R13	0.0	98.58%				
	R14	0.0	98.58%				
	S107-1	0.0	98.33%				
	FORC_004	0.0	98.33%				
Vibrio diabolicus	LMG 3418	0.0	82.50%				
	FDAARGOS_96	0.0	81.98%				
	FDAARGOS_105	0.0	81.98%				

Table 4 *- Results of the BLASTN analysis for the alignment of OmpU sequences of Vibrio parahaemolyticus ATCC 17802.*

Species	Strain	E Value	Identity
Vibrio parahaemolyticus	LVP1	0.0	100.00%
	FORC_072	0.0	100.00%
	2013V-1136	0.0	100.00%
	AM46865	0.0	100.00%
	20151116002-3	0.0	100.00%
	20130629002S01	0.0	100.00%
	160807	0.0	100.00%
	2013V-1244	0.0	100.00%
	AM51552	0.0	100.00%
	UCM-V493	0.0	100.00%
	20140829008-1	0.0	100.00%
	20140722001-1	0.0	100.00%
	20140624012-1	0.0	100.00%
	2010V-1106	0.0	100.00%
	2014V-1125	0.0	100.00%
	2013V-1146	0.0	100.00%
	2012AW-0353	0.0	100.00%
	2014V-1066	0.0	100.00%
	2015AW-0174	0.0	100.00%
	10329	0.0	100.00%
	FDAARGOS_662	0.0	100.00%
	S107-1	0.0	100.00%
	FDAARGOS_51	0.0	100.00%
	MAVP-26	0.0	100.00%
	2012AW-0154	0.0	100.00%
	VPD14	0.0	100.00%
	FDAARGOS_191	0.0	100.00%
	LPV66	0.0	100.00%
	2210633 DNA	0.0	100.00%
	LVP2	0.0	100.00%
	MVP1	0.0	100.00%
	2013V-1181	0.0	100.00%
	MAVP-Q	0.0	100.00%
	FORC_022	0.0	100.00%
	2012AW-0224	0.0	100.00%
	AM43962	0.0	100.00%
	MAVP-R	0.0	100.00%
	20160303005-1	0.0	100.00%
	PB1937	0.0	100.00%
	FORC_018	0.0	100.00%
	FORC_008	0.0	100.00%
	BB22OP	0.0	100.00%
	FORC_023	0.0	100.00%
	2012V-1165	0.0	100.00%
	AM51557	0.0	100.00%
	ATCC 17802	0.0	100.00%
	Isolate R13	0.0	100.00%
	R14	0.0	100.00%
	FORC_006	0.0	100.00%
	CHN25	0.0	100.00%
	FORC_014	0.0	100.00%
	FDAARGOS_667	0.0	100.00%
	19-021-D1	0.0	100.00%
	2012V-1116	0.0	100.00%
	2013V-1174	0.0	100.00%
	O1_Kuk str. FDA_R31	0.0	100.00%
			100.00%

Species	Strains	E Value	Identity
	HN18 porin	0.0	73%
	HN9 porin	0.0	73%
	VpA porin	0.0	73%
	VPL4-90	0.0	73%
Vibrio parahaemolyticus	1.1616 porin	0.0	73%
	HN176	0.0	73%
	1.1615 porin	0.0	73%
	4L	0.0	72%
	ATCC 17802	0.0	73%
	FDAARGOS_96	0.0	100%
	HA2	0.0	98%
	D3112	0.0	98%
	FORC_071	0.0	98%
	Vb0624	0.0	98%
	1682	0.0	97%
Vibrio antiquarius	EX25	0.0	100%
	FA3	0.0	100%
Vibrio diabolicus	LMG 3418	0.0	100%
	FDAARGOS_105	0.0	100%
Vibrio harveyi	345	0.0	81%
	Isolate QT520	0.0	81%
		0.0	
Vibrio alginolyticus	Ym4 DNA	0.0	100%
	YM19 DNA	0.0	100%
	VIO5 DNA	0.0	100%
	138-2 DNA	0.0	100%
	FDAARGOS_110	0.0	100%
	GS_MYPK1	0.0	100%
	2015AW-0011	0.0	100%
	K09K1	0.0	100%
	K08M3	0.0	100%
	K10K4	0.0	100%
	K06K5	0.0	100%
	K05K4	0.0	100%
	K04M5	0.0	100%
	K04M3	0.0	100%
	K04M1	0.0	100%
	K01M1	0.0	100%
Vibrio rotiferianus	AM7 DNA	0.0	81%
	B64D1	0.0	81%
Vibrio jasicida	090810c	0.0	81%

Table 5 - *Results of the BLASTP analysis for the alignment of OmpK sequences of Vibrio parahaemolyticus ATCC 17802.*

Description	E value	Identity	Description	E value	Identity	Description	E value	Identity
OmpK [*Vibrio parahaemolyticus*]	0.0	100.00 %	outer membrane protein K [*Vibrio parahaemolyticus*]	2.00 E-164	88.64%	MULTISPECIES: hypothetical protein [unclassified *Vibrio*]	2.00 E-158	81.82 %
MULTISPECIES: membrane protein [*Vibrio harveyi* group]	0.0	99.63%	hypothetical protein [*Vibrio parahaemolyticus*]	2.00 E-164	88.64%	membrane protein [*Vibrio jasicida*]	3.00 E-157	87.18 %
membrane protein [*Vibrio sp.* AND4]	0.0	97.80%	outer membrane protein K [*Vibrio alginolyticus*]	1.00 E-163	88.64%	membrane protein [*Vibrio breoganii*]	3.00 E-157	82.78 %
hypothetical protein [*Vibrio diabolicus*]	0.0	99.27%	MULTISPECIES: membrane protein [*Vibrio*]	1.00 E-163	83.64%	membrane protein [*Vibrio campbellii*]	4.00 E-157	87.18 %
hypothetical protein [*Vibrio harveyi*]	0.0	95.97%	outer membrane protein K [*Vibrio vulnificus* NBRC 15645 = ATCC 27562]	2.00 E-163	87.13%	hypothetical protein [*Vibrio breoganii*]	3.00 E-156	82.78 %
outer membrane protein K [*Vibrio harveyi*]	0.0	95.60%	membrane protein [*Vibrio vulnificus*]	2.00 E-163	86.81%	MULTISPECIES: membrane protein [*Vibrio*]	8.00 E-156	82.78 %
hypothetical protein [*Vibrio harveyi*]	0.0	95.60%	hypothetical protein [*Vibrio sp. T20*]	3.00 E-163	83.64%	membrane protein [*Photobacterium profundum*]	1.00 E-155	80.36 %
hypothetical protein [*Vibrio* sp. VPAP30]	0.0	93.41%	hypothetical protein [*Vibrio vulnificus*]	4.00 E-163	86.45%	membrane protein [*Vibrio coralliilyticus*]	1.00 E-155	82.42 %
hypothetical protein [*Vibrio bivalvicida*]	1.00 E-180	93.04%	membrane protein [*Vibrio cholerae*]	3.00 E-162	87.91%	membrane protein [*Vibrio cholerae*]	1.00 E-155	86.08 %
hypothetical protein [*Vibrio parahaemolyticus*]	2.00 E-175	88.64%	hypothetical protein [*Vibrio cholerae*]	5.00 E-162	87.91%	nucleoside-specific channel-forming protein [*Vibrio cholerae* RC385]	2.00 E-155	86.08 %
hypothetical protein [*Vibrio lentus*]	8.00 E-174	92.67%	hypothetical protein [*Vibrio cholerae*]	8.00 E-162	87.55%	membrane protein [*Vibrio coralliilyticus*]	2.00 E-155	82.42 %
MULTISPECIES: membrane protein [*Vibrio*]	9.00 E-173	92.31%	membrane protein [*Vibrio cholerae*]	9.00 E-162	87.55%	hypothetical protein [*Vibrio splendidus*]	5.00 E-155	80.95 %
MULTISPECIES: hypothetical protein [*Vibrio*]	1.00 E-171	91.58%	membrane protein [*Vibrio orientalis*]	1.00 E-161	86.08%	MULTISPECIES: hypothetical protein [*Vibrio*]	4.00 E-154	80.59 %
hypothetical protein [*Vibrio cholerae*]	1.00 E-170	86.81%	hypothetical protein [*Vibrio cholerae*]	1.00 E-161	87.91%	hypothetical protein [*Vibrio rotiferianus*]	7.00 E-154	86.81 %
membrane protein [*Vibrio natriegens*]	6.00 E-169	89.01%	hypothetical protein [*Vibrio metoecus*]	2.00 E-161	87.55%	MULTISPECIES: membrane protein [*Vibrio*]	1.00 E-153	85.71 %
hypothetical protein [*Vibrio natriegens*]	4.00 E-168	88.64%	MULTISPECIES: membrane protein [*Vibrio*]	2.00 E-161	87.55%	hypothetical protein [*Vibrio alginolyticus*]	1.00 E-153	80.95 %
membrane protein [*Vibrio owensii*]	5.00 E-167	90.11%	hypothetical protein [*Vibrio cholerae*]	3.00 E-161	87.18%	MULTISPECIES: membrane protein [*Vibrio*]	2.00 E-153	85.71 %
Multispecies	5.00 E-166	89.74%	hypothetical protein [*Vibrio cholerae*]	3.00 E-161	87.18%	MULTISPECIES: hypothetical protein [unclassified *Vibrio*]	3.00 E-153	80.22 %
hypothetical protein [*Vibrio parahaemolyticus*]	6.00 E-166	89.38%	membrane protein [*Vibrio cholerae*]	3.00 E-161	87.18%	membrane protein [*Vibrio anguillarum*]	4.00 E-153	80.95 %
outer membrane protein K [*Vibrio parahaemolyticus*]	9.00 E-166	89.01%	hypothetical protein D5E78_12925 [*Vibrio parahaemolyticus*]	4.00 E-161	88.64%	MULTISPECIES: membrane protein [*Vibrio*]	4.00 E-153	81.82 %
hypothetical protein [*Vibrio parahaemolyticus*]	1.00 E-165	89.01%	hypothetical protein [*Vibrio cholerae*]	4.00 E-161	87.55%	membrane protein [*Vibrio rotiferianus*]	6.00 E-153	85.35 %
MULTISPECIES: membrane protein [*Vibrio*]	1.00 E-165	89.38%	Outer membrane protein OmpK [*Vibrio mimicus* VM603]	4.00 E-161	87.18%	OmpK [*Vibrio alginolyticus*]	7.00 E-153	85.35 %
MULTISPECIES: membrane protein [*Vibrio*]	2.00 E-165	89.01%	outer membrane protein K [*Vibrio parahaemolyticus*]	4.00 E-161	85.82%	MULTISPECIES: hypothetical protein [*Vibrio*]	8.00 E-153	80.51 %
membrane protein [*Vibrio parahaemolyticus*]	2.00 E-165	89.01%	hypothetical protein [*Vibrio cholerae*]	5.00 E-161	87.18%	MULTISPECIES: hypothetical protein [*Vibrio*]	1.00 E-152	79.85 %
hypothetical protein [*Vibrio parahaemolyticus*]	3.00 E-165	88.64%	membrane protein [*Vibrio mimicus*]	6.00 E-161	87.18%	hypothetical protein [*Vibrio splendidus*]	1.00 E-152	80.87 %
membrane protein [*Vibrio parahaemolyticus*]	3.00 E-165	88.64%	MULTISPECIES: outer membrane protein OmpK [*Vibrio*]	8.00 E-161	87.18%	MULTISPECIES: hypothetical protein [*Vibrio*]	1.00 E-152	84.25 %
outer membrane protein [*Vibrio parahaemolyticus*]	3.00 E-165	88.64%	hypothetical protein [*Vibrio cholerae*]	8.00 E-161	87.18%	MULTISPECIES: hypothetical protein [*Vibrio*]	2.00 E-152	80.95 %
hypothetical protein [*Vibrio parahaemolyticus*]	3.00 E-165	88.64%	hypothetical protein [*Vibrio cholerae*]	9.00 E-161	87.18%	hypothetical protein [*Vibrio splendidus*]	2.00 E-152	80.51 %
hypothetical protein [*Vibrio parahaemolyticus*]	4.00 E-165	88.64%	membrane protein [*Vibrio cholerae*]	1.00 E-160	87.18%	outer membrane protein K [*Vibrio alginolyticus*]	2.00 E-152	85.35 %
hypothetical protein [*Vibrio parahaemolyticus*]	4.00 E-165	88.64%	hypothetical protein [*Vibrio sp. PID23_8*]	9.00 E-160	86.45%	hypothetical protein [*Vibrio anguillarum*]	2.00 E-152	80.59 %
hypothetical protein [*Vibrio parahaemolyticus*]	5.00 E-165	88.64%	hypothetical protein [*Aliivibrio fischeri*]	1.00 E-159	80.66%	hypothetical protein [*Vibrio splendidus*]	3.00 E-152	81.68 %
hypothetical protein [*Vibrio parahaemolyticus*]	7.00 E-165	88.64%	hypothetical protein [*Vibrio salilacus*]	2.00 E-159	84.62%	MULTISPECIES: membrane protein [*Vibrio*]	3.00 E-152	81.68 %
MULTISPECIES: membrane protein [*Vibrio*]	9.00 E-165	89.01%						
hypothetical protein [*Vibrio parahaemolyticus*]	1.00 E-164	88.64%						
hypothetical protein [*Vibrio parahaemolyticus*]	1.00 E-164	88.64%						

Table 6 - *Results of the BLASTP analysis for the alignment of OmpW sequences of Vibrio parahaemolyticus ATCC 17802.*

Description	E value	Identity	Description	E value	Identity	Description	E value	Identity
MULTISPECIES: outer membrane protein OmpW [*Vibrio*]	9.00E-156	100.00%	MULTISPECIES: outer membrane protein OmpW [*Vibrio*]	2.00E-150	96.73%	outer membrane protein OmpW [*Vibrio alginolyticus*]	6.00E-143	92.52%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	1.00E-155	99.53%	outer membrane protein OmpW [*Vibrio parahaemolyticus*]	7.00E-150	96.26%	outer membrane protein OmpW [*Vibrio harveyi*]	6.00E-143	91.12%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	1.00E-155	99.53%	outer membrane protein OmpW [*Vibrio parahaemolyticus*]	1.00E-149	96.26%	outer membrane protein OmpW [*Vibrio* natriegens]	2.00E-142	92.52%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	2.00E-155	99.53%	outer membrane protein OmpW [*Vibrio parahaemolyticus*]	1.00E-149	96.26%	outer membrane protein OmpW [*Vibrio* hyugaensis]	2.00E-142	90.65%
MULTISPECIES: outer membrane protein OmpW [*Vibrio*]	3.00E-155	99.53%	outer membrane protein OmpW [*Vibrio parahaemolyticus*]	1.00E-149	95.79%	outer membrane protein OmpW [*Vibrio alginolyticus*]	9.00E-142	92.06%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	3.00E-155	99.53%	outer membrane protein OmpW [*Vibrio parahaemolyticus*]	2.00E-149	96.26%	outer membrane protein OmpW [*Vibrio campbellii*]	1.00E-141	90.65%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	3.00E-155	99.53%	MULTISPECIES: outer membrane protein OmpW [*Vibrio*]	3.00E-149	95.33%	outer membrane protein OmpW [*Vibrio campbellii*]	1.00E-141	90.65%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	3.00E-155	100.00%	MULTISPECIES: outer membrane protein OmpW [*Vibrio harveyi* group]	2.00E-148	94.86%	Outer membrane protein OmpW [*Vibrio diabolicus*]	2.00E-141	94.63%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	3.00E-155	99.53%	MULTISPECIES: outer membrane protein OmpW [*Vibrio*]	2.00E-148	94.86%	outer membrane protein W [*Vibrio alginolyticus*]	2.00E-141	91.59%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	3.00E-155	99.53%	outer membrane protein OmpW [*Vibrio parahaemolyticus*]	1.00E-146	95.33%	outer membrane protein OmpW [*Vibrio parahaemolyticus*]	3.00E-141	91.59%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	3.00E-155	99.53%	outer membrane protein OmpW [*Vibrio* rotiferianus]	1.00E-146	93.46%	outer membrane protein OmpW [*Vibrio campbellii*]	7.00E-141	90.19%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	3.00E-155	99.53%	outer membrane protein OmpW [*Vibrio alginolyticus*]	1.00E-146	93.46%	outer membrane protein OmpW [*Vibrio* sp. OY15]	7.00E-141	91.59%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	3.00E-155	99.53%	outer membrane protein OmpW [*Vibrio parahaemolyticus*]	1.00E-146	96.19%	outer membrane protein OmpW [*Vibrio* hyugaensis]	9.00E-141	90.65%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	3.00E-155	99.53%	outer membrane protein OmpW [*Vibrio* rotiferianus]	3.00E-146	93.46%	outer membrane protein OmpW [*Vibrio campbellii*]	1.00E-140	90.19%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	4.00E-155	99.53%	outer membrane protein OmpW [*Vibrio* rotiferianus]	4.00E-146	93.46%	outer membrane protein OmpW [*Vibrio* hyugaensis]	1.00E-140	89.72%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	6.00E-155	99.07%	major outer membrane protein W [Photobacterium damselae]	4.00E-146	94.86%	outer membrane protein OmpW [*Vibrio campbellii*]	1.00E-140	89.72%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	7.00E-155	99.53%	outer membrane protein OmpW [*Vibrio parahaemolyticus*]	4.00E-146	94.86%	MULTISPECIES: outer membrane protein OmpW [*Vibrio*]	2.00E-140	91.59%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	7.00E-155	99.53%	MULTISPECIES: outer membrane protein OmpW [*Vibrio*]	4.00E-146	92.99%	outer membrane protein OmpW [*Vibrio parahaemolyticus*]	2.00E-140	91.59%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	8.00E-155	99.53%	outer membrane protein OmpW [*Vibrio parahaemolyticus*]	4.00E-146	94.86%	outer membrane protein W [*Vibrio alginolyticus*]	2.00E-148	94.39%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	8.00E-155	99.53%	outer membrane protein W [*Vibrio harveyi*]	5.00E-146	92.99%	outer membrane protein OmpW [*Vibrio alginolyticus*]	5.00E-148	94.39%
hypothetical protein ACX10_22960 [*Vibrio parahaemolyticus*]	1.00E-154	99.53%	outer membrane protein OmpW [*Vibrio parahaemolyticus*]	7.00E-146	94.86%	MULTISPECIES: outer membrane protein OmpW [*Vibrio*]	7.00E-148	94.39%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	1.00E-154	99.53%	outer membrane protein OmpW [*Vibrio harveyi*]	8.00E-146	92.52%	outer membrane protein OmpW [*Vibrio alginolyticus*]	7.00E-148	94.39%
membrane protein [*Vibrio parahaemolyticus*]	2.00E-154	99.53%	outer membrane protein OmpW [*Vibrio harveyi*]	1.00E-145	92.52%	outer membrane protein OmpW [*Vibrio parahaemolyticus*]	7.00E-148	99.51%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	2.00E-154	99.53%	outer membrane protein OmpW [*Vibrio harveyi*]	2.00E-145	92.52%	membrane protein [*Vibrio parahaemolyticus*]	8.00E-148	96.67%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	3.00E-154	100.00%	outer membrane protein OmpW [*Vibrio harveyi*]	2.00E-145	92.52%	outer membrane protein OmpW [*Vibrio parahaemolyticus*]	8.00E-148	95.79%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	3.00E-154	99.07%	outer membrane protein OmpW [*Vibrio harveyi*]	2.00E-145	92.52%	outer membrane protein OmpW [*Vibrio diabolicus*]	1.00E-147	94.84%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	4.00E-154	99.07%	outer membrane protein OmpW [*Vibrio* sp. ES.051]	3.00E-145	92.99%	outer membrane protein OmpW [*Vibrio parahaemolyticus*]	3.00E-147	95.79%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	3.00E-153	100.00%	outer membrane protein OmpW [*Vibrio harveyi*]	4.00E-145	92.52%	outer membrane protein OmpW [*Vibrio alginolyticus*]	4.00E-147	94.39%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	3.00E-152	99.52%	MULTISPECIES: outer membrane protein OmpW [*Vibrio*]	6.00E-145	93.93%	outer membrane protein OmpW [*Vibrio* rotiferianus]	8.00E-147	93.93%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	4.00E-152	98.60%	outer membrane protein OmpW [*Vibrio alginolyticus*]	1.00E-144	93.46%	outer membrane protein OmpW [*Vibrio parahaemolyticus*]	9.00E-147	95.33%
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	9.00E-152	100.00%	outer membrane protein OmpW [*Vibrio* jasicida]	4.00E-144	92.06%			
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	1.00E-151	97.20%	MULTISPECIES: outer membrane protein OmpW [*Vibrio harveyi* group]	6.00E-144	91.59%			
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	7.00E-151	96.26%	outer membrane protein OmpW [*Vibrio* natriegens]	8.00E-144	93.46%			
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	2.00E-150	96.73%	outer membrane protein OmpW [*Vibrio alginolyticus*]	1.00E-143	92.99%			
outer membrane protein OmpW [*Vibrio parahaemolyticus*]	2.00E-150	100.00%	MULTISPECIES: outer membrane protein OmpW [*Vibrio*]	1.00E-143	91.59%			

Table 7 - *Results of the BLASTP analysis for the alignment of OmpV sequences of Vibrio parahaemolyticus ATCC 17802.*

Description	E value	Identity	Description	E Value	Identity	Description	E Value	Identity
MULTISPECIES: MipA/OmpV family protein [*Vibrio*]	0.0	100.00 %	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	membrane protein [*Vibrio parahaemolyticus*]	0.0	99.22%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	98.84%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	98.84%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	98.84%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	98.84%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	98.84%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	98.84%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	98.84%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	98.84%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	98.84%
major outer membrane protein V [Photobacterium damselae]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	Membrane protein [*Vibrio parahaemolyticus* VP766]	0.0	98.84%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MULTISPECIES: MipA/OmpV family protein [*Vibrio*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	98.84%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	98.84%
Membrane protein [*Vibrio parahaemolyticus* VP49]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	98.84%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	98.45%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	98.06%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	98.45%
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%			
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%			
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%			
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.22%			
MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%	MipA/OmpV family protein [*Vibrio parahaemolyticus*]	0.0	99.61%			

Table 8 - *Results of the BLASTP analysis for the alignment of OmpU sequences of Vibrio parahaemolyticus ATCC 17802.*

Description	E value	Identity	Description	E value	Identity	Description	E value	Identity
MULTISPECIES: porin [*Vibrio*]	0.0	100.00%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%
MULTISPECIES: porin [*Vibrio*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.07%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.07%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%	porin [*Vibrio parahaemolyticus*]	0.0	99.07%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	MULTISPECIES: porin [*Vibrio*]	0.0	99.38%	porin [*Vibrio parahaemolyticus*]	0.0	99.07%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%	membrane protein [*Vibrio parahaemolyticus*]	0.0	100.00%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.68%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio alginolyticus*]	0.0	95.67%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%	MULTISPECIES: porin [*Vibrio*]	0.0	95.36%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%	MULTISPECIES: porin [*Vibrio*]	0.0	95.05%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%	MULTISPECIES: porin [*Vibrio*]	0.0	95.36%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio alginolyticus*]	0.0	95.05%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%	porin [*Vibrio alginolyticus*]	0.0	95.05%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%	porin [*Vibrio alginolyticus*]	0.0	94.74%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%	hypothetical protein ACX03_24300 [*Vibrio parahaemolyticus*]	0.0	95.05%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%	porin [*Vibrio* sp. ES.044]	0.0	95.05%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%	MULTISPECIES: porin [*Vibrio*]	0.0	95.05%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio diabolicus*]	0.0	95.05%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%	porin [*Vibrio* antiquarius]	0.0	94.74%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio diabolicus*]	0.0	94.74%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%	LOW QUALITY PROTEIN: putative outer membrane protein [*Vibrio* sp. JCM 19053]	0.0	94.75%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%	MULTISPECIES: porin [*Vibrio diabolicus* subgroup]	0.0	94.74%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%	porin [*Vibrio* sp. ArtGut-C1]	0.0	94.43%
hypothetical protein ACX10_14425 [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%	MULTISPECIES: porin [*Vibrio*]	0.0	94.43%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%	porin [*Vibrio* sp. ES.051]	0.0	91.33%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%	porin [*Vibrio alginolyticus*]	0.0	91.38%
porin [*Vibrio parahaemolyticus*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.38%			
MULTISPECIES: porin [*Vibrio*]	0.0	99.69%	porin [*Vibrio parahaemolyticus*]	0.0	99.69%			