DEVELOPMENT AND IMMUNOLOGICAL EVALUATION OF SARS-COV-2 VIRUS-LIKE PARTICLES (VLP) EXPRESSING THE PREFUSION STABILIZED VARIANTS OF THE SPIKE PROTEIN

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

EMRE MERT İPEKOĞLU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
MOLECULER BIOLOGY AND GENETICS

JANUARY 2022

Approval of the thesis:

DEVELOPMENT AND IMMUNOLOGICAL EVALUATION OF SARS-COV-2 VIRUS-LIKE PARTICLES (VLP) EXPRESSING THE PREFUSION STABILIZED VARIANTS OF THE SPIKE PROTEIN

submitted by **EMRE MERT İPEKOĞLU** in partial fulfillment of the requirements for the degree of **Master of Science** in **Molecular Biology and Genetics, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. Ayşegül Çetin Gözen
Head of the Department, **Biological Sciences** _____

Prof. Dr. Mayda Gürsel
Supervisor, **Biological Sciences, METU** _____

Examining Committee Members:

Asst. Prof. Dr. Tolga Sütlü
Molecular Biology and Genetics, Boğaziçi University _____

Prof. Dr. Mayda Gürsel
Biological Sciences, METU _____

Asst. Prof. Dr. Banu Bayyurt Kocabaş
Biological Sciences, METU _____

Date: 24.01.2022

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name Last name : Emre Mert İpekoğlu

Signature :

ABSTRACT

DEVELOPMENT AND IMMUNOLOGICAL EVALUATION OF SARS-COV-2 VIRUS-LIKE PARTICLES (VLP) EXPRESSING THE PREFUSION STABILIZED VARIANTS OF THE SPIKE PROTEIN

İpekoğlu, Emre Mert
Master of Science, Molecular Biology and Genetics
Supervisor : Prof. Dr. Mayda Gürsel

January 2022, 129 pages

Emergence of the COVID-19 pandemic necessitated rapid development of highly effective vaccines against SARS-CoV-2. One adaptable strategy to produce a vaccine candidate against SARS-CoV-2 is based on the exploitation of the self-assembly feature of the structural proteins of SARS-CoV-2: Spike, Membrane, Envelope and Nucleocapsid, forming Virus-like particles (VLP). Our early studies focused on laboratory-scale production of SARS-CoV-2 VLPs in a mammalian host cell system based on transient expression of viral proteins. Herein, we describe a process development framework for scalable production and purification of SARS-CoV-2 VLPs expressing either the prefusion stabilized (2p) or the thermostable six proline (6p; HexaPro) stabilized Spike proteins. Results showed that suspension culture-adapted HEK 293 cells were able to produce vesicular VLPs that were similar in size to the authentic SARS-CoV-2 virions. Combined use of multimodal chromatography, nucleic acid removal through Denarase treatment and tangential flow filtration,enabled large-scale production of the SARS-CoV-2 VLPs.

In this study, we also produced and affinity purified 6x histidine-tagged recombinant Spike (6p) protein of SARS-CoV-2 to conduct in-house serological assays to monitor VLP vaccine-induced anti-Spike immunoglobulin response in immunized mice. Our results demonstrated that a modified purification protocol was superior to conventional affinity purification technique for the production of the secreted recombinant protein in terms of reproducibility and standardization of ELISA experiments.

Finally, we examined the immunogenicity of 2p and 6p Spike harbouring VLPs combined with vaccine adjuvants K3 CpG ODN and/or Alum. Collectively, our results suggest that Alum adsorbed, K3 CpG adjuvanted high dose 6p VLPs elicited robust humoral and cellular immune responses in vaccinated BALB/c mice as compared to non-adjuvanted, low dose or 2p VLP counterparts, proving the effectiveness of this formulation as a vaccine candidate against SARS-CoV-2.

Keywords: SARS-CoV-2, Virus-like particles, Vaccine, K3 CpG ODN, COVID-19

ÖZ

PREFÜZYON STABİLİZE SPİKE PROTEİNİ VARYANTLARINI TAŞIYAN SARS-COV-2 VİRÜS BENZERİ PARÇACIKLARIN GELİŞTİRİLMESİ VE İMMÜNOLOJİK DEĞERLENDİRİLMESİ

İpekoğlu, Emre Mert
Yüksek Lisans, Moleküler Biyoloji ve Genetik
Tez Yöneticisi: Prof. Dr. Mayda Gürsel

Ocak 2022, 129 sayfa

Coronavirüs hastalığı 2019 (COVID-19), 2019'un sonlarında Çin'in Wuhan kentinde ortaya çıkmıştır ve nedensel ajanı Şiddetli Akut Solunum Sendromu Coronavirüs 2 (SARS-CoV-2) olarak tanımlanmıştır. Bu son derece bulaşıcı virüs, COVID-19 pandemisine yol açan yeni bir pnömoni salgınına neden olmuştur. COVID-19 pandemisinin ortaya çıkması, SARS-CoV-2'ye karşı etkili bir aşıya acilen ihtiyaç duyulmasına sebebiyet vermiştir. SARS-CoV-2'ye karşı bir aşı adayı üretmek için etkili bir strateji, SARS-CoV-2'nin yapısal proteinlerinin (Spike, Membran, Zarf ve Nükleokapsid proteinlerinin) kendi kendine bir araya gelme özelliğinin kullanılarak Virüs benzeri partikülleri (VLP) oluşturmasına dayanmaktadır. İlk çalışmalarımız, memeli konakçı hücre ifadesi için virüsün dört yapısal proteinini kodlayan genlerinin transfeksiyonu yoluyla memeli hücresinden üretilen immünojenik SARS-CoV-2 VLP'lerin laboratuvar ölçeğinde üretimine odaklanmıştır. Burada, prefüzyon stabilize (2p) veya termostabil altı prolin (6p; HexaPro) Spike proteinini barındıran SARS-CoV-2 VLP'lerin ölçeklenebilir üretimi ve saflaştırılması için bir süreç geliştirme çerçevesi açıkladık. Sonuçlarımız, süspansiyon kültüre adapte edilmiş

HEK hücrelerinin, çok-modlu kromatografi ve teğetsel akış filtrasyonu gibi ölçeklenebilir saflaştırma tekniklerinin uygulanmasının ardından, otantik SARS-CoV-2 virionlarına benzer boyutlarda dört yapısal proteini içeren VLP'ler üretebildiğini göstermiştir.

Bu çalışmada ayrıca, farelerde VLP'nin ortaya çıkardığı anti-Spike immünoglobulinlerini izlemek ve serolojik testler yapmak için SARS-CoV-2'nin 6x histidin etiketli rekombinant Spike (6p) proteinini ürettik ve afinite saflaştırması yöntemi ile elde ettik. Sonuçlarımız, ELISA deneylerinin tekrarlanabilirliği ve standardizasyonu açısından, Spike proteininin saflaştırılması için uyarlanmış güncel bir protokolün, salgılanan rekombinant proteinler için kullanılan geleneksel afinite saflaştırma tekniklerinden üstün olduğunu göstermiştir.

Son olarak, K3 CpG ODN ve/veya Alum içeren aşı adjuvanları ile kombinasyon halinde, 2p ve 6p Spike proteinlerini barındıran VLP'lerin bağışıklık koruyucu potansiyelini inceledik. Toplu sonuçlarımız Alum tarafından adsorbe edilmiş, K3 CpG ile adjuvanlanmış yüksek doz 6p VLP'lerin, adjuvanlanmamış, düşük doz veya 2p VLP karşıtlarına kıyasla aşılanmış BALB/c farelerde güçlü hümoral ve hücresel bağışıklık tepkileri ortaya çıkardığını ve SARS-CoV-2'ye karşı bir etkili aşı adayı olabileceğini ortaya koyarken, aynı zamanda bu aşı adayının insan klinik deneylerinde değerlendirilmesi için bir gerekçe sağlamaktadır.

Anahtar Kelimeler: SARS-CoV-2, Virüs benzeri parçacıklar, Aşı, K3 CpG ODN, COVID-19

Ozancanıma

ACKNOWLEDGMENTS

First of all, I would like to express my gratitude to my advisor Prof. Dr. Mayda Gürsel. Her approach to science, life and humanity always inspired me and I consider myself quite lucky to be one of her students and I will always be one of her students. Prof. Dr. İhsan Gürsel was also one of the architects of this process. Without his guidance, I could not survive to the end of this process. Many thanks and those would never be enough.

I also would like to thank my thesis committee members Asst. Prof. Dr. Tolga Sütlü and Asst. Prof. Dr. Banu Bayyurt Kocabaş for their valuable suggestions and comments.

I would like to thank all the MG and IG lab members for the excellent camaraderie that we created together. I will never find this spirit in any other workplace.

Special thanks to Dr. İhsan Cihan Ayanoğlu and Dr. İsmail Cem Yılmaz, my elder brothers, for their contributions to my academic maturation. It was a great pleasure to work with them and thankful for everything they taught me.

I could not be who I am without my friends and my family; those are two concepts I could not separate from each other. Furkan, Nur, Şule, Ayberk, Alper and Koray, you know that I cannot describe my feelings with these poor words, I will just say thank you and I know that you will understand. Thank you Senem and Nuri Akkuş for raising excellent children as Ozancan, Ceren and Azra, my precious brother and sisters. Thank you for surviving us and do not worry; we will survive for you. I owe everything to my parents, my grandparents and my little brother, Nadire Gül İpekoğlu, Hayrettin İpekoğlu, Dudu İpekoğlu, Arif İpekoğlu and Arif Yiğit İpekoğlu. I will always work harder to make you proud. Thank you for believing me.

Finally, very special thanks to my better half, Nilsu, for her endless support and patience. Kissdim...

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

1.1 COVID-19: Pandemic Coronavirus Disease

Coronavirus disease 2019 (COVID-19) has emerged in southeast China, the city of Wuhan in December 2019. The causative agent of COVID-19 was identified as Severe Acute Respiratory Syndrome 2 Virus (SARS-CoV-2). Having been highly transmissible, it has caused an outbreak of novel pneumonia and spread quickly worldwide (Wu et al., 2020). On 12th March of 2020, World Health Organization (WHO) declared this outbreak as a pandemic (WHO, 2020). Herein, characteristics of the virus, its life cycle, pathogenesis and modes of transformation will be briefly examined.

1.1.1 General Characteristics of Severe Acute Respiratory Syndrome 2 Virus (SARS-CoV-2)

Coronaviruses belong to the family of Coronaviridae in the order of Nidoviridales that target mainly respiratory tract in humans and animals. Alpha and beta coronaviruses have been the center of attention among the four genera (alpha, beta, gamma, and delta) due to their ability to cross-contaminate between various host families and most importantly become a threatening pathogen since their emergence (Frutos et al., 2021). Some coronaviruses, including the hCoV-HKU1, hCoV-OC43, hCoV-NL63 and hCoV-229E have been in human circulation for years. Such coronaviruses account for 5-30 % of the common colds with mild symptoms (Sahu et al., 2021). In contrast, beta coronaviruses such as SARS-CoV and MERS-CoV have emerged more recently and were shown to possess high infectivity and caused

high mortality in epidemics described in China and Middle East, respectively. After almost 20 years of the first SARS-CoV epidemic, a novel but related coronavirus, namely the Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2), caused a world-wide pandemic (Hu et al., 2021). Whole-genome sequencing of the SARS-CoV-2 genome showed 79% and 50 % similarity to SARS-CoV and MERS-CoV, respectively (Zhu et al., 2020).

SARS-CoV-2 virions are spherical with a size range of 80 to 160 nanometers (Ke et al., 2020). Lipid enveloped virions consist of four structural proteins: Spike, Envelope, Membrane and Nucleocapsid. The outer envelope surface is decorated with homotrimeric Spike (S) protein (figure 1.1), whereas membrane glycoprotein and envelope proteins are embedded inside the lipidic envelope of the virus particles. Nucleocapsid protein is the only protein located inside the virion, forming a helical structure with genomic single-stranded RNA (positive strand) of the virus (Yang et al., 2020). The size of the genomic RNA of the SARS-CoV-2 is around 30 kilobases (kb) with a unique tertiary structure constituting the longest largest genome of the known RNA viruses (Cao et al., 2021). Single-stranded RNA of the coronavirus genome is capped at its 5' end along with a 3' poly-adenine tail, resembling the mammalian mRNA transcripts. The genomic structure of the SARS-CoV-2 consists of six open reading frames (ORF), including the ORF1a and ORF1b which encode the non-structural proteins (nsp) of the coronavirus (nsp1-16) required for viral replication and life cycle in the host. The rest of the ORFs are responsible for the production of the four structural proteins (S, M, E, N) along with accessory proteins (3a, 3b, 6, 7a, 7b, 8, 9b, 9c 10) that are indispensable for the virion assembly (Chen et al., 2020). The virion and genomic structure of SARS-CoV-2 is illustrated in figures 1.1 and 1.2, respectively.

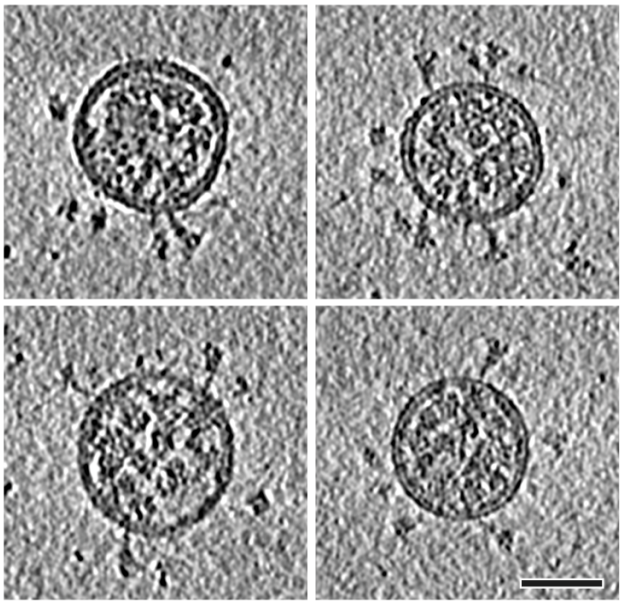


Figure 1.1. Cryo-electron tomography of SARS-CoV-2 Virions

Adapted from (Ke et al., 2020)

Spherical virions with granular density at their center corresponding to genomic RNA-Nucleocapsid complex along with protruding trimeric Spike proteins (scale bar, 50 nm).

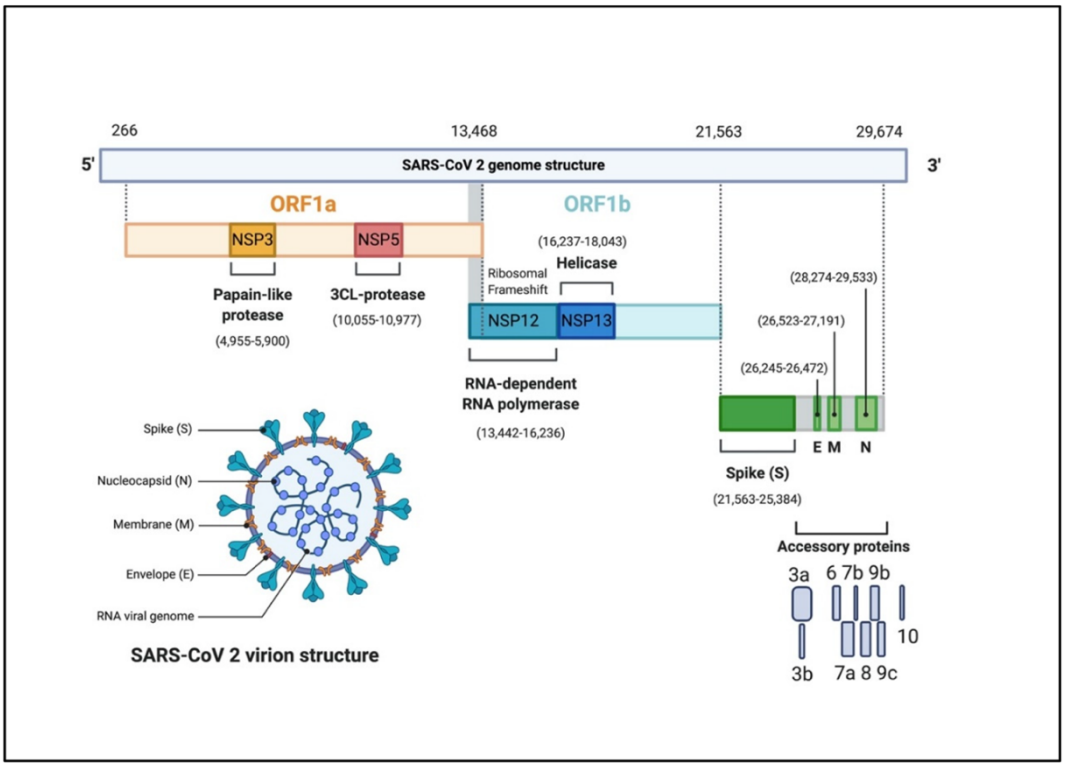


Figure 1.2. Virion structure and genomic organization of SARS-CoV-2.

Adapted from (Alanagreh et al., 2020)

ORF1a (yellow), ORF1b (blue) encode 16 non-structural proteins (NSP1– NSP16). The structural genes encode the structural proteins, spike (S), envelope (E), membrane (M), and nucleocapsid (N) (green). The unique accessory proteins of SARS-CoV-2 are shown in grey.

1.1.1.1 Life Cycle of SARS-CoV-2

Infection and the life cycle of the SARS-CoV-2 starts with the specific binding of the virus to the host cell membrane. This binding is provided by the interaction between the Spike protein of the virus and the host cell surface receptor angiotensin-converting enzyme 2 (ACE2). Viral entry inside the cell is often influenced by the spatiotemporal distribution of the receptor on the cell membrane and host cell facilitators, including the serine protease, TMPRSS2 (Hoffmann et al., 2020).

Following entry, genomic RNA is released into the cytosol and ORF1a and ORF1b expression commences immediately to generate replication-transcription complexes (RTC) through primary translation and polyprotein processing of the non-structural proteins. Replicase complex then starts to synthesize viral RNA in a microenvironment formed by intracellular vesicular structures to protect the nascent genome and the newly synthesized sub-genomic mRNAs. These sub-genomic mRNAs provide the expression of accessory proteins for the inhibition of antiviral host cell responses (Redondo et al., 2021). Subsequently, translated structural proteins translocated to the endoplasmic reticulum membrane to be destined to ER-Golgi intermediate complex (ERGIC) for initializing the virion assembly when nucleocapsid-nascent genomic RNA complex is formed. This interaction is also a signal for budding into the lumen of the secretory vesicles. In the end, virions inside secretory vesicles are released from the infected cell via exocytosis (V'kovski et al., 2021). The intracellular life cycle of the SARS-CoV-2 is summarized in figure 1.3.

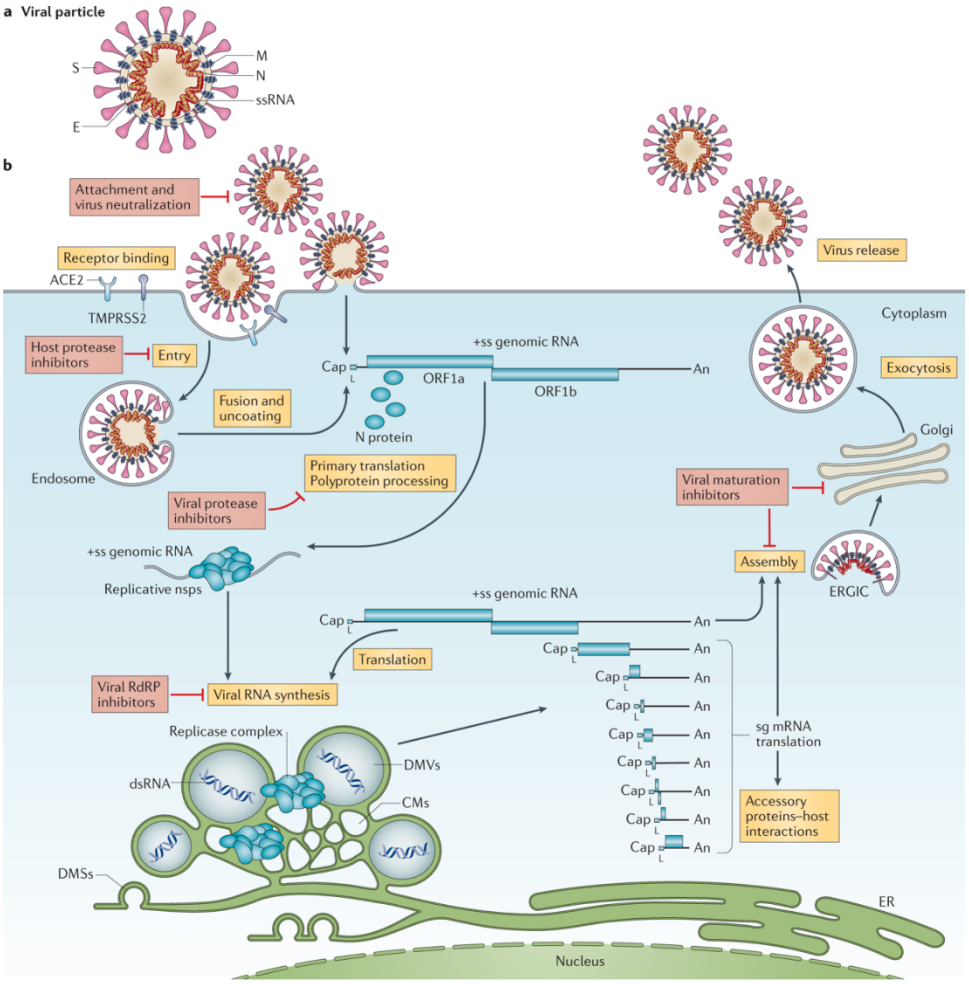


Figure 1.3. The intracellular life cycle of SARS-CoV-2.

Adapted from (V'kovski et al., 2021)

1.1.1.2 Pathogenesis and Modes of Transmission of SARS-CoV-2

Generally, common cold viruses have a tendency to target the upper respiratory tract. In contrast, one of the distinctive features of highly infectious coronaviruses such as SARS-CoV-2 is pulmonary involvement, causing severe pneumonia along with other systemic problems, including renal failure, gastrointestinal disorders and neurological symptoms or long-term effects like myocardial inflammation (D. Wang et al., 2020) (Puntmann et al., 2020). Clinical manifestations of the COVID-19 include respiratory distress, cough, coarse breathing sounds in both lungs and

elevated body temperature. Blood tests of the patients have also illustrated the elevated C-reactive protein and D-dimer levels along with excessive amounts of pro-inflammatory cytokine production, causing a phenomenon known as a 'cytokine storm', along with an increased count of the leukocytes (Ragab et al., 2020). All in all, it has been reported that COVID-19 manifests flu-like symptoms at the initial phase of the infection and could progress into multi-system organ failure and inevitably, to death (Rothan & Byrareddy, 2020).

Several modes of transmission of SARS CoV-2 have been defined, including contact, droplet, airborne, fomite, fecal-oral, bloodborne, mother-to-child, and animal-to-human transmission (World Health Organization Europe (WHO Europe), 2020). The most common transmission route was reported as direct contact and through droplets (Ghinai et al., 2020). Respiratory droplets secreted from an infected individual could cause the transmission of the viral particles directly to the uninfected or these droplets could be taken from the surfaces, also known as fomite-mediated transmission (Pastorino et al., 2020). The presence of replicative virions in aerosols has been proven experimentally, yet, whether or not such aerosols contribute to real-life transmission, remains inconclusive. Therefore, the route of airborne transmission was not elucidated in terms of spreading the viable virions among individuals (Liu et al., 2020). Fecal-oral and bloodborne transmission routes are of low probability (W. Wang et al., 2020) (Chang et al., 2020). Although viral RNA could be detected in mother's breastmilk, it has been observed that there was no replicative virion isolated (Centeno-Tablante et al., 2021). Finally, up-to-date evidence suggests that SARS-CoV-2 could infect other mammals. Nevertheless, it is not completely elucidated whether these animals represent a risk factor for human transmission. SARS-CoV-2 was believed to emerge from a beta coronavirus circulating in the horseshoe bats which then transmitted to humans by an intermediate reservoir. No consensus has been reached by the authorities on the identity of this intermediary reservoir (Hobbs & Reid, 2021).

1.2 Immune responses against SARS-CoV-2

1.2.1 Innate Immunity

Upon infection of the alveolar airway epithelial cells by SARS-CoV-2, the virus is recognized by the immune system's first line of defense, the innate immunity. Pathogen recognition receptors (PRRs) expressed by these epithelial cells and tissue-resident immune cells interact with viral components, resulting in the initiation of an antiviral response. These PRRs could be endosome and lysosome-associated Toll-like receptors or cytosolic sensor proteins that are specialized to recognize pathogen-associated molecular patterns (PAMPs) (Medzhitov & Janeway C., 2000). Upon entry, genomic RNA of the virus could be recognized by Toll-like receptors (TLR) 3, 7 and 8. TLR7 and TLR8 recognize the single-stranded genomic RNA of the virus, whereas TLR3 could interact with double-stranded RNA that is synthesized during the viral replication inside double-membrane vesicles. TLR7 and TLR8 share a common downstream adapter named myeloid differentiation factor 88 (MyD88), that mediates the activation and nuclear translocation of nuclear factor kappa light-chain-enhancer of activated B cells (NF-κB), resulting in the expression of pro-inflammatory cytokines (Kawai & Akira, 2011). In addition, TLR3 sensing of double-stranded RNA results in the activation of the TIR domain-containing adapter protein (TRIF) mediated NF-κB translocation, providing synergistic recognition of viral RNA via endosomal receptors. These pathways are also in cross-talk with interferon regulatory factor-3 (IRF3) and interferon regulatory factor-7 (IRF7) phosphorylation pathway through tank-binding kinase 1 (TBK1) activity, resulting in the production of the type I and type III interferons. Furthermore, cytosolic viral RNAs, including the subgenomic mRNAs can be sensed by cytosolic nucleic acid sensing receptors such as the retinoic acid-inducible gene I (RIG-I) and the melanoma differentiation-associated protein 5 (MDA5), which then stimulate an antiviral response through the mitochondrial antiviral signaling protein (MAVS) activation (Kumar et al., 2011). MAVS recruits other kinases and enables

phosphorylation, homodimerization and nuclear translocation of IRF3 and IRF7, resulting in type I interferon synthesis. The secreted interferons play a role in autocrine signaling of the cells through dimeric interferon receptors (IFNAR1/2), activating JAK1-STAT1/2 dependent transcription of antiviral genes, including chemokines and interferon-stimulated genes (ISGs) for initiating systemic immunity and restriction of virus replication, respectively (Kim & Shin, 2021) (Ricci et al., 2021). Innate immune responses against SARS-CoV-2 infection are illustrated in figure 1.4.

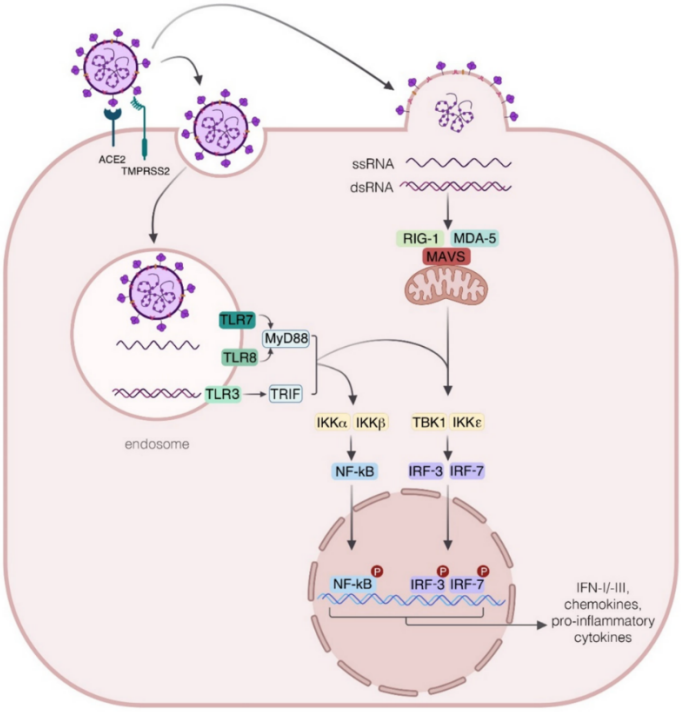


Figure 1.4. Innate immune recognition and response to SARS-CoV-2 infection

Adapted from (Ricci et al., 2021)

These early innate responses are mainly observed in asymptomatic individuals or in individuals with mild symptoms, controlling the disease's progression. In contrast, the severity of the disease was reported to be associated with immune evasion mechanisms of the virus via inhibiting the signaling pathway for type I interferon production through the interaction of accessory or non-structural proteins of the virus with host cell proteins (Lei et al., 2020; Xia et al., 2020).

1.2.2 Adaptive Immunity

Adaptive immunity consists of 2 subdivisions: humoral and cellular immunity. Antigen-specific antibodies are produced by B cells (humoral immunity), whereas, cell-mediated killing of pathogen-infected cells and activation of other immune cells is managed by cytotoxic CD8+ T cells and helper CD4+ T cells, respectively. In the context of SARS-CoV-2 adaptive immune response, phagocytic antigen-presenting cells internalize the virus and/or viral antigens, to present the processed viral epitopes to CD8+ T cells and CD4+ T cells via peptide MHC Class I and MHC Class II complexes, respectively. In turn, T cell receptor (TCR) recognition and signaling is initiated, followed by clonal expansion of T cells bearing antigen-specific TCRs and secretion of the cytokines for helping SARS-CoV-2 specific B cell maturation (Saulle et al., 2021). On the other hand, B cells recognize soluble SARS-CoV-2 structural proteins, including Spike, Receptor-binding domain (RBD) and Nucleocapsid proteins and some accessory proteins according to serosurveys conducted using convalescent plasma from COVID-19 recovered patients (Hachim et al., 2020; Robbiani et al., 2020). This recognition is mediated through B cell receptor-antigen interaction, resulting in the production of soluble IgM, IgG and IgA molecules as a protective adaptive immune response against SARS-CoV-2 infection through somatic hypermutation, selection in germinal centers and affinity maturation of the B cells (Tong et al., 2021). Since neutralizing antibodies against viruses are elicited through the mechanisms mentioned above, most vaccines target mainly B cells. Nevertheless, it has been reported that T-cell help is indispensable for eliciting long-lasting neutralizing and protective antibodies against SARS-CoV-2. Therefore, the orchestration of such an adaptive immune response poses an important bottleneck for vaccine developers (Rydyznski Moderbacher et al., 2020).

1.3 Control and Protection against SARS-CoV-2 Infection

1.3.1 Vaccines and Vaccine Adjuvants

Vaccines are biological agents that induce immune responses specific to a pathogen to provide protection in subsequent encounter with that particular pathogen. Thus, these biological agents consist of pathogen resembling components such as pathogen antigens, along with immunostimulatory molecules known as vaccine adjuvants. While the former is necessary for the activation of the adaptive immune system to elicit a pathogen-specific response, the latter is a need for stimulating the innate immune system in order to provide the danger signal associated responses, mimicking the virtual infection. These signals in turn contribute to the formation and the maturation of adaptive immune response.

In the course of classical vaccination, vaccine antigen is internalized by phagocytic cells such as dendritic cells or tissue-resident macrophages while adjuvant is targeting the same subset of cells to induce innate immune responses through PRR recognition. These cells are then activated and drained through the closest lymph nodes where antigen presentation to T cells occurs. Afterwards, these T cells are clonally expanded while secreting cytokines and chemokines to provide the cytokine milieu for the maturation of effector T cells. Simultaneously, free antigens recognized by the B cells via their receptor on their membrane result in proliferation with the help from previously mentioned T cells and differentiation into memory and the effector B cells (Pollard & Bijker, 2021). The process is shown in figure 1.5.

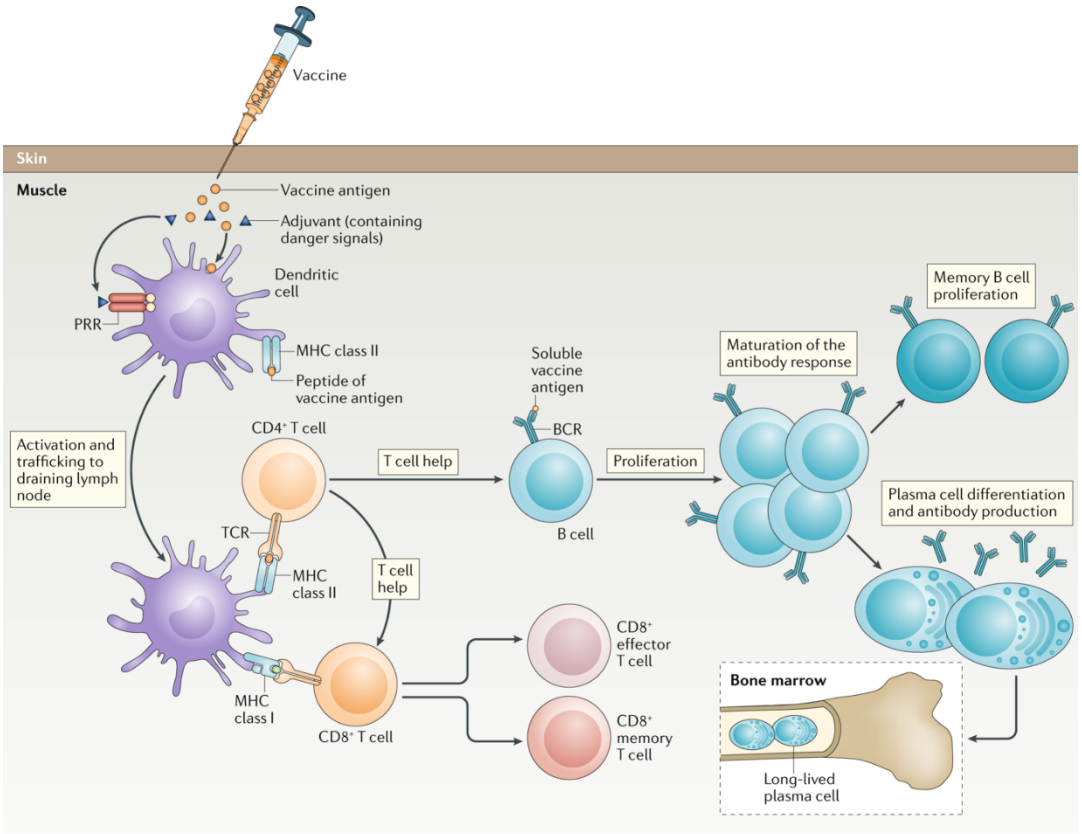


Figure 1.5. Generation of an immune response through vaccination.

Adapted from (Pollard & Bijker, 2021)

To improve immunogenicity, vaccines are generally given in combination with a vaccine adjuvant. The main target of the vaccine adjuvants is the pathogen recognition receptors expressed in antigen-presenting cells. Several experimental and licensed adjuvants include aluminum salts (alum), oil-in-water emulsion MF59, AS01, AS03, AS04 (Pulendran et al., 2021). One of the novel adjuvants that is also used in this study is CpG oligonucleotides. These single-stranded unmethylated cytosine-guanine dinucleotide motifs (CpG) are synthetic TLR9 agonists, mimicking unmethylated bacterial DNA and it has been reported to exert potent immune stimulation (Klinman, 2004). CpG ODNs are classified into four classes based on their differential immunostimulatory effects: A or D-type CpG, B or K-type CpG, C-type CpG, and P-type CpG ODNs (Bode et al., 2011). Their immune activator capabilities are affected by their backbone composition, sequence, 3D conformation

and intracellular localization (Hanagata, 2012). Briefly, K-type ODNs activate B cells and stimulate TNF-α and IL-6 but not interferon-α from plasmacytoid dendritic cells (pDC), whereas D-type ODNs promote mostly interferon-α production in pDC (Gürsel et al., 2002). Despite their excellence in the stimulation of interferon-α production, clinical-grade manufacturing of D-class ODNs are limited due to their propensity to form agglomerated multimers (Gursel & Gursel, 2016). Therefore, K-type CpG ODN stands alone due to its production scalability and potency for immune stimulation.

1.3.1.1 Current Vaccines Against COVID-19

Following the outbreak of COVID-19, several vaccine candidates using different platforms have been developed. These platforms could be described as either traditional (live virus or inactivated) or new generation (recombinant, vector and nucleic acid) vaccines. These platforms (figure 1.6) could be summarized as inactivated virus vaccines in which cell culture propagated SARS-CoV-2 was inactivated through chemical means (c), live-attenuated vaccines consisting of the genetically suppressed virus(d), recombinant protein vaccines targeting full-length Spike protein or Receptor binding domain of the Spike protein to elicit neutralizing antibodies(e-f), virus-like particles that are self-assembled recombinant particles devoid of genomic material of the virus yet include viral structural proteins (g), replication-incompetent or competent viral vectors that infect host cells and express Spike protein (h-i), inactivated viral vectors displaying surface Spike protein (j), DNA vaccines that encode for spike proteins inside host cells (k) and Spike protein encoding RNA-lipid nanoparticles that deliver the stabilized transcript of Spike protein inside host cells for expression of the viral protein (l).

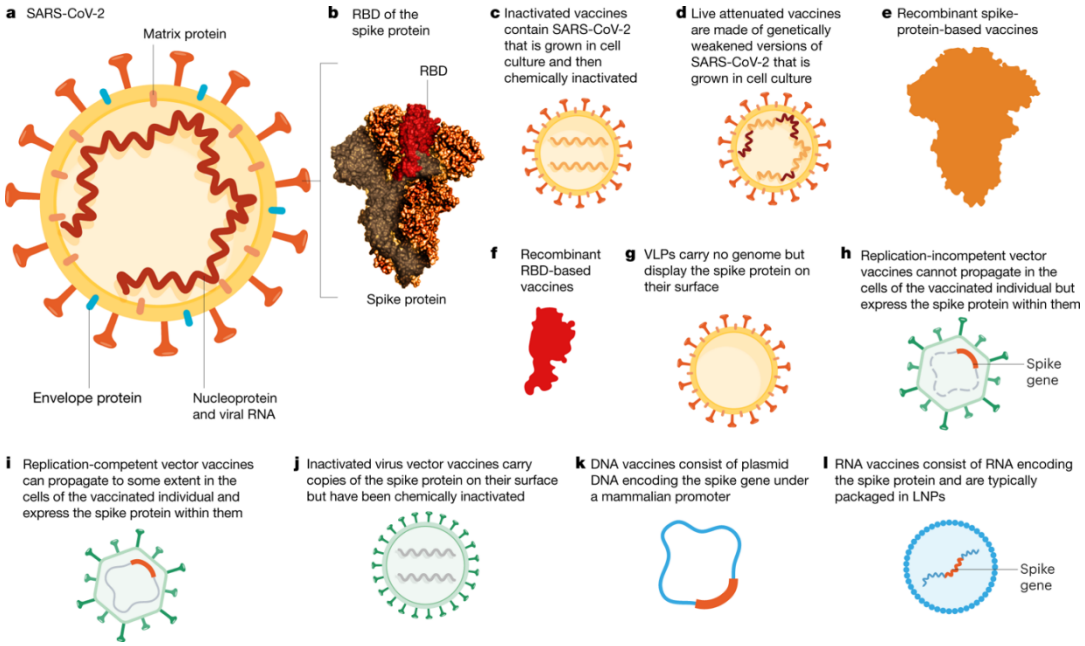


Figure 1.6. Various vaccine platforms developed against SARS-CoV-2.

Adapted from (Krammer, 2020)

According to the COVID-19 vaccine tracker and landscape of WHO (12th October 2021), there are 196 vaccines in preclinical development and 126 clinically registered vaccines listed (World Health Organization, 2021). Subtypes of 126 vaccines in clinical trials are summarized in figure 1.7.

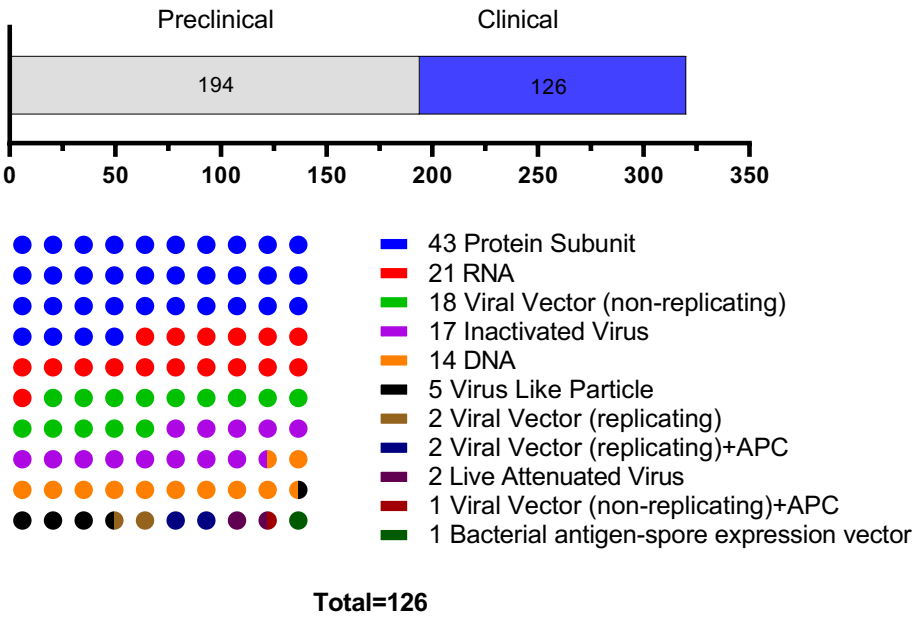


Figure 1.7. COVID-19 vaccine tracker and landscape data summary.

Adapted from (World Health Organization, 2021), Updated on 12th October 2021.

The most prominent vaccines out of the 126 clinically registered include, BioNTech/Pfizer Vaccine (Germany/US), Moderna Vaccine (US), Oxford/Astra Zeneca Vaccine (UK/US), Janssen Vaccine (The Netherlands/US), Novavax Vaccine (US), Sputnik V Vaccine (Russia), CanSino and Sinovac Vaccines (China). Types and efficacies reported for these widely distributed vaccines are summarized in Table 1.1 (Baden et al., 2021; Heath et al., 2021; Logunov et al., 2021; Polack et al., 2020; Tanriover et al., 2021; Voysey et al., 2021).

Table 1.1 . Current COVID-19 Vaccines in worldwide distribution.

Adapted from (Prüβ, 2021),* Efficacy is based on preventing symptomatic COVID-19.

Vaccine	*Institution*	*Country*	*Mechanism*	*Efficacy* *
BNT162b1/BNT162b2	BioNTech/Pfizer	Germany/US	mRNA	95.00%
mRNA-1273	Moderna	US	mRNA	94.50%
ChAdOx1 nCoV-19	University Oxford/Astra Zeneca	UK	Adenovirus vector, chimpanzee	62.1%,90%
Ad26.COV2.S	Janssen/Johnson & Johnson	The Netherlands/US	Adenovirus vector, Ad26	66.90%
NVX-CoV2373	Novavax	US	Protein subunit	89.70%
Sputnik V (Gam-Covid-Vac)	Gamaleya	Russia	Adenovirus vectors	91.60%
CoronaVac	Sinovac Biotech	China	Inactivated SARS CoV-2	51.00%
Ad5-nCOV	CanSino Biologics	China	Adenovirus vector	65.70%

Based on these results, it has been demonstrated that effective vaccination against COVID-19 could prevent symptomatic disease as well as the severity of the disease progression and hospitalization percentage (Kyriakidis et al., 2021). Although some of these vaccines were approved for emergency use and distributed world-wide (Ledford, 2020; Tanne, 2020), vaccine inequity among low and middle-income countries necessitates production of additional innovative, safe and efficient vaccines (Figueroa et al., 2021).

1.3.1.2 Virus-like-particles (VLP) and Examples of VLP Vaccines

Virus-like particles (VLPs) are surrogate nano-scaled viruses that do not contain the genomic material of the virus. They are produced by exploiting the self-assembly features of the structural proteins of the particular enveloped or non-enveloped viruses. Since they are replication-deficient and harbour the structural proteins of the virus itself, they could carry high immunogenicity potential. Production of VLPs could be achieved by transfecting the structural protein-encoding genes of the virus to the host cells such as bacteria, yeast, insect, plant or mammalian cells. When these proteins are expressed at host cells simultaneously, they self-assemble and become ready for purification. Purified VLPs then could be utilized as a vaccine antigen and play a significant role against invader pathogens or cancer. Production, purification and utilization of VLPs are illustrated in Figure 1.8.

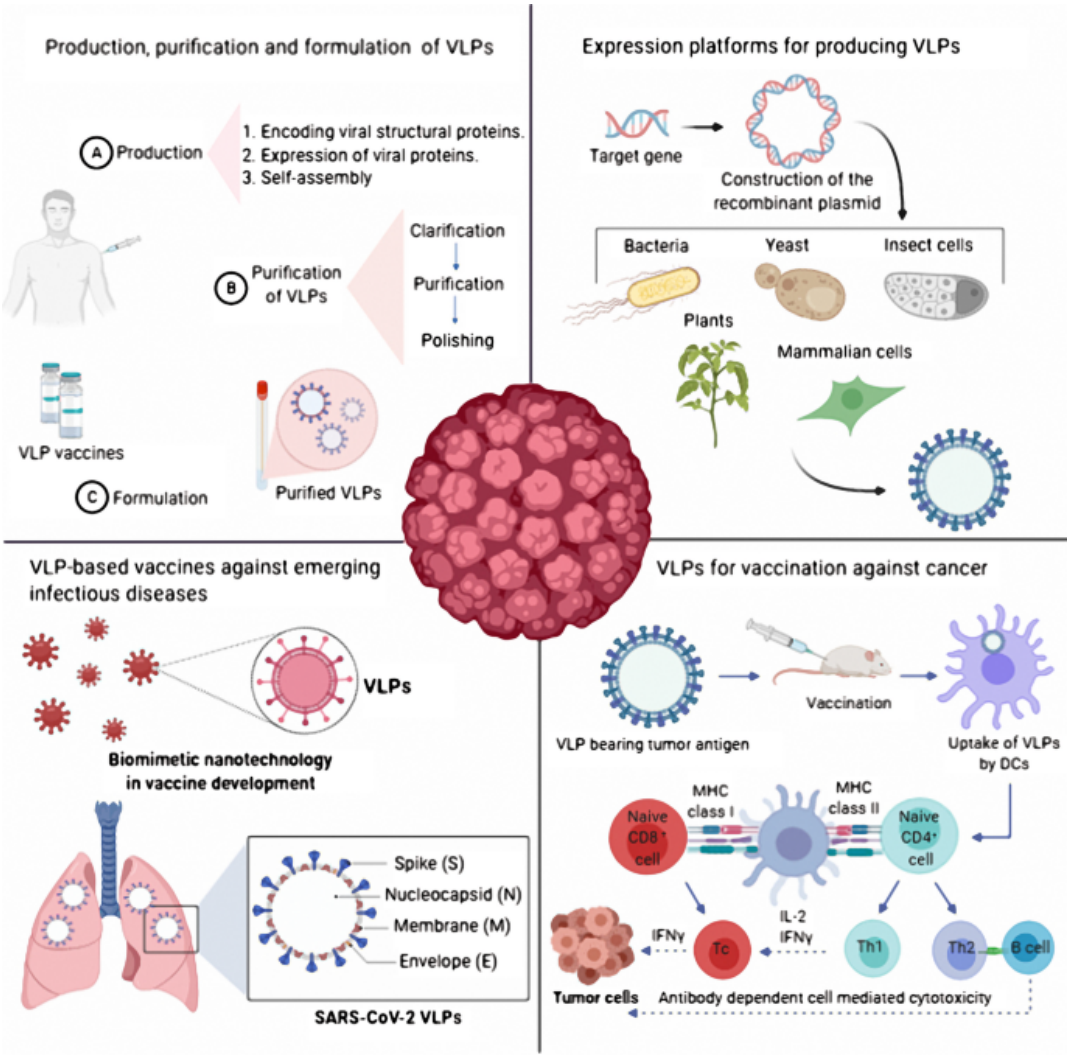


Figure 1.8. Virus-like particle production, purification and utilization scheme.

Adapted from (Nooraei et al., 2021)

Previously, several structural viral proteins have been exploited to produce virus-like particles including those for human immunodeficiency virus (HIV), adeno-associated virus, Hepatitis B virus (HBV) and Hepatitis C virus (HCV) (Joe et al., 2020). Moreover, several vaccines are commercially available against such viruses such as Cervarix and Gardasil, targeting the human papilloma virus (HPV) (Mohsen et al., 2017). Based on this information, SARS-CoV-2 virus-like particles could be used as a robust vaccine antigen for protection against COVID-19. In fact, the first

plant-derived VLP vaccine against SARS-CoV-2 was clinically registered by Medicago Inc. (Ward et al., 2021).

1.4 Aim of the Study

Joint studies carried out in our and Ihsan Gursel's lab from Bilkent University, have first shown the possibility of producing immunogenic mammalian cell-derived SARS-CoV-2 virus-like particles. Herein, we aimed to develop a scalable process for production of mammalian cell-derived SARS-CoV-2 virus-like particles. We have been focusing on the instruments and process development strategies enabling the large-scale production of VLPs that could be applied for GMP manufacturing of VLPs for further clinical trials. The most important aim of this study is to provide the first proof of concept for large scale production, purification and immunogenicity tests of SARS-CoV-2 virus-like particles harbouring the prefusion stabilized (2p) or the thermostable six proline (6p; HexaPro) mutated Spike protein, since such mutated spike proteins were reported to be superior in eliciting robust immune responses against SARS-CoV-2 (Hsieh et al., 2020). Moreover, comparing the effects of various vaccine adjuvantation strategies based on K3 CpG ODN and/or Alum on humoral and cellular immunity was also aimed. Another part of this study concentrates on conducting in-house serosurveys for the assessment of vaccine-elicited humoral responses. Therefore, we aimed to develop a process for mammalian expression and affinity purification of thermostable and prefusion stabilized recombinant six proline (HexaPro) mutated Spike protein of SARS-CoV-2.

CHAPTER 2

MATERIALS & METHODS

2.1 Materials

2.1.1 Suspension HEK293 Cell Line

Suspension HEK293 cell line derived from human embryonic kidney cells was purchased from Florabio A.S.,Turkey. These cells were adapted to grow in suspension cell culture, particularly in shake flasks and bioreactors suitable for production of industry-scale protein(s)-of-interest.

2.1.2 Bacterial Host Strain

NEB Stable competent *E.coli* strain was retrieved from New England Biolabs, USA (cat. # C3040I). This bacterial host has been modified to prohibit self-endonuclease activity to eliminate plasmid degradation and was chosen for high efficiency isolation and propagation of various plasmids. Throughout the study, it has been used for cloning and plasmid propagation purposes.

2.1.3 Cell Culture Media, Buffers and Solutions

Suspension HEK293 cell line was sub-cultured in animal component and serum free Orchid 293CD TFM media (Florabio, Turkey) supplemented with 400 mg L-glutamine (Sigma, USA) per liter. 10% (v/v) DMSO (Sigma, USA) in Orchid 293CD TFM media was used for freezing of the cells. For bacterial cell culture, low salt LB

medium was prepared via dissolving 10 g Tryptone (Sigma, USA, cat. # T7293), 5 g yeast extract (Merck, Germany, cat. #103753) and 5 g NaCl (Isolab, Germany, cat. # 969.036) in 1 liter of distilled water. 10 g of agar (Sigma, USA, cat. # 05040) was added to the former ingredients to prepare the solid low salt LB media. Following autoclaving, hygromycin B Gold (Invivogen, USA, cat. # ant-hg-1) was added to the media at 100 µg/mL final concentration for antibiotic selection. Glycerol stocks of bacteria was prepared in autoclave sterilized 20% (v/v) glycerol solution (Fisher, USA, cat. # BP229-1) Dulbecco's Phosphate buffered saline solution and molecular biology grade water were purchased from Sartorius (Germany, cat. #02-023,1A, 01-869-1A). Elution buffer (50 mM HEPES, 2M NaCl, pH 7.4) for multimodal chromatography was prepared by dissolving 11.9 g HEPES (Sigma, USA, cat. # 54457-250G-F) and 116.9 g NaCl (Isolab, Germany, cat. # 969.036) in distilled water. Following pH adjustment to 7.4, solution was filtered through a 0.22 µm Stericup Quick Release Millipore Express Plus PES filter (Merck, Germany, cat. # S2GPU05RE). Cleaning-in-place (CIP) buffer (1M NaOH in 30% IPA) was prepared by dissolving 40 g of NaOH (Merck, Germany, cat. #1064621000) in 1 liter of 30% (v/v) isopropanol (Sigma, USA, cat. # 34863) in distilled water after 0.22 µm filter sterilization. 1x Running Buffer and 1x Transfer buffer for western blotting were prepared by diluting 100 mL of 10x Tris/Tricine/SDS Buffer (BioRad, USA, cat. # 1610744) and 100 mL of 10x Tris/Glycine Buffer (BioRad, USA, cat. # 1610734) in 900 mL distilled water, respectively. Running buffer for agarose gel electrophoresis (10X Tris-Acetate-EDTA) was prepared by dissolving 48.4 g of Trizma base (Sigma, USA, cat. # T1503-1KG), 3.7 g EDTA disodium salt (Fisher, USA, cat. # BP118-500) and 11.4 mL glacial acetic (Sigma, USA, cat. #27225-M) in final 1L of distilled water. Recipes for all the media, buffers and solutions are available in Appendix A.

2.2 Methods

2.2.1 Mammalian Cell Culture and Maintenance

Suspension HEK293 cell line was sub-cultured every two-three days when the cell count reached 3×10^6 cells / mL in Orchid 293CD TFM media (Florabio, Turkey) by adjusting final concentration to 8×10^5 cells / mL. Depending on the culture volume, cells were passaged in 250 mL, 500 mL or 1 L filter-capped shake flasks (SPL Life Sciences, Korea, cat. # 74250, 74500, 74000). To ensure proper gas exchange, culture volume was adjusted to be 1/5th of the flask volume. Suspension cells were incubated at 37ºC, 8% CO_2 on a reciprocal shaker (130 rpm).

2.2.1.1 Cell Counting and Viability

Cell number and viability were measured by a flow cytometer, Novocyte 3000 (Agilent, USA). Briefly, 200 µL of culture liquid was transferred to 1.5 mL eppendorf tube and 20 µL of propidium iodide (Thermo, USA, cat. #P1304MP) was added to stain non-viable cells. Following 1 minute of incubation, cells were acquired in the flow cytometer. Cells were gated by using forward and side scattering (FSC, SSC) parameters, then viable cell number and viability percentage were analysed by gating the PI negative population in PE channel of the flow cytometer.

.

2.2.1.2 Cryopreservation and Thawing of Cells

Long term storage of cells was achieved through mixing a cell suspension of 1.5×10^6 cells/ mL with an equal volume of freezing medium (Appendix A) in a 2 mL cryogenic vial (Corning, USA). Subsequently, vials were placed in Mr. Frosty™ Freezing Container and stored at -80 ºC refrigerator to provide -1 ºC/minute cooling

rate. Next day, all cryogenic vials were transferred into liquid nitrogen storage for cryopreservation.

While starting a new culture from a liquid nitrogen stock, cells were initially thawed in a 37 ºC water bath for 1-2 minutes. Cell suspension was immediately transferred to fresh TFM medium since presence of cryo-preservatives such as DMSO is extremely toxic near room temperature. The cell suspension was then centrifuged at 100 x g for 10 minutes. After removal of supernatant, cells were resuspended in 1 mL of TFM media for cell counting as described in the section 2.2.1.1. Finally, depending on the cell count, cells were inoculated into a 250 mL shake flask containing TFM medium.

2.2.2 Molecular Cloning of VLP Constructs

2.2.2.1 Design and Cloning

Human codon optimized and BamHI restriction enzyme site flanked 2p-S, HexaPro spike (6p-S), membrane glycoprotein (M) (NCBI Refseq: YP_009724393.1), envelope (E) (NCBI Refseq: YP_009724392.1) and nucleocapsid (N) (NCBI Refseq: YP_009724397.2) SARS-CoV-2 genes were purchased from Integrated DNA Technologies, USA. All genes, except Nucleocapsid, were fused with a CD33 signal sequence to ensure their colocalization for virus-like particle assembly through the standardized secretory pathway (Güler-Gane et al., 2016). To achieve expression of the four structural proteins in suspension HEK293 cells, two different mammalian dual-expression plasmids pVITRO1-hygro-mcs and pVITRO2-hygro-mcs (Invivogen, USA) were employed which have two different multiple cloning sites (MCS). One of the MCS has a BamHI cut site while the other harbors a BglII cut site, enabling the cloning of two BamHI cut site flanked genes in one vector, the former being for BamHI ligation and the latter for BamHI-BglII compatible end ligation. Using this strategy, 2p-S or 6p-S and Envelope genes were cloned sequentially into the pVITRO2 vector. While 2p-S or 6p-S genes were cloned into

BamHI restriction site under the human ferritin heavy chain (hFerH) promoter, the E gene was cloned at BglII site under the human ferritin light chain (hFerL) promoter (pVitro2 S-E). Similarly, the M gene was cloned into pVITRO1 at BamHI site under the rat elongation factor 1 alpha (rEF1) promoter (pVITRO1-M) whereas N gene was cloned into pVITRO1 at BglII site under the mouse elongation factor 1 alpha (mEF1) promoter (pVITRO1- M-N). Since this strategy is open to misdirectional insertion of the genes, HindIII digested clones were verified with agarose gel electrophoresis and next generation sequencing (Intergen, Ankara). All the constructs and genes are illustrated in figure 2.1 and 2.2.

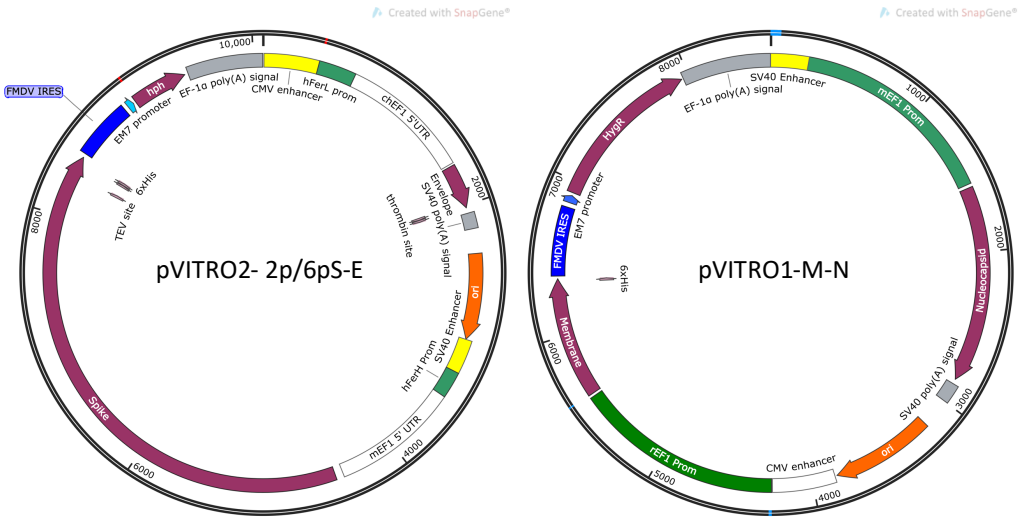


Figure 2.1. Dual Expression Plasmid maps for pVITRO2- 2p/6pS-E and pVITRO1-M-N

pVITRO2) hFerH and hFerL: Human Ferritin heavy and light subunit composite promoters, SV40 enhancer: Simian Virus 40 Enhancer, CMV enhancer: Early enhancer of human cytomegalovirus, SV40 poly(A) signal: Simian Virus polyadenylation signal, ori: pMB1 minimal replication origin, FMDV IRES: Foot and Mouth Disease virus internal ribosome entry site, EM7: Bacterial constitutive promoter, hph: Hygromycin resistance gene, EF1a poly (A) signal: Elongation factor 1 alpha polyadenylation signal

pVITRO1) rEF1 and mEF1 Prom: rat and mouse Elongation factor 1 alpha promoters, SV40 enhancer: Simian Virus 40 Enhancer, CMV enhancer: Early enhancer of human cytomegalovirus, SV40 poly(A) signal: Simian Virus polyadenylation signal, ori: pMB1 minimal replication origin, FMDV IRES: Foot and Mouth Disease virus internal ribosome entry site, EM7: Bacterial constitutive promoter, hph: Hygromycin resistance gene, EF1a poly (A) signal: Elongation factor 1 alpha polyadenylation signal

Construct	Sequence at S1/S2 site	Prefusion stabilizing substitutions
WT	QTQTNSPRRARSVAS 689	--
S – 2p	QTQT---------ILRSVAS 688	K986P, V987P
S – 6p	QTQTNSPGSASSVAS 692	F817P,A892P,A899P,A942P,K986P,V987P

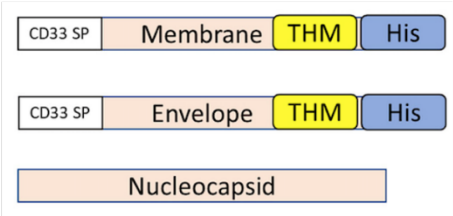


Figure 2.2. Schematic representation of full length Spike protein, Membrane glycoprotein, Envelope protein, Nucleocapsid protein and substitutions of 2p and HexaPro Spike proteins

CD33 SP: CD 33 Signal Peptide, NTD: N-terminus domain, RBD: Receptor-binding-domain, S1/S2: S1-S2 Cleavage site, HR1: Heptad repeat 1, CH: Central Helix, HR2: Heptad repeat 2, TM: Transmembrane domain, CTD: C-terminus domain, TEV: Tobacco Etch Virus protease cleavage site, FD: T4 Fibritin Trimerization Domain, THM: Thrombine Cleavage Site, 6x His: Six histidine amino acid tag

Briefly, all synthesized genes were digested with 1 µL BamHI-HF (New England Biolabs, USA, cat. #R3136S), 2 µL 10X Cutsmart buffer in a 20 µL reaction volume in molecular biology grade water (Sartorius, Germany). Following incubation of reaction mixtures at 37ºC for 60 minutes, 6x Gel Loading Dye was added (HF (New England Biolabs, USA, cat. #B7024S). Digested products were run on 1% (w/v) agarose gel electrophoresis for 45 minutes at 100V. Subsequently, fragments with expected lengths were recovered from agarose via NucleoSpin Gel and PCR Clean-up (Macharey-Nagel, Germany, cat. #740609.50) according to manufacturer's instructions. Measurement of DNA concentration and purity was performed using a BioDrop instrument (Biochrome, UK). In parallel, 1 µg of pVITRO1 and pVITRO2 vector backbones were digested with 1 µL BamHI-HF (New England Biolabs, USA, cat. #R3136S), 2 µL 10X Cutsmart buffer in 20 µL reaction volume. Following incubation of reaction mixtures at 37ºC for 90 minutes, 6x Gel Loading Dye was added (HF (New England Biolabs, USA, cat. #B7024S). Digested products were run on 1% (w/v) agarose gel electrophoresis for 60 minutes at 100V. Then, linearized

backbones were recovered by NucleoSpin Gel and PCR Clean-up (Macharey-Nagel, Germany, cat. #740609.50) according to manufacturer's instructions. In order to prevent self-ligation of the backbones, 10 µL (~1 µg) of both backbones were treated with 1 µL rAPid Alkaline Phospatase (Sigma, USA, cat. #048 9813 3001), 2 µL 10X reaction buffer completed to 20 µL reaction volume with water. Reaction mixes were incubated at 37ºC for 30 minutes, then 2 minutes at 75ºC for enzyme inactivation.

In order to achieve sequential cloning of two genes in a vector, first pVITRO2-2p-S/6p-S and pVITRO1-M plasmids were constructed. Linearized backbones and corresponding genes were mixed at 1:3 (backbone:insert) molar ratio in the presence of 1µL T4 DNA Ligase (New England Biolabs, USA, cat. #M0202S), 1 µL 10X Buffer in 10 µL of reaction volume. Reaction mixes were incubated at room temperature for 2 hours. Subsequently, 10 µL ligation mixes were transformed into NEB Stable competent *E.coli* and plated on low salt LB supplemented with 100 µg/mL hygromycin B Gold (Invivogen, USA, cat. # ant-hg-1). Next day, growing colonies were inoculated into liquid media and plasmids were isolated via NucleoSpin Plasmid (Macharey-Nagel, Germany, cat. #740588.50) to screen correct colonies through restriction enzyme digestion and agarose gel electrophoresis.

Sequentially, pVITRO2-2p-S/6p-S-E and pVITRO1-M-N plasmids were constructed by using pVITRO2-2p-S/6p-S and pVITRO1-M as backbones through repeating the protocols explained above. The only difference is that nascent backbones were digested with BglII (New England Biolabs, USA, cat. #R0144S) restriction enzyme. Final verification of correct constructs was conducted via HindIII digestion and next generation sequencing (Intergen, Ankara).

2.2.3 Expression and Characterization of SARS-CoV-2 Virus-like-particles

2.2.3.1 Transfection optimization of suspension HEK293 cell line

To optimize transient transfection of suspension HEK293 cell line for virus-like particle production, a Polyethylenimine (PEI) based transfection reagent PEIpro®-GMP (Polyplus, France cat. #126-10x10) was utilized (Ehrhardt et al., 2006). Instead of VLP constructs, a green-fluorescent protein encoding plasmid was used to monitor efficiency with a flow cytometer. pCDNA3-GFP was a gift from Martine Roussel (Addgene plasmid # 74165; http://n2t.net/addgene:74165 ; RRID:Addgene_74165). To optimize the transfection cell concentration and PEI to plasmid DNA ratio, a transfection matrix was designed with increasing cell counts such as 1×10^6 cells/mL, 1.5×10^6 cells/mL and 2×10^6 cells/mL and various PEI to plasmid DNA ratio such as 2:1 (v/w), 2.5:1(v/w) and 3:1(v/w). In general, sub-cultured cells were centrifuged at 100 x g and media was replaced with fresh TFM. Plasmid DNA and PEIpro were diluted in null TFM media as in 5% volume of the total culture separately in 15 mL conical centrifuge tubes. Then, PEIpro mix was transferred onto diluted plasmid DNA solution in order to start complexation process. Following 15 minutes of incubation, transfection mix was directly added onto the replenished cell culture for monitoring of the eGFP expression at 24, 48 and 72nd hours. Samples were taken and acquired on a Novocyte 3000 flow cytometer to measure transfection efficiency which was established as percent GFP positive cells with respect to increased signal in FITC channel.

2.2.3.2 Transfection of VLP encoding plasmids to suspension HEK293 cells

1.3×10^6 cells/mL were transfected with 2 µL PEIpro:1 µg total amount of pVITRO2-2p-S or 6p-S-E and pVITRO1-M-N in 8.2: 6.8 molar ratio as described in section 2.2.3.1. After 48 hours, cell media was replenished by centrifuging culture at

100 x g for 5 minutes and resuspending in an equal volume of fresh TFM media to remove the excess PEIpro and stress induced metabolites. Cell counts and viability was monitored daily via Novocyte 3000 flow cytometer until the day of harvest, which was 96-120 hours post transfection.

2.2.3.3 Purification of Virus-like particles

2.2.3.3.1 Harvest and Denarase treatment

VLP producing cells were centrifuged at 1500 x g for 15 minutes to remove cellular debris. Then, cell-free supernatant was microfiltered through 0.22 μm Stericup Quick Release Millipore Express Plus PES filter (Merck, Germany, cat. # S2GPU05RE) to remove macrovesicles. To eliminate host cell-derived nucleic acids, the harvest was treated with 200 active units Denarase (c-LEcta, Germany) per mL of harvest for 1 hour at 37ºC with constant shaking. Depending on the formation of the precipitates due to degraded host-cell derived nucleic acids, Denarase-treated-harvest supernatant was re-filtered with similar microfiltration unit.

2.2.3.3.2 Multimodal Chromatography

Purification of the VLPs was achieved via multimodal chromatography column Hi-Screen Capto Core 700 (Cytiva, U.S.A.) with ÄKTA go fast protein liquid chromatography system (Cytiva, U.S.A.). In summary, Hi-Screen Capto Core 700 column was equilibrated with 5 column volumes (CV) of D-PBS (Biological Industries, U.S.A.). Then, a maximum of 30 CV of the sample was applied to the column while collecting the flow-through fractions containing VLPs. Subsequently, the column was washed with 10 CV of equilibration buffer. Finally, elution of host cell-derived proteins smaller than 700 kDa and small nucleic acid fragments (8-9 bp) was achieved via 5 CV step-elution via 50 mM HEPES buffer supplemented with

2M NaCl (pH 7.4). Fractions from wash and elution steps were also collected for further analysis. The presence of the VLPs in the flow-through fractions and proteins in the wash and eluate was tracked via A280 values. In all the steps, the flow rate was set to 0.75 mL per minutes. Afterwards, cleaning-in-place (CIP) protocol was applied to replenish the column for further use. For this, 2 CV of 1 M NaOH in 30% isopropanol was applied via upwards flow with 0.3 ml per minute flow rate. Following 30 minutes incubation, the column was washed with 5 CV of distilled water. Finally, the column was filled with 20% ethanol and stored at room temperature until next use.

2.2.3.3.3 Tangential Flow Filtration

VLP containing flow-through from multimodal chromatography was concentrated and diafiltrated via Sartoflow® Smart (Sartorius, Germany) with Sartocon® Slice 200 Hydrosart® 100kDA (Sartorius, Germany) cross-flow cassettes. During the process, transmembrane pressure (TMP) was held constant at 0.5 bars. A constant crossflow velocity (280mL/minutes) was aimed to be implemented while having a feed rate at 300 mL/minutes and the feed pressure at 0.95 bars. Following the 5X concentration step of flow-through, diafiltration was initiated until 97-98% of the solution was buffer exchanged with Dulbecco's Phosphate buffered saline solution. Since these purification steps require sterilization, the retentate was micro-filtered through 0.22 µm filters.

2.2.3.4 Characterization of virus-like-particles

2.2.3.4.1 Total Protein Quantification

Total protein quantification of VLP samples was achieved using a Pierce BCA protein assay kit (Thermo Fisher Scientific, USA, cat. # 23225) according to manufacturer's instructions. Briefly, 25 µL of top standard was serially diluted in

PBS as duplicates on a 96 well plate (Corning, USA, cat. #3599) along with 25 µL of VLP samples. Following addition of 200 µl of working reagent onto samples and standards, plate was incubated at 37 °C for 30 minutes, absorbance values at 562 nanometers were obtained through a Multiskan™ FC Microplate Photometer (Thermo, USA). A four-parameter logistic regression curve was constructed to quantify the concentrations of unknown VLP samples.

2.2.3.4.2 SDS-PAGE, Western and dot blotting

VLP samples were denatured in 4x Laemmli buffer at 95ºC for 10 minutes and SDS-PAGE was conducted by using 4–20% Mini-PROTEAN TGX Stain-Free Protein Gel (Bio-Rad, USA) according to manufacturer's protocol. Subsequently, separated proteins were transferred onto a nitrocellulose membrane (Merck, Germany, cat. # HATF00010) via XCell SureLock Mini-Cell transfer system (Thermo, USA). 5% non-fat dry milk in PBS-Tween solution was used for the blocking step for 1 hour at room temperature. Following 3 times washing with PBS-Tween solution for 5 minutes, membranes were incubated in HRP-Conjugated 6-His, His-Tag Antibody solution (Proteintech, Germany cat. # HRP-66005) for 1 hour at room temperature. Then, final wash was conducted and proteins were visualized using the ECL Prime Western Blotting Detection Reagents (Cytiva, USA, cat. #RPN2232) according to manufacturer's instructions.

For the dot blotting, 10 µL of VLP samples were directly spotted onto a nitrocellulose membrane (Merck, Germany, cat. # HATF00010) and incubated until the droplets were air-dried. 5% non-fat dry milk in PBS-T solution was used for the blocking step. Following 3 times washing with PBS-T solution for 5 minutes, membranes were incubated in HRP-Conjugated 6-His, His-Tag Antibody solution (Proteintech, Germany cat. # HRP-66005) anti-His for 1 hour at room temperature. For nucleocapsid blots, membranes were washed with PBS-T 3 times for 5 minutes following incubation with SARS-CoV-2 Nucleoprotein Antibody (ProSci, USA cat. # 35-579) for 1 hour at room temperature. Then, membranes were incubated with

Anti-mouse IgG, HRP-linked Antibody (Cell Signaling Technologies, USA, cat. # #7076) for 1 hour at room temperature. After a final wash step, proteins were visualized with the ECL Prime Western Blotting Detection Reagents (Cytiva, USA, cat. #RPN2232) according to manufacturer's instructions.

2.2.3.4.3 Silver Staining

SDS-PAGE was performed as described in the section 2.2.3.4.2. Then, SDS-PAGE separated proteins were visualized using a Pierce™ Silver Stain for Mass Spectrometry (Thermo Scientific, USA). Following gel wash with ultrapure H2O for 5 minutes twice, the gel was fixed with the fixing Solution for 15 minutes two times. A second ethanol wash step was conducted for 5 minutes twice. Subsequently, the gel was washed with ultrapure water as described above. Sensitizer working solution incubation was performed for 1 minute and 1 minute ultrapure water wash was performed two times. Gel was incubated in enhancer working solution for 5 minutes, followed by washing with ultrapure water for 20 seconds twice. To visualize separated protein bands on the gel, developer working solution was added and images were captured with 30 seconds intervals.

2.2.3.4.4 Nanoparticle analysis via Tunable Resistive Pulse Sensing

Measurements of Tunable-resistive pulse sensing (TRPS) were performed by using the qNano Gold (IZON S/N 601A) system. All the calibration and nanopore related materials such as coating, wetting solutions and calibration particles were used according to the manufacturer's instructions. To measure the size distribution and the particle count of the samples, 35 µl of 1:1000 diluted calibration particles (CPC100, IZON Reagent Kit, RK3-167) or VLP sample liquid was loaded and read at least three times.

2.2.4 Methods for *in vivo* and *ex vivo* experiments

6-10 weeks old, adult BALB/c mice were used for in vivo experiments. All animals used in this study were maintained at Animal Facility of Molecular Biology and Genetics Department at Bilkent University (Ankara, Turkey). Animals were fed *ad libitum* and kept in filtered cages. All the procedures performed throughout the study were approved by animal ethical committee of Bilkent University.

2.2.4.1 Immunization Study

2% Alhydrogel adsorbed, K3-CpG ODN adjuvanted 2p-S VLP and 6p-S VLP (0.4 µg/mouse Low Dose, 4 µg/mouse High Dose) formulations were injected subcutaneously (s.c.) in a total of 200 µL with 2 weeks of intervals (on 0 and 14^{th} days). 2p-S (12 mice/group), or 6p-S expressing VLPs (12 mice/group) were formulated only with Alum (5 µg/mice), only with K3-CpG ODN (20µg/mice) or their combination. Before booster day and 14 days after booster injection, mice were bled from tail vein and blood was collected into a BD microtainer. Following centrifugation at 2500 x g for 10 minutes, upper layer was collected and aliquoted into 1.5 mL centrifuge tubes for storage at -20ºC till next use.

Mice were sacrificed by anesthesia containing the combination of xylazine (5-16 mg/kg) and ketamine (100-200 mg/kg). Following cardiac puncture for blood collection and serum isolation as previously described, spleens were removed surgically and transferred into a six-well tissue culture plate containing 3 mL per well of 10% regular RPMI medium (Appendix A). Subsequently, spleens were homogenized using the sterile end of the syringe plunger and cells were then washed two times at 300 x g for 5 minutes with 2% RPMI wash medium. Final cell pellet was resuspended in 10% regular RPMI medium for cell counting as described in section 2.2.1.1.

2.2.4.2 SARS-CoV-2 antigen specific IgG, IgG1 and IgG2a Enzyme Linked Immunosorbent Assay (ELISA)

2HB ELISA plates (SPL Life Sciences, Korea) were coated overnight at 4°C with 50 µL per well of in-house recombinant proteins of SARS-CoV-2 including recHexaProS (0.5 µg/ml; section 2.2.6.4.1), recombinant nucleocapsid (20 µg/ml) and recombinant RBD (3 µg/ml). Next day, plates were blocked with 200 µl/well blocking buffer at room temperature for 2 hours. Subsequently, plates were washed with wash buffer 5 times. 1/50 diluted mouse sera were transferred on the first row of the plate, then 5-fold serial dilution was performed on vertical wells, followed by incubation at 4°C overnight. Next, plates were again washed with wash buffer 5 times. 1/1000 diluted ALP conjugated anti-mouse IgG (Southern Biotech, USA), anti-mouse IgG1 (Southern Biotech, USA), and anti-mouse IgG2a (Southern Biotech, USA) antibodies were added at 50 µl per well. Plates were incubated at room temperature for 2 hours. Following final washing and rinsing with wash buffer and distilled water, PNPP (Thermo, USA) substrate was added at 50 µl per well for color development. Optical density (OD) values were measured at 405 nanometers using an ELISA plate reader (Molecular Devices, USA). Endpoint titers were calculated using four-parameter logistic regression curve generated on Graphpad Prism 9.0 software.

2.2.4.3 Recall Antigen Assay for Evaluation of T-cell responses

Splenocytes from naive or vaccinated mice were seeded at a density of 1×10^6 per well onto a 96 well U-bottom tissue culture plate. Then, cells were stimulated with in house recHexaProS (5µg/ml) in the presence of costimulatory Ultra-LEAF™ Purified anti-mouse CD28 Antibody (Biolegend, USA, cat. # 122021, 1 µg/mL) for 48 hours at 37°C. Subsequently, cells were centrifuged and supernatants were collected for multiplexed mouse T-helper cytokine secretion assessment.

2.2.4.4 Cytometric Bead Array

T helper cytokine levels were assessed with the LEGENDplex™ MU Th Cytokine Panel (12-plex) w/ VbP V03 kit (Biolegend, U.S.A., cat. # 741044) according to the manufacturer's instructions with modifications. All incubation steps were carried out in a 96-well V-bottom plate at room temperature and protected from light. Centrifugations were performed at 250 x g with a swinging bucket rotor centrifuge. Prior to the assay, reagents were warmed up to room temperature. The Th panel standard cocktail was reconstituted in 250µL assay buffer which was utilized as the top standard in the assay. Four-fold serial dilutions were performed six times starting from the top standard until a final ratio of 1:4096. Assay buffer was used as the zero standard. To perform the assay, 25µL of assay buffer was added to each standard and sample well of the V-bottom plate provided in the panel kit. Next, 25µL of standards and samples were added into each corresponding well followed by the addition of 25µL pre-mixed beads into each well. The V-bottom plate was sealed and incubated at room temperature for 2 hours with continuous rocking to prevent bead sedimentation. Afterwards, the plate was centrifuged at 250 x g for 5 minutes and supernatants were immediately discarded. The plate was dried on a paper towel at an inverted state to get rid of the remaining liquid. The beads were resuspended in 200 µL wash buffer and the plate centrifuged again at 250 x g for 5 minutes. After discarding the wash buffer, the beads were resuspended in 25 µL detection antibody. The plate was sealed and incubated at room temperature for 1 hour with continuous rocking. Next, 25 µL of PE-conjugated streptavidin (SA-PE) was added directly into each well. The plate was incubated for another 30 minutes and washed twice at 250 x g for 5 minutes. Finally, beads were resuspended in 150 µL wash buffer and acquired using a Novocyte 3000 flow cytometer (Agilent, U.S.A.). First, two bead clusters representing two different groups of cytokines were separated according to their FSC (forward scatter) and SSC (side scatter) parameters. Next, each bead population corresponding to a different cytokine, was gated depending on their internal APC (allophycocyanin) intensity. Cytokine levels in experimental groups

were determined based on PE (phycoerythrin) signal intensities and were quantified according to the curve generated by the standard samples. Data analysis was carried out with the LEGENDplex™ Data Analysis Software (Biolegend, U.S.A.).

2.2.5 Molecular Cloning of SARS-CoV-2 Recombinant Spike Construct

2.2.5.1 Design and Cloning

Human codon optimized and BamHI restriction enzyme site flanked 6X- Histidine tagged HexaPro spike (6p-S) gene of SARS-CoV-2 was purchased from Integrated DNA Technologies, USA. In order to increase the secreted version of the recombinant HexaPro Spike protein, transmembrane domain was excluded in the peptide sequence and CD33 signal sequence was included at 5' of the coding sequence. For expression of the recombinant protein in suspension HEK293 cells, mammalian expression plasmid pVITRO2-hygro-mcs was (Invivogen, USA) employed. The gene was cloned into MCS that contains a BamHI cut site. BamHI digested clones were verified with agarose gel electrophoresis and next generation sequencing (Intergen, Ankara). Schematic representation of the protein and the construct is illustrated in figure 2.3.

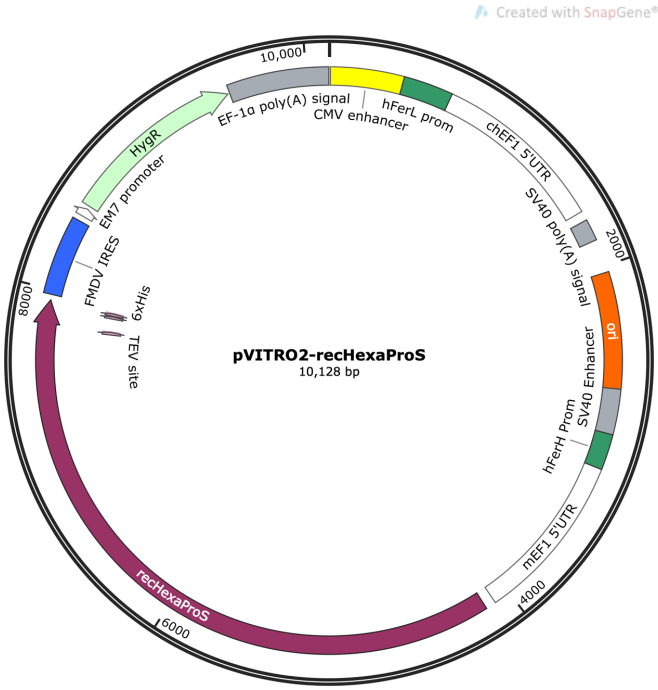


Figure 2.3. Dual Expression Plasmid map for pVITRO2-recHexaProS

pVITRO2) hFerH and hFerL: Human Ferritin heavy and light subunit composite promoters, SV40 enhancer: Simian Virus 40 Enhancer, CMV enhancer: Early enhancer of human cytomegalovirus, SV40 poly(A) signal: Simian Virus polyadenylation signal, ori: pMB1 minimal replication origin, FMDV IRES: Foot and Mouth Disease virus internal ribosome entry site, EM7: Bacterial constitutive promoter, hph: Hygromycin resistance gene, EF1a poly (A) signal: Elongation factor 1 alpha polyadenylation signal

recHexaProS gene was digested with 1 µL BamHI-HF (New England Biolabs, USA, cat. #R3136S), 2 µL 10X Cutsmart buffer in a 20 µL reaction volume finalized with molecular biology grade water (Sartorius, Germany). Following incubation of reaction mixture at 37ºC for 60 minutes, 6x Gel Loading Dye was added (HF (New England Biolabs, USA, cat. #B7024S). Digested products were run on 1% (w/v) agarose gel electrophoresis for 45 minutes at 100V. Subsequently, fragments in expected lengths were recovered from agarose via and NucleoSpin Gel and PCR Clean-up (Macharey-Nagel, Germany, cat. #740609.50) according to manufacturer's instructions. Measurement of DNA concentration and purity was achieved by BioDrop instrument (Biochrome, UK). In parallel, 1 µg of pVITRO2 vector backbone was digested with 1 µL BamHI-HF (New England Biolabs, USA, cat.

#R3136S), 2 µL 10X Cutsmart buffer in 20 µL reaction volume. Following incubation of reaction mixture at 37ºC for 90 minutes, 6x Gel Loading Dye was added (HF (New England Biolabs, USA, cat. #B7024S). Digested products were run on 1% (w/v) agarose gel electrophoresis for 60 minutes at 100V.Then, linearized backbone was recovered by NucleoSpin Gel and PCR Clean-up (Macharey-Nagel, Germany, cat. #740609.50) according to manufacturer's instructions. In order to prevent self-ligation of the backbone, 10 µL(~1 µg) of both backbones were treated with 1 µL rAPid Alkaline Phospatase (Sigma, USA, cat. #048 9813 3001), 2 µL 10X reaction buffer completed to 20 µL reaction volume with water. Reaction mixs was incubated at 37ºC for 30 minutes, then 2 minutes at 75ºC for enzyme inactivation. Linearized backbone and the gene were mixed at 1:3 (backbone:insert) molar ratio in the presence of 1µL T4 DNA Ligase (New England Biolabs, USA, cat. #M0202S), 1 µL 10X Buffer in 10 µL of reaction volume. Reaction mix was incubated at room temperature for 2 hours. Subsequently, 10 µL ligation mix was transformed into NEB Stable competent *E.coli* and plated onto low salt LB supplemented with 100 µg/mL hygromycin B Gold (Invivogen, USA, cat. # ant-hg-1). Next day, growing colonies were inoculated into liquid media and plasmids were isolated via NucleoSpin Plasmid (Macharey-Nagel, Germany, cat. #740588.50) to screen correct colonies through restriction enzyme digestion and agarose gel electrophoresis.

2.2.6 Expression and Characterization of SARS-CoV-2 Recombinant Spike Protein

2.2.6.1 Transfection of SARS-CoV-2 Recombinant Spike encoding plasmid to suspension HEK293 cells

1.3×10^6 cells/mL were transfected with 2 µL PEIpro:1 µg total amount of pVITRO2-recHexaProS as described in section 2.2.3.1. After 48 hours, cell medium was replenished by centrifuging culture at 100 x g for 5 minutes and resuspending in

an equal volume of fresh TFM media to remove the excess PEIpro and stress induced metabolites. Cell counts and viability was monitored daily via Novocyte 3000 flow cytometer until the day of harvest which was 96-120 hours post transfection.

2.2.6.2 Purification of SARS-CoV-2 Recombinant Spike Protein

2.2.6.3 HisTrap Excel Affinity Chromatography

Supernatant containing recHexaProS was harvested as described in the section 2.2.3.3.1 without Denarase treatment. Subsequently, in order to purify the his-tagged proteins from clarified supernatants, affinity chromatography was performed using the ÄKTA go fast protein liquid chromatography system (Cytiva, U.S.A.). Briefly, His-Trap excel 1 mL pre-packed IMAC column (Cytiva, U.S.A.) was equilibrated with 5 column volume (CV) of 20 mM of sodium phosphate supplemented with 0.5 M NaCl (pH 7.2). After the sample application step, all impurities and non-target proteins were washed with the same buffer for 10 CV. To elute the his-tagged proteins, 20 mM of sodium phosphate supplemented with 0.5 M NaCl and 0.5 M imidazole buffer was used for the gradient elution to target 100% concentration of the elution buffer until 20 CV was reached. In all the steps, the flow rate was adjusted to 1 mL per minutes. His-tagged recombinant full-length trimeric Spike of SARS-CoV-2 containing fractions were collected according to their A280 values for further buffer exchange.

2.2.6.3.1 Tangential Flow Filtration-HiScreen Ni-FF Affinity Chromatography

Tangential Flow filtration combined with affinity chromatography was reported to be one of the most efficient protocols for purification of recombinant Spike protein of SARS-CoV-2 (Esposito et al., 2020). To this end, a modified version of the reported protocol was performed to purify secreted recHexaProS from the

supernatant of HEK293 cells. Briefly, harvested supernatant was initially concentrated and diafiltrated via tangential flow filtration with Sartocon® Slice 200 Hydrosart® 100kDA (Sartorius, Germany) as described in section 2.2.3.3.3. Subsequently, affinity chromatography was performed using the ÄKTA go fast protein liquid chromatography system (Cytiva, U.S.A.). Briefly, HiScreen Ni-FF 5 mL pre-packed IMAC column (Cytiva, U.S.A.) was equilibrated with 5 column volume (CV) of D-PBS supplemented with 25mM imidazole. After the sample application step, all impurities and non-target proteins were washed away with the same buffer for 4 CV. Following the orientation change of the column (i.e. reverse flow) 4 CV of imidazole wash was applied in order to remove weakly attached impurities as a gradient from 25 mM to 100 mM imidazole concentration. Then, target protein was eluted from the column with 325 mM step elution by using D-PBS supplemented with 0.5 M imidazole. In all steps, the flow rate was adjusted to 3 mL per minutes. Full-length trimeric Spike of SARS-CoV-2 containing fractions were collected according to A280 values for further buffer exchange.

2.2.6.3.2 Desalting of affinity chromatography purified proteins

To eliminate the imidazole present in recombinant proteins after affinity chromatography for downstream applications, desalting of the protein was performed via HiPrep™ 26/10 Desalting column (Cytiva, U.S.A.). 6 -10 mL of affinity-purified protein was loaded to the HiPrep™ 26/10 Desalting column after 2 CV application of D-PBS (Biological Industries, U.S.A.). Finally, buffer exchange was completed with the isocratic elution step with the D-PBS. For equilibration, sample application and elution step, the flow rate was adjusted to 10 mL per minutes. A280 values of fractions were monitored for protein presence. Concentration with Vivaspin 20 centrifugal concentrator with a 10 kDa molecular weight cutoff (Sartorius, Germany, cat. # Z614610) was performed where it was necessary.

2.2.6.4 Characterization of SARS-CoV-2 Recombinant Spike Protein

To characterize the purified recHexaProS protein, total protein quantification, immunoblotting and silver staining were performed as described in the sections 2.2.3.4.2 and 2.2.3.4.3.

2.2.6.4.1 Optimization of Enzyme Linked Immunosorbent Assay (ELISA) with recHexaProS protein

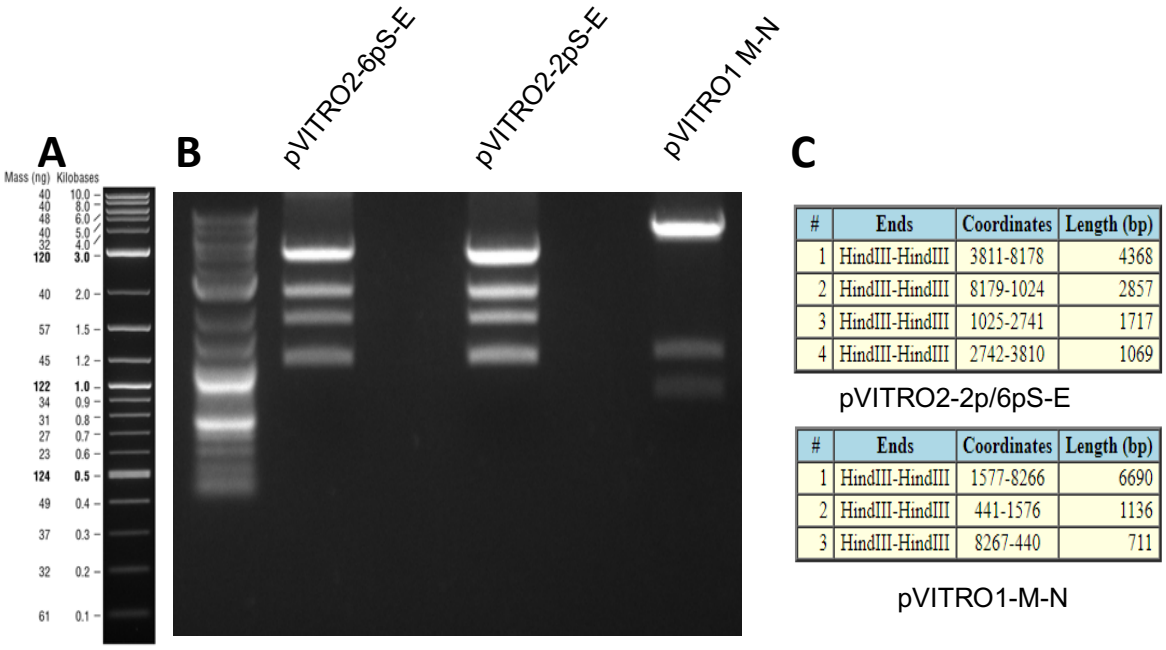
Prior to quantification of the endpoint titers of the sera from vaccinated animals, proper coating concentration of recHexaProS was determined through ELISA. Briefly, 2HB ELISA plates (SPL Life Sciences, Korea) were coated overnight at 4ºC using various concentrations of recHexaProS (0.5-8 µg/mL). Then, Total IgG ELISA was performed as described in section 2.2.4.2 by using sera of at least one placebo group and immunized animals. Performance of the coating concentration was evaluated by comparing the optical density differences versus reciprocal dilutions of the sera and area under the reciprocal dilution curves. All the analyses were based oning four-parameter logistic regression curves generated on Graphpad Prism 9.0 software.

CHAPTER 3

RESULTS AND DISCUSSION

3.1 Verification of VLP and recHexaProS encoding constructs

The cloning strategy for VLP and recombinant trimeric HexaPro Spike protein-encoding plasmids was open for any type of misdirectional insertion or concatamerization of the inserts. Therefore, restriction enzyme digestion was used to verify the inserts' direction and length on the plasmid vector. While VLP encoding plasmids were digested with HindIII enzyme, pVITRO2- recHexaProS was digested with BglII and BamHI. Figure 3.1 and figure 3.2 show that all the fragments from single colony derived plasmids were at their expected lengths (panel C). Moreover, next generation sequencing data showed no mismatch or indels in the coding sequences of the plasmids (Appendix B)



#	Ends	Coordinates	Length (bp)
1	HindIII-HindIII	3811-8178	4368
2	HindIII-HindIII	8179-1024	2857
3	HindIII-HindIII	1025-2741	1717
4	HindIII-HindIII	2742-3810	1069

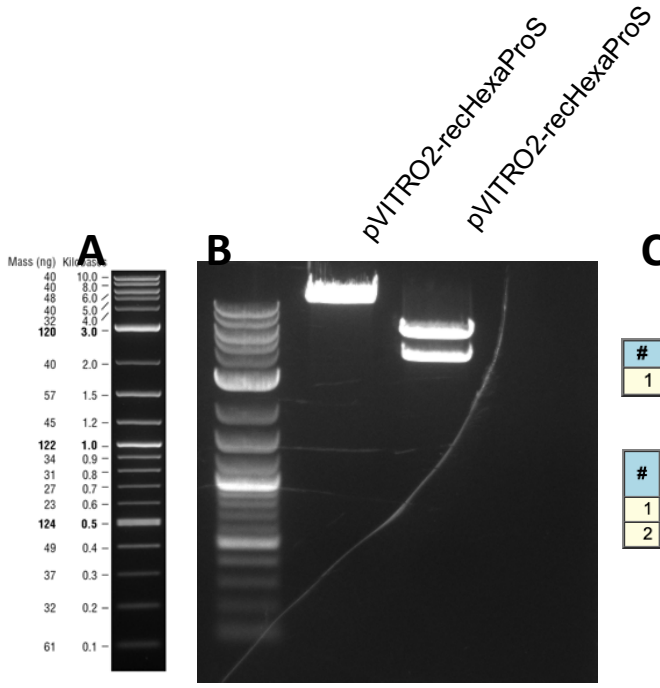
pVITRO2-2p/6pS-E

#	Ends	Coordinates	Length (bp)
1	HindIII-HindIII	1577-8266	6690
2	HindIII-HindIII	441-1576	1136
3	HindIII-HindIII	8267-440	711

pVITRO1-M-N

Figure 3.1. Agarose Gel Electrophoresis image of restriction enzyme digested VLP encoding plasmids.

A. DNA marker, B. Lane 1: NEB 1kb plus DNA ladder, Lane 2:HindIII digested pVITRO2-6pS-E plasmid, Lane 4: HindIII digested pVITRO2-2pS-E plasmid, Lane 6: HindIII digested pVITRO1-M-N plasmid, C. Expected fragment lengths table was generated with NEBCutter V2.0



#	Ends	Coordinates	Length (bp)
1	BglII-BglII	1686-1685	10128

#	Ends	Coordinates	Length (bp)
1	BamHI-BamHI	7944-4136	6321
2	BamHI-BamHI	4137-7943	3807

Figure 3.2. Agarose Gel Electrophoresis image of restriction enzyme digested recombinant HexaProS encoding plasmid.

A. DNA marker, B. Lane 1: NEB 1kb plus DNA ladder, Lane 2: BglII digested pVITRO2-recHexaProS plasmid, Lane 3: BamHI digested pVITRO2-recHexaProS plasmid, C. Expected fragment lengths table was generated with NEBCutter V2.0

Preliminary data (unpublished) and current literature suggest that the configuration of the genes cloned into dual expression plasmids might play a significant role in the assembly and secretion of the SARS-CoV-2 virus-like-particles. It has been reported that only membrane glycoprotein and envelope protein were sufficient to form virus-like-particles (Schoeman & Fielding, 2019). Therefore, keeping the genes encoding for these two proteins in the same dual expression plasmid decreases the probability of Spike or Nucleocapsid incorporating VLPs due to the fact that odds for the co-transfection of two plasmids into a single cell were lower (Gam et al., 2019). To eliminate the presence of M-E VLPs, 2p or 6p Spike proteins and Envelope protein were cloned into pVITRO2 mammalian dual expression plasmid. In contrast, Nucleocapsid and Membrane proteins were cloned into pVITRO1 plasmid.

3.2 Transfection optimization of suspension HEK293 cells

Standard protein expression systems such as *E.coli* fail to produce post-translationally modified or correctly folded proteins. Therefore, mammalian expression systems such as Chinese Hamster Ovary cells (CHO) or HEK293 cells have been utilized over 30 years for the production of viral vectors and proteins (Le Ru et al., 2010) due to their ability to grow in suspension cultures, their high productivities as well as their capability of correct protein folding and post-translational modifications (Dumont et al., 2016). Post-translational modifications of SARS-CoV-2 structural proteins were reported to be indispensable for stimulation of the host immune system (Fung & Liu, 2018). Therefore, a successful protein vaccine candidate against SARS-CoV-2 must be produced in cells that employ post-translational modification machinery. In summary, suspension culture-adapted HEK293 cells have been chosen as the expression system for their high productivity, cellular mimicry of the SARS-CoV-2 replicating host cells and scalability for larger volumes, enabling efficient vaccine manufacturing.

There are several non-viral gene delivery methods for mammalian cells. Nevertheless, none of them could be considered as efficient and scalable as polyethylenimine (PEI) based transfection (Hacker et al., 2013). Thus, Polyethylenimine (PEI) based transfection reagent PEIpro®-GMP was chosen for optimization of efficient delivery of the plasmids to suspension HEK293 cells. Two parameters play a critical role in PEI-based transfection of the HEK293 cells: cell density prior to transfection and PEI to DNA ratio, ensuring proper complexation in the absence of substantial toxicity to the cells. To assess these parameters, cells were transfected in increasing densities (1-1.5-2M cells/mL) and PEIPro to plasmid DNA ratio (2x-2.5x-3x, v/w). Following sampling of the continuous batch cultures 24, 48 and 72 hours post-transfection, samples were acquired in the flow cytometer. Viable cells were gated by using forward and side scattering parameters. GFP positive cells then were evaluated in order to compare the effect of cell density and the PEIPro titration on transfection efficiency at different time points with respect to increase in

their signal in FITC channel. Non-transfected cells were used as a negative control and the baseline for the gating of GFP positive cells.

Figures 3.3 and 3.4 summarize the percent of GFP positive cells 24 (Figure 3.3) or 48 h (Figure 3.4) post-transfection, analyzing the impact of various cell densities and PEIpro®-GMP:plasmid ratios on transfection efficiencies.

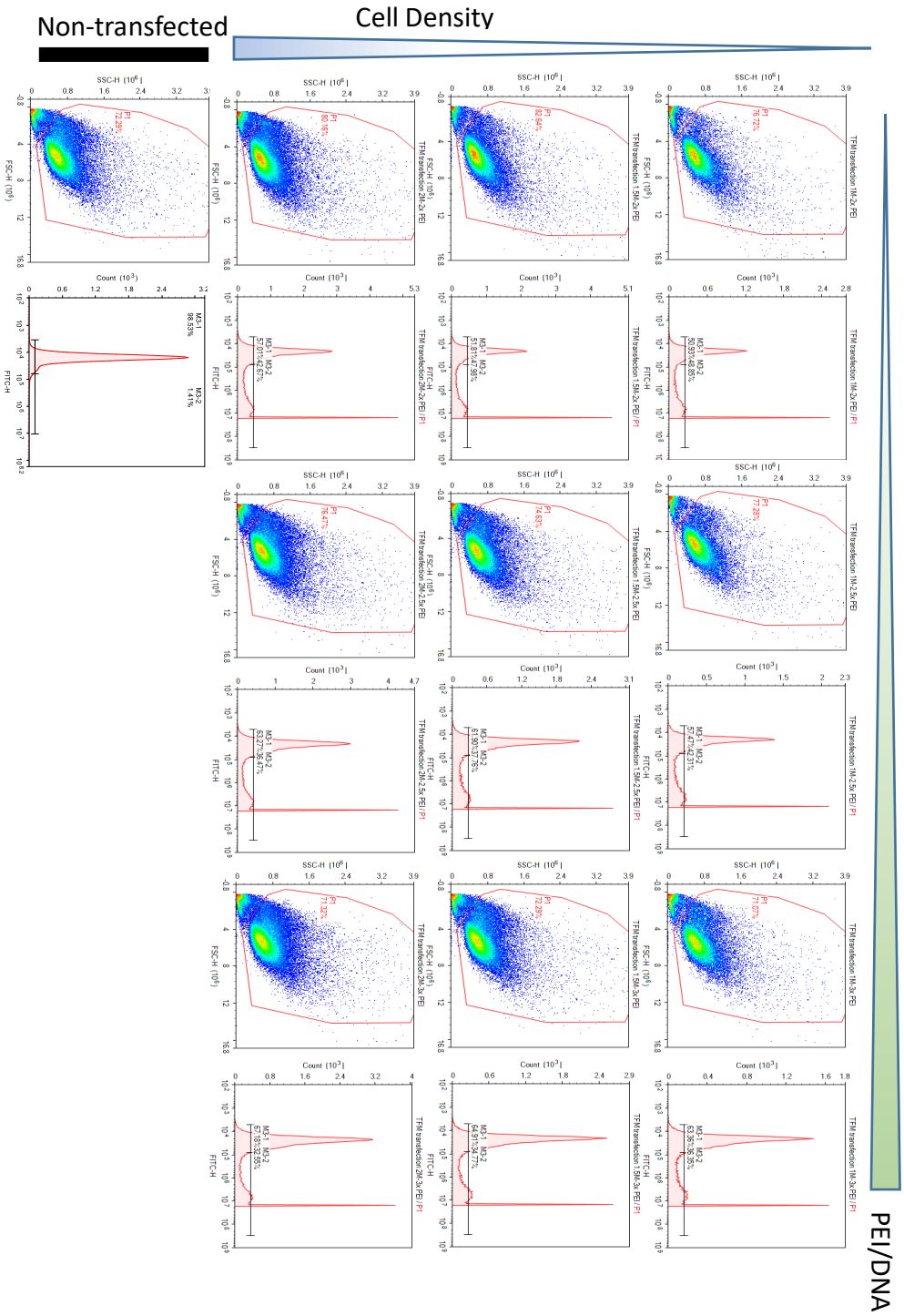


Figure 3.3. Representative dot plots of GFP expressing cells 24 hours post-transfection. P1 gate represents viable cells and M3-2 is GFP+ cells.

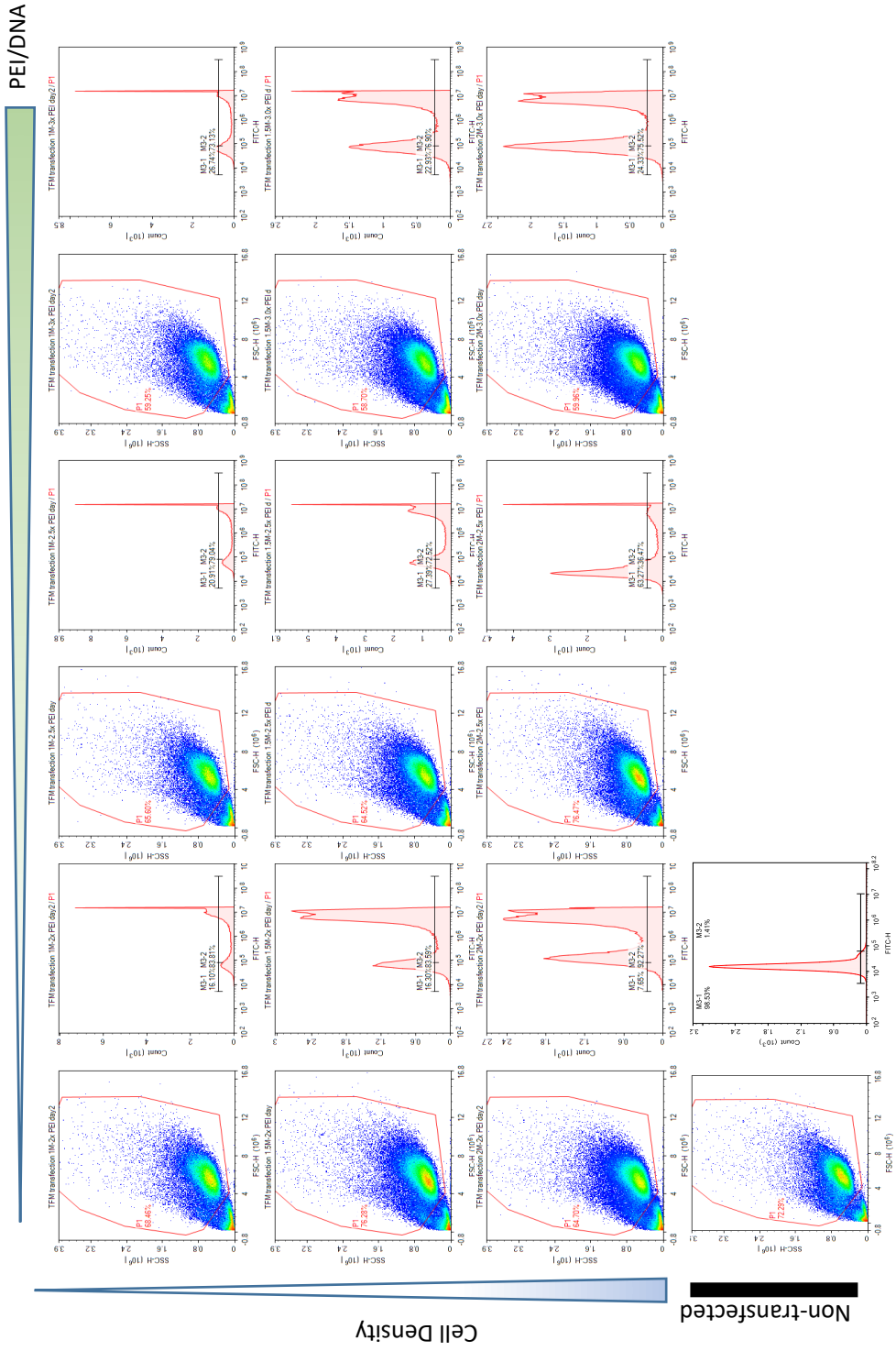


Figure 3.4. Representative dot plots of GFP expressing cells 48 hours post-transfection. P1 gate represents viable cells and M3-2 is GFP+ cells

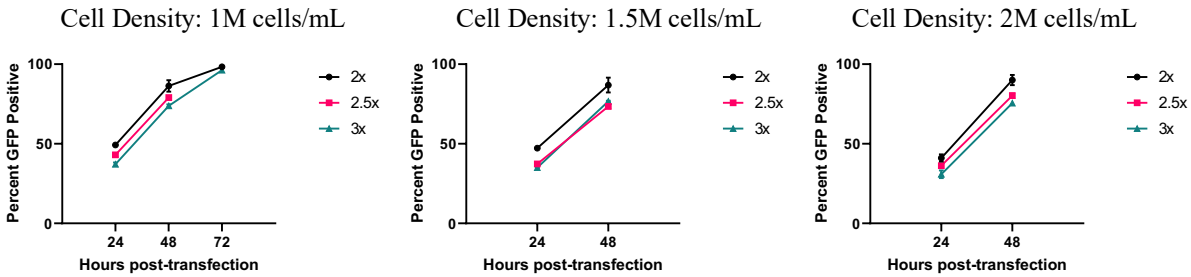


Figure 3.5. Summarized percent transfection efficiencies for various cell densities and PEIPro:DNA ratio.

2x,2.5 and 3x denotes volume to weight (µL/µg) PEIPro:DNA ratio

Table 3.1 Summary of mean percent transfection efficiencies various cell densities and PEIPro:DNA ratio.

Nd.: No data present

		Mean GFP Percent Positive		
Cell Density (cells/mL)	PEIPro:DNA (v/w)	*24 hours post transfection*	*48 hours post transfection*	*72 hours post transfection*
$1x10^6$	2x	49.25	86.37	98.31
	2.5x	43.04	79.04	nd.
	3x	37.14	73.81	96.21
$1.5x10^6$	2x	47.24	86.88	nd.
	2.5x	37.39	73.40	nd.
	3x	35.00	76.61	nd.
$2x10^6$	2x	41.10	90.02	nd.
	2.5x	36.20	80.30	nd.
	3x	30.87	75.52	nd.

Daily monitoring of the transfection efficiencies via tracking the GFP positive cell percentage revealed that 2 µL PEIPro to 1 µg DNA ratio (2x) yielded the highest transfection percentage at each time point (Table 3.1). Moreover, starting cell concentration did not significantly impact transfection efficiency (Figure 3.5).

Therefore, considering that a lower cell-density would be more appropriate for scale-up, a cell concentration of 1.3×10^6 cells/mL was chosen for future VLP and recombinant protein transfection experiments.

3.3 Purification and characterization of 2p-6p prefusion stabilized Spike bearing Virus-like-particles

Several purification methods for viruses and virus-like particles have been described, including ultracentrifugation, size-exclusion chromatography and crossflow filtration (Hillebrandt et al., 2020). Joint studies carried out in our and Ihsan Gursel's laboratory at Bilkent University developed an optimized protocol for VLP purification that involves the following steps: supernatant clarification, degradation of host cell derived nucleic acids via enzymatic treatment, removal of host cell proteins-nucleic acids and media components by multimodal chromatography employing size exclusion and anion exchange capabilities simultaneously, concentration and diafiltration of purified VLPs via tangential flow filtration. The purification flowchart is illustrated in Figure 3.6.

One of the high throughput methods enabling the purification of enveloped or non-enveloped viruses was established by using Capto Core 700 chromatography resin (James et al., 2016). The resin consists of beads that contain an octylamine ligand core and an inactive porous shell. The ligands inside beads have positive charges as well as hydrophobic features, enabling them to capture the molecules under 700 kilodaltons. On the other hand, the inactive core excludes biomolecules exceeding 700 kilodaltons such as viruses that pass in the flow-through. Therefore, Capto Core 700 beads are capable of bi-modal purification based on size exclusion and capturing of <700 kDA anionic/hydrophobic molecules inside of the bead core at the same time. Moreover, application notes from the manufacturer demonstrate that the resin could be effectively employed to purify human papilloma virus-like particles from insect cells (Cytiva application notes 29-1037-62 AA, 29-0003-34 AA, and 29-0983-01 AB) (Zhao et al., 2015). Based on this information, purification of mammalian

cell derived SARS-CoV-2 virus-like particles was achieved using the Capto Core 700 resin. Figure 3.7 describes the structure and working principle of Capto Core 700 beads.

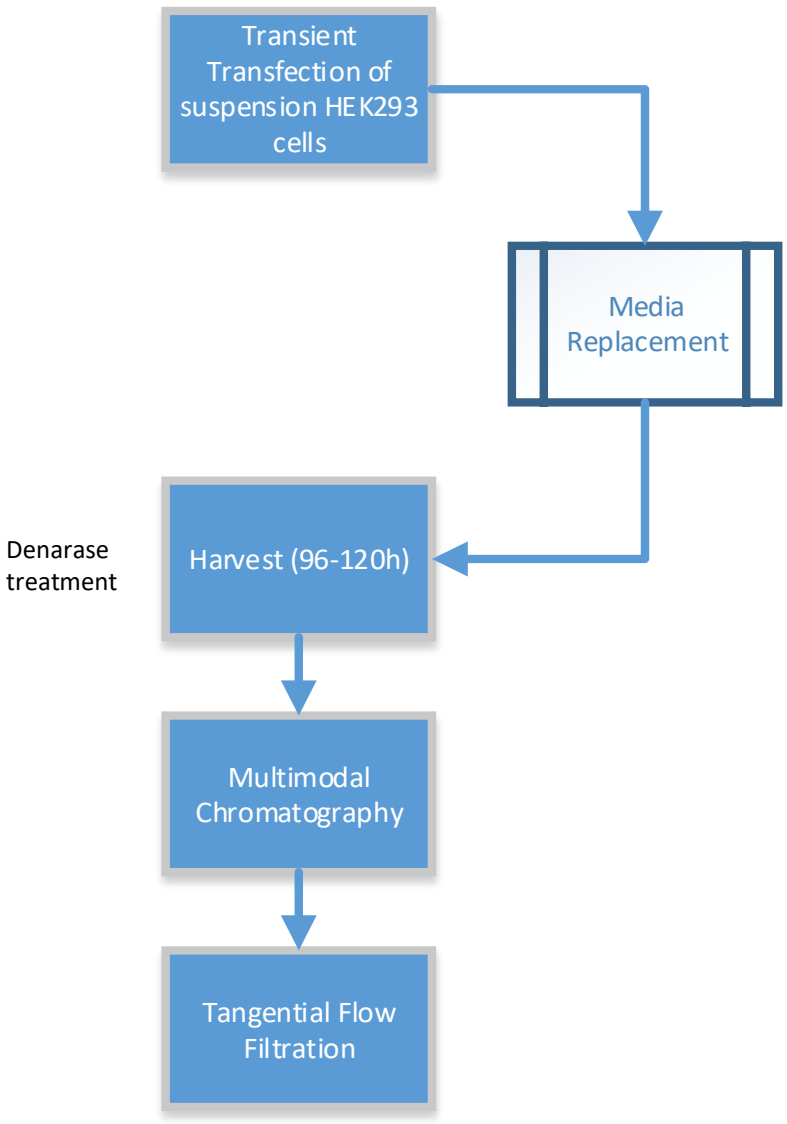


Figure 3.6. Purification flowchart for the SARS-CoV-2 Virus-like-particles

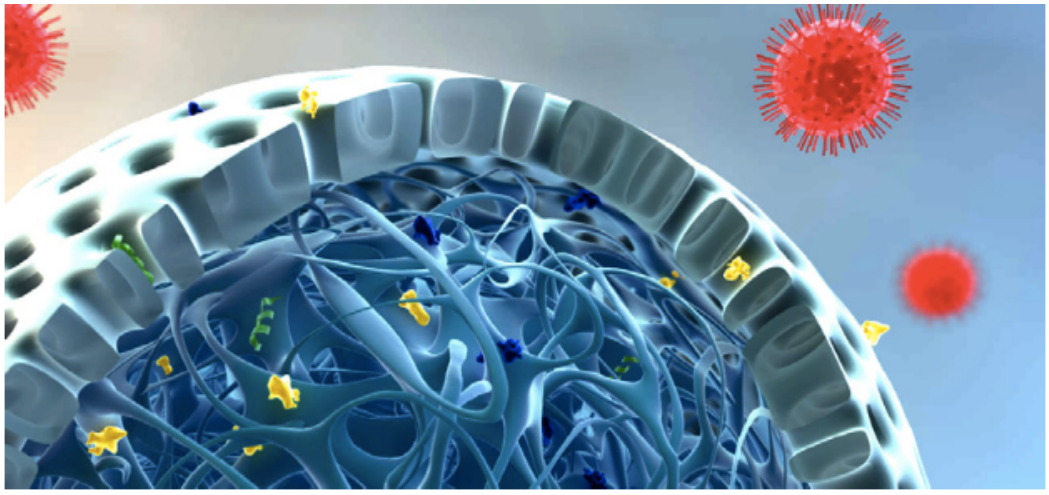


Figure 3.7. Schematic representation of a Capto Core 700 bead.

Reprinted from Capto Core 700 - cytiva. Cytiva Life Sciences. (n.d.). Retrieved October 9, 2021, from https://cdn.cytivalifesciences.com/dmm3bwsv3/AssetStream.aspx?mediaformatid=10061&destinationid=10016&assetid=16124.

The Figure depicts the non-reactive porous outer shell and the ligand containing core. Proteins and impurities (green, yellow and purple) that are smaller than 700 kilodaltons penetrate inside the core and interact with positively charged and hydrophobic ligands. Viruses or virus-like particles (red) are excluded from resin media and pass in the flow-through.

Multimodal chromatography employed herein, was a two-parts process. The first part consisted of sample application onto the prepacked HiScreen Capto Core 700 column in which the flow-through fraction containing the virus-like particles was collected. The second part consisted of the elution of the host cell derived proteins and nucleic acids from inside of the positively charged bead core (James et al., 2016). Throughout the chromatographic steps, the presence of proteins in collected fractions was tracked by miliabsorbance units (mAU) at 280 nm. A representative purification chromatogram was illustrated in figure 3.8.

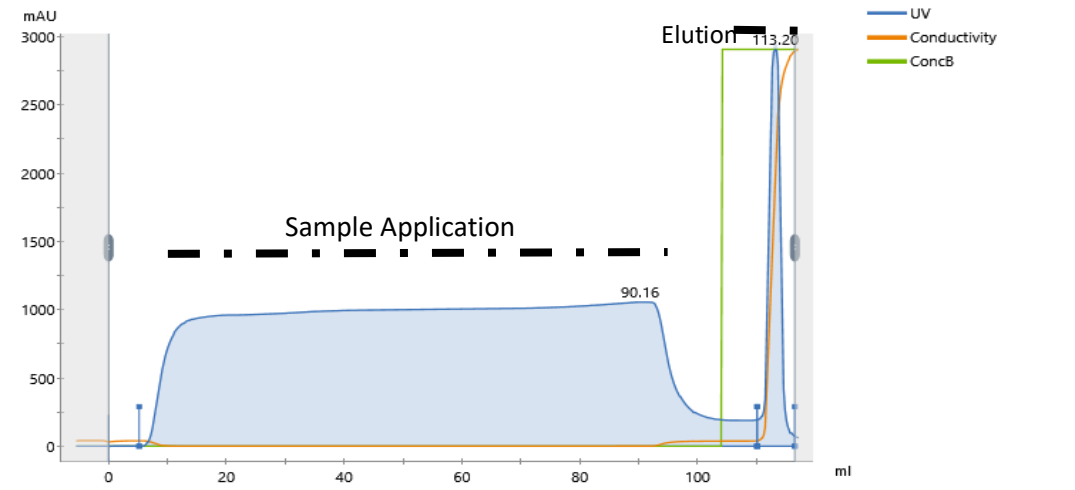


Figure 3.8. Representative chromatogram of multimodal purification of VLPs.

Blue curve represents the absorbance at 280 nm, orange curve is conductivity in mS/cm and green curve is the % concentration of the Elution buffer applied to the column.

Although the tracking depended on protein presence, there was no correlation between the amount of VLPs and the absorbance value of the flow-through fractions owing to the interference of the trace amount of culture media-associated pigment in the flow-through. Thus, multimodal chromatograms were not conclusive in terms of the produced VLP amounts. Instead, they have provided a framework for the assessment of purification success. In figure 3.7, it can be inferred that the absorbance was fixed to a constant value during the sample application phase and elution of the anion exchanged substances was achieved since a large peak was observed. According to previous reports, this pattern was considered as a successful purification (Reiter et al., 2019).

Tangential flow filtration is a method in which a product of interest is concentrated (ultrafiltration) and buffer exchanged (diafiltration) using flat-sheet PVDF or PES membranes with a nominal molecular weight cutoff. A medium containing the product is constantly streamed to the membrane either with constant feed rate or feed pressure where the molecules smaller than the cutoff value are passed to the

permeate, whereas the product of interest is concentrated in feed suspension (retentate). The medium then could be exchanged with another solution by adding it to the reservoir as crossflow continues (Wickramasinghe et al., 2005). There are several critical parameters during a crossflow filtration process of supramolecular structures such as viruses or virus-like particles: transmembrane pressure (TMP) and crossflow velocity (Besnard et al., 2016). TMP is the driving force of liquid through the membrane and crossflow velocity is the rate of the solution flow through the feed and across the membrane. It provides the force for clearing the molecules on the membrane to prevent restriction of permeate flow (Lasareishvili et al., 2021). TMP and crossflow rate were optimized in another study from Ihsan Gursel's lab for concentration and diafiltration of SARS-CoV-2 virus-like particles as 0.5 bars and 300 mL/minutes, respectively. Figure 3.9 and figure 3.10 illustrate the working principle of tangential flow filtration and instrument scheme, respectively.

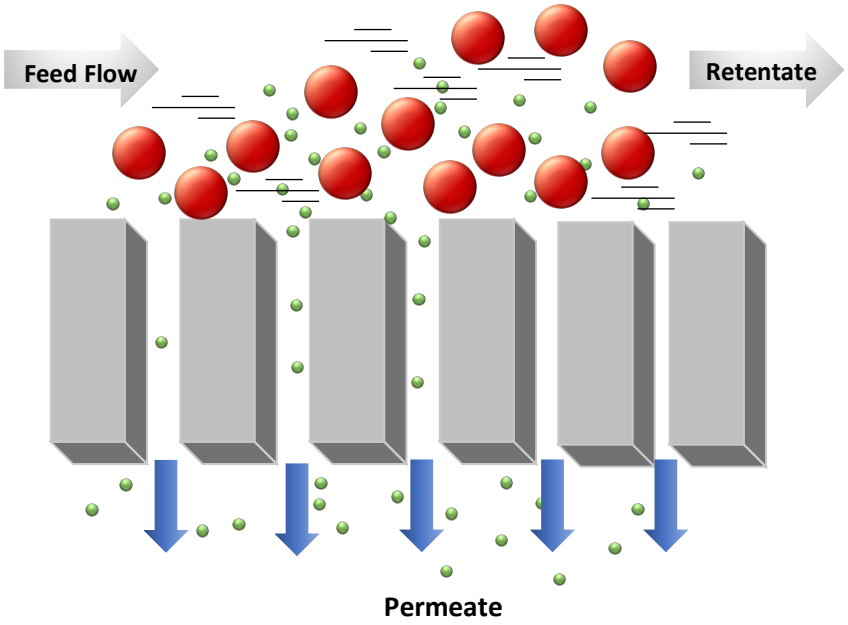


Figure 3.9. Working principle of tangential flow (crossflow) filtration.

Molecules smaller than the cutoff value (green) of the filters(grey) flow through while larger molecules (red) are retained in the stream.

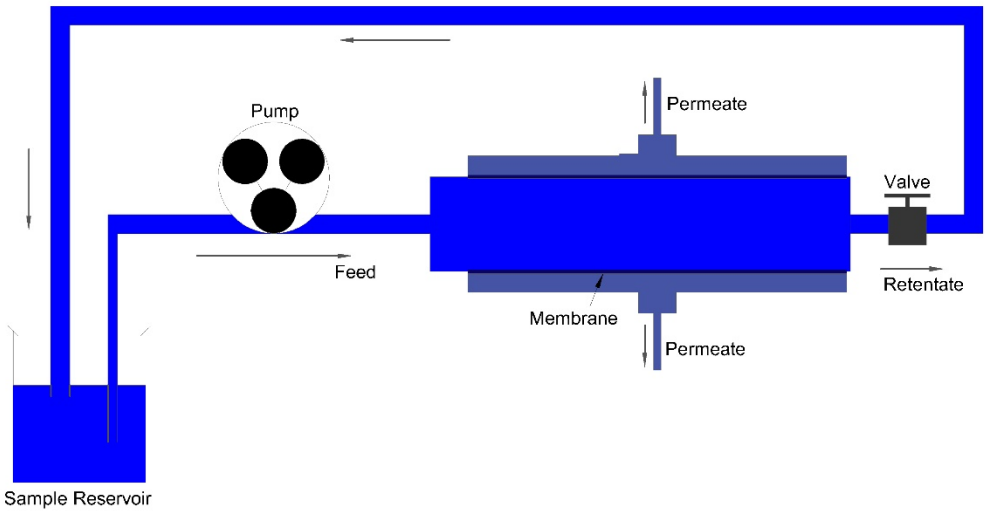


Figure 3.10. Instrument scheme of tangential flow filtration

3.3.1 Evaluation of purification of Virus-like particles via immunoblotting and silver staining

Polyhistidine tagging has been one of the most robust methods for labeling proteins for effortless identification via immunoblotting. For this purpose, Spike, Envelope and Membrane glycoproteins of VLPs were tagged with 6X-Histidine in their C-terminal domain (Figure 2.2). When the purified VLPs were immunoblotted with anti- Histag antibody (figure 3.11), there were three main band signals near 180, 25 and 15 kDa which correspond to relative molecular weights of the Spike, Membrane and Envelope proteins, respectively (Gordon et al., 2020). In addition, because Spike protein was designed to enhance its ability for trimerization by T4 fibritin trimerization domain, the multimeric form of the Spike protein was observed on the membrane. Furthermore, it can be inspected that there were differences in the band intensities of the monomeric and multimeric Spike proteins of 2pS-VLP versus 6pS-VLP when equal amounts of the VLPs were loaded onto the gel and this phenomenon could stem from enhanced stability of the Spike protein in HexaPro configuration (Hsieh et al., 2020).

Additionally, banding pattern change in both medium and high molecular weights could result from the differences in the glycosylation patterns in mutated forms of the Spike protein (Tian et al., 2021). It has been reported that not only was Spike protein affected by changes of the post-translational modifications, but also membrane protein was a strong acceptor of these modifications. Therefore, it could be speculated that the low molecular weight banding pattern changes could be the consequence of glycosylation differences of the membrane protein (Shajahan et al., 2021).

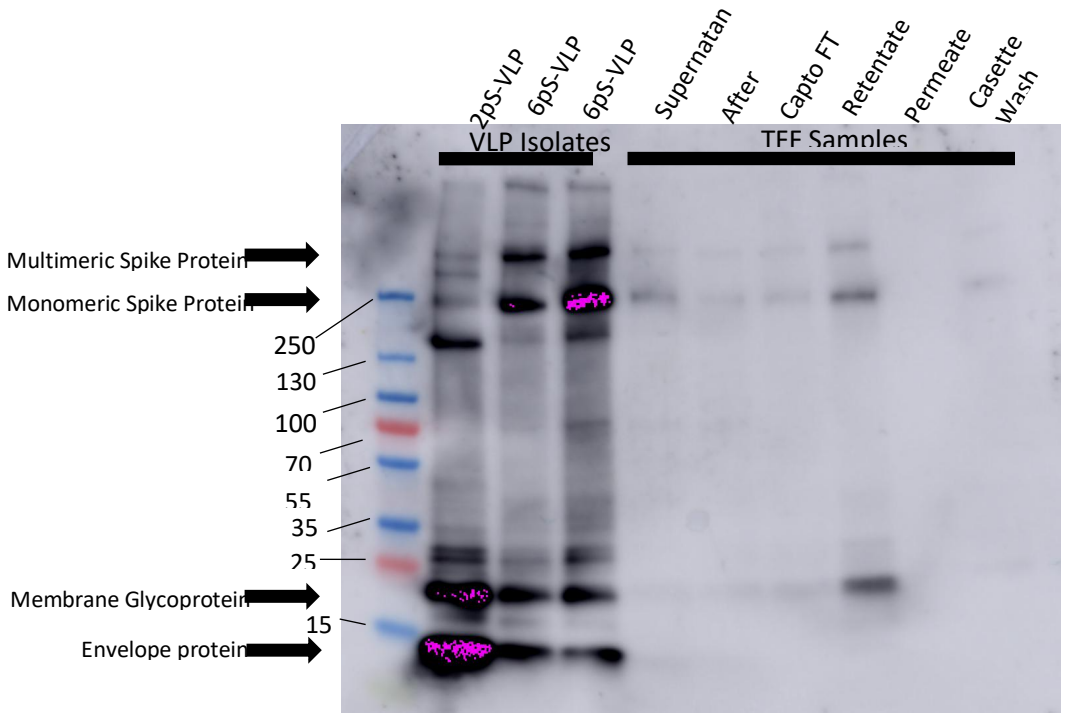


Figure 3.11. Anti-Histag Western blot image of the purified VLPs and samples collected at different stages of the purification/concentration process

Lane 1: 250 kDa Protein Ladder, Lane 2: 2p VLP, Lane 3: 6p VLP Isolate 1, Lane 4: 6p VLP Isolate 2, Lane 5: 6p VLP Harvest Supernatant, Lane 6: 6p VLP Denarase-treated Harvest Supernatant, Lane 7: 6p VLP Capto Flow through, Lane 8: 6p VLP TFF (100 kDa) Retentate, Lane 9: 6p VLP TFF (100 kDa) Permeate, Lane 10: 6p VLP TFF (100 kDa) Cassette Wash

Figure 3.11 also summarizes the VLP mass balance in samples collected through tangential flow filtration, a scalable process for clarifying viruses and virus-like

particles routinely used in vaccine manufacturing (Besnard et al., 2016). It can be observed that there was no signal from the permeate indicating no loss of VLPs through crossflow filtration. Nonetheless, a low signal from the cassette wash step could indicate possible VLP loss owing to the interaction of VLPs and the surface of the TFF cassette. As a result, minor modifications have been implemented to the TFF process as explained in section 2.2.3.3.3 for further large scale purifications, which is a subject of another unpublished study.

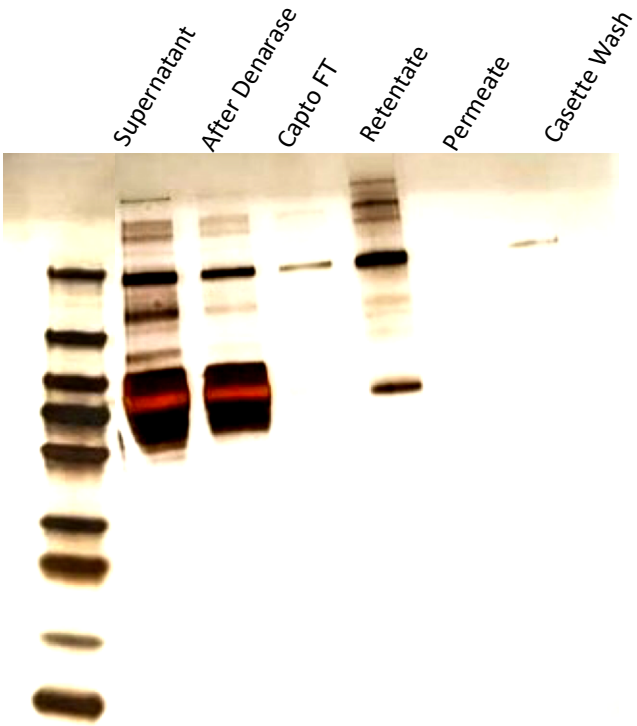


Figure 3.12. Silver stained gel image of samples collected through the TFF process.

Lane 1: 250 kDa Protein Ladder, Lane 2: 6p VLP Harvest Supernatant, Lane 3: 6p VLP Denarase-treated Harvest Supernatant, Lane 4: 6p VLP Capto Flow-through, Lane 5: 6p VLP TFF (100 kDa) Retentate, Lane 6: 6p VLP TFF (100 kDa) Permeate, Lane 7: 6p VLP TFF (100 kDa) Cassette Wash

Total protein silver staining is a commonly used technique for purity assessment of proteins (Chevallet et al., 2006). In figure 3.12, it was clearly shown that most of the

impurities were eliminated through chromatography and tangential flow filtration with minimal loss on the TFF cassette, as previously mentioned. Even though total protein separated via SDS-PAGE was stained with this remarkably sensitive method, some of the expected proteins at low molecular weight were not visible possibly because of their hydrophobic/glycosylated nature (membrane and envelope proteins). However, the elimination of the impurities was easily observed.

Dot blotting was a time-saving method that could be performed instead of western blotting since protein separation via SDS-PAGE step is omitted. Thus, monitoring of the presence of the 6x-Histidine tagged proteins (S, M, E) and Nucleocapsid protein in purified VLPs could be achieved by dot blotting. In figure 3.13 anti-Histag and Anti-Nucleocapsid antibody blots of the purified VLPs were illustrated.

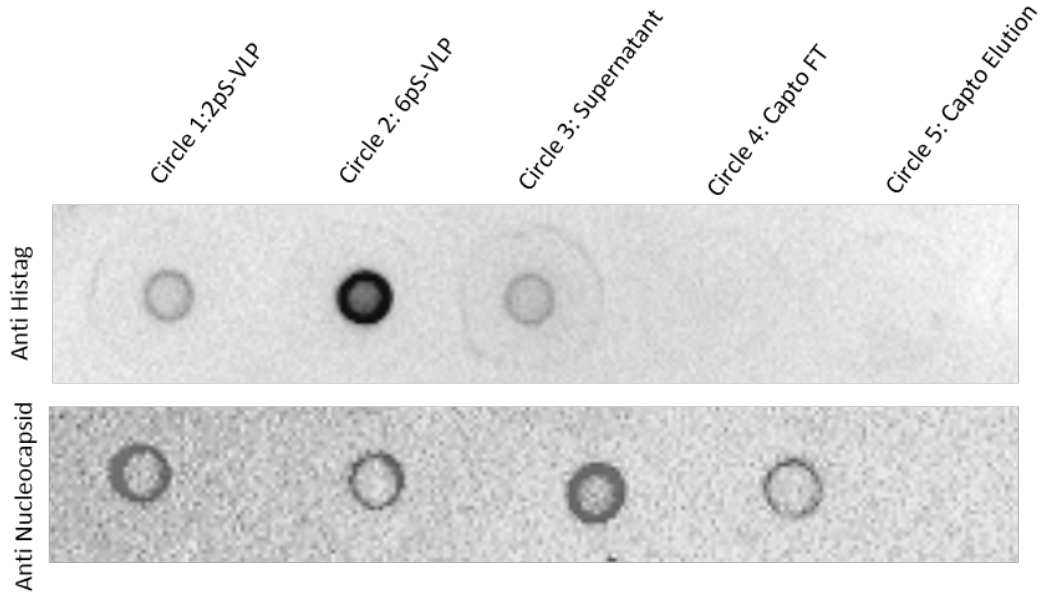


Figure 3.13. Anti-Histag and Anti Nucleocapsid Dot blot images of VLPs.

Upper panel represents Anti-Histag and lower panel shows Anti Nucleocapsid antibody dot blots. Circle 1: 2pS-VLP, Circle 2: 6pS-VLP, Circle 3: Harvest Supernatant, Circle 4: Capto Flow-through, Circle 5: Capto Elution

When equal amounts of VLPs were spotted on the nitrocellulose membrane, it can be seen that as expected, the total signal for 6pS-VLP was higher than 2pS-VLP. Surprisingly, similar signal intensities were acquired when the purified VLPs were

probed with anti-Nucleocapsid antibody. This observation could be explained by the formation of high molecular weight VLP independent Nucleocapsid-host cell RNA condensates (Lu et al., 2021) that are an external source of the Nucleocapsid protein present in the purified VLPs. This external source might create a bias in terms of comparing the Nucleocapsid protein signal on the membrane. On the other hand, the multimodal chromatography elution step sample showed no signal in both blots validating the performance of the chromatography since there was no loss observed.

3.3.2 Characterization of VLP Nanoparticles via Tunable Resistive Pulse Sensing

Tunable resistive pulse sensing (TRPS) is a high precision technique for measuring zeta potential, size distribution and particle counts of nanoparticles. TRPS measurements are performed using a nanopore placed inside of an electrolyte fluidic cell. While sample nanoparticles are flowing through the nanopore by applying voltage and pressure simultaneously, they create a resistive pulse which can be inspected by impedance changes on the nanopore. The physical properties of the resistive pulse are then converted into the particle concentration, size and surface charge by using calibration particles of known properties (Sivakumaran & Platt, 2016). Therefore, following the purification of 6pS/2pS-VLPs, size distribution and particle counts were measured via TRPS. A summary of the TRPS measurements was given in figure 3.14.

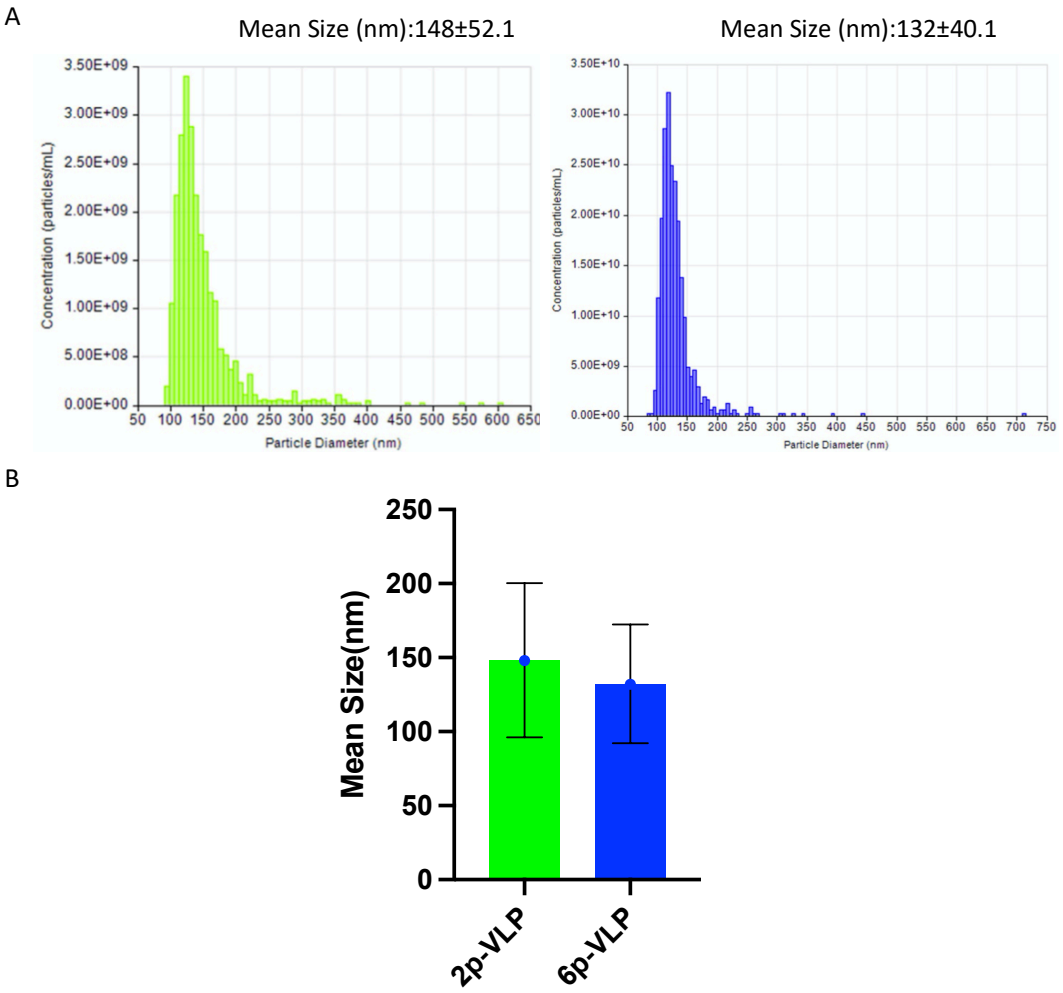


Figure 3.14. TRPS measurements of 6pS-VLPs and 2pS-VLPs.

A: Size distribution of 2pS-VLPs (left) and 6pS-VLPs (right). B: Mean particle size summary of purified VLPs.

Based on these measurements, SARS-CoV-2 virus-like particles resemble the authentic SARS-CoV-2 virions in terms of their mean size and size distribution (Laue et al., 2021).

3.4 Characterization of SARS-CoV-2 Recombinant Spike Protein

3.4.1 Evaluation of purification of recHexaProS protein via silver staining and immunoblotting

There are several methods for purifying histidine-tagged proteins, including immobilized metal affinity chromatography (IMAC) and magnetic beads (Zhang et al., 2020). The former has been employed for more than 50 years to purify histidine tagged proteins following the discovery of binding affinity of transition metals such as cobalt, nickel and zinc to the histidine and cysteine in aqueous solutions (Spriestersbach et al., 2015). Currently, IMAC is one of the most widely applied methods for purifying recombinant proteins ranging from biological function studies to clinical therapeutic applications (Arnau et al., 2006). Briefly, 6x histidine tagged proteins either (N terminus or C terminus) are loaded to the columns containing a resin matrix in which a chelating ligand such as iminodiacetic acid (IDA), nitrilotriacetic acid (NTA) or tris ethylene diamine (TED) form a scaffold for immobilization of transition metals. Following the binding of the target proteins, impurities are eliminated through washing of the resin. Imidazole containing buffers are used to elute the target proteins on the resin since free imidazole substitutes for the imidazole ring on the histidine residue. Most commercially available resins contain Nickel ion as a binding metal for the imidazole rings of the histidine amino acids. Based on this information, to purify recHexaProS from the cell free supernatant, IMAC using HisTrap Excel and Ni-FF resins from Cytiva were performed.

3.4.1.1 HisTrap Excel Method

HisTrap Excel is a Ni Sepharose excel resin-based and ready to use immobilized metal affinity chromatography column specifically used to purify histidine-tagged proteins from the cell free supernatant. It is highly advantageous since its unique

matrix properties make direct loading of the cell free supernatants onto the column possible without nickel shedding (Riguero et al., 2020). Thus, it was aimed to purify recHexaProS by using the HisTrap Excel column. Figure 3.12 describes the silver staining and the anti-Histag blotting of the purified recHexaProS protein to evaluate its identity and purity.

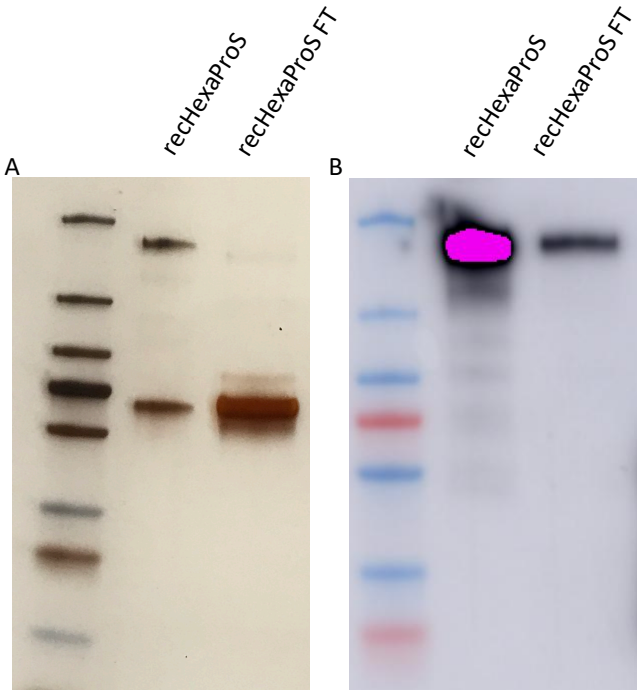


Figure 3.15. Silver staining (A) and the Anti Histag blotting (B) of the HisTrap Excel column purified recHexaProS.

Lane 1: 250 kDa protein ladder, Lane 2: Final recHexaProS product, Lane 3: HisTrap excel column flow-through.

Although purification of the recHexaProS was achieved through the HisTrap Excel IMAC method, several impurities were observed. Those were identified in the current literature through mass spectroscopy such as granulins, lamins and heparan sulfate proteoglycan (Johari et al., 2021). Also, the main co-purified contaminating protein was identified as endoplasmic reticulum chaperone BiP precursor protein by N-terminal sequencing (Cai et al., 2020). Not only were co-purified impurities the

reason for implementing updated purification methods, but also it was the significant amount of target protein presence in the column flow-through meaning decrease in protein yield as it can be observed in the panel B of figure 3.15. In order to improve purification while minimizing product loss, the method described in section 2.2.6.3.1 was executed.

3.4.1.2 Tangential Flow Filtration-Affinity Chromatography Method

To standardize functional serology assays such as ELISA against structural proteins of the SARS-CoV-2, especially Spike protein was the main target of the recent research since pre-existing anti-Spike IgG antibody was strongly associated with protection against SARS-CoV-2 (Addetia et al., 2020). Therefore, a scalable and high yield method for the purification of Spike protein was an urgent need. To this end, a current protocol describing a low impurity, high yield production of the Spike protein was adapted with minor modifications (Esposito et al., 2020). In this method, cell free supernatant was first concentrated through tangential flow filtration. This step was reported to provide for initial removal of the host cell derived impurities which could decrease the binding efficiency of the target proteins to the affinity column. Further wash steps were included with an optimized imidazole gradient to eliminate weakly attached impurities while keeping the target protein bound to the resin. The elution step was also optimized to a fixed imidazole concentration providing the maximum retrieval of the target protein. Since this method was reported to be one of the most efficient protocols for purifying Spike protein of SARS-CoV-2, it has been adapted for conducting serological assays for our vaccination experiments. Following purification of recS protein using this method, protein quality and identity evaluations were given in figure 3.16.

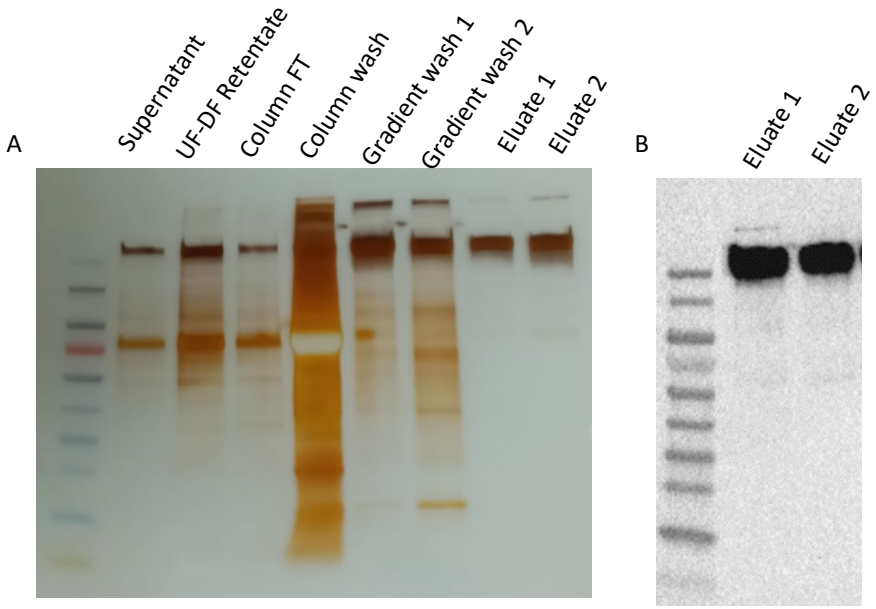


Figure 3.16. Silver staining (A) and the Anti Histag blotting (B) of the TFF-Ni FF column purified recHexaProS.

A)Lane 1: 180 kDa protein ladder, Lane 2: Harvest Supernatant, Lane 3: Retentate, Lane 4: Column Flow-through, Lane 5:Column Wash, Lane 6: 25mM-175mM gradient imidazole wash early fraction, Lane 7: 25mM-175mM gradient imidazole wash late fraction, Lane 8: Elution fraction 1, Lane 9: Elution fraction 2

B) Lane 1: 180 kDa protein ladder, Lane 2: Elution fraction 1, Lane 3: Elution fraction 2

Figure 3.16-A shows that most of the impurities were eliminated through the wash step with 25 mM imidazole containing buffer, as seen in Lane 5. Subsequently, contaminant proteins with a low affinity against immobilized metals on the column were further removed by gradually increasing the imidazole concentration during the wash step. Finally, a step elution with 325 mM was optimized to elute recHexaProS from the column as observed in Lane 8 and Lane 9. When these fractions were probed with anti-Histag antibody as seen in panel B, collected signals were identified as the Spike protein since their position on the gel were correct and monomeric and trimeric conformations were observable. In summary, using the method described above, we were able to obtain high purity recombinant spike protein to be employed in serological assays.

3.5 Determination of Immunogenicity of 2pS-VLP and 6pS-VLP formulations

To determine the immunogenicity of SARS-CoV-2 virus-like particles, Alhydrogel adsorbed, K3-CpG ODN adjuvanted 2p-S VLP and 6p-S VLP (0.4 μg/mouse Low Dose, 4 μg/mouse High Dose) formulations were injected subcutaneously (s.c.) in a total volume of 200 μL with 2 weeks of intervals (on 0 and 14th days). 2p-S (12 mice/group), or 6p-S expressing VLPs (12 mice/group) were formulated only with Alum (5 μg/mice), only with K3-CpG ODN (20μg/mice) or their combination (Table 3.2). Before booster day (day 14) and 14 days after booster injection (day 28), mice were bled from the tail vein to assess primary and secondary humoral responses, respectively.

Table 3.2. Design of vaccination groups concerning antigen type, adjuvantation and dosing

Antigen	*Adjuvantation*	*Dose*
-	Alum	-
-	CpG	-
6pS-VLP	-	Low
6pS-VLP	CpG	Low
6pS-VLP	Alum	Low
6pS-VLP	Alum + CpG	Low
2pS-VLP	Alum + CpG	Low
6pS-VLP	-	High
6pS-VLP	CpG	High
6pS-VLP	Alum	High
6pS-VLP	Alum + CpG	High
2pS-VLP	Alum + CpG	High

Vaccination groups were designed to understand the effect of three parameters in humoral immunity against SARS-CoV-2: the administered dose of VLPs, effect of adjuvantation strategy using Alum and/or K3 CpG ODN and differential

immunogenicity of two proline (2pS-VLP) and six proline (6pS-VLP) mutated Spike harbouring VLPs. Based on this, low dose and high dose VLP formulations were injected to mice to evaluate the difference of early (primary bleeding) and late (secondary bleeding) humoral responses in terms of total IgG, IgG1 and IgG2a levels in mice against the Spike protein of SARS-CoV-2. Furthermore, non-adjuvanted, single adjuvanted (Alum or CpG) and dual adjuvanted (Alum+CpG) VLP formulations were compared in terms of their potential to elicited antibody subtypes especially for IgG2a, since IgG2a in mice is known to be the most effective immunoglobulin subtype for viral clearance (Rubtsova et al., 2013). Furthermore, a comparison of dual adjuvanted 2pS-VLP and 6pS-VLP elicited antibody titers were evaluated to understand whether *in vitro* enhanced stability of Spike proteins decorated on the VLP surface increases the immunogenicity of the vaccine formulations.

To assess SARS-CoV-2-specific VLP-induced humoral responses, Spike-, nucleocapsid- or RBD-specific total IgG, IgG1 and IgG2a levels from the sera of the immunized mice were measured by ELISA. Following color development, OD values were utilized to calculate the endpoint antibody titers based on the Frey method (Frey et al., 1998) in which reciprocal dilution value, where the OD decreases to 3 standard deviations plus mean of background (placebo) is considered as the endpoint titer.

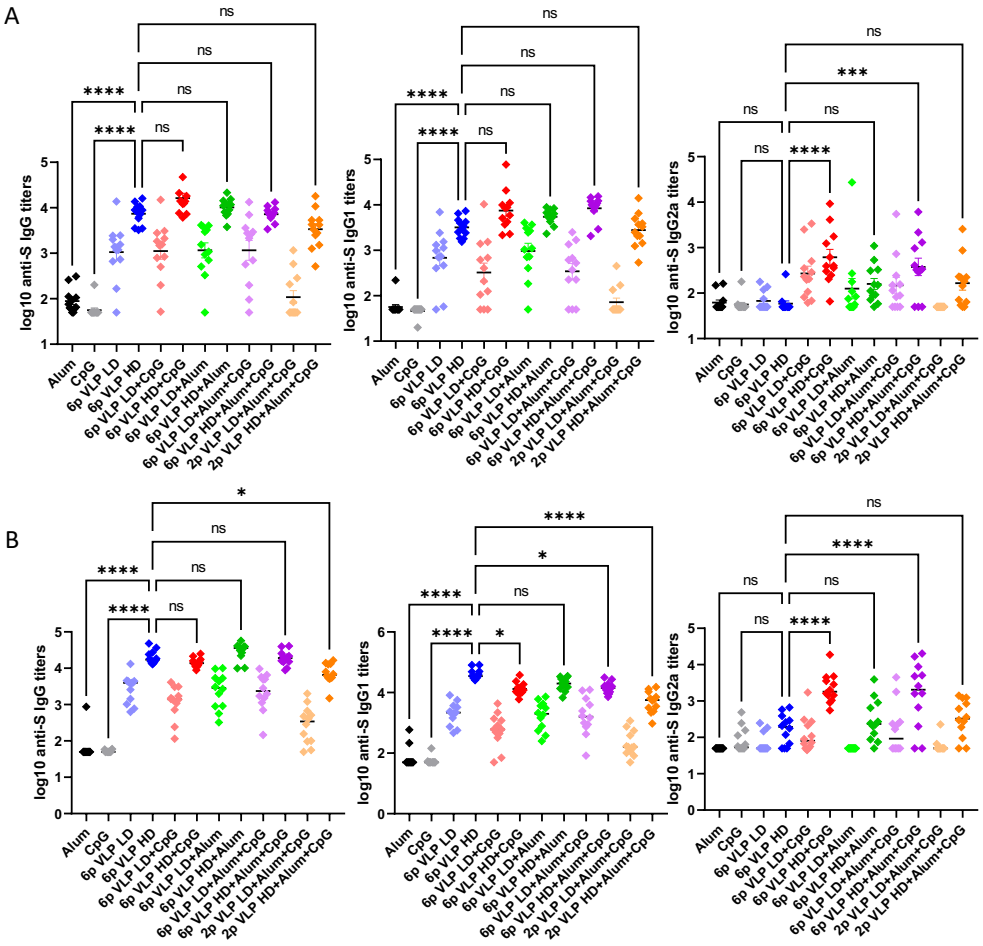


Figure 3.17. Assessment of SARS-CoV-2 S-specific IgG, IgG1, and IgG2a titers in vaccinated mice

Panel A: Sera were collected 2 weeks post-prime and Panel B: 2 weeks post-boost (B). Vaccinated groups were compared by one-way ANOVA with Dunnett's multiple comparisons test. *P < .05, **P < .01, ***P < .001, ****P < .0001. Data are presented as GMT ± geometric SEM. LD: Low Dose, HD: High Dose

Figure 3.17-A shows that all vaccinated groups displayed detectable levels of antibodies against Spike protein even after post-prime bleeding (day 14). Detectable post-prime antibody levels were predominantly observed in high-dose VLP groups, proving the rationale for using 4 µg/mouse for further vaccination experiments. These antibody titers were elevated to even higher values after booster injection as seen in panel B. Figure 3.18 further illustrates the fold inductions in total IgG levels in mice vaccinated with high dose or low dose VLP formulations.

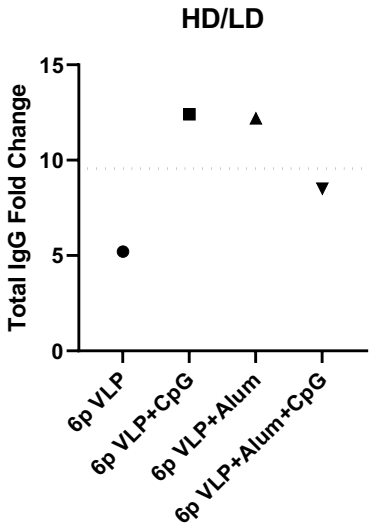


Figure 3.18. Total IgG Fold change against S protein (HD/LD : High dose over Low dose titer) in different vaccination groups.

Dashed line shows Mean Fold Change: 9.55

Fold induction of low dose and high dose VLP formula elicited total IgG antibodies were described as the ratio of high dose to low dose VLP elicited total IgG titer. These ratios demonstrate that when the high dose VLP containing formulations were injected to mice, an average of 9.55 times higher total IgG against S protein was elicited as compared to the low dose VLP vaccinated groups, proving the advantage of administering of high dose VLP for enhanced humoral immunity.

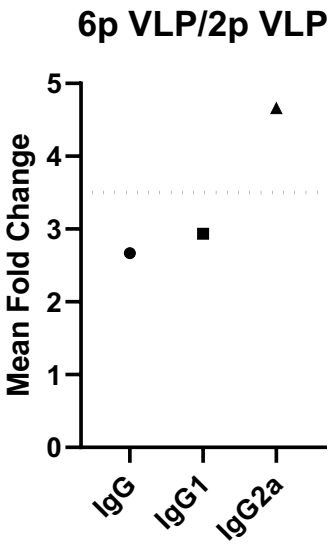


Figure 3.19. Fold change of total IgG, IgG1 and IgG2a antibodies against S protein of 6p-S VLP over 2p-S VLP injected mice.

Dashed line shows Mean Fold Change: 3.5.

Figure 3.19 shows that when the mice were immunized with 6p VLP + CpG + Alum, fold change over 2p VLP + CpG + Alum elicited total IgG, IgG1 and IgG2a antibodies were 2.7, 3 and 4.7, respectively. These fold changes indicate the superior immunogenicity of prefusion stabilized hexaproline mutated Spike protein over its 2-proline mutated counterpart.

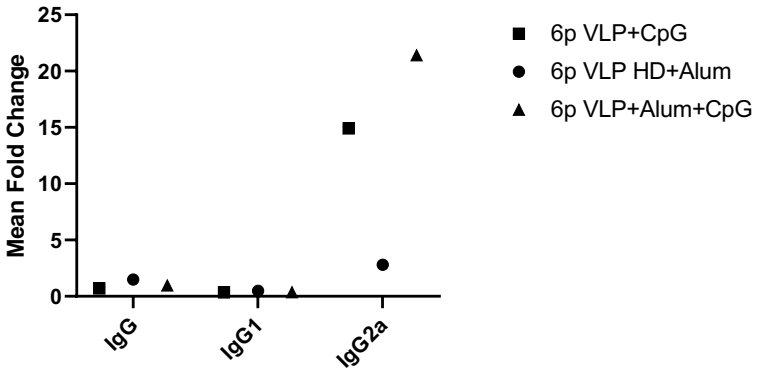


Figure 3.20. Fold change of total IgG, IgG1 and IgG2a antibodies against S protein of adjuvanted VLP formulations over non-adjuvanted VLP injected mice.

Figure 3.20 demonstrates that compared to their non-adjuvanted counterparts, CpG only or CpG + Alum adjuvanted high dose VLPs triggered 14- and 21- fold more IgG2a antibody titers while there was no significant fold change observed in total IgG and IgG1 levels. Here, it could be pointed out that elevated IgG2a titers was a result of T-helper 1 type of immune response as a consequence of using the K3-CpG oligonucleotide. As a result, based on previous findings demonstrating the effect of T-helper 1 type of immune response on the immune protection against viral infections (Swain et al., 2012), K3 CpG adjuvantation poses a significant advantage over non-adjuvantation or single Alum adsorption of VLPs.

Anti-Nucleocapsid non-neutralizing antibodies and RBD directed antibodies can prevent the cell-to-cell spread of the virus (Zeng et al., 2021) and binding of the virus to ACE2, respectively (Trougakos et al., 2021). Therefore, levels of total IgG, IgG1 and IgG2a titers against Receptor binding domain of Spike protein (RBD) and Nucleocapsid protein were evaluated from the sera of the vaccinated mice following secondary bleeding.

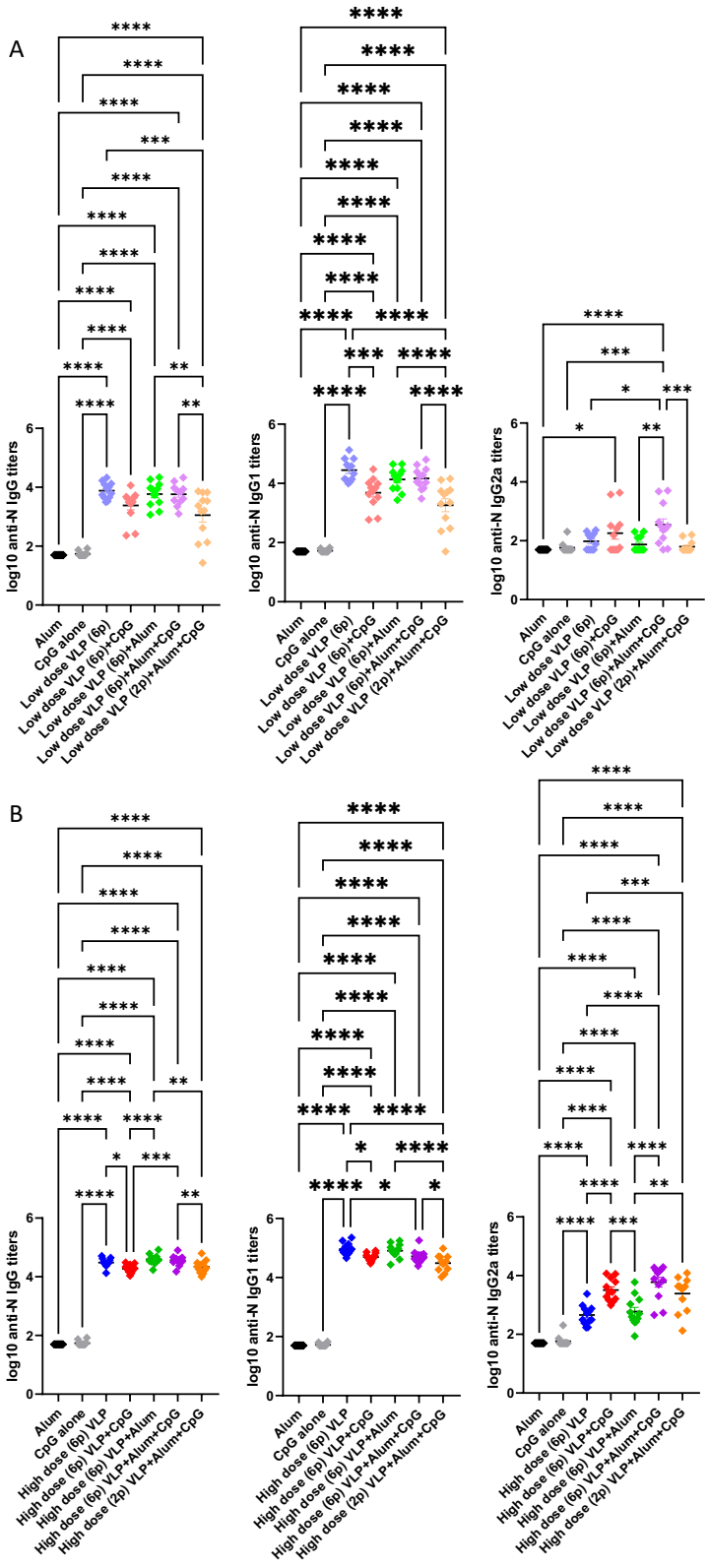


Figure 3.21. Assessment of SARS-CoV-2 N-specific IgG, IgG1, and IgG2a titers in vaccinated mice

Panel A: Low dose VLP injected groups' sera 2 weeks post-boost, Panel B: High dose VLP injected groups' sera 2 weeks post-boost (B). Vaccinated groups were compared by one-way ANOVA with Dunnett's multiple comparisons test. *P < .05, **P < .01, ***P < .001, ****P < .0001. Data are presented as GMT ± geometric SEM.

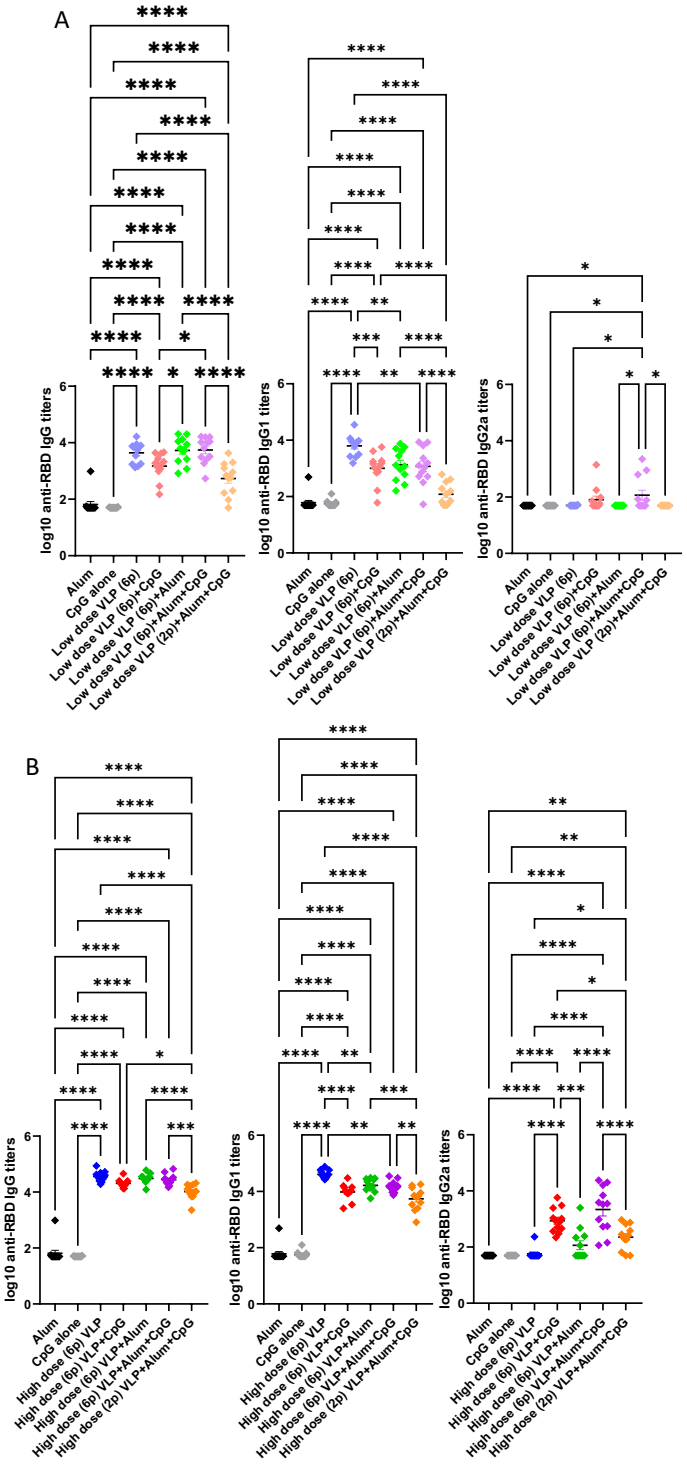


Figure 3.22. Assessment of SARS-CoV-2 RBD-specific IgG, IgG1, and IgG2a titers in vaccinated mice

Panel A: Low dose VLP injected groups' sera 2 weeks post-boost, Panel B: High dose VLP injected groups' sera 2 weeks post-boost (B). Vaccinated groups were compared by one-way ANOVA with Dunnett's multiple comparisons test. *P < .05, **P < .01, ***P < .001, ****P < .0001. Data are presented as GMT ± geometric SEM.

Results presented in Figures 3.21 and 3.22 suggest that all VLP formulations elicited robust humoral responses against RBD and N proteins similar to secondary responses observed against the Spike protein. Moreover, only the high dose VLP + CpG + Alum immunized mice produced IgG2a subtype of antibodies against the Nucleocapsid and the RBD, re-emphasizing the importance of adjuvantation using K3 CpG oligonucleotide.

3.6 Determination of T-helper cell responses in vaccinated animals

LEGENDplex™ MU Th Cytokine Panel (12-plex) is a multiplexed bead-based assay that employs a similar working principle as a sandwich ELISA to capture the secreted cytokines from cellular supernatants. In this assay, bead populations with different sizes and differing levels of internal APC fluorescence are conjugated with capture antibodies for cytokines secreted mainly from helper T cells. When the supernatants from the cells and the beads are incubated together, these cytokines are captured by the antibodies immobilized on the beads. Subsequently, another biotinylated antibody specific for each cytokine is incubated to create the "sandwich" formation. In order to amplify the signal coming from the captured cytokines, Phycoerythrin (PE) conjugated streptavidin is added to the bead suspension to provide binding of streptavidin and biotin conjugated detection antibodies. Finally, beads are acquired on a flow cytometer and collected data are analyzed through a software created by the manufacturer. Bead classes are differentiated through the FSC and SSC parameters and one type of cytokine capturing bead cluster is determined by its APC signal since each type of cytokine is captured by a single bead population. While bead class A (smaller beads) represents IL-5, IL-13, IL-2, IL-6, IL-9, IL-10, bead class B (larger beads) represents IFN-γ, TNF-α, IL-17A, IL-17F,

IL-4, IL-22 cytokines. Cytokine-specific beads are then discriminated with respect to their increasing fluorescent intensities at the APC channel for bead class A or B. PE signal present in a particular bead class is directly correlated with the amount of the specific cytokine captured by the antibodies attached to the bead. Quantifying cytokine concentration is achieved by using the standard cytokine cocktail in which known concentrations of particular recombinant cytokines are run through similar steps to construct a standard curve for each cytokine. Conversion of the PE signal of the unknown sample to cytokine concentration is performed by constructing a five-parameter logistic regression curve using the PE signals of the standards on the Legendplex software. The assay principle is summarized in figure 3.20.

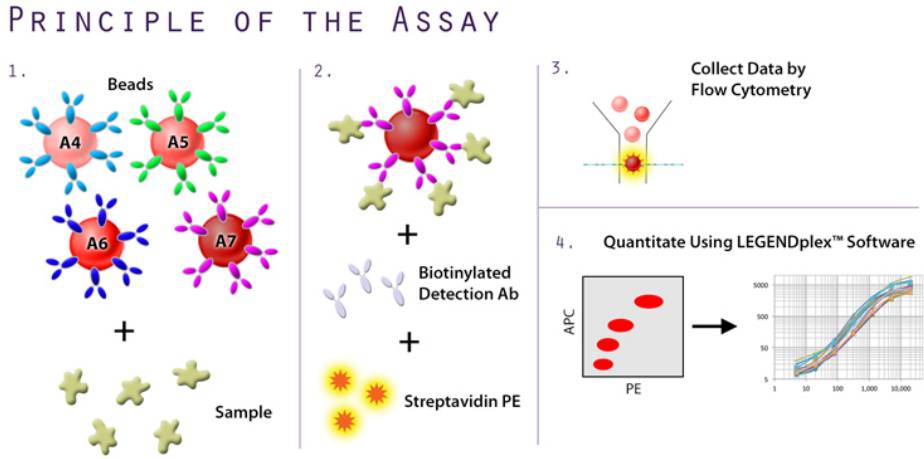


Figure 3.23. LEGENDplex™ MU Th Cytokine Panel (12-plex) assay principle.

In the light of this information, supernatants from antigen recall assay in which splenocytes extracted from vaccinated animals were stimulated with recHexaProS in the presence of anti-CD28 antibody for costimulation. Antigen-recognizing T-cells, especially CD4+ T cells then produce and secrete various cytokines. Helper T cell polarity (Th1/Th2) established by vaccination was determined from the spectrum of released cytokines through Legendplex Th cytokine panel. Hence, to assess the profile of the secreted cytokines from the splenocytes of the vaccinated animals, the Legendplex T helper cytokine kit was utilized.

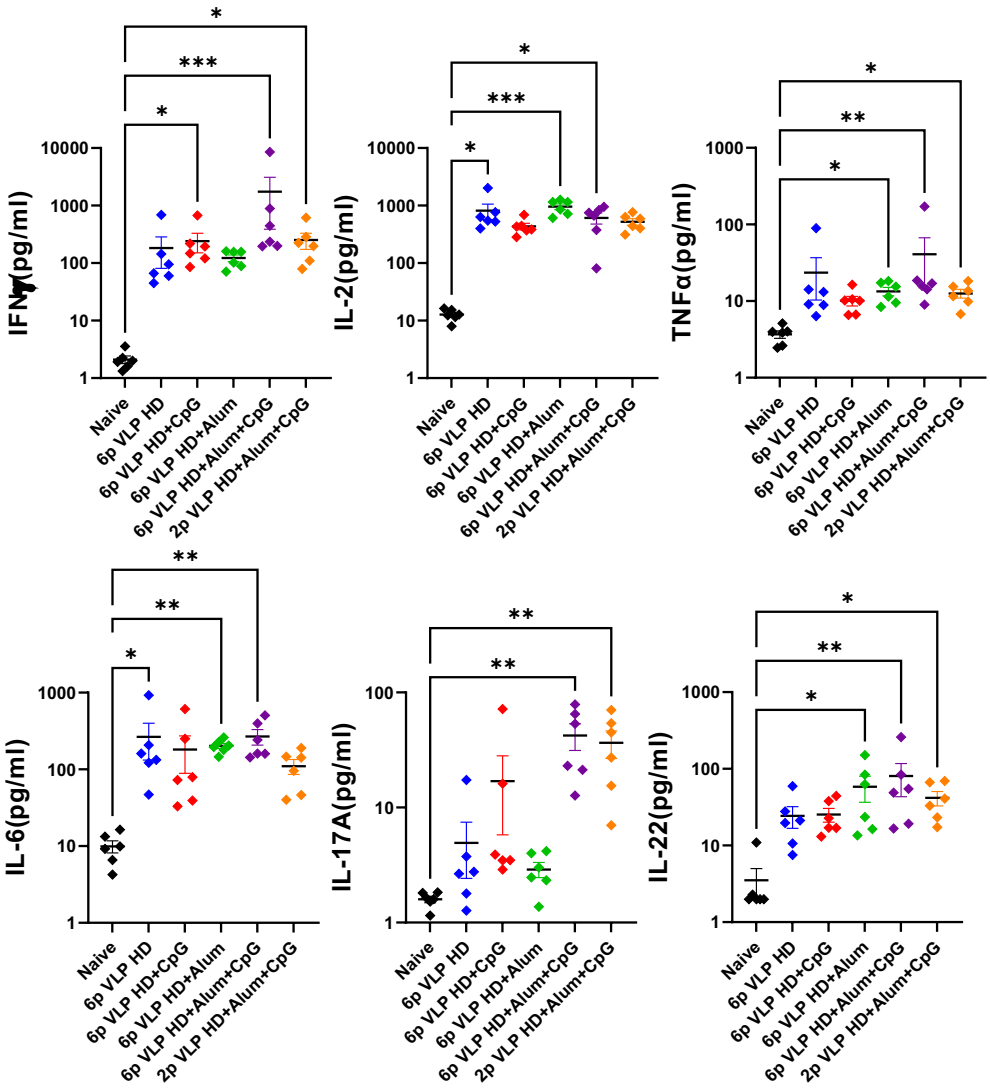


Figure 3.24. Th1 associated cytokine response from recHexaProS stimulated splenocytes.

Vaccinated groups were compared by one-way ANOVA with Dunnett's multiple comparisons test. *P < .05, **P < .01, ***P < .001, ****P < .0001.

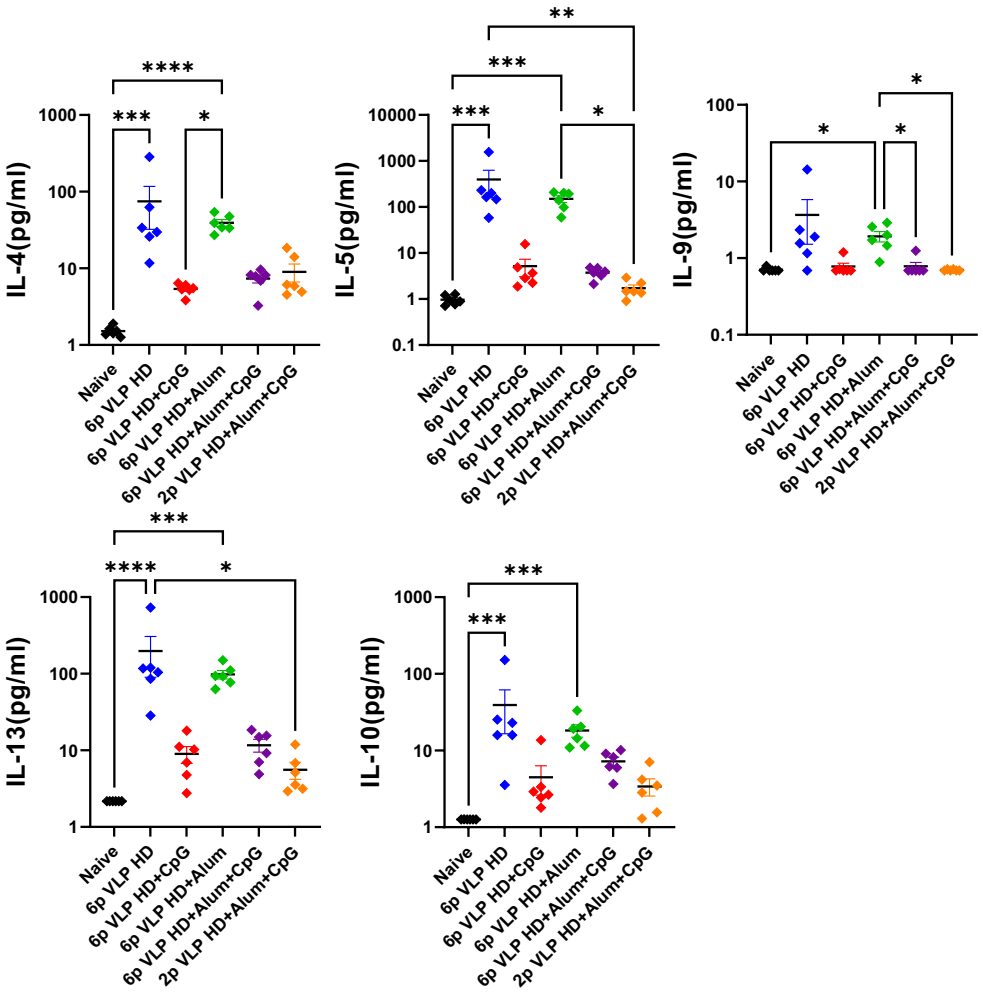


Figure 3.25. Th2 associated cytokine response from recHexaProS stimulated splenocytes.

Vaccinated groups were compared by one-way ANOVA with Dunnett's multiple comparisons test.
*P < .05, **P < .01, ***P < .001, ****P < .0001.

Figures 3.24 and 3.25 show that splenocytes from mice immunized with 6p-VLP or 6p-VLP+ Alum, secreted significant amounts of Th2 cytokines IL-4, IL-5, IL-13, and IL-10. In contrast, only the K3-CpG or K3-CpG+Alum-adjuvanted VLPs induced a Th1-biased IFN-γ response but no Th2-associated cytokines. All in all, this data suggests that tempering the potential Alum induced immunopathology could be achieved by 6p-VLP/Alum/K3-CpG vaccination to avoid Th2-dependent vaccine-associated enhanced respiratory disease (VAERD) (Bolles et al., 2011). To assess the vaccine-induced T helper cell subclass responses more efficiently, data

were also summarized as pie charts (Figure 3.26). Results demonstrate that Th1 dominating response (red) could be observed in the groups K3-CpG or K3-CpG+Alum-adjuvanted VLPs while repressing Th2 associated cytokines (green).

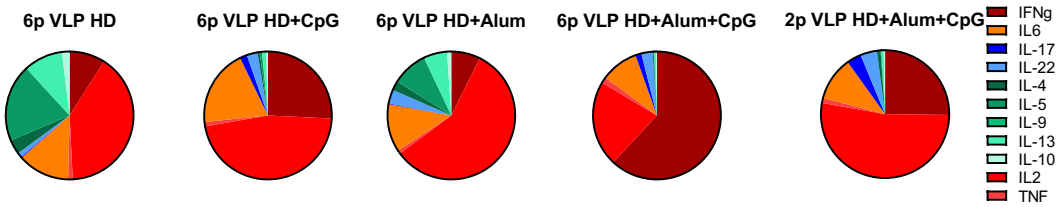


Figure 3.26. Proportions of individual secreted S-specific T helper cell cytokines.

Collectively, evaluation of humoral and cellular responses of vaccinated mice suggests that Alum adsorbed, K3 CpG adjuvanted high dose 6p-S VLP formulation is the most potent and promising vaccine candidate to confer protection against SARS-CoV-2 infection.

CHAPTER 4

CONCLUSION AND FUTURE PERSPECTIVES

Our initial joint studies with Bilkent University were focused on producing SARS-CoV-2 virus-like particles containing four structural proteins harbouring the wild-type Spike protein. In these early studies, we aimed to produce laboratory-scale VLPs in an adherent HEK293 cell line with co-transfecting dual expression plasmids using common transfection reagents including calcium phosphate or liposome-based reagents. The purification strategy was similar to the isolation of extracellular vesicles using ultracentrifugation. These initial studies demonstrated the possibility of producing mammalian cell-derived immunogenic VLPs for protective immunity against SARS-CoV-2. In this thesis, we aimed to improve the immunogenicity of the VLPs as well as to develop a scalable process for large-scale manufacturing of the SARS-CoV-2 virus-like particles for clinical use.

To increase the immunogenicity of the VLPs, we designed two new Spike genes of SARS-CoV-2 to ensure its prefusion conformation and enhanced thermostability with two (2p) and six proline (6p; HexaPro) substitutions (figure 2.2) (Hsieh et al., 2020), respectively. These genes and other structural protein-encoding genes including Envelope, Membrane and Nucleocapsid were human codon-optimized. We have cloned these four genes into two separate dual expression plasmids rather than employing four different constructs to increase the probability of co-transfection of the genes to a single cell for obtaining fully assembled VLPs following simultaneous expression of four structural proteins in a single cell. Furthermore, through partitioning of the 2 vital genes for VLP production (M and E) into separate plasmids, we made sure that the cells would assemble VLPs only if they were transfected with the 2 plasmids simultaneously, precluding the possibility

of VLP generation devoid of the S protein. Other configurations of genes in the plasmids are still under investigation to increase and fine-tune the production of the fully assembled VLPs. Parameters affecting the efficiency of expression of VLP proteins, including differential transcription due to various mammalian promoters on the vectors could be elucidated in further studies. Yet, based on the results obtained from gel image of restriction enzyme digested constructs and next-generation sequencing, we have verified the successful cloning of the structural protein-encoding genes of SARS-CoV-2 into pVITRO1 and pVITRO2 mammalian dual expression plasmids.

Another step to increase the scalability of this process was the transfection optimization of suspension HEK cells. As mentioned before, laboratory-scale HEK293 cells are adherent cells that grow by attaching onto flask surfaces. This limits cell propagation in large volumes and numbers due to their high space necessity. Therefore, suspension HEK293 cell culture is more advantageous than using adherent culture due to its ease of scaling up to higher cell numbers and large volumes as well as their enhanced productivity. Moreover, they are ready to adapt to even larger culture volumes up to 1000 L by utilizing industrial bioreactors.

Widely used gene delivery methods employing cationic liposomes or calcium phosphate have a number of drawbacks, including low transfection efficiency on large scales and incompatibility with GMP production pipelines for clinical use of vaccine antigens. To circumvent this problem, we employed a GMP production-compatible, high-efficiency transfection reagent, PEIpro. Since the transfection protocol vary depending on the cell type and density, plasmid type, transfection reagent to plasmid amount ratio, we first optimized the starting cell density and PEIpro to plasmid amount (with constant DNA amount) ratio during the time of transfection. For ease of monitoring, we utilized a GFP encoding plasmid for transfection, followed by Fluorescence-activated cell sorting (FACS) analysis of continuous batch cultures. Our results revealed that 2 µL PEIPro to 1 µg DNA ratio (2x) yielded the highest transfection percentage at each time point with no significant impact of starting cell density. Therefore, we decided to continue with the lowest

starting cell density possible (1.3×10^6 cells/mL) with 2x PEIpro to DNA ratio for subsequent VLP transfection experiments. The rest of the parameters concerning the upstream optimization of transfection including the day of harvest, comparison of shake flask and bioreactor productions and batch-to-batch variations due to culture conditions such as proper gas exchange, pH and supplementary media components of the culture were also investigated in our joint studies with Nobel Biotechnology during the GMP production of VLPs, but the data was outside the scope of this thesis and hence were not included here.

Next, we devised a purification strategy that is also applicable for large-scale purification of VLPs. Our strategy was solely based on removing host cell-derived nucleic acids, proteins and media components present in harvest supernatant. Following clarification of the supernatant via centrifugation to remove cellular debris, for elimination of large nucleic acid fragments, we employed a GMP production compatible endonuclease, Denarase, which is capable of cleaving both RNA and DNA into 5-8 bases long fragments. Widely used chromatographic techniques for purification of supramolecular structures, include size exclusion chromatography (SEC) in combination with anion exchange chromatography. These multistep approaches have several disadvantages such as the risk of introducing adventitious organisms due to a non-sterile working environment (chromatography systems, columns and resin) and a high probability of product loss in-between chromatographic steps. Thus, we selected a one-step multimodal chromatography resin called Capto Core 700 (figure 3.7). This resin is capable of SEC and anion exchange at the same time, ensuring the removal of host cell-derived proteins, small fragment nucleic acids and media components via anion exchange while size excluding the supramolecular structures such as VLPs, since these are larger than 700 kDa. The resin's time-saving physical features such as resistance to high pressures and high flow velocities make large-scale purifications possible for manufacturing purposes. Using this resin, we focused on the removal of the impurities which can then be eluted, following completion of VLP collection in flow-through fractions. A large peak at 280 nm appears during the elution step which is

associated with host cell-derived proteins, nucleic acids and most of the pigments of the culture medium as observed in a representative chromatogram (Figure 3.8).

Tangential flow filtration is also another scalable process for concentrating and buffer exchanging of the VLP containing flow-through fractions to obtain the injectable volumes of VLP antigens for preclinical and clinical use. Our initial studies were based on the use of ultracentrifugation and centrifugal ultrafiltration units with a particular molecular weight cutoff, which was time-consuming, resulted in low yields and was not suitable for handling large scale volumes. Therefore, we employed a tangential flow filtration process to concentrate the product further while buffer exchanging the product into an injectable solution (DPBS). All parameters concerning TFF, including TMP, crossflow velocity and volume concentration factor were optimized in joint studies with Bilkent University. Again, removal of the impurities was monitored until the OD value of the permeate reached an OD value of the DPBS at 280 nm (data not shown), indicating that input and output contents are similar and successful buffer exchange is completed.

Next, we aimed to show the identities of the VLP forming proteins via Western and dot blotting and for the assessment of impurities, samples were silver stained. When polyhistidine-tagged proteins of VLPs (S, E, M) were blotted with anti-Histag antibody, we observed the signals corresponding to relative expected molecular weights of the proteins. A semi-quantitative comparison of 2p and 6p VLPs indicated that 6p VLPs generated enhanced monomeric Spike band intensities, suggestive of enhanced stabilization and possibly increased immunopotency of the modified spike. In addition, including an exogenous trimerization signal for both Spike sequences resulted in elevated signals of bands, possibly corresponding to the multimeric forms of the Spike protein, especially in 6p VLP isolate. Based on previous findings demonstrating the effect of large antigens with repetitive structures for enhanced B cell receptor signaling (Avalos & Ploegh, 2014), observing more multimeric forms of the Spike protein in 6p VLP isolates is beneficial for induction of a robust humoral response. Furthermore, possible glycosylated forms of Spike and membrane proteins might be beneficial for host immune response following vaccination (Tian et al.,

2021), yet this observed phenomenon needs further elucidation. Analysis of TFF samples indicated a minimal loss on the membrane. Silver-stained VLP samples remained inconclusive to show the identities of the VLP forming proteins, yet, consistent with western blot results, we concluded that most of the impurities were removed through multimodal chromatography and no significant loss was observed during the TFF process. Finally, dot blotting was conducted to demonstrate the presence of nucleocapsid protein in the VLP samples. Results suggested that some VLP-unassociated nucleoprotein/nucleic acid condensates might exist during VLP production, but following Denarase treatment and purification process, this amount is minimized.

Next, we re-confirmed the success of isolation of VLP nanoparticles by TRPS measurements, where we observed that the mean size of our VLP isolates are in parallel with the size range of the authentic SARS-CoV-2 virion isolates (Laue et al., 2021).

Another objective of this study was to produce recombinant Spike protein of SARS-CoV-2 to conduct in-house serological assays for monitoring of VLP elicited anti-Spike immunoglobulins in mice. Thus, we designed a human codon-optimized six proline substituted (6p; HexaPro) Spike protein-encoding gene flanked with 5' CD33 signal sequence and 3' 6x histidine tag and successfully cloned into the pVITRO2 mammalian expression plasmid based on the results of digested construct's gel image and next-generation sequencing. Furthermore, we excluded the transmembrane domain of the Spike protein (figure 2.2-TM) in order to enhance the secretion of the protein to the culture supernatant. Purification of the secreted proteins are more advantageous than cell-retained recombinant proteins since cell lysis reagents could introduce denaturation of the target protein, less amount of host cell-derived proteins are present as impurities in the culture supernatant and low concentration of cell proteases are secreted that prevents proteolytic degradation of the protein of interest. To purify these secreted proteins following similar transfection protocols for VLPs, we first focused on utilizing the HisTrap Excel IMAC column prepacked with a particular resin optimized for loading non-clarified

cell supernatants without nickel shedding due to chelators present in the supernatant. Also, this resin is resistant to high pressures, making it suitable for the rapid purification of recombinant proteins under high flow velocities. Nevertheless, we have observed low yield and purity of our recombinant Spike isolates. To overcome this problem, we have adapted an up-to-date affinity chromatography protocol optimized for purifying specifically recombinant Spike protein of SARS-CoV-2. Characterization of the isolates displayed an excellent purity and yield and very high performance in the serosurveys conducted with the sera of the immunized mice. This high performance resulted in obtaining a very high purity recombinant Spike protein of SARS-CoV-2 showing minimal background values in ELISA experiments. Adapting this protocol enabled us to produce reproducible high purity recombinant Spike protein to standardize our in-house serological assays conducted with immune sera obtained from various species including immunized mice, rats, ferrets, humans and human convalescent plasma samples throughout the joint studies we have been conducting in our and Ihsan Gursel' lab in Bilkent University.

Next, we have elucidated the humoral and cellular responses against various 2p and 6p VLP formulations. To determine humoral responses, 2 different doses of both types of VLPs were formulated with Alum and/or CpG (Table 3.2). These formulations were administered subcutaneously to mice with a two week interval. To determine whether VLP formulations triggered an early humoral immune response (day 14), sera were collected and S-specific IgG ELISA was performed. Primarily, mice that received a high dose of VLP vaccine, displayed detectable antibody levels against the Spike protein. Moreover, secondary responses (day 28) of the high dose VLP immunized mice demonstrated a 9.50 fold higher total anti-Spike IgG titers. It was the first *in vivo* study where a standardized VLP purification strategy was followed and thus, it has been concluded that high dose VLP (4µg/mice) should be considered immunogenic (Rhodes et al., 2019).

Another focus of this immunization study was to compare 6p and 2p VLP-elicited antibody titers, given the *in vitro* super-stability of the 6p over 2p mutated Spike protein. Our results showed a 3.5-mean fold more total anti-S IgG, anti-S IgG1 and

anti-S IgG2a titers in 6p VLP injected mice when both VLPs were formulated with K3 CpG and Alum, demonstrating that hexaproline substituted (6p) Spike harboring VLPs were superior to their two proline (2p) counterpart in terms of immunogenicity.

Determination of the most suitable adjuvant to formulate the SARS-CoV-2 VLPs was a prominent aspect of this immunization study to elicit a Th1 dominated response for both humoral and cellular immunity to confer protection against SARS-CoV-2 based on previous findings demonstrating the effect of Th1 dominated immunity on durable protection against viral infections (Swain et al., 2012). Therefore, IgG2a titers against Spike protein in mice were evaluated due to the essential effector functions of IgG2a antibodies in viral clearance mechanisms (Huber et al., 2006). We observed that K3 CpG adjuvanted 6p VLP formulation immunized mice displayed on average 18- fold more IgG2a titers compared to their non-adjuvanted counterparts. Furthermore, only Alum adsorbed K3 CpG adjuvanted high dose 6p VLP formulation elicited the highest titers of anti-Nucleocapsid and anti-RBD IgG2a antibodies, which could play a role in the prevention of cell-to-cell spreading and canonical neutralization of the virus, respectively (Trougakos et al., 2021; Zeng et al., 2021).

Next, we evaluated the cellular immune responses of the vaccinated animals via an antigen recall assay followed by multiplexed cytokine measurement from the supernatants of recombinant Spike protein stimulated splenocytes. Our data demonstrated that secretion of Th2 associated cytokines including IL-4, IL-5, IL-10 and IL-13 were observed only in VLP and VLP + Alum injected animals. These unwanted effector cytokines that are associated with a potential to develop VAERD, were suppressed using K3 CpG in the vaccine formulation, through potentiation of Th1 cytokines, including IFN-γ. These results verified our rationale to use CpG ODN to confer cellular immunity against SARS-CoV-2 infection.

In conclusion, CpG mediated Th1 responses and antibody repertoire against SARS-CoV-2 proteins are broadened by the use of fully assembled VLPs since all four

structural proteins possess T-cell epitopes (Grifoni et al., 2020; Matchett et al., 2021). Furthermore use of bidirectionally functioning adjuvant combination; CpG and Alum, has increased the immunoprotective potential of the vaccine since CpG suppress the Alum induced immunopathology and confers Th1 dominating response, whereas alum adsorption facilitates the simultaneous delivery of the CpG and VLPs to the draining lymph nodes (Davis, 2008; Lopez et al., 2006). All in all, we concluded that Alum adsorbed CpG adjuvanted high dose 6p VLP is the most promising vaccine candidate against SARS-CoV-2, generating a robust and well-orchestrated humoral and cellular response.

We have so far conducted further studies to prove the protective efficacy of the vaccine candidate. We developed virus-like particles decorated with the Alpha variant's Spike protein following the emergence of this SARS-CoV-2 variants of concern (Abdool Karim & de Oliveira, 2021). Similar experiments with Wuhan and Alpha 6p VLP formulations were conducted with other species, including rats and ferrets to assess cellular and humoral immunity along with live virus neutralization capability of vaccine-elicited antibodies. Moreover, to evaluate the immunoprotectivity of the candidate formula, K18-hACE2 transgenic mice were immunized for intranasal challenge with authentic SARS-CoV-2. Vaccinated groups displayed a low virus titer with no lung pathology as compared to placebo groups, further demonstrating the immunoprotective potential of 6p VLP + K3 CpG + Alum formulation (Yilmaz et al., 2021).

Next, we have produced the bulk VLP antigen under GMP conditions in collaboration with Nobel Biotechnology to register for human clinical trials. GMP produced and formulated 6p Wuhan and Alpha SARS-CoV-2 VLPs were administered to volunteers in a phase I and a Phase II clinical trials: "Study of a Severe Acute Respiratory Syndrome CoV-2 (SARS-CoV-2) Virus-like Particle (VLP) Vaccine in Healthy Adults (COVID-19)" (ClinicalTrials.gov identifier- Phase I: NCT04818281, Phase II: NCT04962893). Safety, efficacy and immunogenicity data from these ongoing clinical trials will be published following their completion.

In summary, we have developed a scalable process and a safe platform for mammalian expression and purification of SARS-CoV-2 virus-like particles that could be rapidly extended to produce vaccine antigens against newly emerging SARS-CoV-2 variants of concern. Furthermore, this platform is highly adaptable for the production of immunogenic VLPs for other enveloped viruses including the influenza virus and HIV as well as chimeric virus-like particles for producing multivalent vaccines against various pathogens. Current studies have been focused on preclinical evaluation and GMP preparation of 6p Delta VLPs for further phase III clinical trials.

REFERENCES

Abdool Karim, S. S., & de Oliveira, T. (2021). New SARS-CoV-2 Variants — Clinical, Public Health, and Vaccine Implications. *New England Journal of Medicine*, *384*(19). https://doi.org/10.1056/nejmc2100362

Addetia, A., Crawford, K. H. D., Dingens, A., Zhu, H., Roychoudhury, P., Huang, M. L., Jerome, K. R., Bloom, J. D., & Greninger, A. L. (2020). Neutralizing antibodies correlate with protection from SARS-CoV-2 in humans during a fishery vessel outbreak with a high attack rate. *Journal of Clinical Microbiology*, *58*(11). https://doi.org/10.1128/JCM.02107-20

Alanagreh, L., Alzoughool, F., & Atoum, M. (2020). The human coronavirus disease covid-19: Its origin, characteristics, and insights into potential drugs and its mechanisms. In *Pathogens* (Vol. 9, Issue 5). https://doi.org/10.3390/pathogens9050331

Arnau, J., Lauritzen, C., Petersen, G. E., & Pedersen, J. (2006). Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. In *Protein Expression and Purification* (Vol. 48, Issue 1). https://doi.org/10.1016/j.pep.2005.12.002

Avalos, A. M., & Ploegh, H. L. (2014). Early BCR events and antigen capture, processing, and loading on MHC class II on B cells. In *Frontiers in Immunology* (Vol. 5, Issue MAR). https://doi.org/10.3389/fimmu.2014.00092

Baden, L. R., El Sahly, H. M., Essink, B., Kotloff, K., Frey, S., Novak, R., Diemert, D., Spector, S. A., Rouphael, N., Creech, C. B., McGettigan, J., Khetan, S., Segall, N., Solis, J., Brosz, A., Fierro, C., Schwartz, H., Neuzil, K., Corey, L., … Zaks, T. (2021). Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *New England Journal of Medicine*, *384*(5). https://doi.org/10.1056/nejmoa2035389

Besnard, L., Fabre, V., Fettig, M., Gousseinov, E., Kawakami, Y., Laroudie, N., Scanlan, C., & Pattnaik, P. (2016). Clarification of vaccines: An overview of

filter based technology trends and best practices. In *Biotechnology Advances* (Vol. 34, Issue 1). https://doi.org/10.1016/j.biotechadv.2015.11.005

Bode, C., Zhao, G., Steinhagen, F., Kinjo, T., & Klinman, D. M. (2011). CpG DNA as a vaccine adjuvant. In *Expert Review of Vaccines* (Vol. 10, Issue 4). https://doi.org/10.1586/erv.10.174

Bolles, M., Deming, D., Long, K., Agnihothram, S., Whitmore, A., Ferris, M., Funkhouser, W., Gralinski, L., Totura, A., Heise, M., & Baric, R. S. (2011). A Double-Inactivated Severe Acute Respiratory Syndrome Coronavirus Vaccine Provides Incomplete Protection in Mice and Induces Increased Eosinophilic Proinflammatory Pulmonary Response upon Challenge. *Journal of Virology*, *85*(23). https://doi.org/10.1128/jvi.06048-11

Cai, Y., Zhang, J., Xiao, T., Peng, H., Sterling, S. M., Walsh, R. M., Rawson, S., Rits-Volloch, S., & Chen, B. (2020). Distinct conformational states of SARS-CoV-2 spike protein. *Science*, *369*(6511). https://doi.org/10.1126/science.abd4251

Cao, C., Cai, Z., Xiao, X., Rao, J., Chen, J., Hu, N., Yang, M., Xing, X., Wang, Y., Li, M., Zhou, B., Wang, X., Wang, J., & Xue, Y. (2021). The architecture of the SARS-CoV-2 RNA genome inside virion. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-22785-x

Centeno-Tablante, E., Medina-Rivera, M., Finkelstein, J. L., Rayco-Solon, P., Garcia-Casal, M. N., Rogers, L., Ghezzi-Kopel, K., Ridwan, P., Peña-Rosas, J. P., & Mehta, S. (2021). Transmission of SARS-CoV-2 through breast milk and breastfeeding: a living systematic review. In *Annals of the New York Academy of Sciences* (Vol. 1484, Issue 1). https://doi.org/10.1111/nyas.14477

Chang, L., Zhao, L., Gong, H., Gong, H., Gong, H., Wang, L., Wang, L., Wang, L., Yang, Z., & Xu, B. (2020). Severe Acute Respiratory Syndrome Coronavirus 2 RNA Detected in Blood Donations. *Emerging Infectious Diseases*, *26*(7). https://doi.org/10.3201/eid2607.200839

Chen, Y., Liu, Q., & Guo, D. (2020). Emerging coronaviruses: Genome structure,

replication, and pathogenesis. In *Journal of Medical Virology* (Vol. 92, Issue 4). https://doi.org/10.1002/jmv.25681

Chevallet, M., Luche, S., & Rabilloud, T. (2006). Silver staining of proteins in polyacrylamide gels. *Nature Protocols*, *1*(4). https://doi.org/10.1038/nprot.2006.288

Davis, H. L. (2008). Novel vaccines and adjuvant systems: The utility of animal models for predicting immunogenicity in humans. In *Human Vaccines* (Vol. 4, Issue 3). https://doi.org/10.4161/hv.4.3.5318

Dumont, J., Euwart, D., Mei, B., Estes, S., & Kshirsagar, R. (2016). Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. In *Critical Reviews in Biotechnology* (Vol. 36, Issue 6). https://doi.org/10.3109/07388551.2015.1084266

Ehrhardt, C., Schmolke, M., Matzke, A., Knoblauch, A., Will, C., Wixler, V., & Ludwig, S. (2006). Polyethylenimine, a cost-effective transfection reagent. *Signal Transduction*, *6*(3). https://doi.org/10.1002/sita.200500073

Esposito, D., Mehalko, J., Drew, M., Snead, K., Wall, V., Taylor, T., Frank, P., Denson, J. P., Hong, M., Gulten, G., Sadtler, K., Messing, S., & Gillette, W. (2020). Optimizing high-yield production of SARS-CoV-2 soluble spike trimers for serology assays. *Protein Expression and Purification*, *174*. https://doi.org/10.1016/j.pep.2020.105686

Figueroa, J. P., Bottazzi, M. E., Hotez, P., Batista, C., Ergonul, O., Gilbert, S., Gursel, M., Hassanain, M., Kim, J. H., Lall, B., Larson, H., Naniche, D., Sheahan, T., Shoham, S., Wilder-Smith, A., Strub-Wourgaft, N., Yadav, P., & Kang, G. (2021). Urgent needs of low-income and middle-income countries for COVID-19 vaccines and therapeutics. In *The Lancet* (Vol. 397, Issue 10274). https://doi.org/10.1016/S0140-6736(21)00242-7

Frey, A., Di Canzio, J., & Zurakowski, D. (1998). A statistically defined endpoint titer determination method for immunoassays. *Journal of Immunological Methods*, *221*(1–2). https://doi.org/10.1016/S0022-1759(98)00170-7

Frutos, R., Serra-Cobo, J., Pinault, L., Lopez Roig, M., & Devaux, C. A. (2021). Emergence of Bat-Related Betacoronaviruses: Hazard and Risks. In *Frontiers in Microbiology* (Vol. 12). https://doi.org/10.3389/fmicb.2021.591535

Fung, T. S., & Liu, D. X. (2018). Post-translational modifications of coronavirus proteins: Roles and function. In *Future Virology* (Vol. 13, Issue 6). https://doi.org/10.2217/fvl-2018-0008

Gam, J. J., DiAndreth, B., Jones, R. D., Huh, J., & Weiss, R. (2019). A "poly-transfection" method for rapid, one-pot characterization and optimization of genetic systems. *Nucleic Acids Research*, *47*(18). https://doi.org/10.1093/nar/gkz623

Ghinai, I., McPherson, T. D., Hunter, J. C., Kirking, H. L., Christiansen, D., Joshi, K., Rubin, R., Morales-Estrada, S., Black, S. R., Pacilli, M., Fricchione, M. J., Chugh, R. K., Walblay, K. A., Ahmed, N. S., Stoecker, W. C., Hasan, N. F., Burdsall, D. P., Reese, H. E., Wallace, M., … Layden, J. E. (2020). First known person-to-person transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in the USA. *The Lancet*, *395*(10230). https://doi.org/10.1016/S0140-6736(20)30607-3

Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., O'Meara, M. J., Rezelj, V. V., Guo, J. Z., Swaney, D. L., Tummino, T. A., Hüttenhain, R., Kaake, R. M., Richards, A. L., Tutuncuoglu, B., Foussard, H., Batra, J., Haas, K., Modak, M., … Krogan, N. J. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*, *583*(7816). https://doi.org/10.1038/s41586-020-2286-9

Grifoni, A., Weiskopf, D., Ramirez, S. I., Mateus, J., Dan, J. M., Moderbacher, C. R., Rawlings, S. A., Sutherland, A., Premkumar, L., Jadi, R. S., Marrama, D., de Silva, A. M., Frazier, A., Carlin, A. F., Greenbaum, J. A., Peters, B., Krammer, F., Smith, D. M., Crotty, S., & Sette, A. (2020). Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell*, *181*(7).

https://doi.org/10.1016/j.cell.2020.05.015

Güler-Gane, G., Kidd, S., Sridharan, S., Vaughan, T. J., Wilkinson, T. C. I., & Tigue, N. J. (2016). Overcoming the refractory expression of secreted recombinant proteins in mammalian cells through modification of the signal peptide and adjacent amino acids. *PLoS ONE*, *11*(5). https://doi.org/10.1371/journal.pone.0155340

Gursel, M., & Gursel, I. (2016). Development of CpG ODN based vaccine adjuvant formulations. In *Methods in Molecular Biology* (Vol. 1404). https://doi.org/10.1007/978-1-4939-3389-1_20

Gürsel, M., Verthelyi, D., Gürsel, I., Ishii, K. J., & Klinman, D. M. (2002). Differential and competitive activation of human immune cells by distinct classes of CpG oligodeoxynucleotide. *Journal of Leukocyte Biology*, *71*(5). https://doi.org/10.1189/jlb.71.5.813

Hachim, A., Kavian, N., Cohen, C. A., Chin, A. W. H., Chu, D. K. W., Mok, C. K. P., Tsang, O. T. Y., Yeung, Y. C., Perera, R. A. P. M., Poon, L. L. M., Peiris, M. J. S., & Valkenburg, S. A. (2020). Beyond the Spike: Identification of viral targets of the antibody response to SARS-CoV-2 in COVID-19 patients. *Nature Immunology*. https://doi.org/10.1101/2020.04.30.20085670

Hacker, D. L., Kiseljak, D., Rajendra, Y., Thurnheer, S., Baldi, L., & Wurm, F. M. (2013). Polyethyleneimine-based transient gene expression processes for suspension-adapted HEK-293E and CHO-DG44 cells. In *Protein Expression and Purification* (Vol. 92, Issue 1). https://doi.org/10.1016/j.pep.2013.09.001

Hanagata, N. (2012). Structure-dependent immunostimulatory effect of CpG oligodeoxynucleotides and their delivery system. In *International Journal of Nanomedicine* (Vol. 7). https://doi.org/10.2147/IJN.S30197

Heath, P. T., Galiza, E. P., Baxter, D. N., Boffito, M., Browne, D., Burns, F., Chadwick, D. R., Clark, R., Cosgrove, C., Galloway, J., Goodman, A. L., Heer, A., Higham, A., Iyengar, S., Jamal, A., Jeanes, C., Kalra, P. A., Kyriakidou, C., McAuley, D. F., … Toback, S. (2021). Safety and Efficacy of

NVX-CoV2373 Covid-19 Vaccine. *New England Journal of Medicine*, *385*(13). https://doi.org/10.1056/nejmoa2107659

Hillebrandt, N., Vormittag, P., Bluthardt, N., Dietrich, A., & Hubbuch, J. (2020). Integrated Process for Capture and Purification of Virus-Like Particles: Enhancing Process Performance by Cross-Flow Filtration. *Frontiers in Bioengineering and Biotechnology*, *8*. https://doi.org/10.3389/fbioe.2020.00489

Hobbs, E. C., & Reid, T. J. (2021). Animals and SARS-CoV-2: Species susceptibility and viral transmission in experimental and natural conditions, and the potential implications for community transmission. In *Transboundary and Emerging Diseases* (Vol. 68, Issue 4). https://doi.org/10.1111/tbed.13885

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T. S., Herrler, G., Wu, N. H., Nitsche, A., Müller, M. A., Drosten, C., & Pöhlmann, S. (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*, *181*(2). https://doi.org/10.1016/j.cell.2020.02.052

Hsieh, C. L., Goldsmith, J. A., Schaub, J. M., DiVenere, A. M., Kuo, H. C., Javanmardi, K., Le, K. C., Wrapp, D., Lee, A. G., Liu, Y., Chou, C. W., Byrne, P. O., Hjorth, C. K., Johnson, N. V., Ludes-Meyers, J., Nguyen, A. W., Park, J., Wang, N., Amengor, D., … McLellan, J. S. (2020). Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science*, *369*(6509). https://doi.org/10.1126/SCIENCE.ABD0826

Hu, B., Guo, H., Zhou, P., & Shi, Z. L. (2021). Characteristics of SARS-CoV-2 and COVID-19. In *Nature Reviews Microbiology* (Vol. 19, Issue 3). https://doi.org/10.1038/s41579-020-00459-7

Huber, V. C., McKeon, R. M., Brackin, M. N., Miller, L. A., Keating, R., Brown, S. A., Makarova, N., Perez, D. R., MacDonald, G. H., & McCullers, J. A. (2006). Distinct contributions of vaccine-induced immunoglobulin G1 (IgG1) and IgG2a antibodies to protective immunity against influenza. *Clinical and*

Vaccine Immunology, *13*(9). https://doi.org/10.1128/CVI.00156-06

James, K. T., Cooney, B., Agopsowicz, K., Trevors, M. A., Mohamed, A., Stoltz, D., Hitt, M., & Shmulevitz, M. (2016). Novel High-throughput Approach for Purification of Infectious Virions. *Scientific Reports*, *6*. https://doi.org/10.1038/srep36826

Joe, C. C. D., Chatterjee, S., Lovrecz, G., Adams, T. E., Thaysen-Andersen, M., Walsh, R., Locarnini, S. A., Smooker, P., & Netter, H. J. (2020). Glycoengineered hepatitis B virus-like particles with enhanced immunogenicity. *Vaccine*, *38*(22). https://doi.org/10.1016/j.vaccine.2020.03.007

Johari, Y. B., Jaffé, S. R. P., Scarrott, J. M., Johnson, A. O., Mozzanino, T., Pohle, T. H., Maisuria, S., Bhayat-Cammack, A., Lambiase, G., Brown, A. J., Tee, K. L., Jackson, P. J., Wong, T. S., Dickman, M. J., Sargur, R. B., & James, D. C. (2021). Production of trimeric SARS-CoV-2 spike protein by CHO cells for serological COVID-19 testing. *Biotechnology and Bioengineering*, *118*(2). https://doi.org/10.1002/bit.27615

Kawai, T., & Akira, S. (2011). Toll-like Receptors and Their Crosstalk with Other Innate Receptors in Infection and Immunity. In *Immunity* (Vol. 34, Issue 5). https://doi.org/10.1016/j.immuni.2011.05.006

Ke, Z., Oton, J., Qu, K., Cortese, M., Zila, V., McKeane, L., Nakane, T., Zivanov, J., Neufeldt, C. J., Cerikan, B., Lu, J. M., Peukes, J., Xiong, X., Kräusslich, H. G., Scheres, S. H. W., Bartenschlager, R., & Briggs, J. A. G. (2020). Structures and distributions of SARS-CoV-2 spike proteins on intact virions. *Nature*, *588*(7838). https://doi.org/10.1038/s41586-020-2665-2

Kim, Y. M., & Shin, E. C. (2021). Type I and III interferon responses in SARS-CoV-2 infection. In *Experimental and Molecular Medicine* (Vol. 53, Issue 5). https://doi.org/10.1038/s12276-021-00592-0

Klinman, D. M. (2004). Immunotherapeutic uses of CpG oligodeoxynucleotides. In *Nature Reviews Immunology* (Vol. 4, Issue 4). https://doi.org/10.1038/nri1329

Krammer, F. (2020). SARS-CoV-2 vaccines in development. In *Nature* (Vol. 586, Issue 7830). https://doi.org/10.1038/s41586-020-2798-3

Kumar, H., Kawai, T., & Akira, S. (2011). Pathogen recognition by the innate immune system. *International Reviews of Immunology*, *30*(1). https://doi.org/10.3109/08830185.2010.529976

Kyriakidis, N. C., López-Cortés, A., González, E. V., Grimaldos, A. B., & Prado, E. O. (2021). SARS-CoV-2 vaccines strategies: a comprehensive review of phase 3 candidates. In *npj Vaccines* (Vol. 6, Issue 1). https://doi.org/10.1038/s41541-021-00292-w

Lasareishvili, B., Shi, H., Wang, X., Hillstead, K. D., Tediashvili, M., Jaiani, E., & Tarabara, V. V. (2021). Virus recovery by tangential flow filtration: A model to guide the design of a sample concentration process. *Biotechnology Progress*, *37*(1). https://doi.org/10.1002/btpr.3080

Laue, M., Kauter, A., Hoffmann, T., Möller, L., Michel, J., & Nitsche, A. (2021). Morphometry of SARS-CoV and SARS-CoV-2 particles in ultrathin plastic sections of infected Vero cell cultures. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-82852-7

Le Ru, A., Jacob, D., Transfiguracion, J., Ansorge, S., Henry, O., & Kamen, A. A. (2010). Scalable production of influenza virus in HEK-293 cells for efficient vaccine manufacturing. *Vaccine*, *28*(21). https://doi.org/10.1016/j.vaccine.2010.03.029

Ledford, H. (2020). Moderna COVID vaccine becomes second to get US authorization. *Nature*. https://doi.org/10.1038/d41586-020-03593-7

Lei, X., Dong, X., Ma, R., Wang, W., Xiao, X., Tian, Z., Wang, C., Wang, Y., Li, L., Ren, L., Guo, F., Zhao, Z., Zhou, Z., Xiang, Z., & Wang, J. (2020). Activation and evasion of type I interferon responses by SARS-CoV-2. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-17665-9

Liu, Y., Ning, Z., Chen, Y., Guo, M., Liu, Y., Gali, N. K., Sun, L., Duan, Y., Cai,

J., Westerdahl, D., Liu, X., Xu, K., Ho, K. fai, Kan, H., Fu, Q., & Lan, K. (2020). Aerodynamic analysis of SARS-CoV-2 in two Wuhan hospitals. *Nature*, *582*(7813). https://doi.org/10.1038/s41586-020-2271-3

Logunov, D. Y., Dolzhikova, I. V., Shcheblyakov, D. V., Tukhvatulin, A. I., Zubkova, O. V., Dzharullaeva, A. S., Kovyrshina, A. V., Lubenets, N. L., Grousova, D. M., Erokhova, A. S., Botikov, A. G., Izhaeva, F. M., Popova, O., Ozharovskaya, T. A., Esmagambetov, I. B., Favorskaya, I. A., Zrelkin, D. I., Voronina, D. V., Shcherbinin, D. N., … Gintsburg, A. L. (2021). Safety and efficacy of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine: an interim analysis of a randomised controlled phase 3 trial in Russia. *The Lancet*, *397*(10275). https://doi.org/10.1016/S0140-6736(21)00234-8

Lopez, A. M., Hecker, R., Mutwiri, G., van Drunen Littel-van den Hurk, S., Babiuk, L. A., & Townsend, H. G. G. (2006). Formulation with CpG ODN enhances antibody responses to an equine influenza virus vaccine. *Veterinary Immunology and Immunopathology*, *114*(1–2). https://doi.org/10.1016/j.vetimm.2006.07.013

Lu, S., Ye, Q., Singh, D., Cao, Y., Diedrich, J. K., Yates, J. R., Villa, E., Cleveland, D. W., & Corbett, K. D. (2021). The SARS-CoV-2 nucleocapsid phosphoprotein forms mutually exclusive condensates with RNA and the membrane-associated M protein. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-020-20768-y

Matchett, W. E., Joag, V., Stolley, J. M., Shepherd, F. K., Quarnstrom, C. F., Mickelson, C. K., Wijeyesinghe, S., Soerens, A. G., Becker, S., Thiede, J. M., Weyu, E., O'Flanagan, S. D., Walter, J. A., Vu, M. N., Menachery, V. D., Bold, T. D., Vezys, V., Jenkins, M. K., Langlois, R. A., & Masopust, D. (2021). Cutting Edge: Nucleocapsid Vaccine Elicits Spike-Independent SARS-CoV-2 Protective Immunity. *The Journal of Immunology*, *207*(2). https://doi.org/10.4049/jimmunol.2100421

Medzhitov, R., & Janeway C., J. (2000). Advances in immunology: Innate immunity. *New England Journal of Medicine*, *343*(5). https://doi.org/10.1056/NEJM200008033430506

Mohsen, M. O., Zha, L., Cabral-Miranda, G., & Bachmann, M. F. (2017). Major findings and recent advances in virus–like particle (VLP)-based vaccines. In *Seminars in Immunology* (Vol. 34). https://doi.org/10.1016/j.smim.2017.08.014

Nooraei, S., Bahrulolum, H., Hoseini, Z. S., Katalani, C., Hajizade, A., Easton, A. J., & Ahmadian, G. (2021). Virus-like particles: preparation, immunogenicity and their roles as nanovaccines and drug nanocarriers. In *Journal of Nanobiotechnology* (Vol. 19, Issue 1). https://doi.org/10.1186/s12951-021-00806-7

Pastorino, B., Touret, F., Gilles, M., de Lamballerie, X., & Charrel, R. N. (2020). Prolonged Infectivity of SARS-CoV-2 in Fomites. In *Emerging infectious diseases* (Vol. 26, Issue 9). https://doi.org/10.3201/eid2609.201788

Polack, F. P., Thomas, S. J., Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S., Perez, J. L., Pérez Marc, G., Moreira, E. D., Zerbini, C., Bailey, R., Swanson, K. A., Roychoudhury, S., Koury, K., Li, P., Kalina, W. V., Cooper, D., Frenck, R. W., Hammitt, L. L., … Gruber, W. C. (2020). Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New England Journal of Medicine*, *383*(27). https://doi.org/10.1056/nejmoa2034577

Pollard, A. J., & Bijker, E. M. (2021). A guide to vaccinology: from basic principles to new developments. In *Nature Reviews Immunology* (Vol. 21, Issue 2). https://doi.org/10.1038/s41577-020-00479-7

Prüβ, B. M. (2021). Current state of the first covid-19 vaccines. In *Vaccines* (Vol. 9, Issue 1). https://doi.org/10.3390/vaccines9010030

Pulendran, B., S. Arunachalam, P., & O'Hagan, D. T. (2021). Emerging concepts in the science of vaccine adjuvants. In *Nature Reviews Drug Discovery* (Vol. 20, Issue 6). https://doi.org/10.1038/s41573-021-00163-y

Puntmann, V. O., Carerj, M. L., Wieters, I., Fahim, M., Arendt, C., Hoffmann, J., Shchendrygina, A., Escher, F., Vasa-Nicotera, M., Zeiher, A. M., Vehreschild, M., & Nagel, E. (2020). Outcomes of Cardiovascular Magnetic Resonance Imaging in Patients Recently Recovered from Coronavirus Disease 2019 (COVID-19). *JAMA Cardiology*, *5*(11). https://doi.org/10.1001/jamacardio.2020.3557

Ragab, D., Salah Eldin, H., Taeimah, M., Khattab, R., & Salem, R. (2020). The COVID-19 Cytokine Storm; What We Know So Far. In *Frontiers in Immunology* (Vol. 11). https://doi.org/10.3389/fimmu.2020.01446

Redondo, N., Zaldívar-López, S., Garrido, J. J., & Montoya, M. (2021). SARS-CoV-2 Accessory Proteins in Viral Pathogenesis: Knowns and Unknowns. In *Frontiers in Immunology* (Vol. 12). https://doi.org/10.3389/fimmu.2021.708264

Reiter, K., Aguilar, P. P., Wetter, V., Steppert, P., Tover, A., & Jungbauer, A. (2019). Separation of virus-like particles and extracellular vesicles by flow-through and heparin affinity chromatography. *Journal of Chromatography A*, *1588*. https://doi.org/10.1016/j.chroma.2018.12.035

Rhodes, S. J., Knight, G. M., Kirschner, D. E., White, R. G., & Evans, T. G. (2019). Dose finding for new vaccines: The role for immunostimulation/immunodynamic modelling. In *Journal of Theoretical Biology* (Vol. 465). https://doi.org/10.1016/j.jtbi.2019.01.017

Ricci, D., Etna, M. P., Rizzo, F., Sandini, S., Severa, M., & Coccia, E. M. (2021). Innate immune response to sars-cov-2 infection: From cells to soluble mediators. In *International Journal of Molecular Sciences* (Vol. 22, Issue 13). https://doi.org/10.3390/ijms22137017

Riguero, V., Clifford, R., Dawley, M., Dickson, M., Gastfriend, B., Thompson, C., Wang, S. C., & O'Connor, E. (2020). Immobilized metal affinity chromatography optimization for poly-histidine tagged proteins. *Journal of Chromatography A*, *1629*. https://doi.org/10.1016/j.chroma.2020.461505

Robbiani, D. F., Gaebler, C., Muecksch, F., Lorenzi, J. C. C., Wang, Z., Cho, A., Agudelo, M., Barnes, C. O., Gazumyan, A., Finkin, S., Hägglöf, T., Oliveira, T. Y., Viant, C., Hurley, A., Hoffmann, H. H., Millard, K. G., Kost, R. G., Cipolla, M., Gordon, K., … Nussenzweig, M. C. (2020). Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature*, *584*(7821). https://doi.org/10.1038/s41586-020-2456-9

Rothan, H. A., & Byrareddy, S. N. (2020). The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. In *Journal of Autoimmunity* (Vol. 109). https://doi.org/10.1016/j.jaut.2020.102433

Rubtsova, K., Rubtsov, A. V., Van Dyk, L. F., Kappler, J. W., & Marrack, P. (2013). T-box transcription factor T-bet, a key player in a unique type of B-cell activation essential for effective viral clearance. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(34). https://doi.org/10.1073/pnas.1312348110

Rydyznski Moderbacher, C., Ramirez, S. I., Dan, J. M., Grifoni, A., Hastie, K. M., Weiskopf, D., Belanger, S., Abbott, R. K., Kim, C., Choi, J., Kato, Y., Crotty, E. G., Kim, C., Rawlings, S. A., Mateus, J., Tse, L. P. V., Frazier, A., Baric, R., Peters, B., … Crotty, S. (2020). Antigen-Specific Adaptive Immunity to SARS-CoV-2 in Acute COVID-19 and Associations with Age and Disease Severity. *Cell*, *183*(4). https://doi.org/10.1016/j.cell.2020.09.038

Sahu, A. K., Sreepadmanabh, M., Rai, M., & Chande, A. (2021). SARS-CoV-2: phylogenetic origins, pathogenesis, modes of transmission, and the potential role of nanotechnology. In *VirusDisease* (Vol. 32, Issue 1). https://doi.org/10.1007/s13337-021-00653-y

Saulle, I., Vicentini, C., Clerici, M., & Biasin, M. (2021). Antigen presentation in SARS-CoV-2 infection: the role of class I HLA and ERAP polymorphisms. *Human Immunology*, *82*(8). https://doi.org/10.1016/j.humimm.2021.05.003

Schoeman, D., & Fielding, B. C. (2019). Coronavirus envelope protein: Current knowledge. In *Virology Journal* (Vol. 16, Issue 1).

https://doi.org/10.1186/s12985-019-1182-0

Shajahan, A., Pepi, L. E., Rouhani, D. S., Heiss, C., & Azadi, P. (2021). Glycosylation of SARS-CoV-2: structural and functional insights. In *Analytical and Bioanalytical Chemistry*. https://doi.org/10.1007/s00216-021-03499-x

Sivakumaran, M., & Platt, M. (2016). Tunable resistive pulse sensing: Potential applications in nanomedicine. In *Nanomedicine* (Vol. 11, Issue 16, pp. 2197–2214). https://doi.org/10.2217/nnm-2016-0097

Spriestersbach, A., Kubicek, J., Schäfer, F., Block, H., & Maertens, B. (2015). Purification of His-Tagged Proteins. In *Methods in Enzymology* (Vol. 559). https://doi.org/10.1016/bs.mie.2014.11.003

Swain, S. L., McKinstry, K. K., & Strutt, T. M. (2012). Expanding roles for CD4 + T cells in immunity to viruses. In *Nature Reviews Immunology* (Vol. 12, Issue 2). https://doi.org/10.1038/nri3152

Tanne, J. H. (2020). Covid-19: FDA panel votes to approve Pfizer BioNTech vaccine. *BMJ (Clinical Research Ed.)*, *371*. https://doi.org/10.1136/bmj.m4799

Tanriover, M. D., Doğanay, H. L., Akova, M., Güner, H. R., Azap, A., Akhan, S., Köse, Ş., Erdinç, F. Ş., Akalın, E. H., Tabak, Ö. F., Pullukçu, H., Batum, Ö., Şimşek Yavuz, S., Turhan, Ö., Yıldırmak, M. T., Köksal, İ., Taşova, Y., Korten, V., Yılmaz, G., … Aksu, K. (2021). Efficacy and safety of an inactivated whole-virion SARS-CoV-2 vaccine (CoronaVac): interim results of a double-blind, randomised, placebo-controlled, phase 3 trial in Turkey. *The Lancet*, *398*(10296). https://doi.org/10.1016/S0140-6736(21)01429-X

Tian, W., Li, D., Zhang, N., Bai, G., Yuan, K., Xiao, H., Gao, F., Chen, Y., Wong, C. C. L., & Gao, G. F. (2021). O-glycosylation pattern of the SARS-CoV-2 spike protein reveals an “O-Follow-N” rule. In *Cell Research*. https://doi.org/10.1038/s41422-021-00545-2

Tong, P., Gautam, A., Windsor, I. W., Travers, M., Chen, Y., Garcia, N., Whiteman, N. B., McKay, L. G. A., Storm, N., Malsick, L. E., Honko, A. N., Lelis, F. J. N., Habibi, S., Jenni, S., Cai, Y., Rennick, L. J., Duprex, W. P., McCarthy, K. R., Lavine, C. L., … Wesemann, D. R. (2021). Memory B cell repertoire for recognition of evolving SARS-CoV-2 spike. *Cell*. https://doi.org/10.1016/j.cell.2021.07.025

Trougakos, I. P., Terpos, E., Zirou, C., Sklirou, A. D., Apostolakou, F., Gumeni, S., Charitaki, I., Papanagnou, E. D., Bagratuni, T., Liacos, C. I., Scorilas, A., Korompoki, E., Papassotiriou, I., Kastritis, E., & Dimopoulos, M. A. (2021). Comparative kinetics of SARS-CoV-2 anti-spike protein RBD IgGs and neutralizing antibodies in convalescent and naïve recipients of the BNT162b2 mRNA vaccine versus COVID-19 patients. *BMC Medicine*, *19*(1). https://doi.org/10.1186/s12916-021-02090-6

V'kovski, P., Kratzel, A., Steiner, S., Stalder, H., & Thiel, V. (2021). Coronavirus biology and replication: implications for SARS-CoV-2. In *Nature Reviews Microbiology* (Vol. 19, Issue 3). https://doi.org/10.1038/s41579-020-00468-6

Voysey, M., Clemens, S. A. C., Madhi, S. A., Weckx, L. Y., Folegatti, P. M., Aley, P. K., Angus, B., Baillie, V. L., Barnabas, S. L., Bhorat, Q. E., Bibi, S., Briner, C., Cicconi, P., Collins, A. M., Colin-Jones, R., Cutland, C. L., Darton, T. C., Dheda, K., Duncan, C. J. A., … Zuidewind, P. (2021). Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *The Lancet*, *397*(10269). https://doi.org/10.1016/S0140-6736(20)32661-1

Wang, D., Hu, B., Hu, C., Zhu, F., Liu, X., Zhang, J., Wang, B., Xiang, H., Cheng, Z., Xiong, Y., Zhao, Y., Li, Y., Wang, X., & Peng, Z. (2020). Clinical Characteristics of 138 Hospitalized Patients with 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA - Journal of the American Medical Association*, *323*(11). https://doi.org/10.1001/jama.2020.1585

Wang, W., Xu, Y., Gao, R., Lu, R., Han, K., Wu, G., & Tan, W. (2020). Detection of SARS-CoV-2 in Different Types of Clinical Specimens. In *JAMA - Journal of the American Medical Association* (Vol. 323, Issue 18). https://doi.org/10.1001/jama.2020.3786

Ward, B. J., Gobeil, P., Séguin, A., Atkins, J., Boulay, I., Charbonneau, P. Y., Couture, M., D'Aoust, M. A., Dhaliwall, J., Finkle, C., Hager, K., Mahmood, A., Makarkov, A., Cheng, M. P., Pillet, S., Schimke, P., St-Martin, S., Trépanier, S., & Landry, N. (2021). Phase 1 randomized trial of a plant-derived virus-like particle vaccine for COVID-19. *Nature Medicine*, *27*(6). https://doi.org/10.1038/s41591-021-01370-1

WHO. (2020). WHO/Europe | Coronavirus disease (COVID-19) outbreak - WHO announces COVID-19 outbreak a pandemic. In *WHO/Europe*.

Wickramasinghe, S. R., Kalbfuß, B., Zimmermann, A., Thorn, V., & Reichl, U. (2005). Tangential flow microfiltration and ultrafiltration for human influenza A virus concentration and purification. *Biotechnology and Bioengineering*, *92*(2). https://doi.org/10.1002/bit.20599

World Health Organization. (2021). *COVID-19 vaccine tracker and landscape*. World Health Organization.

World Health Organization Europe (WHO Europe). (2020). Transmission of SARS-CoV-2: implications for infection prevention precautions. Scientific brief, 09 July 2020. *Who*, *March*.

Wu, J. T., Leung, K., & Leung, G. M. (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet*, *395*(10225). https://doi.org/10.1016/S0140-6736(20)30260-9

Xia, H., Cao, Z., Xie, X., Zhang, X., Chen, J. Y. C., Wang, H., Menachery, V. D., Rajsbaum, R., & Shi, P. Y. (2020). Evasion of Type I Interferon by SARS-CoV-2. *Cell Reports*, *33*(1). https://doi.org/10.1016/j.celrep.2020.108234

Yang, Y., Xiao, Z., Ye, K., He, X., Sun, B., Qin, Z., Yu, J., Yao, J., Wu, Q., Bao, Z., & Zhao, W. (2020). SARS-CoV-2: characteristics and current advances in research. In *Virology Journal* (Vol. 17, Issue 1). https://doi.org/10.1186/s12985-020-01369-z

Yilmaz, I. C., Ipekoglu, E. M., Bulbul, A., Turay, N., Yildirim, M., Evcili, I., Yilmaz, N. S., Guvencli, N., Aydin, Y., Gungor, B., Saraydar, B., Bartan, A. G., Ibibik, B., Bildik, T., Baydemir, İ., Sanli, H. A., Kayaoglu, B., Ceylan, Y., Yildirim, T., … Gursel, M. (2021). Development and preclinical evaluation of virus-like particle vaccine against COVID-19 infection. *Allergy: European Journal of Allergy and Clinical Immunology*. https://doi.org/10.1111/all.15091

Zeng, C., Evans, J. P., King, T., Zheng, Y.-M., Oltz, E. M., Whelan, S. P. J., Saif, L., Peeples, M. E., & Liu, S.-L. (2021). SARS-CoV-2 Spreads through Cell-to-Cell Transmission. *BioRxiv : The Preprint Server for Biology*. https://doi.org/10.1101/2021.06.01.446579

Zhang, Y., Zhao, J., He, S., & Cao, X. (2020). Soluble Expression of Recombinant Human Cystatin C and Comparison of the Ni Column and Magnetic Bead Purification. *Protein Journal*, *39*(1). https://doi.org/10.1007/s10930-019-09873-0

Zhao, D., Sun, B., Jiang, H., Sun, S., Kong, F. T., Ma, Y., Jiang, L., Bai, L., Chen, X., Yang, P., Liu, C., Xu, Y., Su, W., Kong, W., Xu, F., & Jiang, C. (2015). Enterovirus71 virus-like particles produced from insect cells and purified by multistep chromatography elicit strong humoral immune responses in mice. *Journal of Applied Microbiology*, *119*(4). https://doi.org/10.1111/jam.12922

Zhu, Z., Lian, X., Su, X., Wu, W., Marraro, G. A., & Zeng, Y. (2020). From SARS and MERS to COVID-19: A brief summary and comparison of severe acute respiratory infections caused by three highly pathogenic human coronaviruses. In *Respiratory Research* (Vol. 21, Issue 1). https://doi.org/10.1186/s12931-020-01479-w

APPENDICES

A. Recipes of Various Media and Buffers

Low Salt LB Media

- 10 g Tryptone
- 5 g NaCl
- 5 g Yeast Extract
- 10 g Agar (For Solid Media)

  Finalize to 1L with distilled water and autoclave. Add Hygromycin B gold at 100 µg/mL final concentration for antibiotic selection.

10 % Regular RPMI Medium (500 mL)

- 375.5 mL RPMI 1640 with L-Glutamine (Biological Industries, USA, cat. #01-100-1A)
- 10 mL heat inactivated Fetal bovine serum (Biological Industries, USA, cat. #04-127-1A)
- 5 mL Penicillin/Streptomycin Solution (Biological Industries, USA, cat. #03-031-1B)
- 5 mL MEM Non-Essential Amino Acids Solution 100X (Biological Industries, USA, cat. #03-340-1B)
- 5 mL Sodium Pyruvate Solution (Biological Industries, USA, cat. #03-042-1B)

10X TAE Buffer

- 48.4 g Trizma Base
- 3.7 g EDTA disodium salt
- 11.4 Ml glacial acetic acid

  Finalize to 1L with distilled water and store at room temperature.

Multimodal Chromatography Equilibration Buffer (1X DPBS)

- 0.2 g Potassium Chloride (KCl)
- 0.2 g Potassium Phosphate Monobasic (KH2PO4)
- 8 g Sodium Chloride (NaCl)
- 1.15 Sodium Phosphate Dibasic (Na2HPO4)

  Adjust pH to 7.2 and finalize to 1 L with distilled water. 0.22 µm filter sterilize and store at room temperature.

Multimodal Chromatography Elution Buffer

- 11.9 g HEPES (50 mM)
- 116.9 g NaCl (2M)

Adjust pH to 7.2, finalize to 1L with distilled water and 0.22 µm filter sterilize.

Multimodal Chromatography Cleaning-In-Place (CIP) Solution

- 40 g NaOH (1M)
- 300 mL isopropanol
- 700 mL distilled water

0.22 µm filter sterilize and store at room temperature.

Affinity Chromatography Equilibration Buffer

- 0.2 g Potassium Chloride (KCl)
- 0.2 g Potassium Phosphate Monobasic (KH2PO4)
- 8 g Sodium Chloride (NaCl)
- 1.15 Sodium Phosphate Dibasic (Na2HPO4)
- 1.36 g Imidazole (25 mM)

  Adjust pH to 7.4 and finalize to 1 L with distilled water. 0.22 µm filter sterilize and store at room temperature.

Affinity Chromatography Elution Buffer

- 0.2 g Potassium Chloride (KCl)
- 0.2 g Potassium Phosphate Monobasic (KH2PO4)
- 8 g Sodium Chloride (NaCl)
- 1.15 Sodium Phosphate Dibasic (Na2HPO4)
- 34 g Imidazole (500 mM)

  Adjust pH to 7.4 and finalize to 1 L with distilled water. 0.22 µm filter sterilize and store at room temperature.

4X Laemmli Buffer

- 62.5 mM Tris-HCl, pH 6.8
- 10% glycerol
- 1% SDS
- 0.005% Bromophenol Blue

  For reducing conditions add 100 µl of 2-mercaptoethanol per 900 µl (final concentration of 355 mM).

Phosphate buffered saline- Tween Solution (1X PBS-T)

- 0.2 g Potassium Chloride (KCl)
- 0.2 g Potassium Phosphate Monobasic (KH2PO4)
- 8 g Sodium Chloride (NaCl)
- 1.15 Sodium Phosphate Dibasic (Na2HPO4)
- 1 mL Tween 20

  Adjust pH to 7.4 and finalize to 1 L with distilled water.

Immunoblotting Blocking Solution

- 5 g Non-Fat dry milk (5%)

  Dissolve in 100 mL 1X PBS-T solution, store at 4 ºC.

10X DPBS for ELISA

- 80 g NaCl
- 2 g KCl
- 15.2 g $Na_2HPO_4.2H_2O$
- 2 g KH_2PO_4

Adjust pH to 6.8 finalize to 1L with distilled water and autoclave.

ELISA Blocking Buffer

- 500 mL 1X DPBS
- 25 g Bovine Serum Albumin
- 250 µL Tween 20

Store at -20 ºC.

ELISA Wash Buffer

- 500 mL 10X DPBS
- 2.5 mL Tween 20,
- 4.5 L distilled water

B. Pairwise alignments of putative DNA sequences and next generation sequencing results

Pairwise alignment of 6pS encoding DNA Sequence (Upper: Putative, Lower:Sequencing)

```
1ATGCCTTTGCTTTTGTTGCTTCCTCTTCTCTGGGCTGGGGCTTTGGCCAAGCTTGTGAAC
1ATGCCTTTGCTTTTGTTGCTTCCTCTTCTCTGGGCTGGGGCTTTGGCCAAGCTTGTGAAC
1************************************************************

61CTTACCACACGCACCCAATTGCCCCCAGCTTACACAAACAGCTTTACTAGGGGCGTCTAC
61CTTACCACACGCACCCAATTGCCCCCAGCTTACACAAACAGCTTTACTAGGGGCGTCTAC
61************************************************************

121TATCCCGACAAGGTGTTTCGGAGTAGCGTGCTGCACTCTACTCAAGACCTCTTCCTGCCC
121TATCCCGACAAGGTGTTTCGGAGTAGCGTGCTGCACTCTACTCAAGACCTCTTCCTGCCC
121************************************************************

181TTCTTTTCCAATGTCACTTGGTTCCACGCCATCCATGTCAGCGGTACCAATGGGACCAAA
181TTCTTTTCCAATGTCACTTGGTTCCACGCCATCCATGTCAGCGGTACCAATGGGACCAAA
181************************************************************

241AGGTTTGACAACCCAGTGTTGCCATTTAACGATGGCGTGTATTTCGCCAGTACTGAAAAG
241AGGTTTGACAACCCAGTGTTGCCATTTAACGATGGCGTGTATTTCGCCAGTACTGAAAAG
241************************************************************

301AGCAATATCATCAGGGGCTGGATCTTCGGTACTACCCTGGATAGCAAGACCCAGTCTTTG
301AGCAATATCATCAGGGGCTGGATCTTCGGTACTACCCTGGATAGCAAGACCCAGTCTTTG
301************************************************************

361CTGATTGTGAACAATGCTACAAATGTGGTAATCAAAGTTTGTGAGTTTCAATTCTGCAAT
361CTGATTGTGAACAATGCTACAAATGTGGTAATCAAAGTTTGTGAGTTTCAATTCTGCAAT
361************************************************************

421GACCCATTCCTGGGCGTCTATTACCACAAGAATAACAAAAGTTGGATGGAATCCGAGTTT
421GACCCATTCCTGGGCGTCTATTACCACAAGAATAACAAAAGTTGGATGGAATCCGAGTTT
421************************************************************

481CGGGTGTACAGCTCCGCTAACAACTGCACATTCGAATATGTGTCCCAGCCGTTCCTGATG
481CGGGTGTACAGCTCCGCTAACAACTGCACATTCGAATATGTGTCCCAGCCGTTCCTGATG
481************************************************************

541GACCTTGAGGGGAAGCAGGGAAATTTCAAGAACCTGCGGGAGTTCGTATTTAAAAATATC
541GACCTTGAGGGGAAGCAGGGAAATTTCAAGAACCTGCGGGAGTTCGTATTTAAAAATATC
541************************************************************

601GACGGCTACTTTAAAATCTATAGTAAGCACACCCCCATTAACCTGGTCCGAGACCTTCCG
601GACGGCTACTTTAAAATCTATAGTAAGCACACCCCCATTAACCTGGTCCGAGACCTTCCG
601************************************************************
```

```
661CAGGGATTTTCCGCTCTGGAGCCTCTGGTTGACCTCCCTATAGGCATTAACATCACCAGA
661CAGGGATTTTCCGCTCTGGAGCCTCTGGTTGACCTCCCTATAGGCATTAACATCACCAGA
661************************************************************

721TTCCAGACCCTCCTGGCCCTGCACCGCTCATATTTGACCCCCGGGGACTCAAGTAGTGGA
721TTCCAGACCCTCCTGGCCCTGCACCGCTCATATTTGACCCCCGGGGACTCAAGTAGTGGA
721************************************************************

781TGGACAGCCGGGGCTGCTGCTTATTATGTAGGCTATCTGCAGCCTCGGACATTCCTGCTG
781TGGACAGCCGGGGCTGCTGCTTATTATGTAGGCTATCTGCAGCCTCGGACATTCCTGCTG
781************************************************************

841AAGTACAACGAAAACGGGACTATCACTGATGCCGTCGACTGCGCCCTTGACCCTCTTTCT
841AAGTACAACGAAAACGGGACTATCACTGATGCCGTCGACTGCGCCCTTGACCCTCTTTCT
841************************************************************
901GAGACCAAGTGTACACTTAAATCATTCACCGTGGAGAAAGGTATCTACCAGACCTCTAAT
901GAGACCAAGTGTACACTTAAATCATTCACCGTGGAGAAAGGTATCTACCAGACCTCTAAT
901************************************************************

961TTCCGAGTCCAGCCTACAGAGTCCATCGTGAGGTTTCCGAACATTACAAATCTTTGTCCC
961TTCCGAGTCCAGCCTACAGAGTCCATCGTGAGGTTTCCGAACATTACAAATCTTTGTCCC
961************************************************************
1021TTTGGCGAAGTGTTTAATGCCACTAGATTTGCAAGTGTGTACGCTTGGAACCGAAAGAGA
1021TTTGGCGAAGTGTTTAATGCCACTAGATTTGCAAGTGTGTACGCTTGGAACCGAAAGAGA
1021************************************************************

1081ATCTCCAACTGTGTCGCTGACTATTCTGTGCTCTACAACAGCGCTAGCTTTAGCACATTT
1081ATCTCCAACTGTGTCGCTGACTATTCTGTGCTCTACAACAGCGCTAGCTTTAGCACATTT
1081************************************************************

1141AAGTGTTACGGAGTTTCCCCCACCAAGCTGAATGACCTGTGCTTTACCAACGTCTATGCA
1141AAGTGTTACGGAGTTTCCCCCACCAAGCTGAATGACCTGTGCTTTACCAACGTCTATGCA
1141************************************************************

1201GACTCCTTTGTGATTAGAGGCGATGAAGTGCGACAAATTGCACCTGGCCAAACGGGCAAA
1201GACTCCTTTGTGATTAGAGGCGATGAAGTGCGACAAATTGCACCTGGCCAAACGGGCAAA
1201************************************************************

1261ATTGCTGATTATAACTACAAGCTGCCCGACGATTTCACAGGCTGCGTTATCGCCTGGAAC
1261ATTGCTGATTATAACTACAAGCTGCCCGACGATTTCACAGGCTGCGTTATCGCCTGGAAC
1261************************************************************

1321AGCAACAACCTGGACAGTAAGGTTGGCGGAAATTATAACTACCTGTATCGGCTGTTCAGG
1321AGCAACAACCTGGACAGTAAGGTTGGCGGAAATTATAACTACCTGTATCGGCTGTTCAGG
1321************************************************************

1381AAGAGCAATCTTAAGCCCTTTGAAAGAGACATTAGTACCGAAATATACCAGGCCGGCTCT
1381AAGAGCAATCTTAAGCCCTTTGAAAGAGACATTAGTACCGAAATATACCAGGCCGGCTCT
1381************************************************************

1441ACCCCATGCAATGGCGTTGAGGGCTTCAATTGCTATTTTCCCCTTCAGTCATATGGGTTT
1441ACCCCATGCAATGGCGTTGAGGGCTTCAATTGCTATTTTCCCCTTCAGTCATATGGGTTT
1441************************************************************
```

```
1501CAGCCAACCAATGGGGTGGGGTACCAGCCTTATCGCGTAGTGGTACTGAGCTTTGAGCTG
1501CAGCCAACCAATGGGGTGGGGTACCAGCCTTATCGCGTAGTGGTACTGAGCTTTGAGCTG
1501************************************************************

1561TTGCACGCACCAGCAACAGTGTGCGGGCCGAAGAAATCTACAAATCTCGTGAAGAACAAA
1561TTGCACGCACCAGCAACAGTGTGCGGGCCGAAGAAATCTACAAATCTCGTGAAGAACAAA
1561************************************************************

1621TGTGTCAATTTTAACTTTAACGGTCTGACCGGCACGGGGGTCCTGACAGAATCTAACAAG
1621TGTGTCAATTTTAACTTTAACGGTCTGACCGGCACGGGGGTCCTGACAGAATCTAACAAG
1621************************************************************

1681AAATTTCTCCCCTTTCAGCAGTTCGGAAGAGATATCGCAGATACAACTGATGCTGTGCGC
1681AAATTTCTCCCCTTTCAGCAGTTCGGAAGAGATATCGCAGATACAACTGATGCTGTGCGC
1681************************************************************

1741GATCCTCAGACGCTTGAGATTTTGGATATTACACCTTGCTCTTTCGGGGGGGTCAGCGTC
1741GATCCTCAGACGCTTGAGATTTTGGATATTACACCTTGCTCTTTCGGGGGGGTCAGCGTC
1741************************************************************
1801ATCACTCCCGGAACAAATACAAGCAACCAGGTGGCAGTCCTGTACCAGGACGTAAACTGC
1801ATCACTCCCGGAACAAATACAAGCAACCAGGTGGCAGTCCTGTACCAGGACGTAAACTGC
1801************************************************************

1861ACCGAAGTGCCCGTTGCAATTCATGCTGACCAGCTTACTCCCACATGGAGGGTGTACTCC
1861ACCGAAGTGCCCGTTGCAATTCATGCTGACCAGCTTACTCCCACATGGAGGGTGTACTCC
1861************************************************************

1921ACCGGATCTAATGTCTTTCAGACAAGAGCTGGCTGCCTGATCGGTGCCGAACATGTGAAT
1921ACCGGATCTAATGTCTTTCAGACAAGAGCTGGCTGCCTGATCGGTGCCGAACATGTGAAT
1921************************************************************

1981AATAGCTACGAGTGTGACATTCCCATCGGAGCCGGCATTTGTGCATCCTACCAAACTCAG
1981AATAGCTACGAGTGTGACATTCCCATCGGAGCCGGCATTTGTGCATCCTACCAAACTCAG
1981************************************************************
2041ACGAACAGCCCCGGCAGCGCCAGCAGTGTGGCCTCCCAGAGCATTATCGCATACACTATG
2041ACGAACAGCCCCGGCAGCGCCAGCAGTGTGGCCTCCCAGAGCATTATCGCATACACTATG
2041************************************************************

2101TCTCTGGGTGCTGAAAACTCCGTGGCATATTCAAACAACTCAATCGCCATTCCAACTAAC
2101TCTCTGGGTGCTGAAAACTCCGTGGCATATTCAAACAACTCAATCGCCATTCCAACTAAC
2101************************************************************

2161TTTACCATTTCAGTCACAACCGAAATCCTGCCAGTGTCTATGACCAAGACAAGCGTGGAT
2161TTTACCATTTCAGTCACAACCGAAATCCTGCCAGTGTCTATGACCAAGACAAGCGTGGAT
2161************************************************************

2221TGCACTATGTATATTTGCGGCGACTCCACAGAATGCTCTAATTTGCTTCTGCAGTACGGC
2221TGCACTATGTATATTTGCGGCGACTCCACAGAATGCTCTAATTTGCTTCTGCAGTACGGC
2221************************************************************

2281TCCTTTTGCACTCAATTGAATAGGGCATTGACTGGCATTGCTGTGGAGCAAGATAAAAAT
2281TCCTTTTGCACTCAATTGAATAGGGCATTGACTGGCATTGCTGTGGAGCAAGATAAAAAT
2281************************************************************
```

```
2341ACTCAGGAGGTATTTGCTCAGGTGAAACAAATCTACAAAACTCCACCAATTAAAGACTTC
2341ACTCAGGAGGTATTTGCTCAGGTGAAACAAATCTACAAAACTCCACCAATTAAAGACTTC
2341************************************************************

2401GGAGGTTTTAATTTCTCACAGATTCTGCCTGATCCATCAAAGCCTTCTAAGCGGAGTCCC
2401GGAGGTTTTAATTTCTCACAGATTCTGCCTGATCCATCAAAGCCTTCTAAGCGGAGTCCC
2401************************************************************

2461ATCGAAGACTTGCTCTTTAACAAAGTCACTCTGGCTGACGCCGGCTTTATCAAGCAGTAC
2461ATCGAAGACTTGCTCTTTAACAAAGTCACTCTGGCTGACGCCGGCTTTATCAAGCAGTAC
2461************************************************************

2521GGTGACTGCTTGGGCGACATCGCTGCTCGCGACCTGATCTGCGCTCAGAAGTTCAATGGC
2521GGTGACTGCTTGGGCGACATCGCTGCTCGCGACCTGATCTGCGCTCAGAAGTTCAATGGC
2521************************************************************

2581CTGACCGTCCTGCCACCACTGCTGACAGACGAAATGATTGCCCAGTACACCTCTGCATTG
2581CTGACCGTCCTGCCACCACTGCTGACAGACGAAATGATTGCCCAGTACACCTCTGCATTG
2581************************************************************

2641CTTGCCGGGACCATCACATCTGGGTGGACCTTCGGAGCTGGACCCGCTCTCCAGATCCCC
2641CTTGCCGGGACCATCACATCTGGGTGGACCTTCGGAGCTGGACCCGCTCTCCAGATCCCC
2641************************************************************

2701TTCCCCATGCAAATGGCCTATCGCTTCAATGGCATCGGGGTCACCCAGAATGTGCTGTAC
2701TTCCCCATGCAAATGGCCTATCGCTTCAATGGCATCGGGGTCACCCAGAATGTGCTGTAC
2701************************************************************

2761GAAAACCAGAAACTTATTGCCAATCAGTTTAATAGCGCGATCGGCAAGATTCAGGACAGT
2761GAAAACCAGAAACTTATTGCCAATCAGTTTAATAGCGCGATCGGCAAGATTCAGGACAGT
2761************************************************************

2821CTGTCTTCTACTCCCTCTGCCCTGGGAAAGTTGCAGGATGTGGTGAATCAGAATGCCCAG
2821CTGTCTTCTACTCCCTCTGCCCTGGGAAAGTTGCAGGATGTGGTGAATCAGAATGCCCAG
2821************************************************************

2881GCTCTGAATACACTGGTCAAGCAACTCAGCAGCAACTTCGGCGCAATTTCTTCCGTTCTG
2881GCTCTGAATACACTGGTCAAGCAACTCAGCAGCAACTTCGGCGCAATTTCTTCCGTTCTG
2881************************************************************

2941AACGATATCCTGTCTAGACTCGACCCACCAGAAGCTGAGGTACAGATTGATAGGCTTATC
2941AACGATATCCTGTCTAGACTCGACCCACCAGAAGCTGAGGTACAGATTGATAGGCTTATC
2941************************************************************



3001ACAGGCCGGCTCCAATCACTGCAAACATACGTGACTCAGCAGCTGATCCGGGCTGCCGAA
3001ACAGGCCGGCTCCAATCACTGCAAACATACGTGACTCAGCAGCTGATCCGGGCTGCCGAA
3001************************************************************

3061ATTAGGGCCAGTGCAAACCTGGCTGCGACTAAGATGTCTGAGTGCGTTCTGGGTCAGTCC
3061ATTAGGGCCAGTGCAAACCTGGCTGCGACTAAGATGTCTGAGTGCGTTCTGGGTCAGTCC
3061************************************************************
```

```
3121AAGAGGGTAGATTTTTGTGGCAAGGGGTACCATCTCATGTCTTTTCCTCAGAGTGCACCA
3121AAGAGGGTAGATTTTTGTGGCAAGGGGTACCATCTCATGTCTTTTCCTCAGAGTGCACCA
3121************************************************************

3181CATGGAGTGGTCTTCCTCCACGTGACCTACGTGCCCGCCCAAGAGAAAAATTTTACAACC
3181CATGGAGTGGTCTTCCTCCACGTGACCTACGTGCCCGCCCAAGAGAAAAATTTTACAACC
3181************************************************************

3241GCTCCAGCTATCTGCCACGATGGAAAAGCTCATTTCCCGCGGGAGGGCGTGTTCGTCAGT
3241GCTCCAGCTATCTGCCACGATGGAAAAGCTCATTTCCCGCGGGAGGGCGTGTTCGTCAGT
3241************************************************************

3301AATGGAACGCACTGGTTCGTGACCCAGAGAAACTTTTACGAGCCCCAGATAATTACCACC
3301AATGGAACGCACTGGTTCGTGACCCAGAGAAACTTTTACGAGCCCCAGATAATTACCACC
3301************************************************************

3361GATAATACATTTGTTAGCGGCAACTGCGATGTCGTGATCGGGATAGTGAATAACACAGTG
3361GATAATACATTTGTTAGCGGCAACTGCGATGTCGTGATCGGGATAGTGAATAACACAGTG
3361************************************************************

3421TACGACCCACTCCAGCCCGAACTCGACTCATTTAAGGAAGAGCTTGACAAGTACTTTAAG
3421TACGACCCACTCCAGCCCGAACTCGACTCATTTAAGGAAGAGCTTGACAAGTACTTTAAG
3421************************************************************

3481AATCACACCAGCCCAGATGTGGACCTGGGAGATATCAGCGGCATTAATGCCTCAGTGGTG
3481AATCACACCAGCCCAGATGTGGACCTGGGAGATATCAGCGGCATTAATGCCTCAGTGGTG
3481************************************************************

3541AATATTCAGAAGGAAATAGATAGACTGAATGAAGTGGCAAAGAATCTGAACGAGTCCCTC
3541AATATTCAGAAGGAAATAGATAGACTGAATGAAGTGGCAAAGAATCTGAACGAGTCCCTC
3541************************************************************
3601ATTGATCTGCAGGAGCTCGGCAAGTACGAGCAATATATTAAATGGCCGTGGTACATTTGG
3601ATTGATCTGCAGGAGCTCGGCAAGTACGAGCAATATATTAAATGGCCGTGGTACATTTGG
3601************************************************************

3661CTGGGATTTATCGCAGGACTCATTGCCATCGTGATGGTGACTATCATGCTGTGTTGTATG
3661CTGGGATTTATCGCAGGACTCATTGCCATCGTGATGGTGACTATCATGCTGTGTTGTATG
3661************************************************************

3721ACCAGCTGCTGCAGCTGTTTGAAAGGTTGTTGCAGCTGTGGTTCCTGTTGCAAGTTCGAT
3721ACCAGCTGCTGCAGCTGTTTGAAAGGTTGTTGCAGCTGTGGTTCCTGTTGCAAGTTCGAT
3721************************************************************

3781GAAGATGATAGCGAGCCAGTTCTGAAGGGCGTTAAGCTGCACTATACTGAAAACCTGTAT
3781GAAGATGATAGCGAGCCAGTTCTGAAGGGCGTTAAGCTGCACTATACTGAAAACCTGTAT
3781************************************************************

3841TTTCAGAGCGGATCTGGGTACATCCCCGAGGCACCAAGAGATGGTCAGGCATATGTTAGG
3841TTTCAGAGCGGATCTGGGTACATCCCCGAGGCACCAAGAGATGGTCAGGCATATGTTAGG
3841************************************************************

3901AAGGATGGGGAATGGGTCCTCCTGAGCACATTTTTGGGAGGAAGCCTGGTACCCAGAGGC
3901AAGGATGGGGAATGGGTCCTCCTGAGCACATTTTTGGGAGGAAGCCTGGTACCCAGAGGC
3901************************************************************
```

```
3961TCTCCACACCATCATCACCACCACTAA
3961TCTCCACACCATCATCACCACCACTAA
3961**************************
```

Pairwise alignment of 2pS protein encoding DNA Sequence (Upper: Putative, Lower:Sequencing)

```
1ATGCCTTTGCTTTTGTTGCTTCCTCTTCTCTGGGCTGGGGCTTTGGCCAAGCTTGTGAAC
1ATGCCTTTGCTTTTGTTGCTTCCTCTTCTCTGGGCTGGGGCTTTGGCCAAGCTTGTGAAC
1************************************************************

61CTTACCACACGCACCCAATTGCCCCCAGCTTACACAAACAGCTTTACTAGGGGCGTCTAC
61CTTACCACACGCACCCAATTGCCCCCAGCTTACACAAACAGCTTTACTAGGGGCGTCTAC
61************************************************************

121TATCCCGACAAGGTGTTTCGGAGTAGCGTGCTGCACTCTACTCAAGACCTCTTCCTGCCC
121TATCCCGACAAGGTGTTTCGGAGTAGCGTGCTGCACTCTACTCAAGACCTCTTCCTGCCC
121************************************************************

181TTCTTTTCCAATGTCACTTGGTTCCACGCCATCCATGTCAGCGGTACCAATGGGACCAAA
181TTCTTTTCCAATGTCACTTGGTTCCACGCCATCCATGTCAGCGGTACCAATGGGACCAAA
181************************************************************

241AGGTTTGACAACCCAGTGTTGCCATTTAACGATGGCGTGTATTTCGCCAGTACTGAAAAG
241AGGTTTGACAACCCAGTGTTGCCATTTAACGATGGCGTGTATTTCGCCAGTACTGAAAAG
241************************************************************

301AGCAATATCATCAGGGGCTGGATCTTCGGTACTACCCTGGATAGCAAGACCCAGTCTTTG
301AGCAATATCATCAGGGGCTGGATCTTCGGTACTACCCTGGATAGCAAGACCCAGTCTTTG
301************************************************************

361CTGATTGTGAACAATGCTACAAATGTGGTAATCAAAGTTTGTGAGTTTCAATTCTGCAAT
361CTGATTGTGAACAATGCTACAAATGTGGTAATCAAAGTTTGTGAGTTTCAATTCTGCAAT
361************************************************************

421GACCCATTCCTGGGCGTCTATTACCACAAGAATAACAAAAGTTGGATGGAATCCGAGTTT
421GACCCATTCCTGGGCGTCTATTACCACAAGAATAACAAAAGTTGGATGGAATCCGAGTTT
421************************************************************

481CGGGTGTACAGCTCCGCTAACAACTGCACATTCGAATATGTGTCCCAGCCGTTCCTGATG
481CGGGTGTACAGCTCCGCTAACAACTGCACATTCGAATATGTGTCCCAGCCGTTCCTGATG
481************************************************************

541GACCTTGAGGGGAAGCAGGGAAATTTCAAGAACCTGCGGGAGTTCGTATTTAAAAATATC
541GACCTTGAGGGGAAGCAGGGAAATTTCAAGAACCTGCGGGAGTTCGTATTTAAAAATATC
541************************************************************

601GACGGCTACTTTAAAATCTATAGTAAGCACACCCCCATTAACCTGGTCCGAGACCTTCCG
601GACGGCTACTTTAAAATCTATAGTAAGCACACCCCCATTAACCTGGTCCGAGACCTTCCG
601************************************************************
```

```
661CAGGGATTTTCCGCTCTGGAGCCTCTGGTTGACCTCCCTATAGGCATTAACATCACCAGA
661CAGGGATTTTCCGCTCTGGAGCCTCTGGTTGACCTCCCTATAGGCATTAACATCACCAGA
661************************************************************

721TTCCAGACCCTCCTGGCCCTGCACCGCTCATATTTGACCCCCGGGGACTCAAGTAGTGGA
721TTCCAGACCCTCCTGGCCCTGCACCGCTCATATTTGACCCCCGGGGACTCAAGTAGTGGA
721************************************************************

781TGGACAGCCGGGGCTGCTGCTTATTATGTAGGCTATCTGCAGCCTCGGACATTCCTGCTG
781TGGACAGCCGGGGCTGCTGCTTATTATGTAGGCTATCTGCAGCCTCGGACATTCCTGCTG
781************************************************************

841AAGTACAACGAAAACGGGACTATCACTGATGCCGTCGACTGCGCCCTTGACCCTCTTTCT
841AAGTACAACGAAAACGGGACTATCACTGATGCCGTCGACTGCGCCCTTGACCCTCTTTCT
841************************************************************

901GAGACCAAGTGTACACTTAAATCATTCACCGTGGAGAAAGGTATCTACCAGACCTCTAAT
901GAGACCAAGTGTACACTTAAATCATTCACCGTGGAGAAAGGTATCTACCAGACCTCTAAT
901************************************************************



961TTCCGAGTCCAGCCTACAGAGTCCATCGTGAGGTTTCCGAACATTACAAATCTTTGTCCC
961TTCCGAGTCCAGCCTACAGAGTCCATCGTGAGGTTTCCGAACATTACAAATCTTTGTCCC
961************************************************************

1021TTTGGCGAAGTGTTTAATGCCACTAGATTTGCAAGTGTGTACGCTTGGAACCGAAAGAGA
1021TTTGGCGAAGTGTTTAATGCCACTAGATTTGCAAGTGTGTACGCTTGGAACCGAAAGAGA
1021************************************************************

1081ATCTCCAACTGTGTCGCTGACTATTCTGTGCTCTACAACAGCGCTAGCTTTAGCACATTT
1081ATCTCCAACTGTGTCGCTGACTATTCTGTGCTCTACAACAGCGCTAGCTTTAGCACATTT
1081************************************************************

1141AAGTGTTACGGAGTTTCCCCCACCAAGCTGAATGACCTGTGCTTTACCAACGTCTATGCA
1141AAGTGTTACGGAGTTTCCCCCACCAAGCTGAATGACCTGTGCTTTACCAACGTCTATGCA
1141************************************************************

1201GACTCCTTTGTGATTAGAGGCGATGAAGTGCGACAAATTGCACCTGGCCAAACGGGCAAA
1201GACTCCTTTGTGATTAGAGGCGATGAAGTGCGACAAATTGCACCTGGCCAAACGGGCAAA
1201************************************************************

1261ATTGCTGATTATAACTACAAGCTGCCCGACGATTTCACAGGCTGCGTTATCGCCTGGAAC
1261ATTGCTGATTATAACTACAAGCTGCCCGACGATTTCACAGGCTGCGTTATCGCCTGGAAC
1261************************************************************

1321AGCAACAACCTGGACAGTAAGGTTGGCGGAAATTATAACTACCTGTATCGGCTGTTCAGG
1321AGCAACAACCTGGACAGTAAGGTTGGCGGAAATTATAACTACCTGTATCGGCTGTTCAGG
1321************************************************************

1381AAGAGCAATCTTAAGCCCTTTGAAAGAGACATTAGTACCGAAATATACCAGGCCGGCTCT
1381AAGAGCAATCTTAAGCCCTTTGAAAGAGACATTAGTACCGAAATATACCAGGCCGGCTCT
1381************************************************************
```

```
1441ACCCCATGCAATGGCGTTGAGGGCTTCAATTGCTATTTTCCCCTTCAGTCATATGGGTTT
1441ACCCCATGCAATGGCGTTGAGGGCTTCAATTGCTATTTTCCCCTTCAGTCATATGGGTTT
1441************************************************************

1501CAGCCAACCAATGGGGTGGGGTACCAGCCTTATCGCGTAGTGGTACTGAGCTTTGAGCTG
1501CAGCCAACCAATGGGGTGGGGTACCAGCCTTATCGCGTAGTGGTACTGAGCTTTGAGCTG
1501************************************************************

1561TTGCACGCACCAGCAACAGTGTGCGGGCCGAAGAAATCTACAAATCTCGTGAAGAACAAA
1561TTGCACGCACCAGCAACAGTGTGCGGGCCGAAGAAATCTACAAATCTCGTGAAGAACAAA
1561************************************************************

1621TGTGTCAATTTTAACTTTAACGGTCTGACCGGCACGGGGGTCCTGACAGAATCTAACAAG
1621TGTGTCAATTTTAACTTTAACGGTCTGACCGGCACGGGGGTCCTGACAGAATCTAACAAG
1621************************************************************

1681AAATTTCTCCCCTTTCAGCAGTTCGGAAGAGATATCGCAGATACAACTGATGCTGTGCGC
1681AAATTTCTCCCCTTTCAGCAGTTCGGAAGAGATATCGCAGATACAACTGATGCTGTGCGC
1681************************************************************

1741GATCCTCAGACGCTTGAGATTTTGGATATTACACCTTGCTCTTTCGGGGGGGTCAGCGTC
1741GATCCTCAGACGCTTGAGATTTTGGATATTACACCTTGCTCTTTCGGGGGGGTCAGCGTC
1741************************************************************

1801ATCACTCCCGGAACAAATACAAGCAACCAGGTGGCAGTCCTGTACCAGGACGTAAACTGC
1801ATCACTCCCGGAACAAATACAAGCAACCAGGTGGCAGTCCTGTACCAGGACGTAAACTGC
1801************************************************************

1861ACCGAAGTGCCCGTTGCAATTCATGCTGACCAGCTTACTCCCACATGGAGGGTGTACTCC
1861ACCGAAGTGCCCGTTGCAATTCATGCTGACCAGCTTACTCCCACATGGAGGGTGTACTCC
1861************************************************************



1921ACCGGATCTAATGTCTTTCAGACAAGAGCTGGCTGCCTGATCGGTGCCGAACATGTGAAT
1921ACCGGATCTAATGTCTTTCAGACAAGAGCTGGCTGCCTGATCGGTGCCGAACATGTGAAT
1921************************************************************

1981AATAGCTACGAGTGTGACATTCCCATCGGAGCCGGCATTTGTGCATCCTACCAAACTCAG
1981AATAGCTACGAGTGTGACATTCCCATCGGAGCCGGCATTTGTGCATCCTACCAAACTCAG
1981************************************************************

2041ACGATTCTGAGGAGTGTGGCCTCCCAGAGCATTATCGCATACACTATGTCTCTGGGTGCT
2041ACGATTCTGAGGAGTGTGGCCTCCCAGAGCATTATCGCATACACTATGTCTCTGGGTGCT
2041************************************************************

2101GAAAACTCCGTGGCATATTCAAACAACTCAATCGCCATTCCAACTAACTTTACCATTTCA
2101GAAAACTCCGTGGCATATTCAAACAACTCAATCGCCATTCCAACTAACTTTACCATTTCA
2101************************************************************

2161GTCACAACCGAAATCCTGCCAGTGTCTATGACCAAGACAAGCGTGGATTGCACTATGTAT
2161GTCACAACCGAAATCCTGCCAGTGTCTATGACCAAGACAAGCGTGGATTGCACTATGTAT
2161************************************************************
```

```
2221ATTTGCGGCGACTCCACAGAATGCTCTAATTTGCTTCTGCAGTACGGCTCCTTTTGCACT
2221ATTTGCGGCGACTCCACAGAATGCTCTAATTTGCTTCTGCAGTACGGCTCCTTTTGCACT
2221************************************************************

2281CAATTGAATAGGGCATTGACTGGCATTGCTGTGGAGCAAGATAAAAATACTCAGGAGGTA
2281CAATTGAATAGGGCATTGACTGGCATTGCTGTGGAGCAAGATAAAAATACTCAGGAGGTA
2281************************************************************

2341TTTGCTCAGGTGAAACAAATCTACAAAACTCCACCAATTAAAGACTTCGGAGGTTTTAAT
2341TTTGCTCAGGTGAAACAAATCTACAAAACTCCACCAATTAAAGACTTCGGAGGTTTTAAT
2341************************************************************

2401TTCTCACAGATTCTGCCTGATCCATCAAAGCCTTCTAAGCGGAGTTTTATCGAAGACTTG
2401TTCTCACAGATTCTGCCTGATCCATCAAAGCCTTCTAAGCGGAGTTTTATCGAAGACTTG
2401************************************************************

2461CTCTTTAACAAAGTCACTCTGGCTGACGCCGGCTTTATCAAGCAGTACGGTGACTGCTTG
2461CTCTTTAACAAAGTCACTCTGGCTGACGCCGGCTTTATCAAGCAGTACGGTGACTGCTTG
2461************************************************************

2521GGCGACATCGCTGCTCGCGACCTGATCTGCGCTCAGAAGTTCAATGGCCTGACCGTCCTG
2521GGCGACATCGCTGCTCGCGACCTGATCTGCGCTCAGAAGTTCAATGGCCTGACCGTCCTG
2521************************************************************

2581CCACCACTGCTGACAGACGAAATGATTGCCCAGTACACCTCTGCATTGCTTGCCGGGACC
2581CCACCACTGCTGACAGACGAAATGATTGCCCAGTACACCTCTGCATTGCTTGCCGGGACC
2581************************************************************

2641ATCACATCTGGGTGGACCTTCGGAGCTGGAGCAGCTCTCCAGATCCCCTTCGCCATGCAA
2641ATCACATCTGGGTGGACCTTCGGAGCTGGAGCAGCTCTCCAGATCCCCTTCGCCATGCAA
2641************************************************************

2701ATGGCCTATCGCTTCAATGGCATCGGGGTCACCCAGAATGTGCTGTACGAAAACCAGAAA
2701ATGGCCTATCGCTTCAATGGCATCGGGGTCACCCAGAATGTGCTGTACGAAAACCAGAAA
2701************************************************************

2761CTTATTGCCAATCAGTTTAATAGCGCGATCGGCAAGATTCAGGACAGTCTGTCTTCTACT
2761CTTATTGCCAATCAGTTTAATAGCGCGATCGGCAAGATTCAGGACAGTCTGTCTTCTACT
2761************************************************************

2821GCTTCTGCCCTGGGAAAGTTGCAGGATGTGGTGAATCAGAATGCCCAGGCTCTGAATACA
2821GCTTCTGCCCTGGGAAAGTTGCAGGATGTGGTGAATCAGAATGCCCAGGCTCTGAATACA
2821************************************************************

2881CTGGTCAAGCAACTCAGCAGCAACTTCGGCGCAATTTCTTCCGTTCTGAACGATATCCTG
2881CTGGTCAAGCAACTCAGCAGCAACTTCGGCGCAATTTCTTCCGTTCTGAACGATATCCTG
2881************************************************************

2941TCTAGACTCGACCCACCAGAAGCTGAGGTACAGATTGATAGGCTTATCACAGGCCGGCTC
2941TCTAGACTCGACCCACCAGAAGCTGAGGTACAGATTGATAGGCTTATCACAGGCCGGCTC
2941************************************************************

3001CAATCACTGCAAACATACGTGACTCAGCAGCTGATCCGGGCTGCCGAAATTAGGGCCAGT
3001CAATCACTGCAAACATACGTGACTCAGCAGCTGATCCGGGCTGCCGAAATTAGGGCCAGT
3001************************************************************
```

```
3061GCAAACCTGGCTGCGACTAAGATGTCTGAGTGCGTTCTGGGTCAGTCCAAGAGGGTAGAT
3061GCAAACCTGGCTGCGACTAAGATGTCTGAGTGCGTTCTGGGTCAGTCCAAGAGGGTAGAT
3061************************************************************

3121TTTTGTGGCAAGGGGTACCATCTCATGTCTTTTCCTCAGAGTGCACCACATGGAGTGGTC
3121TTTTGTGGCAAGGGGTACCATCTCATGTCTTTTCCTCAGAGTGCACCACATGGAGTGGTC
3121************************************************************

3181TTCCTCCACGTGACCTACGTGCCCGCCCAAGAGAAAAATTTTACAACCGCTCCAGCTATC
3181TTCCTCCACGTGACCTACGTGCCCGCCCAAGAGAAAAATTTTACAACCGCTCCAGCTATC
3181************************************************************

3241TGCCACGATGGAAAAGCTCATTTCCCGCGGGAGGGCGTGTTCGTCAGTAATGGAACGCAC
3241TGCCACGATGGAAAAGCTCATTTCCCGCGGGAGGGCGTGTTCGTCAGTAATGGAACGCAC
3241************************************************************

3301TGGTTCGTGACCCAGAGAAACTTTTACGAGCCCCAGATAATTACCACCGATAATACATTT
3301TGGTTCGTGACCCAGAGAAACTTTTACGAGCCCCAGATAATTACCACCGATAATACATTT
3301************************************************************

3361GTTAGCGGCAACTGCGATGTCGTGATCGGGATAGTGAATAACACAGTGTACGACCCACTC
3361GTTAGCGGCAACTGCGATGTCGTGATCGGGATAGTGAATAACACAGTGTACGACCCACTC
3361************************************************************

3421CAGCCCGAACTCGACTCATTTAAGGAAGAGCTTGACAAGTACTTTAAGAATCACACCAGC
3421CAGCCCGAACTCGACTCATTTAAGGAAGAGCTTGACAAGTACTTTAAGAATCACACCAGC
3421************************************************************

3481CCAGATGTGGACCTGGGAGATATCAGCGGCATTAATGCCTCAGTGGTGAATATTCAGAAG
3481CCAGATGTGGACCTGGGAGATATCAGCGGCATTAATGCCTCAGTGGTGAATATTCAGAAG
3481************************************************************

3541GAAATAGATAGACTGAATGAAGTGGCAAAGAATCTGAACGAGTCCCTCATTGATCTGCAG
3541GAAATAGATAGACTGAATGAAGTGGCAAAGAATCTGAACGAGTCCCTCATTGATCTGCAG
3541************************************************************

3601GAGCTCGGCAAGTACGAGCAATATATTAAATGGCCGTGGTACATTTGGCTGGGATTTATC
3601GAGCTCGGCAAGTACGAGCAATATATTAAATGGCCGTGGTACATTTGGCTGGGATTTATC
3601************************************************************

3661GCAGGACTCATTGCCATCGTGATGGTGACTATCATGCTGTGTTGTATGACCAGCTGCTGC
3661GCAGGACTCATTGCCATCGTGATGGTGACTATCATGCTGTGTTGTATGACCAGCTGCTGC
3661************************************************************

3721AGCTGTTTGAAAGGTTGTTGCAGCTGTGGTTCCTGTTGCAAGTTCGATGAAGATGATAGC
3721AGCTGTTTGAAAGGTTGTTGCAGCTGTGGTTCCTGTTGCAAGTTCGATGAAGATGATAGC
3721************************************************************

3781GAGCCAGTTCTGAAGGGCGTTAAGCTGCACTATACTGAAAACCTGTATTTTCAGAGCGGA
3781GAGCCAGTTCTGAAGGGCGTTAAGCTGCACTATACTGAAAACCTGTATTTTCAGAGCGGA
3781************************************************************

3841TCTGGGTACATCCCCGAGGCACCAAGAGATGGTCAGGCATATGTTAGGAAGGATGGGGAA
3841TCTGGGTACATCCCCGAGGCACCAAGAGATGGTCAGGCATATGTTAGGAAGGATGGGGAA
3841************************************************************
```

```
3901TGGGTCCTCCTGAGCACATTTTTGGGAGGAAGCCTGGTACCCAGAGGCTCTCCACACCAT
3901TGGGTCCTCCTGAGCACATTTTTGGGAGGAAGCCTGGTACCCAGAGGCTCTCCACACCAT
3901************************************************************

3961CATCACCACCACTAA
3961CATCACCACCACTAA
3961***************
```

Pairwise alignment of Envelope protein encoding DNA Sequence (Upper: Putative, Lower:Sequencing)

```
1AAAATGCCTTTGCTTTTGTTGCTTCCTCTTCTCTGGGCTGGGGCTTTGGCCAAGCTTATG
1AAAATGCCTTTGCTTTTGTTGCTTCCTCTTCTCTGGGCTGGGGCTTTGGCCAAGCTTATG
1************************************************************

61TATAGCTTTGTGTCAGAGGAGACTGGCACACTTATAGTTAATAGCGTCCTCCTGTTCCTC
61TATAGCTTTGTGTCAGAGGAGACTGGCACACTTATAGTTAATAGCGTCCTCCTGTTCCTC
61************************************************************

121GCATTCGTAGTCTTTTTGCTCGTCACTCTCGCGATCTTGACGGCTTTGAGGTTGTGCGCT
121GCATTCGTAGTCTTTTTGCTCGTCACTCTCGCGATCTTGACGGCTTTGAGGTTGTGCGCT
121************************************************************

181TACTGCTGCAACATTGTGAATGTCAGTCTCGTAAAACCTTCTTTCTACGTCTACTCCAGA
181TACTGCTGCAACATTGTGAATGTCAGTCTCGTAAAACCTTCTTTCTACGTCTACTCCAGA
181************************************************************

241GTAAAAAATCTTAATAGTTCCCGCGTGCCAGACTTGCTGGTAGGAAGCCTGGTACCCAGA
241GTAAAAAATCTTAATAGTTCCCGCGTGCCAGACTTGCTGGTAGGAAGCCTGGTACCCAGA
241************************************************************

301GGCTCTCCACACCATCATCACCACCACTAA
301GGCTCTCCACACCATCATCACCACCACTAA
301****************************
```

Pairwise alignment of Nucleocapsid protein encoding DNA Sequence (Upper: Putative, Lower:Sequencing)

```
1ATGAGCGACAACGGCCCACAAAATCAACGGAACGCCCCCCGGATAACCTTCGGTGGGCCA
1ATGAGCGACAACGGCCCACAAAATCAACGGAACGCCCCCCGGATAACCTTCGGTGGGCCA
1************************************************************

61TCTGACAGTACTGGCTCTAATCAGAACGGAGAGCGCTCCGGAGCAAGAAGCAAGCAACGA
61TCTGACAGTACTGGCTCTAATCAGAACGGAGAGCGCTCCGGAGCAAGAAGCAAGCAACGA
61************************************************************

121CGGCCGCAAGGACTGCCCAATAATACTGCCAGTTGGTTCACAGCACTGACCCAGCACGGG
121CGGCCGCAAGGACTGCCCAATAATACTGCCAGTTGGTTCACAGCACTGACCCAGCACGGG
121************************************************************
```

```
181AAAGAGGATCTCAAATTTCCGCGCGGTCAAGGTGTCCCCATTAACACTAATTCATCTCCC
181AAAGAGGATCTCAAATTTCCGCGCGGTCAAGGTGTCCCCATTAACACTAATTCATCTCCC
181************************************************************

241GATGATCAAATCGGGTACTATCGACGAGCAACGCGCCGAATAAGGGGCGGCGACGGAAAA
241GATGATCAAATCGGGTACTATCGACGAGCAACGCGCCGAATAAGGGGCGGCGACGGAAAA
241************************************************************

301ATGAAGGACCTGTCCCCAAGGTGGTACTTTTACTATCTTGGCACCGGGCCGGAAGCGGGT
301ATGAAGGACCTGTCCCCAAGGTGGTACTTTTACTATCTTGGCACCGGGCCGGAAGCGGGT
301************************************************************

361CTTCCGTATGGAGCAAACAAGGACGGAATAATTTGGGTCGCGACCGAGGGTGCTCTTAAT
361CTTCCGTATGGAGCAAACAAGGACGGAATAATTTGGGTCGCGACCGAGGGTGCTCTTAAT
361************************************************************

421ACTCCCAAAGATCACATTGGTACTCGCAATCCTGCTAACAACGCAGCCATCGTTCTCCAG
421ACTCCCAAAGATCACATTGGTACTCGCAATCCTGCTAACAACGCAGCCATCGTTCTCCAG
421************************************************************

481CTGCCGCAGGGAACAACGCTGCCCAAGGGTTTTTACGCGGAAGGCTCAAGGGGTGGAAGC
481CTGCCGCAGGGAACAACGCTGCCCAAGGGTTTTTACGCGGAAGGCTCAAGGGGTGGAAGC
481************************************************************

541CAGGCGTCAAGTCGCAGCTCCAGCAGAAGTAGAAACAGTTCACGAAATAGCACCCCAGGC
541CAGGCGTCAAGTCGCAGCTCCAGCAGAAGTAGAAACAGTTCACGAAATAGCACCCCAGGC
541************************************************************

601TCCTCACGCGGTACGAGTCCAGCTAGAATGGCAGGAAACGGAGGGGACGCTGCTCTTGCC
601TCCTCACGCGGTACGAGTCCAGCTAGAATGGCAGGAAACGGAGGGGACGCTGCTCTTGCC
601************************************************************

661TTGCTGCTTCTGGATCGGCTCAACCAGCTCGAAAGCAAGATGAGTGGTAAGGGGCAGCAA
661TTGCTGCTTCTGGATCGGCTCAACCAGCTCGAAAGCAAGATGAGTGGTAAGGGGCAGCAA
661************************************************************

721CAACAGGGTCAGACCGTAACGAAAAAGTCTGCTGCAGAGGCTTCAAAGAAACCTAGGCAG
721CAACAGGGTCAGACCGTAACGAAAAAGTCTGCTGCAGAGGCTTCAAAGAAACCTAGGCAG
721************************************************************

781AAGCGGACCGCAACCAAAGCCTACAATGTTACCCAAGCCTTCGGGCGGAGAGGTCCGGAA
781AAGCGGACCGCAACCAAAGCCTACAATGTTACCCAAGCCTTCGGGCGGAGAGGTCCGGAA
781************************************************************

841CAAACGCAAGGGAATTTCGGTGATCAGGAACTGATTCGGCAAGGAACCGATTACAAGCAT
841CAAACGCAAGGGAATTTCGGTGATCAGGAACTGATTCGGCAAGGAACCGATTACAAGCAT
841************************************************************

901TGGCCACAGATTGCTCAATTCGCCCCAAGCGCGAGTGCTTTCTTTGGGATGTCACGCATC
901TGGCCACAGATTGCTCAATTCGCCCCAAGCGCGAGTGCTTTCTTTGGGATGTCACGCATC
901************************************************************

961GGGATGGAAGTAACTCCATCCGGAACGTGGTTGACGTATACAGGCGCGATAAAGCTGGAC
961GGGATGGAAGTAACTCCATCCGGAACGTGGTTGACGTATACAGGCGCGATAAAGCTGGAC
961************************************************************
```

```
1021GATAAAGACCCTAATTTCAAGGATCAGGTTATTCTGTTGAATAAACATATAGACGCGTAT
1021GATAAAGACCCTAATTTCAAGGATCAGGTTATTCTGTTGAATAAACATATAGACGCGTAT
1021************************************************************

1081AAAACCTTCCCGCCTACCGAGCCCAAGAAAGATAAGAAGAAAAAAGCCGACGAAACCCAG
1081AAAACCTTCCCGCCTACCGAGCCCAAGAAAGATAAGAAGAAAAAAGCCGACGAAACCCAG
1081************************************************************

1141GCTCTCCCTCAGAGGCAAAAGAAACAGCAAACGGTTACTCTCCTCCCTGCCGCCGATTTG
1141GCTCTCCCTCAGAGGCAAAAGAAACAGCAAACGGTTACTCTCCTCCCTGCCGCCGATTTG
1141************************************************************

1201GACGACTTTAGCAAACAGTTGCAGCAGAGTATGTCCAGTGCAGACTCCACTCAGGCATAA
1201GACGACTTTAGCAAACAGTTGCAGCAGAGTATGTCCAGTGCAGACTCCACTCAGGCATAA
1201************************************************************
```

Pairwise alignment of Membrane protein encoding DNA Sequence (Upper: Putative, Lower:Sequencing)

```
1AAAATGCCTTTGCTTTTGTTGCTTCCTCTTCTCTGGGCTGGGGCTTTGGCCAAGCTTATG
1AAAATGCCTTTGCTTTTGTTGCTTCCTCTTCTCTGGGCTGGGGCTTTGGCCAAGCTTATG
1************************************************************

61GCGGACAGTAATGGAACTATTACGGTGGAGGAACTTAAGAAACTGCTGGAACAATGGAAC
61GCGGACAGTAATGGAACTATTACGGTGGAGGAACTTAAGAAACTGCTGGAACAATGGAAC
61************************************************************

121CTGGTAATAGGTTTCCTGTTCCTCACATGGATCTGCCTCTTGCAATTCGCGTACGCGAAC
121CTGGTAATAGGTTTCCTGTTCCTCACATGGATCTGCCTCTTGCAATTCGCGTACGCGAAC
121************************************************************

181CGGAATAGATTTCTCTACATAATAAAGCTGATTTTCCTCTGGCTCCTGTGGCCCGTCACA
181CGGAATAGATTTCTCTACATAATAAAGCTGATTTTCCTCTGGCTCCTGTGGCCCGTCACA
181************************************************************

241CTCGCCTGTTTCGTACTCGCAGCAGTGTACAGGATTAACTGGATTACCGGGGGAATTGCG
241CTCGCCTGTTTCGTACTCGCAGCAGTGTACAGGATTAACTGGATTACCGGGGGAATTGCG
241************************************************************

301ATTGCAATGGCTTGTCTGGTGGGACTGATGTGGCTTAGCTATTTTATTGCTAGTTTTCGA
301ATTGCAATGGCTTGTCTGGTGGGACTGATGTGGCTTAGCTATTTTATTGCTAGTTTTCGA
301************************************************************

361CTTTTCGCAAGAACTCGAAGCATGTGGTCCTTCAATCCTGAGACGAATATACTGTTGAAT
361CTTTTCGCAAGAACTCGAAGCATGTGGTCCTTCAATCCTGAGACGAATATACTGTTGAAT
361************************************************************

421GTGCCTTTGCATGGAACAATTTTGACCCGACCGCTTCTTGAAAGCGAGCTTGTAATTGGG
421GTGCCTTTGCATGGAACAATTTTGACCCGACCGCTTCTTGAAAGCGAGCTTGTAATTGGG
421************************************************************

481GCCGTGATACTTAGAGGTCACTTGCGGATCGCAGGCCACCACCTTGGGCGGTGTGATATC
481GCCGTGATACTTAGAGGTCACTTGCGGATCGCAGGCCACCACCTTGGGCGGTGTGATATC
481************************************************************
```

```
541AAGGATCTTCCTAAGGAAATCACGGTCGCAACGAGCCGCACGTTGAGCTACTATAAACTC
541AAGGATCTTCCTAAGGAAATCACGGTCGCAACGAGCCGCACGTTGAGCTACTATAAACTC
541************************************************************

601GGCGCGAGCCAGCGGGTTGCTGGTGATTCAGGTTTCGCCGCTTACTCTAGATACAGGATT
601GGCGCGAGCCAGCGGGTTGCTGGTGATTCAGGTTTCGCCGCTTACTCTAGATACAGGATT
601************************************************************

661GGTAATTACAAACTCAATACAGACCACAGTTCCTCTTCCGATAACATTGCTCTTCTGGTA
661GGTAATTACAAACTCAATACAGACCACAGTTCCTCTTCCGATAACATTGCTCTTCTGGTA
661************************************************************

721CAGGGAAGCCTGGTACCCAGAGGCTCTCCACACCATCATCACCACCACTAA
721CAGGGAAGCCTGGTACCCAGAGGCTCTCCACACCATCATCACCACCACTAA
721**************************************************
```

Pairwise alignment of recombinant Spike protein encoding DNA Sequence (Upper: Putative, Lower:Sequencing)

```
1ATGCCTTTGCTTTTGTTGCTTCCTCTTCTCTGGGCTGGGGCTTTGGCCAAGCTTGTGAAC
1ATGCCTTTGCTTTTGTTGCTTCCTCTTCTCTGGGCTGGGGCTTTGGCCAAGCTTGTGAAC
1************************************************************

61CTTACCACACGCACCCAATTGCCCCCAGCTTACACAAACAGCTTTACTAGGGGCGTCTAC
61CTTACCACACGCACCCAATTGCCCCCAGCTTACACAAACAGCTTTACTAGGGGCGTCTAC
61************************************************************

121TATCCCGACAAGGTGTTTCGGAGTAGCGTGCTGCACTCTACTCAAGACCTCTTCCTGCCC
121TATCCCGACAAGGTGTTTCGGAGTAGCGTGCTGCACTCTACTCAAGACCTCTTCCTGCCC
121************************************************************

181TTCTTTTCCAATGTCACTTGGTTCCACGCCATCCATGTCAGCGGTACCAATGGGACCAAA
181TTCTTTTCCAATGTCACTTGGTTCCACGCCATCCATGTCAGCGGTACCAATGGGACCAAA
181************************************************************

241AGGTTTGACAACCCAGTGTTGCCATTTAACGATGGCGTGTATTTCGCCAGTACTGAAAAG
241AGGTTTGACAACCCAGTGTTGCCATTTAACGATGGCGTGTATTTCGCCAGTACTGAAAAG
241************************************************************

301AGCAATATCATCAGGGGCTGGATCTTCGGTACTACCCTGGATAGCAAGACCCAGTCTTTG
301AGCAATATCATCAGGGGCTGGATCTTCGGTACTACCCTGGATAGCAAGACCCAGTCTTTG
301************************************************************

361CTGATTGTGAACAATGCTACAAATGTGGTAATCAAAGTTTGTGAGTTTCAATTCTGCAAT
361CTGATTGTGAACAATGCTACAAATGTGGTAATCAAAGTTTGTGAGTTTCAATTCTGCAAT
361************************************************************

421GACCCATTCCTGGGCGTCTATTACCACAAGAATAACAAAAGTTGGATGGAATCCGAGTTT
421GACCCATTCCTGGGCGTCTATTACCACAAGAATAACAAAAGTTGGATGGAATCCGAGTTT
421************************************************************
```

```
481CGGGTGTACAGCTCCGCTAACAACTGCACATTCGAATATGTGTCCCAGCCGTTCCTGATG
481CGGGTGTACAGCTCCGCTAACAACTGCACATTCGAATATGTGTCCCAGCCGTTCCTGATG
481************************************************************

541GACCTTGAGGGGAAGCAGGGAAATTTCAAGAACCTGCGGGAGTTCGTATTTAAAAATATC
541GACCTTGAGGGGAAGCAGGGAAATTTCAAGAACCTGCGGGAGTTCGTATTTAAAAATATC
541************************************************************

601GACGGCTACTTTAAAATCTATAGTAAGCACACCCCCATTAACCTGGTCCGAGACCTTCCG
601GACGGCTACTTTAAAATCTATAGTAAGCACACCCCCATTAACCTGGTCCGAGACCTTCCG
601************************************************************

661CAGGGATTTTCCGCTCTGGAGCCTCTGGTTGACCTCCCTATAGGCATTAACATCACCAGA
661CAGGGATTTTCCGCTCTGGAGCCTCTGGTTGACCTCCCTATAGGCATTAACATCACCAGA
661************************************************************

721TTCCAGACCCTCCTGGCCCTGCACCGCTCATATTTGACCCCCGGGGACTCAAGTAGTGGA
721TTCCAGACCCTCCTGGCCCTGCACCGCTCATATTTGACCCCCGGGGACTCAAGTAGTGGA
721************************************************************

781TGGACAGCCGGGGCTGCTGCTTATTATGTAGGCTATCTGCAGCCTCGGACATTCCTGCTG
781TGGACAGCCGGGGCTGCTGCTTATTATGTAGGCTATCTGCAGCCTCGGACATTCCTGCTG
781************************************************************

841AAGTACAACGAAAACGGGACTATCACTGATGCCGTCGACTGCGCCCTTGACCCTCTTTCT
841AAGTACAACGAAAACGGGACTATCACTGATGCCGTCGACTGCGCCCTTGACCCTCTTTCT
841************************************************************
901GAGACCAAGTGTACACTTAAATCATTCACCGTGGAGAAAGGTATCTACCAGACCTCTAAT
901GAGACCAAGTGTACACTTAAATCATTCACCGTGGAGAAAGGTATCTACCAGACCTCTAAT
901************************************************************

961TTCCGAGTCCAGCCTACAGAGTCCATCGTGAGGTTTCCGAACATTACAAATCTTTGTCCC
961TTCCGAGTCCAGCCTACAGAGTCCATCGTGAGGTTTCCGAACATTACAAATCTTTGTCCC
961************************************************************

1021TTTGGCGAAGTGTTTAATGCCACTAGATTTGCAAGTGTGTACGCTTGGAACCGAAAGAGA
1021TTTGGCGAAGTGTTTAATGCCACTAGATTTGCAAGTGTGTACGCTTGGAACCGAAAGAGA
1021************************************************************

1081ATCTCCAACTGTGTCGCTGACTATTCTGTGCTCTACAACAGCGCTAGCTTTAGCACATTT
1081ATCTCCAACTGTGTCGCTGACTATTCTGTGCTCTACAACAGCGCTAGCTTTAGCACATTT
1081************************************************************

1141AAGTGTTACGGAGTTTCCCCCACCAAGCTGAATGACCTGTGCTTTACCAACGTCTATGCA
1141AAGTGTTACGGAGTTTCCCCCACCAAGCTGAATGACCTGTGCTTTACCAACGTCTATGCA
1141************************************************************

1201GACTCCTTTGTGATTAGAGGCGATGAAGTGCGACAAATTGCACCTGGCCAAACGGGCAAA
1201GACTCCTTTGTGATTAGAGGCGATGAAGTGCGACAAATTGCACCTGGCCAAACGGGCAAA
1201************************************************************

1261ATTGCTGATTATAACTACAAGCTGCCCGACGATTTCACAGGCTGCGTTATCGCCTGGAAC
1261ATTGCTGATTATAACTACAAGCTGCCCGACGATTTCACAGGCTGCGTTATCGCCTGGAAC
1261************************************************************
```

```
1321AGCAACAACCTGGACAGTAAGGTTGGCGGAAATTATAACTACCTGTATCGGCTGTTCAGG
1321AGCAACAACCTGGACAGTAAGGTTGGCGGAAATTATAACTACCTGTATCGGCTGTTCAGG
1321************************************************************

1381AAGAGCAATCTTAAGCCCTTTGAAAGAGACATTAGTACCGAAATATACCAGGCCGGCTCT
1381AAGAGCAATCTTAAGCCCTTTGAAAGAGACATTAGTACCGAAATATACCAGGCCGGCTCT
1381************************************************************

1441ACCCCATGCAATGGCGTTGAGGGCTTCAATTGCTATTTTCCCCTTCAGTCATATGGGTTT
1441ACCCCATGCAATGGCGTTGAGGGCTTCAATTGCTATTTTCCCCTTCAGTCATATGGGTTT
1441************************************************************

1501CAGCCAACCAATGGGGTGGGGTACCAGCCTTATCGCGTAGTGGTACTGAGCTTTGAGCTG
1501CAGCCAACCAATGGGGTGGGGTACCAGCCTTATCGCGTAGTGGTACTGAGCTTTGAGCTG
1501************************************************************

1561TTGCACGCACCAGCAACAGTGTGCGGGCCGAAGAAATCTACAAATCTCGTGAAGAACAAA
1561TTGCACGCACCAGCAACAGTGTGCGGGCCGAAGAAATCTACAAATCTCGTGAAGAACAAA
1561************************************************************

1621TGTGTCAATTTTAACTTTAACGGTCTGACCGGCACGGGGGTCCTGACAGAATCTAACAAG
1621TGTGTCAATTTTAACTTTAACGGTCTGACCGGCACGGGGGTCCTGACAGAATCTAACAAG
1621************************************************************

1681AAATTTCTCCCCTTTCAGCAGTTCGGAAGAGATATCGCAGATACAACTGATGCTGTGCGC
1681AAATTTCTCCCCTTTCAGCAGTTCGGAAGAGATATCGCAGATACAACTGATGCTGTGCGC
1681************************************************************

1741GATCCTCAGACGCTTGAGATTTTGGATATTACACCTTGCTCTTTCGGGGGGGTCAGCGTC
1741GATCCTCAGACGCTTGAGATTTTGGATATTACACCTTGCTCTTTCGGGGGGGTCAGCGTC
1741************************************************************
1801ATCACTCCCGGAACAAATACAAGCAACCAGGTGGCAGTCCTGTACCAGGACGTAAACTGC
1801ATCACTCCCGGAACAAATACAAGCAACCAGGTGGCAGTCCTGTACCAGGACGTAAACTGC
1801************************************************************

1861ACCGAAGTGCCCGTTGCAATTCATGCTGACCAGCTTACTCCCACATGGAGGGTGTACTCC
1861ACCGAAGTGCCCGTTGCAATTCATGCTGACCAGCTTACTCCCACATGGAGGGTGTACTCC
1861************************************************************

1921ACCGGATCTAATGTCTTTCAGACAAGAGCTGGCTGCCTGATCGGTGCCGAACATGTGAAT
1921ACCGGATCTAATGTCTTTCAGACAAGAGCTGGCTGCCTGATCGGTGCCGAACATGTGAAT
1921************************************************************

1981AATAGCTACGAGTGTGACATTCCCATCGGAGCCGGCATTTGTGCATCCTACCAAACTCAG
1981AATAGCTACGAGTGTGACATTCCCATCGGAGCCGGCATTTGTGCATCCTACCAAACTCAG
1981************************************************************

2041ACGAACAGCCCCGGCAGCGCCAGCAGTGTGGCCTCCCAGAGCATTATCGCATACACTATG
2041ACGAACAGCCCCGGCAGCGCCAGCAGTGTGGCCTCCCAGAGCATTATCGCATACACTATG
2041************************************************************


2101TCTCTGGGTGCTGAAAACTCCGTGGCATATTCAAACAACTCAATCGCCATTCCAACTAAC
2101TCTCTGGGTGCTGAAAACTCCGTGGCATATTCAAACAACTCAATCGCCATTCCAACTAAC
2101************************************************************
```

```
2161TTTACCATTTCAGTCACAACCGAAATCCTGCCAGTGTCTATGACCAAGACAAGCGTGGAT
2161TTTACCATTTCAGTCACAACCGAAATCCTGCCAGTGTCTATGACCAAGACAAGCGTGGAT
2161************************************************************

2221TGCACTATGTATATTTGCGGCGACTCCACAGAATGCTCTAATTTGCTTCTGCAGTACGGC
2221TGCACTATGTATATTTGCGGCGACTCCACAGAATGCTCTAATTTGCTTCTGCAGTACGGC
2221************************************************************

2281TCCTTTTGCACTCAATTGAATAGGGCATTGACTGGCATTGCTGTGGAGCAAGATAAAAAT
2281TCCTTTTGCACTCAATTGAATAGGGCATTGACTGGCATTGCTGTGGAGCAAGATAAAAAT
2281************************************************************

2341ACTCAGGAGGTATTTGCTCAGGTGAAACAAATCTACAAAACTCCACCAATTAAAGACTTC
2341ACTCAGGAGGTATTTGCTCAGGTGAAACAAATCTACAAAACTCCACCAATTAAAGACTTC
2341************************************************************

2401GGAGGTTTTAATTTCTCACAGATTCTGCCTGATCCATCAAAGCCTTCTAAGCGGAGTCCC
2401GGAGGTTTTAATTTCTCACAGATTCTGCCTGATCCATCAAAGCCTTCTAAGCGGAGTCCC
2401************************************************************

2461ATCGAAGACTTGCTCTTTAACAAAGTCACTCTGGCTGACGCCGGCTTTATCAAGCAGTAC
2461ATCGAAGACTTGCTCTTTAACAAAGTCACTCTGGCTGACGCCGGCTTTATCAAGCAGTAC
2461************************************************************

2521GGTGACTGCTTGGGCGACATCGCTGCTCGCGACCTGATCTGCGCTCAGAAGTTCAATGGC
2521GGTGACTGCTTGGGCGACATCGCTGCTCGCGACCTGATCTGCGCTCAGAAGTTCAATGGC
2521************************************************************

2581CTGACCGTCCTGCCACCACTGCTGACAGACGAAATGATTGCCCAGTACACCTCTGCATTG
2581CTGACCGTCCTGCCACCACTGCTGACAGACGAAATGATTGCCCAGTACACCTCTGCATTG
2581************************************************************

2641CTTGCCGGGACCATCACATCTGGGTGGACCTTCGGAGCTGGACCCGCTCTCCAGATCCCC
2641CTTGCCGGGACCATCACATCTGGGTGGACCTTCGGAGCTGGACCCGCTCTCCAGATCCCC
2641************************************************************

2701TTCCCCATGCAAATGGCCTATCGCTTCAATGGCATCGGGGTCACCCAGAATGTGCTGTAC
2701TTCCCCATGCAAATGGCCTATCGCTTCAATGGCATCGGGGTCACCCAGAATGTGCTGTAC
2701************************************************************

2761GAAAACCAGAAACTTATTGCCAATCAGTTTAATAGCGCGATCGGCAAGATTCAGGACAGT
2761GAAAACCAGAAACTTATTGCCAATCAGTTTAATAGCGCGATCGGCAAGATTCAGGACAGT
2761************************************************************

2821CTGTCTTCTACTCCCTCTGCCCTGGGAAAGTTGCAGGATGTGGTGAATCAGAATGCCCAG
2821CTGTCTTCTACTCCCTCTGCCCTGGGAAAGTTGCAGGATGTGGTGAATCAGAATGCCCAG
2821************************************************************

2881GCTCTGAATACACTGGTCAAGCAACTCAGCAGCAACTTCGGCGCAATTTCTTCCGTTCTG
2881GCTCTGAATACACTGGTCAAGCAACTCAGCAGCAACTTCGGCGCAATTTCTTCCGTTCTG
2881************************************************************
```

```
2941AACGATATCCTGTCTAGACTCGACCCACCAGAAGCTGAGGTACAGATTGATAGGCTTATC
2941AACGATATCCTGTCTAGACTCGACCCACCAGAAGCTGAGGTACAGATTGATAGGCTTATC
2941************************************************************


3001ACAGGCCGGCTCCAATCACTGCAAACATACGTGACTCAGCAGCTGATCCGGGCTGCCGAA
3001ACAGGCCGGCTCCAATCACTGCAAACATACGTGACTCAGCAGCTGATCCGGGCTGCCGAA
3001************************************************************

3061ATTAGGGCCAGTGCAAACCTGGCTGCGACTAAGATGTCTGAGTGCGTTCTGGGTCAGTCC
3061ATTAGGGCCAGTGCAAACCTGGCTGCGACTAAGATGTCTGAGTGCGTTCTGGGTCAGTCC
3061************************************************************

3121AAGAGGGTAGATTTTTGTGGCAAGGGGTACCATCTCATGTCTTTTCCTCAGAGTGCACCA
3121AAGAGGGTAGATTTTTGTGGCAAGGGGTACCATCTCATGTCTTTTCCTCAGAGTGCACCA
3121************************************************************

3181CATGGAGTGGTCTTCCTCCACGTGACCTACGTGCCCGCCCAAGAGAAAAATTTTACAACC
3181CATGGAGTGGTCTTCCTCCACGTGACCTACGTGCCCGCCCAAGAGAAAAATTTTACAACC
3181************************************************************

3241GCTCCAGCTATCTGCCACGATGGAAAAGCTCATTTCCCGCGGGAGGGCGTGTTCGTCAGT
3241GCTCCAGCTATCTGCCACGATGGAAAAGCTCATTTCCCGCGGGAGGGCGTGTTCGTCAGT
3241************************************************************

3301AATGGAACGCACTGGTTCGTGACCCAGAGAAACTTTTACGAGCCCCAGATAATTACCACC
3301AATGGAACGCACTGGTTCGTGACCCAGAGAAACTTTTACGAGCCCCAGATAATTACCACC
3301************************************************************

3361GATAATACATTTGTTAGCGGCAACTGCGATGTCGTGATCGGGATAGTGAATAACACAGTG
3361GATAATACATTTGTTAGCGGCAACTGCGATGTCGTGATCGGGATAGTGAATAACACAGTG
3361************************************************************

3421TACGACCCACTCCAGCCCGAACTCGACTCATTTAAGGAAGAGCTTGACAAGTACTTTAAG
3421TACGACCCACTCCAGCCCGAACTCGACTCATTTAAGGAAGAGCTTGACAAGTACTTTAAG
3421************************************************************

3481AATCACACCAGCCCAGATGTGGACCTGGGAGATATCAGCGGCATTAATGCCTCAGTGGTG
3481AATCACACCAGCCCAGATGTGGACCTGGGAGATATCAGCGGCATTAATGCCTCAGTGGTG
3481************************************************************

3541AATATTCAGAAGGAAATAGATAGACTGAATGAAGTGGCAAAGAATCTGAACGAGTCCCTC
3541AATATTCAGAAGGAAATAGATAGACTGAATGAAGTGGCAAAGAATCTGAACGAGTCCCTC
3541************************************************************

3601ATTGATCTGCAGGAGCTCGGCAAGTACGAGCAAGAAAACCTGTATTTTCAGAGCGGATCT
3601ATTGATCTGCAGGAGCTCGGCAAGTACGAGCAAGAAAACCTGTATTTTCAGAGCGGATCT
3601************************************************************

3661GGGTACATCCCCGAGGCACCAAGAGATGGTCAGGCATATGTTAGGAAGGATGGGGAATGG
3661GGGTACATCCCCGAGGCACCAAGAGATGGTCAGGCATATGTTAGGAAGGATGGGGAATGG
3661************************************************************
```

```
3721GTCCTCCTGAGCACATTTTTGGGAGGAAGCCTGGTACCCAGAGGCTCTCCACACCATCAT
3721GTCCTCCTGAGCACATTTTTGGGAGGAAGCCTGGTACCCAGAGGCTCTCCACACCATCAT
3721************************************************************

3781CACCACCACTAA
3781CACCACCACTAA
3781************
```