(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
07 October 2021 (07.10.2021)

WIPO | PCT

(10) International Publication Number
WO 2021/198376 A1

(51) International Patent Classification:
A61K 39/12 (2006.01) *A61P 31/14* (2006.01)

(21) International Application Number:
PCT/EP2021/058513

(22) International Filing Date:
31 March 2021 (31.03.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
20167405.8 31 March 2020 (31.03.2020) EP

(71) Applicant: **VAXIMM AG** [CH/CH]; Hochbergerstrasse 60c, 4057 Basel (CH).

(72) Inventors: **LUBENAU, Heinz**; An der Eselshaut 12, 67435 Neustadt an der Weinstrasse (DE). **MANSOUR, Marc**; 42 Stone gate Drive, Halifax, Nova Scotia B3N3J2 (CA).

(74) Agent: **WALLINGER RICKER SCHLOTTER TOSTMANN**; Patent- und Rechtsanwälte, Partnerschaft mbB, Zweibrückenstraße 5-7, 80331 Munich (DE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

WO 2021/198376 A1

(54) Title: NOVEL SALMONELLA-BASED CORONAVIRUS VACCINE

(57) Abstract: The present invention relates to a DNA vaccine comprising a *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding at least a COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof. In particular, the present invention relates to the DNA vaccine for use in the prevention and/ or the treatment of coronavirus disease 2019 (COVID-19) or a SARS-CoV-2 infection.

NOVEL SALMONELLA-BASED CORONAVIRUS VACCINE

FIELD OF THE INVENTION

[0001] The present invention relates to a DNA vaccine comprising a *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding at least a COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof. In particular, the present invention relates to said DNA vaccine for use in the prevention and/ or the treatment of coronavirus disease 2019 (COVID-19) or a SARS-CoV-2 infection.

BACKGROUND OF THE INVENTION

[0002] At the end of December 2019, Chinese public health authorities reported several cases of acute respiratory syndrome in Wuhan City, Hubei province, China. Chinese scientists soon identified a novel coronavirus as the main causative agent. The disease is now referred to as coronavirus disease 2019 (COVID-19), and the causative virus is called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). It is a new strain of coronavirus that has not been previously identified in humans.

[0003] The initial outbreak in Wuhan spread rapidly, affecting other parts of China. Cases were soon detected in several other countries. Outbreaks and clusters of the disease have since been observed in Asia, Europe, Australia, Africa and America.

[0004] The WHO in its first emergency meeting estimated the fatality rate of COVID-19 to be around 4%. Although the fatality rate seems to vary between countries and may not be accurate due to an unknown number of unreported cases the spread of SARS-CoV-2 (originally referred to as 2019 novel Coronavirus (2019-nCoV)) has become a worldwide thread and treatment of and/or vaccination against COVID-19 is desperately needed to stop further spreading of the virus.

[0005] Coronaviruses are positive-sense single-stranded RNA viruses belonging to the family Coronaviridae. These viruses mostly infect animals, including birds and mammals. In

humans, coronaviruses typically cause mild respiratory infections. Since 2003 two highly pathogenic human Coronaviruses, Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) and Middle East Respiratory Syndrome Coronavirus (MERS-CoV), have led to global epidemics with high morbidity and mortality. Both endemics were caused by zoonotic coronaviruses that belong to the genus Betacoronavirus within Coronaviridae.

[0006] Like SARS-CoV and MERS-CoV, the new SARS-CoV-2 belongs to the Betacoronavirus genus. As reported by Zhou et al. (Cell Discovery (2020) 6:14) SARS-CoV-2 shares the highest nucleotide sequence identity with SARS-CoV (79.7%). Specifically, the envelope and nucleocapsid proteins of SARS-CoV-2 are two evolutionarily conserved regions, with sequence identities of 96% and 89.6%, respectively, compared to SARS-CoV. The spike protein was reported to exhibit the lowest sequence conservation (sequence identity of 77%) between SARS-CoV-2 and SARS-CoV, while the spike protein of SARS-CoV-2 only has 31.9% sequence identity with the spike protein of MERS-CoV.

[0007] Various reports relating to SARS-CoV suggest a protective role of both humoral and cell-mediated immune responses. The S protein is the most exposed protein and antibody responses against the SARS-CoV S protein have been shown to protect from SARS-CoV infection in a mouse model. While being effective antibody responses may be short-lived. In contrast, T cell responses have been shown to provide long-term protection against SARS-CoV. Thus, vaccines capable of eliciting humoral as well as cell-mediated immune responses are most promising.

[0008] Several national and international research groups are working on the development of vaccines to prevent and treat the 2019-nCoV/SARS-CoV-2, but effective vaccines are not available yet. Thus, there remains an imminent need for an effective therapeutic and/or prophylactic vaccine that can be developed and approved in a short period of time.

SUMMARY OF THE INVENTION

[0009] In view of the current understanding of the novel corona virus and the worldwide epidemic caused by SARS-CoV-2, it is an object of the present invention to provide a novel oral DNA vaccine for prevention and/or the treatment of coronavirus disease 2019 (COVID-

19) or a SARS-CoV-2 infection. The DNA vaccine according to the present invention comprises a *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding at least a COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof. This vaccine is based on a live attenuated Salmonella typhi strain referred to as *Salmonella typhi* Ty21a that serves as a carrier and adjuvant for the DNA molecule encoding the immunogenic antigen for expression within the host cells. This Salmonella-based carrier comprising the DNA molecule encoding the antigen can be developed and produced in a short period of time at large scale and may be adapted to potential mutations occurring in the virus if required.

[0010] Furthermore, the live, attenuated *S. typhi* Ty21a strain used as a carrier is the active component of Typhoral L®, also known as Vivotif® (manufactured by Berna Biotech Ltd., a Crucell Company, Switzerland), the only licensed live oral vaccine against typhoid fever. This vaccine has been extensively tested and has proved to be safe regarding patient toxicity as well as transmission to third parties (Wahdan et al., J. Infectious Diseases 1982, 145:292-295). The vaccine is licensed in more than 40 countries and has been used in millions of individuals including thousands of children for prophylactic vaccination against typhoid fever. It has an unparalleled safety track record. The carrier used in the DNA vaccine of the present invention is therefore suited for getting approval and the product on the market in a short period of time.

[0011] The DNA vaccine according to the present invention therefore has several advantages that makes it particularly suitable for the challenge of providing an effective vaccine against COVID-19 and/or SARS-CoV-2 infection.

[0012] Provided herein is a DNA vaccine comprising a *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding at least a COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof. In certain embodiments the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof comprises (a) a SARS-CoV-2 full-length S protein; (b) a SARS-CoV-2 S protein ectodomain; (c) a SARS-CoV-2 S protein subunit S1; (d) a SARS-CoV-2 S protein receptor binding domain (RBD); or (d) at least 3 immune-dominant epitopes of SARS-CoV-2 S protein.

[0013] In one embodiment the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein is a SARS-CoV-2 full-length S protein. The SARS-CoV-2 full-length S protein may comprise an amino acid sequence of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with SEQ ID NO: 1. The SARS-CoV-2 full-length S protein may also be the full-length S protein of a variant of SARS-CoV-2, such as lineage B.1.1.7, B.1.351 or B.1.1.28 (renamed P.1). The SARS-CoV-2 full-length S protein may also be a prefusion-stabilized form of the SARS-CoV-2 full-length S protein, such as comprising two or more stabilizing mutations. In one embodiment the prefusion-stabilized form of the SARS-CoV-2 full-length S protein comprises two stabilizing mutations to proline corresponding to amino acid position K986 and V987 in the amino acid sequence of SEQ ID NO: 1.

[0014] In certain embodiments the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof comprises the SARS-CoV-2 S protein ectodomain. The SARS-CoV-2 S protein ectodomain has an amino acid sequence of amino acid residues 1-1208 of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with amino acid residues 1-1208 of SEQ ID NO: 1. The SARS-CoV-2 S protein ectodomain may also be the S protein ectodomain of a variant of SARS-CoV-2, such as lineage B.1.1.7, B.1.351 or P.1. The SARS-CoV-2 S protein or a portion thereof may also comprise a prefusion-stabilized form of the SARS-CoV-2 S protein ectodomain comprising two or more stabilizing mutations. In one embodiment the prefusion-stabilized form of the SARS-CoV-2 S protein ectodomain comprises two stabilizing mutations to proline corresponding to amino acid position K986 and V987 in the amino acid sequence of amino acid residues 1 to 1208 of SEQ ID NO: 1.

[0015] In certain embodiments the SARS-CoV-2 S protein or a portion thereof has an amino acid sequence of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with SEQ ID NO: 1, comprising two stabilizing mutations K986P and V987P. In certain alternative embodiments SARS-CoV-2 S protein or a portion thereof comprises an amino acid sequence of amino acid residues 1-1208 of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with amino acid residues 1-1208 of SEQ ID NO: 1, comprising two stabilizing mutations K986P and V987P.

[0016] In certain embodiments the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof comprises the SARS-CoV-2 S protein subunit S1. The SARS-CoV-2 protein

subunit S1 may comprise an amino acid sequence of amino acid residues 1-681 of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with amino acid residues 1-681 of SEQ ID NO: 1. The SARS-CoV-2 S protein subunit S1 may also be the S protein subunit S1 of a variant of SARS-CoV-2, such as lineage B.1.1.7, B.1.351 or P.1

[0017] In certain embodiments the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof comprises the SARS-CoV-2 S protein receptor binding domain (RBD). The SARS-CoV-2 protein RBD may comprise an amino acid sequence of amino acid residues 319-541 of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with amino acid residues 319-541 of SEQ ID NO: 1. The SARS-CoV-2 S protein RBD may also be the S protein RBD of a variant of SARS-CoV-2, such as lineage B.1.1.7, B.1.351 or P.1

[0018] The DNA vaccine according to the invention may comprise a DNA molecule encoding the SARS-CoV-2 S protein or a portion thereof and optionally further encoding another SARS-CoV-2 protein or a portion thereof, preferably a SARS-CoV-2 N protein. In certain embodiments the eukaryotic expression cassette encodes the SARS-CoV-2 S protein or a portion thereof and further encodes another SARS-CoV-2 protein or a portion thereof, such as a SARS-CoV-2 N protein or a portion thereof.

[0019] The DNA vaccine according to the invention may further comprise one or more pharmaceutically acceptable excipients. In certain embodiments the DNA vaccine is an oral dosage form, such as an enteric coated capsule, a lyophilized powder or a suspension. The DNA vaccine according to the invention may further comprising one or more adjuvants.

[0020] Also provided herein is the DNA vaccine according to the invention for use in the treatment and/or the prevention of coronavirus disease 2019 (COVID-19) or a SARS-CoV-2 infection.

[0021] Also provided herein is a method for treating and/or preventing coronavirus disease 2019 (COVID-19) or a SARS-CoV-2 infection comprising administering the DNA vaccine according to the invention to a patient in need thereof. In preferred embodiments the DNA vaccine is administered orally. In certain embodiments a single dose of the DNA vaccine comprises the *Salmonella typhi* Ty21a strain at about 1×10^6 to about 1×10^9 colony forming

units (CFU), and/or the DNA vaccine is to be administered 2 to 4 times in one week for priming, optionally followed by at least one boosting dose. In one embodiment the DNA vaccine is to be administered 2 to 4 times within the first week, followed by one or more single dose boosting each at least 2 weeks later, preferably each at least 4 weeks later.

SHORT DESCRIPTION OF THE FIGURES

Figure 1: Amino acid sequence of SARS-CoV-2 Spike protein (SEQ ID NO: 1) with amino acid residues 1-1208 marked as underlined and residues K986, V987, R682G, R683S and R685S in bold.

Figure 2: Plasmid map of pVAX10.SCV-1

Figure 3: SARS-CoV-2 constructs for cloning into pVAX10, with the X indicating the presence of the domain in the order from N-terminal (left) to C-terminal (right), The following abbreviations are used; S FL (full-length S protein, SEQ ID NO: 1; *indicates signal domain (Met1-SER12 of SEQ ID NO: 1) replaced with that of invariant chain (Met1-Arg29 of SEQ ID NO: 15)), S ecto: (S protein ectodomain), S1 (S protein S1 subunit), RBD (receptor binding domain), T4 trimer (T4 fibritin trimerization motif), 3C3d (enhancer sequence comprising three copies of the C3d protein), 2A (2A peptide, such as T2a or P2a), Ubi. (ubiquitin), N (N protein), S2 (S protein S2 subunit) and SV40 DTS (SV40 DNA nuclear targeting sequence).

Figure 4: Immune responses elicited by VXM-SCV-3 in healthy mice. The serum of vaccinated mice was analysed for antibodies against SARS-CoV spike protein (see Example 5). The assay background lies at 400 endpoint titer, as indicated by the dotted straight line.

Figure 5: Immune responses elicited by VXM-SCV-30 in healthy mice. The serum of vaccinated mice was analysed for antibodies towards SARS-CoV spike protein (see Example 6). The assay background lies at 400 endpoint titer, as indicated by the straight line.

Figure 6: Immune responses elicited by VXM-SCV-42 in healthy mice. The serum of vaccinated mice was analysed for antibodies towards SARS-CoV spike protein (see Example 7). The assay background lies at 400 endpoint titer, as indicated by the dotted straight line.

Figure 7: Immune responses elicited by VXM-SCV-53 in healthy mice. The serum of vaccinated mice was analysed for antibodies towards SARS-CoV spike protein (see Example 8). The assay background lies at 400 endpoint titer, as indicated by the dotted straight line.

DETAILED DESCRIPTION OF THE INVENTION

[0022] Provided herein is a DNA vaccine comprising a *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding at least a COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof.

[0023] According to the invention, the *Salmonella typhi* Ty21a strain functions as the bacterial carrier of the DNA molecule comprising a eukaryotic expression cassette encoding at least a COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof for the delivery of said DNA molecule into a target cell. Thus, the DNA molecule is delivered to a host cell and the S protein or a portion thereof is expressed by the host cell. The strain *Salmonella typhi* Ty21a is an attenuated *Salmonella typhi* strain and the DNA vaccine according to the invention comprises the live attenuated *Salmonella typhi* strain *Salmonella typhi* Ty21a.

[0024] In the context of the present invention, the term “attenuated” refers to a bacterial strain of reduced virulence compared to the parental bacterial strain, not harboring the attenuating mutation. Attenuated bacterial strains have preferably lost their virulence but retained their ability to induce protective immunity. Attenuation can be accomplished by deletion of various genes, including virulence, regulatory, and metabolic genes. Attenuated bacteria may be found naturally or they may be produced artificially in the laboratory, for example by adaptation to a new medium or cell culture or they may be produced by recombinant DNA technology. Administration of about 10^{11} CFU of the attenuated strain of Salmonella according to the present invention preferably causes Salmonellosis in less than 5%, more preferably less than 1%, most preferably less than 1‰ of subjects.

[0025] The term “comprises” or “comprising” means “including, but not limited to”. The term is intended to be open-ended, to specify the presence of any stated features, elements, integers, steps or components, but not to preclude the presence or addition of one or more

other features, elements, integers, steps, components or groups thereof. The term "comprising" thus includes the more restrictive terms "consisting of" and "essentially consisting of". In one embodiment the term "comprising" may be individually replaced by the term "consisting of". With regard to sequences the terms "having an amino acid sequence of" and "comprising an amino acid of" are used interchangeably and include the embodiment "consisting of the amino acid sequence of". The term "a" as used herein may include the plural and hence includes, but is not limited, to "one".

[0026] The term "SARS-CoV-2 S protein or a portion thereof" or "another SARS-CoV-2 protein or a portion thereof" as used herein refers to the SARS-CoV-2 S protein or an immunogenic portion thereof or another SARS-CoV-2 protein and an immunogenic portion thereof. An immunogenic portion of a protein may comprise one or more domain(s) of the immunogenic protein. However, it is also encompassed by the present invention that the immunogenic portion comprises only the immunogenic part of a domain, such as the receptor binding domain or the ectodomain. The term "immunogenic" as used herein refers to a part of protein that elicits an immune response, such as a B cell and/or T cell response.

[0027] A DNA molecule comprising at least one eukaryotic expression cassette may also be referred to as a recombinant DNA molecule, i.e. an engineered DNA construct, preferably composed of DNA pieces of different origin. The DNA molecule can be a linear nucleic acid or a circular nucleic acid. Preferably the DNA molecule is a plasmid, more preferably an expression plasmid. The plasmid may be generated by introducing an open reading frame encoding at least the SARS-CoV-2 S protein or a portion thereof into a eukaryotic expression cassette of a plasmid. A plasmid comprising a eukaryotic expression cassette may also be referred to as eukaryotic expression plasmid.

[0028] In the context of the present invention, the term "expression cassette" refers to a nucleic acid unit comprising at least one open reading frame (ORF) under the control of regulatory sequences controlling its expression. Preferably the expression cassette also comprises a transcription termination signal. Expression cassettes can preferably mediate transcription of the included open reading frame encoding at least the SARS-CoV-2 S protein or a portion thereof in a target cell. Eukaryotic expression cassettes

typically comprise a promoter, at least one open reading frame and a transcription termination signal, which allow expression in a eukaryotic target cell.

[0029] Coronaviruses are positive-sense single-stranded RNA viruses belonging to the family Coronaviridae. These viruses mostly infect animals, including birds and mammals. In humans, coronaviruses typically cause mild respiratory infections. Since 2003 two highly pathogenic human Coronaviruses including Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) and Middle East Respiratory Syndrome Coronavirus (MERS-CoV) have led to global epidemics with high morbidity and mortality. Both endemics were caused by zoonotic coronaviruses that belong to the genus Betacoronavirus within Coronaviridae.

[0030] Like SARS-CoV and MERS-CoV, the new coronavirus SARS-CoV-2 belongs to the Betacoronavirus genus. The genome of SARS-CoV-2 has about 30 kilobase and encodes for multiple structural and non-structural proteins. The structural proteins include the spike (S) protein, the envelope (E) protein, the membrane (M) protein, and the nucleocapsid (N) protein. As reported by Zhou et al. (Cell Discovery (2020) 6:14) SARS-CoV-2 shares the highest nucleotide sequence identity with SARS-CoV (79.7%). Specifically, the envelope and nucleocapsid proteins of SARS-CoV-2 are two evolutionarily conserved regions, with sequence identities of 96% and 89.6%, respectively, compared to SARS-CoV. The spike protein was reported to exhibit the lowest sequence conservation (sequence identity of 77%) between SARS-CoV-2 and SARS-CoV, while the spike protein of SARS-CoV-2 only has 31.9% sequence identity with the spike protein of MERS-CoV. Several non-structural proteins were predicted for SARS-CoV-2 which are coded for by the open reading frames ORF 1ab, ORF 3a, ORF3b, ORF6, ORF 7a, ORF7b, ORF8, ORF9a, ORF9b, and ORF10 (Srinivasan et al. Viruses (2020) 12:360). In the meantime, several variants of SARS-CoV-2 were identified, for instance, the SARS-CoV-2 lineage B.1.1.7 first reported in the UK, the B.1.351 lineage first reported in South Africa and the B.1.1.28 subclade first reported in Brazil which was renamed as P.1 (Galloway et al., MMWR Morb Mortal Wkly Rep. 2021 Jan 22; 70(3): 95–99.). According to Galloway et al. these variants carry a constellation of genetic mutations, including in the S protein receptor-binding domain, which is essential for binding to the host cell angiotensin-converting enzyme-2 (ACE-2) receptor to facilitate virus entry. It seems that these variants spread more efficiently.

[0031] Various reports related to SARS-CoV suggest a protective role of both humoral and cell-mediated immune response. The S protein is the most exposed protein and antibody responses against the SARS-CoV S protein have been shown to protect from SARS-CoV infection in a mouse model. While being effective antibody responses may be short-lived. In contrast, T cell responses have been shown to provide long-term protection. In addition, multiple studies have shown that antibodies are generated against the N protein of SARS-CoV and by extension to SARS-CoV-2, the N protein is considered to be a highly immunogenic and abundantly expressed protein during infection. Further, of the structural proteins, T cell responses against the S and N proteins have been reported to be the most dominant and long-lasting (Ahmed et al. Viruses (2020) 12:254). The attenuated strain of *Salmonella*, *Salmonella typhi* Ty21a, is of the species *Salmonella enterica*. Attenuated derivatives of *Salmonella enterica* are attractive vehicles for the delivery of heterologous antigens to the mammalian immune system, since *S. enterica* strains can potentially be delivered via mucosal routes of immunization, i.e. orally or nasally, which offers advantages of simplicity and safety compared to parenteral administration. Furthermore, *Salmonella* strains elicit strong humoral and cellular immune responses at the level of both systemic and mucosal compartments. Batch preparation costs are low and formulations of live bacterial vaccines are highly stable. Attenuation can be accomplished by deletion of various genes, including virulence, regulatory, and metabolic genes.

[0032] Several *Salmonella typhimurium* strains attenuated by *aro* mutations have been shown to be safe and effective delivery vehicles for heterologous antigens in animal models.

[0033] The attenuated strain *Salmonella typhi* Ty21a has been shown to be safe and effective as a vaccine against typhoid fever and as a delivery vehicle for heterologous antigens for vaccination in humans, primarily for vaccination against tumor antigens and/or stroma antigens.

[0034] The live, attenuated *S. typhi* Ty21a strain is the active component of Typhoral L®, also known as Vivotif® (manufactured by Berna Biotech Ltd., a Crucell Company, Switzerland). It is currently the only licensed live oral vaccine against typhoid fever. This vaccine has been extensively tested and has proved to be safe regarding patient toxicity as well as transmission to third parties (Wahdan et al., J. Infectious Diseases 1982, 145:292-295). The

vaccine is licensed in more than 40 countries and has been used in millions of individuals including thousands of children for prophylactic vaccination against typhoid fever. The Marketing Authorization number of Typhoral L® is PL 15747/0001 dated 16 December 1996. One dose of vaccine contains at least 2 x 10^9 viable *S. typhi* Ty21a colony forming units and at least 5 x 10^9 non-viable S. typhi Ty21a cells.

[0035] This well-tolerated, live oral vaccine against typhoid fever was derived by chemical mutagenesis of the wild-type virulent bacterial isolate *S. typhi* Ty2 and harbors a loss-of-function mutation in the *gal*E gene resulting in its inability to metabolize galactose. The attenuated bacterial strain is also not able to reduce sulfate to sulfide which differentiates it from the wild-type *Salmonella typhi* Ty2 strain. With regard to its serological characteristics, the *Salmonella typhi* Ty21a strain contains the O9-antigen which is a polysaccharide of the outer membrane of the bacteria and lacks the O5-antigen which is in turn a characteristic component of *Salmonella typhimurium*. This serological characteristic supports the rationale for including the respective test in a panel of identity tests for batch release.

[0036] SARS-CoV-2 S protein is a glycoprotein with 66 N-linked glycosylation sites per trimer. The protein also comprises O-linked glycans at residues S673, T678 and S686. Furthermore, the S protein contains two functional domains: a receptor binding domain, and a second domain which contains sequences that mediate fusion of the viral and cell membranes. The S glycoprotein must be cleaved by cell proteases to enable exposure of the fusion sequences and hence is needed for cell entry. Protein sequence of the S glycoprotein of SARS-CoV-2 reveals the presence of a furin cleavage sequence (PRRARS|V) at residues 681-687 due to an insertion of the sequence PRRA. Because furin proteases are abundant in the respiratory tract, it is possible that SARS-CoV-2 S glycoprotein is cleaved upon exit from epithelial cells and consequently can efficiently infect other cells.

[0037] The expression cassette used for the DNA vaccine according to the invention is a eukaryotic expression cassette. In the context of the present invention, the term “eukaryotic expression cassette” refers to an expression cassette which allows for expression of the open reading frame in a eukaryotic cell. It has been shown that the amount of heterologous antigen required to induce an adequate immune response may be toxic for the bacterium and may result in cell death, over-attenuation or loss of expression of the heterologous

antigen. Using a eukaryotic expression cassette that is not expressed in the bacterial vector but only in the target cell may overcome this toxicity problem and the protein expressed typically exhibits a eukaryotic glycosylation pattern.

[0038] A eukaryotic expression cassette comprises regulatory sequences that are able to control the expression of an open reading frame in a eukaryotic cell, preferably a promoter and a polyadenylation signal. Promoters and polyadenylation signals included in the recombinant DNA molecules comprised by the attenuated strain of Salmonella of the present invention are preferably selected to be functional within the cells of the subject to be immunized. Examples of suitable promoters, especially for the production of a DNA vaccine for humans, include but are not limited to promoters from Cytomegalovirus (CMV), such as the strong CMV immediate early promoter, Simian Virus 40 (SV40), Mouse Mammary Tumor Virus (MMTV), Human Immunodeficiency Virus (HIV), such as the HIV Long Terminal Repeat (LTR) promoter, Moloney virus, Epstein Barr Virus (EBV), and from Rous Sarcoma Virus (RSV), the synthetic CAG promoter composed of the CMV early enhancer element, the promoter, the first exon and the first intron of chicken beta-actin gene and the splice acceptor of the rabbit beta globin gene, as well as promoters from human genes such as human actin, human myosin, human hemoglobin, human muscle creatine, and human metallothionein. In a particular embodiment, the eukaryotic expression cassette contains the CMV promoter. In the context of the present invention, the term “CMV promoter” refers to the strong immediate-early cytomegalovirus promoter.

[0039] Examples of suitable polyadenylation signals, especially for the production of a DNA vaccine for humans, include but are not limited to the bovine growth hormone (BGH) polyadenylation site, SV40 polyadenylation signals and LTR polyadenylation signals. In a particular embodiment, the eukaryotic expression cassette included in the recombinant DNA molecule comprised by the attenuated strain of Salmonella of the present invention comprises the BGH polyadenylation site.

[0040] In addition to the regulatory elements required for expression of the heterologous SARS-CoV-2 S protein or a portion thereof, like a promoter and a polyadenylation signal, other elements can also be included in the recombinant DNA molecule. Such additional elements include enhancers. The enhancer can be, for example, the enhancer of human

actin, human myosin, human hemoglobin, human muscle creatine and viral enhancers such as those from CMV, RSV and EBV.

[0041] In the context of the present invention it is generally advantageous to use a gene (or open reading frame) encoding the SARS-CoV-2 S protein or a portion thereof (as well as an optional further SARS-CoV-2 protein or a portion thereof, such as the SARS-CoV-2 N protein or a portion thereof) that it codon-optimized for mammalian expression, particularly for human expression. Thus, in certain embodiments the eukaryotic expression cassette comprises at least a codon-optimized sequence encoding COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof.

[0042] The COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof encoded by the DNA vaccine according to the invention comprises without being limited thereto (a) a SARS-CoV-2 full-length S protein; (b) a SARS-CoV-2 S protein ectodomain; (c) a SARS-CoV-2 protein subunit S1; (d) a SARS-CoV-2 receptor binding domain (RBD) or (e) at least 3 immune-dominant epitopes of SARS-CoV-2 S protein.

[0043] In certain embodiments the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein is a SARS-CoV-2 full-length S protein. The SARS-CoV-2 full-length S protein may comprise an amino acid sequence of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with SEQ ID NO: 1. In a preferred embodiment the SARS-CoV-2 full-length S protein has an amino acid sequence having at least 96%, at least 97%, at least 98%, or at least 99% sequence identity with SEQ ID NO: 1. In one embodiment the SARS-CoV-2 full-length S protein has an amino acid sequence having at least 98% to 100% sequence identity with SEQ ID NO: 1. In a specific embodiment the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein is a SARS-CoV-2 full-length S protein consisting of an amino acid sequence of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with SEQ ID NO: 1. The amino acid sequence of SEQ ID NO: 1 has the GenBank accession number MN_908947 and has been published by Wu et al. (Nature 2020, 579: 265-269). In a specific embodiment the SARS-CoV-2 full-length S protein may also be the full-length S protein of a variant of SARS-CoV-2, such as lineage B.1.1.7, B.1.351 or P.1.

[0044] We compared different S protein sequences of SARS-CoV-2 available at GenBank in an alignment of the sequences of the GenBank accession numbers (protein-id): MN_908947 (QHD434616.1), MN_988668 (QHQ62107.1), NC_045512 (YP_009724390.1), MN_938384.1 (QHN73795.1), MN_975262.1 (QHN73810.1), MN_985325.1 (QHQ60594.1), MN_988713.1 (QHQ62877.1), MN_994467.1 (QHQ71963.1), MN_994468.1 (QHQ71973.1), and MN997409.1 (QHQ82464.1) and found no differences. However, minor variations have previously been reported in the SARS-CoV-2 S protein. For example the following substitutions have been described by Wrapp et al. (Science, 2020, 367: 1260-1263) in clinical isolates F32I, H49Y, S247R, N354D, D364Y, V367F, D614G, V1129L and E1262G. Moreover the substitutions H49Y and V860Q have been reported by Wang et al. (J. Med. Virol. March 13, 2020: 1-8). Further homology analysis of the published SARS-CoV-2 sequences by the same authors revealed a nucleotide homology of the S protein of 99.82% to 100% and an amino acid homology of the S protein of 99.53% to 100%. The identified variants B.1.1.7, B.1.351 and P.1 carry several mutations. The B.1.1.7 variant S protein has the deletions 69-70 HV and 144 Y and the following mutations: N501Y, A570D, D614G, P681H, T761I, S982A, D1118H. The variant B.1.351 carries the following mutations in the S protein: K417N, E484K, N501Y, D614G and A701V. The P.1 variant carries a L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y and T1027I mutation in the S protein (Galloway et al., MMWR Morb Mortal Wkly Rep. 2021 Jan 22; 70(3): 95–99.). However, further substitutions or variants may occur or be identified over time.

[0045] The SARS-CoV-2 full-length S protein may also be a prefusion-stabilized form of the SARS-CoV-2 full-length S protein, such as comprising two or more stabilizing mutations. In certain embodiments the prefusion-stabilized form of the SARS-CoV-2 full-length S protein comprises two stabilizing mutations to proline corresponding to amino acid position K986 and V987 in the amino acid sequence of SEQ ID NO: 1.

[0046] Prefusion-stabilized forms of SARS-CoV-2 S protein have been described by Wrapp et al. (Science, 2020, 367: 1260-1263) by adding two stabilizing proline mutations at residues 986 and 987 in the C-terminal S2 fusion machinery using a previously stabilizing strategy that proved effective for other betacoronavirus S proteins. Furthermore, Wrapp et al. (Science, 2020, 367: 1260-1263) described a “GSAS” mutation in the furin cleavage site at

residues 682-685, replacing the RRAR sequence at this position. Both these mutations stabilize the protein and hence prevent fusion. This may not only improve stability and expression of the S protein, but also improve safety by preventing cell fusion. In certain embodiments the prefusion-stabilized form of the SARS-CoV-2 full-length S protein comprises two stabilizing mutations to proline corresponding to amino acid position K986 and V987 in the amino acid sequence of SEQ ID NO: 1 and/or a mutation of the furin cleavage sequence (PRRARS|V) corresponding to residues 681-687 of SEQ ID NO: 1, such as a R682G, R683S and R685S mutation. Preferably the SARS-CoV-2 full-length S protein has an amino acid sequence of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with SEQ ID NO: 1, further comprising two stabilizing mutations K986P and V987P; or furin cleavage sequence mutations R682G, R683S and R685S, or two stabilizing mutations K986P and V987P and furin cleavage sequence mutations R682G, R683S and R685S. Alternatively amino acids of the furin cleavage sequence may be deleted such as amino acids 680-683. Thus, in one embodiment the SARS-CoV-2 full-length S protein has an amino acid sequence of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with SEQ ID NO: 1, further comprising a deletion in the furin cleavage sequence, such as a deletion comprising or consisting of amino acids S680-R683. Other amino acid substitutions or amino acid deletions resulting in a pre-fusion stabilized form of the S protein may also be employed.

[0047] In certain embodiments the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof comprises the SARS-CoV-2 S protein ectodomain. The term “ectodomain” refers to the extracellular portion of the transmembrane protein SARS-CoV-2 S protein, i.e., lacking the transmembrane domain and the cytoplasmic domain. The ectodomain comprises the membrane distal subunit S1 comprising the receptor binding domain and the membrane proximate subunit S2. The SARS-CoV-2 S protein ectodomain comprises an amino acid sequence of amino acid residues 1-1208 of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with amino acid residues 1-1208 of SEQ ID NO: 1. However the SARS-CoV-2 S protein ectodomain as used herein may be a sequence corresponding at least to amino acid residues 1 to 1208 of SEQ ID NO: 1 or may be slightly longer, such as up to the N-terminal 1213 amino acid residues of SEQ ID NO: 1 or a sequence having at least

95% sequence identity with amino acid residues 1-1213 of SEQ ID NO: 1. In a preferred embodiment the SARS-CoV-2 S protein or a portion thereof comprises the SARS-CoV-2 S protein ectodomain having an amino acid sequence having at least 96%, at least 97%, at least 98%, or at least 99% sequence identity with the sequence of amino acid residues 1-1208 of SEQ ID NO: 1. In one embodiment the SARS-CoV-2 S protein ectodomain has an amino acid sequence having at least 98% to 100% sequence identity with amino acid residues 1-1208 of SEQ ID NO: 1. In a specific embodiments the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof is the SARS-CoV-2 S protein ectodomain having or consisting of an amino acid sequence of amino acid residues 1-1208 of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with amino acid residues 1-1208 of SEQ ID NO: 1. In a further specific embodiment the SARS-CoV-2 S protein ectodomain may also be the S protein ectodomain of a variant of SARS-CoV-2, such as lineage B.1.1.7, B.1.351 or P.1.

[0048] The SARS-CoV-2 S protein or a portion thereof may also comprise a prefusion-stabilized form of the SARS-CoV-2 S protein ectodomain comprising two or more stabilizing mutations. In one embodiment the prefusion-stabilized form of the SARS-CoV-2 S protein ectodomain comprises two stabilizing mutations to proline corresponding to amino acid position K986 and V987 in the amino acid sequence of amino acid residues 1 to 1208 of SEQ ID NO: 1.

[0049] In certain embodiments the SARS-CoV-2 S protein or a portion thereof comprises an amino acid sequence of amino acid residues 1-1208 of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with amino acid residues 1-1208 of SEQ ID NO: 1, further comprising two stabilizing mutations K986P and V987P.

[0050] In certain embodiments the prefusion-stabilized form of the SARS-CoV-2 S protein ectodomain comprises two stabilizing mutations to proline corresponding to amino acid position K986 and V987 in the amino acid sequence of amino acid residues 1-1208 of SEQ ID NO: 1 and/or a mutation of the furin cleavage sequence (PRRARS|V) corresponding to residues 681-687 of the amino acid sequence of amino acid residues 1-1208 of SEQ ID NO: 1, such as a R682G, R683S and R685S mutation. Preferably the SARS-CoV-2 S protein ectodomain has an amino acid sequence of amino acid residues 1-1208 of SEQ ID NO: 1 or

an amino acid sequence having at least 95% sequence identity with the amino acid sequence of amino acid residues 1-1208 of SEQ ID NO: 1, comprising two stabilizing mutations K986P and V987P; or furin cleavage sequence mutations R682G, R683S and R685S, or two stabilizing mutations K986P and V987P and furin cleavage sequence mutations R682G, R683S and R685S. Alternatively amino acids of the furin cleavage sequence may be deleted such as amino acids 680-683. Thus, in one embodiment the SARS-CoV-2 full-length S protein has an amino acid sequence of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with SEQ ID NO: 1, comprising a deletion in the furin cleavage sequence, such as a deletion comprising or consisting of amino acids S680-R683. Other amino acid substitutions or amino acid deletions resulting in a pre-fusion stabilized form of the S protein ectodomain may also be employed.

[0051] The SARS-CoV-2 ectodomain may further comprise a fusion domain for stabilization and/or improved expression and/or improved secretion. The fusion domain may also be a trimerization domain, such as a C-terminal T4 fibritin timerization motif. The trimerization domain of the bacteriophage T4 fibritin, termed “foldon”, has the amino acid sequence GYIPEAPRDGQAYVRKDGEWVLLSTFL (SEQ ID NO: 10) corresponding to amino acid residues aa 457-483 of the fibritin protein.

[0052] The sequence encoding the SARS-CoV-2 S protein or a portion thereof preferably comprises a signaling sequence encoding a signaling peptide. The signaling peptide of the SARS-CoV-2 S protein has for example the amino acid sequence: MFVFLVLLPLVSSQC (SEQ ID NO: 3) corresponding to amino acid residues 1-15 of SEQ ID NO: 1 or an equivalent functional signaling peptide having at least 80% sequence identity, preferably at least 90% sequence identity, with the amino acid sequence of SEQ ID NO: 3. In one embodiment the signaling peptide of the SARS-CoV-2 S protein the signal peptide of the invariant chain, wherein in a preferred embodiment amino acid residues 1-12 of SEQ ID NO: 1 is replaced with amino acid residues 1-29 of SEQ ID NO: 15.

[0053] In certain embodiments the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof comprises the SARS-CoV-2 S protein subunit S1. The SARS-CoV-2 S protein subunit S1 comprises an amino acid sequence of amino acid residues 1-681 of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with amino acid

residues 1-681 of SEQ ID NO: 1. In a preferred embodiment the SARS-CoV-2 S protein or a portion thereof comprises the SARS-CoV-2 S protein subunit S1 having an amino acid sequence having at least 96%, at least 97%, at least 98%, or at least 99% sequence identity with the sequence of amino acid residues 1-681 of SEQ ID NO: 1. In one embodiment the SARS-CoV-2 S protein subunit S1 has an amino acid sequence having at least 98% to 100% sequence identity with amino acid residues 1-681 of SEQ ID NO: 1. In a specific embodiments the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof is the SARS-CoV-2 S protein subunit S1 having or consisting of an amino acid sequence of amino acid residues 1-681 of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with amino acid residues 1-681 of SEQ ID NO: 1. In a further specific embodiment the SARS-CoV-2 S protein subunit S1 may also be the S protein subunit S1 of a variant of SARS-CoV-2, such as lineage B.1.1.7, B.1.351 or P.1.

[0054] In certain embodiments the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof comprises the SARS-CoV-2 S protein receptor binding domain (RBD). The SARS-CoV-2 S protein RBD comprises an amino acid sequence of amino acid residues 319-541 of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with amino acid residues 319-541 of SEQ ID NO: 1. In a preferred embodiment the SARS-CoV-2 S protein or a portion thereof comprises the SARS-CoV-2 S protein RBD having an amino acid sequence having at least 96%, at least 97%, at least 98%, or at least 99% sequence identity with the sequence of amino acid residues 319-541 of SEQ ID NO: 1. In one embodiment the SARS-CoV-2 S protein RBD has an amino acid sequence having at least 98% to 100% sequence identity with amino acid residues 319-541 of SEQ ID NO: 1. In a specific embodiments the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof is the SARS-CoV-2 S protein RBD having or consisting of an amino acid sequence of amino acid residues 319-541 of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with amino acid residues 319-541 of SEQ ID NO: 1. In a specific embodiment the SARS-CoV-2 S protein RBD may also be the S protein RBD of a variant of SARS-CoV-2, such as lineage B.1.1.7, B.1.351 or P.1.

[0055] One advantage of using the SARS-CoV-2 full-length S protein, the SARS-CoV-2 S protein ectodomain, the SARS-CoV-2 protein subunit S1 or the SARS-CoV-2 RBD is that it

provides for a polyclonal humoral immune response (including a neutralizing antibody response) maintaining efficacy over mutating SARS-CoV-2 and that the humoral as well as the cellular immune response is not MHC restricted and hence limited to patients with certain HLA type.

[0056] In the context of the present invention, the term “at least 95% sequence identity with” refers to a protein that may differ in the amino acid sequence and/or the nucleic acid sequence encoding the amino acid sequence of the reference sequence, such as the amino acid sequence of SEQ ID NO: 1 or the amino acid sequence of amino acid residues 1-1208, amino acid residues 1-681 or amino acid residues 319-541 of SEQ ID NO: 1 (also referred to as the corresponding portion thereof). The S protein or the portion thereof may be of natural origin, e.g. a mutant version or a variation of the S protein of SARS-CoV-2 having the amino acid sequence of SEQ ID NO: 1 or an engineered protein, e.g. an engineered glycoprotein derivative, which has been modified by introducing site directed mutations or cloning, or a combination thereof. It is known that the usage of codons is different between species. Thus, when expressing a heterologous protein in a target cell, it may be necessary, or at least helpful, to adapt the nucleic acid sequence to the codon usage of the target cell. Methods for designing and constructing derivatives of a given protein are well known to the person skill in the art. Adapting the nucleic acid sequence to the codon usage of the target cell is also known as codon-optimization.

[0057] The S protein or a portion thereof that shares at least about 95% sequence identity with the amino acid sequence of SEQ ID NO: 1 or the corresponding portion thereof may contain one or more mutations comprising an addition, a deletion and/or a substitution of one or more amino acids. According to the teaching of the present invention, said deleted, added and/or substituted amino acids may be consecutive amino acids or may be interspersed over the length of the amino acid sequence of the S protein or the portion thereof that shares at least about 95% sequence identity with the amino acid sequence of SEQ ID NO: 1 or the corresponding portion thereof. According to the teaching of the present invention, any number of amino acids may be added, deleted, and/or substitutes, as long as the amino acid sequence identity with the amino acid sequence of SEQ ID NO: 1 or the corresponding portion thereof is at least about 95%. In particular embodiments, the sequence identity of the

amino acid sequence of the S protein or a portion thereof with the amino acid sequence of SEQ ID NO: 1 or the corresponding portion thereof is at least 95%, at least 96%, at least 97%, at least 98%, or preferably at least 99 %. All percentages are in relation to the amino acid sequence of SEQ ID NO: 1 or the corresponding portion thereof (such as amino acid residues 1-1208, amino acid residues 1-681 or amino acid residues 329-541). Methods and algorithms for determining sequence identity including the comparison of a parental protein and its derivative having deletions, additions and/or substitutions relative to a parental sequence, are well known to the practitioner of ordinary skill in the art. On the DNA level, the nucleic acid sequences encoding the S protein or a portion thereof that shares at least about 95% sequence identity with the amino acid sequence of SEQ ID NO: 1 may differ to a larger extent due to the degeneracy of the genetic code and the optional codon-optimization.

[0058] According to the invention, the DNA vaccine may comprise in certain embodiments a *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding from N-terminal to C-terminal at least a SARS-CoV-2 S protein or a portion thereof and an enhancer sequence, such as a complement peptide sequence, more preferably three copies of complement protein C3d (SEQ ID NO: 4) preferably each of the three C3d separated by a GS linker (3C3d; SEQ ID NO: 5). Such sequences have been described to enhance humoral immune responses, particularly eliciting a stronger antibody response. In case the SARS-CoV-2 S protein or a portion thereof comprises the SARS-CoV-2 S protein ectodomain, the SARS-CoV-2 S protein subunit S1 or the SARS-CoV-2 S protein RBD, the eukaryotic expression cassette may further encode a trimerization domain, such as a C-terminal T4 fibritin trimerization motif (SEQ ID NO: 10), preferably fused to the SARS-CoV-2 S protein portion. Thus, in certain embodiments, the DNA vaccine may also comprise a *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding from N-terminal to C-terminal at least the SARS-CoV-2 protein or a portion thereof comprising the SARS-CoV-2 S protein ectodomain, the SARS-CoV-2 S protein subunit S1 or the SARS-CoV-2 S protein RBD (preferably the SARS-CoV-2 S protein ectodomain), a trimerization domain and optionally an enhancer sequence, such as a complement peptide sequence.

[0059] Exemplary enhancers sequences such as ubiquitin peptide sequences or complement peptide sequences to promote presentation of antigens in MHC class I or II molecules, respectively, are known in the art. Plasmid vectors encoding MHC class I antigens and ubiquitin peptides delivered by Salmonella typhimurium to murine have demonstrated enhanced antigen-specific T cell responses and tumour control in a B16 tumour challenge model (Xiang et al, PNAS, 2000). Antibody responses to B cell epitopes encoded by DNA vectors have been shown to be enhanced by introduction of three copies of peptides of complement protein C3d, which binds to the CR2 (CD21) receptor found on B cells and follicular dendritic cells to enhance antigen-specific B cell activation (Moveseyan, J Neuroimmunol, 2008; Yang, Virus Res, 2010; Hou, Virology J, 2019). Thus, in order to enhance B cell responses, complement peptides sequences such as three copies of complement protein C3d (KFLTTAKDKNRWEDPGKQLYNVEATSYA; SEQ ID NO: 4) may be added C-terminally to the sequence encoding the SARS-CoV-2 S protein or a portion thereof. Preferably the three 28 amino acid peptides are separated by a GS linker, such as $GS(G_4S)_2GS$ as in SEQ ID NO: 5 (3C3d). Further, to improve nuclear import of the DNA molecule (such as a plasmid) comprising the eukaryotic expression cassette encoding at least a SARS-CoV-2 S protein or a portion thereof from the cytoplasm, the DNA molecule may further comprise a DNA nuclear targeting sequence, such as one or more copies of the SV40 DNA nuclear targeting sequence (DTS; SEQ ID NO: 16), preferably two or more copies of the DTS.

[0060] The DNA vaccine according to the invention may further encode another SARS-CoV-2 protein or a portion thereof, preferably a SARS-CoV-2 N protein or a portion thereof. In preferred embodiments the SARS-CoV-2 N protein or a portion thereof comprises the sequence of SEQ ID NO: 8 or a portion thereof or a sequence having at least 95% sequence identity with SEQ ID NO: 8 or a corresponding portion thereof. Preferably the SARS-CoV-2 N protein or a portion thereof has an amino acid sequence having at least 96%, at least 97%, at least 98%, or at least 99% sequence identity with the sequence of SEQ ID NO: 8. In one embodiment the SARS-CoV-2 N protein or a portion thereof has an amino acid sequence having at least 98% to 100% sequence identity with the sequence of SEQ ID NO: 8 or the corresponding portion thereof. In a further embodiment, the SARS-CoV-2 N protein or a

portion thereof may also have the amino acid sequence of a variant of SARS-CoV-2, such as lineage B.1.1.7, B.1.351 or P.1.

[0061] The another SARS-CoV 2 protein or a portion thereof may be expressed by a further DNA vaccine comprising a Salmonella typhi Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding at least a COVID-19 coronavirus (SARS-CoV-2) protein other than the spike (S) protein or a portion thereof. The two DNA vaccines may be co-administered to induce an immune response against the SARS-CoV-2 S protein and the another SARS-CoV-2 protein. Alternatively the another SARS-CoV 2 protein or a portion thereof may be expressed by the DNA vaccine according to the invention further comprising a second DNA molecule encoding the another SARS-CoV-2 protein. Thus, the DNA vaccine comprises a Salmonella typhi Ty21a strain comprising a first DNA molecule comprising a eukaryotic expression cassette encoding at least a COVID-19 coronavirus (SARS-CoV-2) protein spike (S) protein or a portion thereof and a second DNA molecule comprising a eukaryotic expression cassette encoding at least a COVID-19 coronavirus (SARS-CoV-2) protein other than the spike (S) protein or a portion thereof. Preferably the first and the second DNA molecules are plasmids, preferably expression plasmids. More preferably the plasmids have the same vector backbone, such as a pVAX10 backbone. It is also contemplated that the another SARS-CoV-2 protein or a portion thereof is expressed by the same DNA molecule comprising a first expression cassette encoding the SARS-CoV-2 S protein or a portion thereof and a second expression cassette encoding another SARS-CoV-2 protein or a portion thereof. All these embodiments may be freely combined with the embodiments referred to previously, particularly further defining the expression cassette encoding at least the SARS-CoV-2 S protein or a portion thereof optionally comprising an enhancer sequence and/or a trimerization domain.

[0062] It is further contemplated that the DNA molecule comprises a eukaryotic expression cassette encoding the SARS-CoV-2 S protein or a portion thereof and the another SARS-CoV-2 protein or a portion thereof. Thus, in certain embodiments the DNA vaccine comprises a *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding at least a COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof and another COVID-19 coronavirus (SARS-CoV-2) protein

(structural or non-structural). Preferably the SARS-CoV-2 S protein or a portion thereof is N-terminally expressed and the another SARS-CoV-2 protein or a portion thereof is C-terminally expressed. The following embodiments may be freely combined with the embodiments referred to previously, particularly further defining the expression cassette encoding at least the SARS-CoV-2 S protein or a portion thereof optionally comprising an enhancer sequence and/or a trimerization domain. In a preferred embodiment the DNA vaccine comprises a *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding at least a COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof and a COVID-19 coronavirus (SARS-CoV-2) N protein or a portion thereof. The SARS-CoV-2 N protein or a portion thereof may comprise the sequence of SEQ ID NO: 8 or a portion thereof or a sequence having at least 95% sequence identity with SEQ ID NO: 8 or a corresponding portion thereof. Preferably the SARS-CoV-2 N protein or a portion thereof has an amino acid sequence having at least 96%, at least 97%, at least 98%, or at least 99% sequence identity with the sequence of SEQ ID NO: 8. In one embodiment the SARS-CoV-2 N protein or a portion thereof has an amino acid sequence having at least 98% to 100% sequence identity with the sequence of SEQ ID NO: 8 or the corresponding portion thereof. In one embodiment the SARS-CoV-2 N protein or a portion thereof may also have the amino acid sequence of a variant of SARS-CoV-2, such as lineage B.1.1.7, B.1.351 or P.1. The SARS-CoV-2 S protein or a portion thereof may be linked to the another SARS-CoV-2 protein via a 2A self-cleaving peptide (2A peptide) or an internal ribosomal entry site (IRES), preferably a 2A peptide. Examples of 2A peptides are P2a with the amino acid sequence of GSGATNFSLLKQAGDVEENPGP (SEQ ID NO: 6) or T2a with the amino acid sequence of GSGEGRGSLLTCGDVEENPGP (SEQ ID NO: 7).

[0063] According to the invention, the DNA vaccine may comprise a *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding from N-terminal to C-terminal at least a SARS-CoV-2 S protein or a portion thereof, a 2A peptide or an IRES sequence and another SARS-CoV-2 protein or a portion thereof, preferably a SARS-CoV-2 N protein or a portion thereof. The another SARS-CoV-2 protein or a portion thereof may further be followed by a SARS-CoV-2 protein subunit S2, particularly if the SARS-CoV-2 S protein or a portion thereof is the SARS-CoV-2 protein subunit S1. In

certain embodiments, the SARS-CoV-2 protein subunit S2 comprises amino acid residues 686-1208 of SEQ ID NO: 1 or a sequence having at least 95% identity with amino acid residues 686-1208 of SEQ ID NO:1. In one embodiment subunit S2 comprises amino acid residues 686-1273 of SEQ ID NO: 1 or a sequence having at least 95% identity with amino acid residues 686-1273 of SEQ ID NO: 1.

[0064] The another SARS-CoV-2 protein or a portion thereof may further be preceded by an enhancer sequence, such as an ubiquitin sequence. Ubiquitin is conserved between mouse and human and has the amino acid sequence MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKE STLHLVLRLRG (SEQ ID NO: 9). Without being bound by theory, a N-terminal ubiquitin sequence may enhance T cell responses of antigens. Thus, also contemplated is a DNA vaccine comprising a *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding from N-terminal to C-terminal at least a SARS-CoV-2 S protein or a portion thereof, a 2A peptide or an IRES sequence, an ubiquitin sequence and another SARS-CoV-2 protein or a portion thereof, preferably a SARS-CoV-2 N protein or a portion thereof, optionally followed by the SARS-CoV-2 protein subunit S2.

[0065] The N protein is considered to mainly elicit a T cell response. Plasmid vectors encoding MHC class I antigens and ubiquitin peptides delivered by Salmonella typhimurium to murine have demonstrated enhanced antigen-specific T cell responses and tumour control in a B16 tumour challenge model (Xiang et al, PNAS, 2000). Thus, T cell enhancing sequences may be fused, preferably N-terminally, to the another SARS-CoV-2 protein or a portion thereof, such as the SARS-CoV-2 N protein or a portion thereof.

[0066] The term “2A self-cleaving peptides”, “2A cleavage site” or “2A peptides” are used synonymously herein and refer to a class of 18–22 aa-long peptides, which can induce the cleaving of the recombinant protein in a cell. 2A peptides are originally found in the 2A region in a viral genome of virus and have been adopted as tool to express polypeptides in one expression cassette. The 2A-peptide-mediated cleavage occurs after the translation and the cleavage is trigged by breaking of peptide bond between the Proline (P) and Glycine (G) in

C-terminal of 2A peptide. Sequences encoding 2A peptide linker are known in the art, such as provided in SEQ ID NOs: 6 or 7.

[0067] The term “internal ribosome entry site”, abbreviated IRES, as used herein is an RNA element that allows for translation initiation in a cap-independent manner and hence translation in an mRNA comprising an IRES sequences is also initiated at the IRES sequence.

[0068] In another embodiment the DNA vaccine comprising a *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding at least a COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof, wherein the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof comprises at least 3 immune-dominant epitopes of SARS-CoV-2 S protein. In one embodiment the expression cassette encodes at least 3 immune-dominant epitopes of SARS-CoV-2 S protein and an enhancer sequence, such as a complement peptide sequence as described above.

[0069] The term “at least 3 immune-dominant epitopes of SARS-CoV-2 S protein” as used herein refers to one polypeptide or more than one polypeptide comprising together 3 or more immune-dominant epitopes of SARS-CoV-2 S protein. Whether the three or more immune-dominant epitopes of SARS-CoV-2 S protein are part of the same or different polypeptides is not relevant. The three or more immune-dominant epitopes of SARS-CoV-2 S protein may therefore be expressed as one polypeptide or as more than one polypeptide. In one embodiment the eukaryotic expression cassette encodes at least one polypeptide comprising at least 3 immune-dominant epitopes of SARS-CoV-2 S protein. The immune-dominant epitopes comprised within the at least one or more polypeptide(s) are 3 or more, 5 or more, 10 or more, 20 or more, 30 or more, 50 or more, or even more than 50 immune-dominant epitopes. In the context of the *Salmonella typhi* Ty21a strain as used herein, the eukaryotic expression cassette encoding the at least 3 immune-dominant epitopes of SARS-CoV-2 S protein may encode one polypeptide comprising up to 50 immune-dominant epitopes or even more, such as up to 300. Antigens presented as peptides on MHC class I or II (in humans HLA) are typically from 11 to 30 amino acids long for MHC II (CD4 antigens) and from 8 to 10 amino acids for MHC I (CD8 antigens).

Thus, preferred ranges for immune-dominant epitopes to be contained within the at least one polypeptide are 3 to 300, 5 to 300, 10 to 300, 20 to 300 or 50 to 300 immune-dominant epitopes. Thus, the polypeptide may further comprise immune-dominant epitopes form other structural proteins of SARS-CoV-2, such as of the E protein, the M protein or the N protein, preferably of the N-protein. Preferred ranges for immune-dominant epitopes of SARS-CoV-2 S protein to be expressed by the eukaryotic expression cassette or to be contained within the at least one polypeptide are 3 to 25, 3 to 20 or 5 to 15. Each polypeptide comprising fused immune-dominant epitopes is proteolytically cleaved into the epitopes inside antigen presenting cells and presented via HLA to elicit a T-cell response.

[0070] Given the close genetic similarity between the S protein of SARS-CoV-2 with SARS-CoV (76%), SARS-CoV-2 T and B epitopes may be predicted using pre-existing immunological studies of SARS-CoV (Ahmed et al, Viruses, 2020). T and B cell epitopes may also be predicted using bioinformatic approaches with validated algorithms to recognize amino acid motifs that bind to MHC class I and class II proteins of various HLA molecule (Grifoni et al, Cell, 2020). Public resources such as Immune Epitope Database and Analysis Resource (IEDB), NetMHCPan, and NetMHCIIPan can be used to generate putative T and B cell epitopes. Using these approaches, a multi-epitope vaccine may be designed to encompass sections of the S protein that are rich in epitopes. One region of particular interest is the Receptor Binding Motif (RBM) of the S protein which interacts with the angiotensin-converting enzyme 2 (ACE2) receptor on human target cells to facilitate viral entry. Antibodies towards the RBM of SARS-CoV are neutralizing, however the RBM of SRS-CoV and SARS-CoC-2 has only 50% shared identity and the antibodies do not cross-neutralize (Ju et al, BioRxiv, 2020 – submitted; Walls et al, Cell, 2020).

[0071] According to the invention, the at least 3 immune-dominant epitopes of SARS-CoV-2 S protein may comprise CD8 T cell antigens and/or CD4 T cell antigens. Preferably, the at least 3 immune-dominant epitopes of SARS-CoV-2 S comprise CD8 T cell antigens and CD4 T cells antigens.

[0072] An immune-dominant epitopes is typically a peptide having 8 to 30 amino acids, preferably 8 to 20, more preferably 8 to 12 amino acids.

[0073] For vaccine comprising immune-dominant epitopes of SARS-CoV-2 S protein it is beneficial if the vaccine targets multiple immune-dominant epitopes of the S protein, preferably additionally even of further structural proteins, such as the N protein, as this reduces the risk of immune-evasion due to mutations in the S proteins.

[0074] Alternatively in certain embodiments the DNA vaccine comprises a *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding from N-terminal to C-terminal at least three immune-dominant epitopes of SARS-CoV-2 S protein and optionally an enhancer sequence, a 2A peptide or an IRES sequence, an optional ubiquitin sequence and another SARS-CoV-2 protein or a portion thereof, preferably a SARS-CoV-2 N protein or a portion thereof. The portion of the SARS-CoV-2 N protein may be at least three immune-dominant epitopes of SARS-CoV-2 N protein.

[0075] Advantage of DNA vaccine according to the invention comprising *Salmonella typhi* Ty21a, as carrier for the at least SARS-CoV-2 S protein or a portion thereof (such as the 3 immune-dominant epitopes of SARS-CoV-2 S protein, the full-length S protein, the S protein ectodomain, the S protein subunit S1 or the S protein RBD) are the established quality control assay, the individual differences of the plasmid only in the insert encoding the antigen, no need for expansion and no requirements with regard to sterility testing due to oral administration. Furthermore, expression plasmids suitable for transformation as well as the *Salmonella typhi* Ty21a strain as carrier allow a large insert such as the full-length S protein or a high number of immune-dominant epitopes. It further allows to further introduce another SARS-CoV-2 protein or a portion thereof, such as the SARS-CoV-2 N protein or a portion thereof linked via a 2A peptide or an IRES sequence to the SARS-CoV-2 S protein or a portion thereof.

[0076] The immune-dominant epitopes of SARS-CoV-2 S protein (or optionally also N protein) may be inserted into the plasmid as a string of beads (expressed as one or more polypeptides), optionally separated by a linker. The linker may be, without being limited

thereto, a GS linker, a 2A cleavage site, or an IRES sequence. Due to the fast generation and only limited need for quality control, the time for generating the *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising at least one eukaryotic expression cassette encoding the SARS-CoV-2 S protein or a portion thereof is short and can for example be achieved within 15 days, preferably within 14 days or less after identification of the antigen, including immune-dominant epitopes or new clinical isolates or mutants. Overnight fermentation is sufficient and no upscaling is required due to high yield of bacteria with a net yield in the range of 10^{11} colony forming units (CFU) in a 1L culture. This allows for a short manufacturing time, as well as the low manufacturing costs. Furthermore, the drug product was shown to be stable for at least three years. Thus, this DNA vaccine is suitable for fast development and production of an effective SARS-CoV-2 prophylactic and/ or therapeutic vaccine for use in a large number of subjects in need thereof. Moreover, it is easy to store and does not need medical trained personal for administration.

[0077] DNA sequences encoding at least a SARS-CoV-2 S protein or a portions thereof may be separated from DNA sequences encoding the another SARS-CoV-2 protein or a portions thereof with the use of a linker which may be, without being limited thereto, a GS linker, a 2A cleavage site, or an IRES sequence.

[0078] Methods for detecting immune-dominant epitopes in a protein and reliably predicting or determining those peptides with high-affinity binding of autologous human leukocyte antigen (HLA) molecules are known in the art. Peptides are then selected that are predicted to likely bind to autologous HLA-A or HLA-B proteins of the patient or which is predominant in the population. This may be confirmed, e.g., by *ex vivo* interferon γ enzyme-linked immunospot (ELISPOT).

[0079] In certain embodiments, the DNA molecule or the DNA molecule comprising the at least one eukaryotic expression cassette comprises an antibiotic resistance gene, such as the kanamycin antibiotic resistance gene, an ori, such as the pMB1 ori or the pUC, and a strong promoter, such as a CMV promoter. In particular embodiments, the

DNA molecule or the DNA molecule comprising the at least one eukaryotic expression cassette is a plasmid, such as a plasmid based on or derived from the commercially available pVAX1™ expression plasmid (Invitrogen, San Diego, California).

[0080] This expression vector may be modified by replacing the high copy pUC origin of replication by the low copy pMB1 origin of replication of pBR322. The low copy modification was made in order to reduce the metabolic burden and to render the construct more stable. The generated expression vector backbone was designated pVAX10.

[0081] The expression vector may also be designed to contain enhancers such as ubiquitin or complement to promote presentation of antigens in MHC class I or II molecules. Plasmid vectors encoding MHC class I antigens and ubiquitin delivered by Salmonella typhimurium to murine have demonstrated enhanced antigen-specific T cell responses and tumour control in a B16 tumour challenge model (Xiang et al, PNAS, 2000). Antibody responses to B cell epitopes encoded by DNA vectors have been shown to be enhanced by inclusion of three copies of complement protein C3d (SEQ ID NO: 4), which binds to the CR2 (CD21) receptor found on B cells and follicular dendritic cells to enhance antigen-specific B cell activation (Moveseyan, J Neuroimmunol, 2008; Yang, Virus Res, 2010; Hou, Virology J, 2019).

[0082] Several methods have been used to facilitate translation of multiple genes using a single plasmid vector, including inserting a Internal Ribosome Entry Site (IRES) (Ma et al, Hum Vaccin Immunother, 2013) or 2A peptides between peptide gene sequences (Liu et al, Scientific Reports, 2017).

[0083] In particular embodiments, the expression plasmid comprises the DNA molecule of SEQ ID NO: 2 (vector backbone pVAX10), which correlates to the sequence of expression vector pVAX10 without the portion of the multiple cloning site which is located between the restriction sites *Nhe*I and *Xho*I. In one embodiment the expression plasmid comprises a nucleic acid sequence of SEQ ID NO: 2 and a sequence encoding the

amino acid sequence of SEQ ID NO:1 or a portion thereof or an amino acid sequence that has at least 95% sequence identity with SEQ ID NO: 1 or a portion thereof.

[0084] Inserting SARS-CoV-2 S protein encoding ORF with a nucleic acid sequence encoding SEQ ID NO: 1 into this expression vector backbone via *Nhe*I/*Xho*I yielded the expression plasmid. The expression plasmid pVAX10.SCV-1 is schematically depicted in Figure 2.

[0085] The DNA vaccine according to the invention may be in the form of a pharmaceutical composition. Thus, in certain embodiments the DNA vaccine comprising a *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding at least a COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof further comprise one or more pharmaceutically acceptable excipients. In certain embodiments the DNA vaccine is an oral dosage form. The DNA vaccine of the present invention may be in the form of a solution, a suspension or any other form suitable for the intended oral use. Alternative dosage forms are an enteric coated capsule or a lyophilized powder. Typically, the DNA vaccine according to the present invention is provided as drinking solution, preferably as a suspension, more preferably as an aqueous suspension. This embodiment offers the advantage of improved patient compliance and allows for rapid, feasible and affordable mass vaccination programs, especially in poor geographies.

[0086] The invention also provides a pharmaceutical composition comprising the DNA vaccine according to the invention.

[0087] In the context of the present invention, the term “excipient” refers to a natural or synthetic substance formulated alongside the active ingredient of a medication. Suitable excipients include solvents, anti-adherents, binders, coatings, disintegrants, flavors, colors, lubricants, glidants, sorbents, preservatives and sweeteners.

[0088] In the context of the present invention, the term “pharmaceutically acceptable” refers to molecular entities and other ingredients of pharmaceutical compositions such as a DNA vaccine that are physiologically tolerable and do not typically produce untoward reactions when administered to a mammal (e.g., human). The term “pharmaceutically acceptable” may also mean approved by a regulatory agency of a Federal or a state government or listed in

the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in mammals, and, more particularly, in humans.

[0089] In certain embodiments, the DNA vaccine or the pharmaceutical composition according to the present invention is in the form of an enteric coated capsule, a lyophilized powder or a suspension. Suitable suspensions comprise means to neutralize gastric acids at least to a certain degree, i.e. to bring the pH of the gastric juice closer to a pH of 7. Thus, in certain embodiment the suspension is a buffered suspension obtained by suspending the attenuated strain of Salmonella according to the present invention in a suitable buffer, preferably in a buffer that neutralizes gastric acids to at least a certain degree, preferably in a buffer containing 2.6 g sodium hydrogen carbonate, 1.7 g L-ascorbic acid, 0.2 g lactose monohydrate and 100 ml of drinking water.

[0090] In certain embodiments, the DNA vaccine of the pharmaceutical composition according to the invention further comprises one or more adjuvants.

[0091] In the context of the present invention, the term “adjuvant” refers to an agent that modifies the effect of an active ingredient, i.e. the attenuated strain of Salmonella according to the present invention. Adjuvants may boost the immune response to an antigen, thereby allowing to minimize the amount of administered antigen.

[0092] In the context of the present invention, the term “vaccine” refers to an agent which is able to induce an immune response in a subject upon administration. A vaccine can preferably prevent, ameliorate or treat a disease. A vaccine in accordance with the present invention comprises the live attenuated strain of Salmonella typhi, S. typhi Ty21a. The vaccine in accordance with the present invention is a DNA vaccine and hence further comprises at least one copy of a DNA molecule comprising a eukaryotic expression cassette, encoding at least a COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof.

[0093] The term “DNA vaccine” or “DNA vaccination” as used herein refers to a vaccine for protecting against or treating a disease or infection by delivery of genetically engineered linear DNA or preferably plasmid(s) containing the DNA sequence encoding the antigen(s), such as the SARS-CoV-2 S protein or a portion thereof, against which an immune response

is sought to target cells of the patient in need thereof. Thus, the antigen is produced by target cells and induces an immune response. DNA vaccines have potential advantages over conventional vaccines, including the ability to induce a wider range of immune response types, such as a humoral and/or cell-mediated immune response. The plasmid can be delivered to the tissue by several methods, including the use of injection in saline, gene gun, liposomes or via carriers, such as bacterial and viral vectors. The DNA vaccine according to the invention comprises a *Salmonella typhi* Ty21a strain as carrier for delivery of the DNA molecule comprising a eukaryotic expression cassette encoding at least a COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof. Preferably the DNA molecule delivered by the live attenuated Salmonella typhi Ty21a strain is a plasmid.

[0094] The live attenuated Salmonella strain according to the present invention stably carries a DNA molecule encoding at least a COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof. It can be used as a vehicle for the oral delivery of this DNA molecule. Such a delivery vector comprising a DNA molecule encoding a heterologous antigen, such as SARS-CoV-2 S protein or a portion thereof, is referred to as DNA vaccine in the context of the present invention.

[0095] Genetic immunization might be advantageous over conventional vaccination. The target DNA can be detected for a considerable period of time thus acting as a depot of the antigen. Sequence motifs in some plasmids, like GpC islands, are immunostimulatory and can function as adjuvants furthered by the immunostimulation due to LPS and other bacterial components.

[0096] Live attenuated Salmonella vectors, such as Salmonella typhi Ty21a, produce their own immunomodulatory factors such as lipopolysaccharides (LPS) *in situ* which may constitute an advantage over other forms of administration such as microencapsulation. Moreover, the mucosal DNA vaccine according to the present invention uses the natural entry site of Coronaviruses, which may prove to be of benefit. The mucosal vaccination has an intra-lymphatic mode of action. After ingestion of the attenuated vaccine according to the present invention, macrophages and other cells in Peyer's patches of the gut are invaded by the modified bacteria. The bacteria are taken up by these phagocytic cells. Due to their attenuating mutations, bacteria of the *Salmonella typhi* Ty21 strain are not able to persist in

these phagocytic cells but die at this time point. The DNA molecules are released from the bacterium and the endosome and are subsequently transferred into the cytosol of the phagocytic immune cells, either via a specific transport system or by endosomal leakage. Finally, the recombinant DNA molecules enter the nucleus, where they are transcribed, leading to massive SARS-CoV-2 S protein expression within the phagocytic cells. The infected cells undergo apoptosis, loaded with the S protein antigen, and are taken up and processed by the gut's immune system. The danger signals of the bacterial infection serve as a strong adjuvant in this process, leading to strong antigen specific CD8+ T-cell and antibody responses at the level of both systemic and mucosal compartments. The intra-lymphatic mucosal vaccination route is especially useful for mass vaccinations, and for pathogens that use a mucosal route of entry, such as Coronaviruses.

[0097] Salmonella vaccines containing eukaryotic plasmids can generate B cell responses to the antigens encoded by the plasmid. In mice immunized orally with Salmonella typhimurium containing the pCMVb eukaryotic expressed vector encoding antigens listeriolysin or ActA, antigen-specific antibodies could be detected in blood serum by 4 weeks post immunization (Darji et al, Cell, 1997; Darji et al, FEMS Immunol Med Microbiol, 2000).

[0098] The vaccine strain Salmonella typhi Ty21a, has an unparalleled safety track record. There is no data available indicating that *Salmonella typhi* Ty21a is able to enter the bloodstream systemically. The live attenuated *Salmonella typhi* Ty21a vaccine strain thus allows specific targeting of the immune system in the gut, while being safe and well-tolerated. In contrast, adenovirus-based DNA vaccines might bear an inherent risk of unintended virus replication. In addition, preexisting immunity against adenoviruses was shown to limit vaccine efficacy in humans.

[0099] Also provided herein is the DNA vaccine according to the invention for use in the treatment and/or the prevention of coronavirus disease 2019 (COVID-19) or a SARS-CoV-2 infection. Also provided herein is a method for treating and/or preventing coronavirus disease 2019 (COVID-19) or a SARS-CoV-2 infection comprising administering the DNA vaccine according to the invention to a patient in need thereof.

[00100] Should adverse events occur that resemble hypersensitivity reactions mediated by histamine, leukotrienes, or cytokines, treatment options for fever, anaphylaxis, blood pressure instability, bronchospasm, and dyspnoea are available. Treatment options in case of unwanted T-cell derived auto-aggression are derived from standard treatment schemes in acute and chronic graft vs. host disease applied after stem cell transplantation. Cyclosporin and glucocorticoids are proposed as treatment options.

[00101] In the unlikely case of systemic *Salmonella typhi* Ty21a type infection, appropriate antibiotic therapy is recommended, for example with fluoroquinolones including ciprofloxacin or ofloxacin. Bacterial infections of the gastrointestinal tract are to be treated with respective agents, such as rifaximin.

[00102] In preferred embodiments, the DNA vaccine comprising the *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding at least a COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof according to the invention is administered orally. Oral administration is simpler, safer and more comfortable than parenteral administration. Although the DNA vaccine of the present invention may also be administered by any other suitable route, the oral route is preferred. Preferably, a therapeutically effective dose is administered to the subject, and this dose may depend on the particular application, particularly whether the DNA vaccine is for therapeutic or prophylactic use, the subject's weight, age, sex and state of health, the manner of administration and the formulation, etc. Administration may be single or multiple, as required.

[00103] The DNA vaccine according to the present invention may be provided in the form of a solution, a suspension, lyophilisate, an enteric coated capsule, or any other suitable form. Typically, the attenuated strain of Salmonella according to the present invention is formulated as drinking solution. This embodiment offers the advantage of improved patient compliance. Preferably, the drinking solution comprises means to neutralize gastric acids at least to a certain degree, i.e. to bring the pH of the gastric juice closer to a pH of 7. Preferably, the drinking solution is a buffered suspension comprising the attenuated strain of Salmonella according to the present invention. In a particular embodiment, the buffered suspension is obtained by suspending the attenuated strain of Salmonella according to the

present invention in a suitable buffer, preferably containing 2.6 g sodium hydrogen carbonate, 1.7 g L-ascorbic acid, 0.2 g lactose monohydrate and 100 ml of drinking water.

[00104] In particular embodiments, the treatment and/or prevention of COVID-19 or a SARS-CoV-2 infection may further comprises administration of a further SARS-CoV-2 vaccine or an anti-SARS-CoV-2 treatment. The treatment and/or prevention of COVID-19 and/or a SARS-CoV-2 infection may further comprises a DNA vaccine comprising a *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding at least another SARS-CoV 2 protein or a portion thereof, such as a COVID-19 coronavirus (SARS-CoV-2) envelope (E) protein, membrane (M) protein, or nucleocapsid (N) protein or a portion thereof, preferably a SARS-CoV-2 N protein or a portion thereof. The two DNA vaccines may be co-administered or may be administered subsequently, preferably the two DNA vaccines are co-administered.

[00105] In certain embodiments, the treatment and/or the prevention of COVID-19 and/or a SARS-CoV-2 infection comprises a prime/boost vaccination against SARS-CoV-2. In the context of the present invention, the term “prime/boost vaccination” refers to an immunization regimen that comprises immunizing a subject with a prime vaccination and subsequently with at least one boost vaccination. In preferred embodiments, the prime vaccine and the boost vaccine are the same; i.e. the prime/boost vaccination represents a homologous prime/boost vaccination. Particularly, the DNA vaccine according to the present invention is administered as prime vaccine and as boost vaccine. In other embodiments, the prime vaccine and the boost vaccine represent different types of vaccines against the same pathogen; i.e. the prime/boost vaccination represents a heterologous prime/boost vaccination. In certain embodiments, the DNA vaccine according to the present invention may be administered as prime vaccine and a further SARS-CoV-2 vaccine is administered as boost vaccine. In particular other embodiments, the further betacoronavirus vaccine is administered as prime vaccine and the attenuated Salmonella strain according to the present invention is administered as boost vaccine. The prime/boost vaccination may elicit superior immune responses than vaccination with a single prime vaccination alone. Improved initial T-cell responses, antibody responses and/or longevity of the immune responses may be achieved by prime/boost vaccination.

[00106] In certain embodiments, administration of the prime and the boost DNA vaccine according to the invention occurs within eight consecutive weeks, more particularly within three to six consecutive weeks. Prime vaccine and boost vaccine may be administered via the same route or via different routes. Preferably the prime and the boost DNA vaccine according to the invention are administered via the same route, more preferably the prime and the boost DNA vaccine are administered orally. Also, the DNA vaccine according to the invention may be administered one or several times at the same or different dosages. It is within the ability of the person skilled in the art to optimize prime/boost vaccination regimes, including optimization of the timing and dose of vaccine administration.

[00107] In particular embodiments, a single dose of the DNA vaccine comprises the *Salmonella typhi* Ty21a strain according to the invention at about 10^5 to about 10^{11} or at about 1 x 10^6 to about 1 x 10^{10}, more preferably at about 1 x 10^6 to about 1 x 10^9, at about 1 x 10^6 to about 1 x 10^8, or at about 1 x 10^6 to about 1 x 10^7 colony forming units (CFU). In one embodiment, a single dose of DNA vaccine comprises the *Salmonella typhi* Ty21a strain at about 1 x 10^6 to about 1 x 10^9 colony forming units (CFU). Administration of low doses of this live attenuated bacterial DNA vaccine minimizes the risk of excretion and thus of transmission to third parties. It has previously been shown no excretion is detectable below 1 x 10^9 CFU.

[00108] In this context, the term “about” or “approximately” means within a factor of 3, alternatively within a factor of 2, including within a factor of 1.5 of a given value or range.

[00109] In certain embodiments, the treatment and/or the prevention of COVID-19 or a SARS-CoV-2 infection comprises multiple administrations of the DNA vaccine according to the present invention. The single dose of the DNA vaccine administrations may be the same or different, preferably the single dose is the same and comprises the *Salmonella typhi* Ty21a strain at about 1 x 10^6 to about 1 x 10^9 colony forming units (CFU). In particular, the treatment and/or the prevention of COVID-19 or a SARS-CoV-2 infection comprises 1, 2, 3, 4, 5 or 6 administrations of the DNA vaccine according to the present invention. Preferably, the treatment and/or prevention of COVID-19 or a SARS-CoV-2 infection comprises that the DNA vaccine is to be administered two to four time in one week for priming (as prime vaccination), optionally followed by one or more single dose boosting. In certain

embodiments the DNA vaccine is to be administered 2 to 4 times within the first week (as prime vaccination), followed by one or more single dose boosting each at least 2 weeks later (as boost vaccination), i.e., prime vaccination in the first week and a single dose boost vaccination in week three or later, optionally followed by one (or more) further single dose boost vaccination at least 2 weeks later. In an alternative embodiment the DNA vaccine is to be administered 2 to 4 times within the first week (as prime vaccination), followed by one or more single dose boosting each at least 4 weeks later (as boost vaccination), i.e., prime vaccination in the first week and a single dose boost vaccination in week five or later, optionally followed by one (or more) further single dose boost vaccination at least 4 weeks later.

EXAMPLES

Example 1: Preparation of recombinant plasmid pVAX10.SCV-1

[00110] DNA encoding SARS-CoV-2 S protein (1273 aa, SEQ ID NO: 1) is cloned into the pVAX10 backbone derived of pVAX10.VR2-1 (WO 2013/091898). S protein DNA fragments are generated by double-strand gene synthesis, where oligonucleotides are linked together using a thermostable ligase. The obtained ligation products are amplified by PCR. Upon amplification, the *in vitro* synthesized S protein DNA fragment is cloned into the pVAX10 backbone via *Nhe*I/*Xho*I (the VEGFR-2 coding region of recombinant plasmid pVAX10.VR2-1 is replaced by the S protein coding region). For quality control, the entire plasmid is sequenced and aligned to the respective reference sequence after transformation into *E. coli* to show that it proves to be free of errors. The resulting plasmid is designated pVAX10.SCV-1 (Figure 2). Other suitable constructs are shown in Figure 3.

Example 2: Transformation of attenuated Salmonella strains with the recombinant plasmid pVAX10.SCV-1

[00111] S. typhi Ty 21a is transformed with plasmid pVAX10.SCV-1. The transformation is performed by electroporation.

Preparation of competent Salmonella cells:

[00112] Glycerol cultures of S. typhi Ty21a were inoculated on LB plates (animal component free [ACF] soy peptone). The plates were incubated at 37°C overnight. One colony was used for overnight-liquid-preculture. 3 ml LB medium (ACF soy peptone) inoculated with one colony was incubated at 37°C and 180 rpm overnight. To prepare competent cells, 2 x 300 ml of LB medium (ACF soy peptone) were inoculated with 3 ml of the overnight culture and incubated at 37°C and 180 rpm up to an OD_{600} of about 0.5. The cultures were then put on ice for 10 minutes. Subsequently, the bacteria were centrifuged for 10 minutes at 3000xg at 4°C and each pellet was resuspended in 500 mL of ice cold H_2O_{dest}. After a new centrifugation step, the bacterial pellets were washed twice in 10% ice cold glycerol. Both pellets were put together in 2 ml of 10% glycerol and finally frozen in aliquots of 50 µL on dry ice. The used glycerol did not contain any animal ingredients (Sigma Aldrich, G5150).

Transformation of competent Salmonella cells:

[00113] For each transformation reaction, a 50 µl aliquot of competent S. typhi Ty21a cells are thawed on ice for 10 minutes. After adding 3-5 µL of plasmid DNA pVAX10.SCV-1 the mixtures is incubated on ice for five minutes. Subsequently, the mixtures are transferred to pre-cooled cuvettes (1 mm thickness). The electric pulse is carried out at 12.5 kV/cm. Immediately afterwards, 1 ml of LB medium (ACF soy peptone) is added to the cells, the cells are transferred into a 2 ml Eppendorf tube and shaken for 1 hour at 37°C. After a short centrifugation step on a bench centrifuge (16600 rcf, 20 s), the bacterial pellet is resuspended in 200 µl of LB (ACF soy peptone) antibiotic-free medium. The mixtures is applied with a Drigalski spatula on LB plates (ACF soy peptone) containing kanamycin (concentration = 25 µg/ml or 50 µg/ml). The plates are incubated at 37°C overnight.

Plasmid preparation of recombinant Salmonella clones:

[00114] Three clones of the recombinant Salmonella typhi Ty21a strain are incubated overnight in 3 ml of LB medium (ACF soy peptone) containing kanamycin (50 µg/ml) at 37°C.

The bacterial culture is then pelleted by centrifugation (16600 rcf, 30 s). Plasmid isolation is performed using the NucleoSpin Plasmid Kit from Macherey-Nagel. The plasmid DNA is eluted from the silica gel columns with 50 µl water. 5 µl of the eluate is used in agarose gel electrophoresis for control.

[00115] For long-term storage, 1 ml glycerol cultures of the positive clones are produced. For this purpose, 172 µl glycerol (no animal ingredients) are added to 828 µl medium of a logarithmically growing 3 ml culture in a 1 low ml screw microtube. The samples are stored at -70°C until further use.

Complete sequencing of recombinant plasmid DNA isolated from Salmonella:

[00116] 3ml of liquid LB-Kan medium (ACF soy peptone) are inoculated with one colony of recombinant Salmonella (S. typhi Ty21a harboring pVAX10.SCV-1) and incubated overnight at 37°C and 180 rpm. The overnight culture is pelleted by centrifugation at 1300 rpm for 30 s on a bench centrifuge (Biofuge pico, Heraeus). The plasmid isolation is performed with the NucleoSpin Plasmid Kit from Macherey-Nagel. After alkaline lysis and precipitation of high molecular weight genomic DNA and cellular components, the plasmid DNA is loaded onto columns with a silica membrane. After a washing step, the plasmids are eluted from the column with 50 µl of sterile water and sequenced. The sequences are then compared with the respective reference sequence by clone specific alignments, i.e. the plasmid sequences of each Salmonella clone is one by one aligned with the reference sequence to check whether all sequences are in line with the respective reference sequences. The recombinant Salmonella strain is designated VXM-SCV-1 (S. typhi Ty21a harboring plasmid pVAX10.SCV-1).

Example 3: Large-scale production of VXM-SCV-1

[00117] Bacterial fermentation is carried out as described in WO 2013/091898. Down-stream processing consists of diafiltration, dilution and filling. One 100 l fermentation run yields approximately 5 liters of 1-10x10^{10} CFU/ml of vaccine. The vaccine is further diluted into suitable aliquots and stored at -70°C. The aliquots can be shipped on dry ice. On site,

the aliquots are diluted into an application buffer to yield the ready to use vaccine (a 100 ml drinking solution, prepared in bulk).

Example 4: Preclinical study design - Assessing immune responses elicited by VXM-SCV-1 in healthy mice

[00118] Immune responses against SARS-CoV-2 in healthy C57Bl/6, BALBc or CD1 mice are evaluated by antibody ELISA. Mice are vaccinated with Salmonella typhimurium containing plasmid pVAX10.SCV-1 (10^8-10^9 CFU/dose). Salmonella typhimurium containing plasmid pVAX10.SCV-1 are prepared as described above for Salmonella typhi Ty21a. As negative control, a vector control group (10^8-10^{10} CFU/dose Salmonella typhimurium containing no expression plasmid) is included in the study setup to discriminate the desired immunologic effect from any unspecific background stimulation caused by Salmonella empty vector. Immune monitoring is carried out at one or more post-vaccination time points.

1. Animal maintenance

[00119] Healthy female mice, 6 weeks old at reception, are observed for 7 days in a specific-pathogen-free (SPF) animal care unit before starting the procedure. Animals are maintained in rooms under controlled conditions of temperature (23 ± 2°C), humidity (45 ± 10%), photoperiod (12 h light/12 h dark) and air exchange. Animals are maintained in SPF conditions. Room temperature and humidity are continuously monitored. The air handling system is programmed for 14 air changes/hour, with no recirculation. Fresh outside air is passed through a series of filters, before being diffused evenly into each room. A positive pressure (20 ± 4 Pa) is maintained in the experimentation room to prevent contamination or the spread of pathogens within a rodent colony. Animals are housed in polycarbonate cages (Techniplast, Limonest, France) that are equipped to provide food and water. The standard area cages used are 800 cm^2 with a maximum of 10 mice per cage (from the same group). Bedding for animals is sterile corn cob bedding (ref: LAB COB 12, SERLAB, Cergy-Pontoise, France), replaced twice a week. Animal food is purchased from DIETEX (Saint-Gratien, France). Irradiated RM1 is used as sterile controlled granules. Food is provided *ad libitum* from water bottles equipped with rubber stoppers and sipper tubes. Water bottles are

sterilized by sterile filtration and replaced twice a week. At D0, mice are distributed according to their individual body weight into 2 groups using Vivo manager® software (Biosystemes, Couternon, France). The mean body weight of the two groups (which are then divided into groups 1 to 5 and of groups 6 to 10, respectively) is not statistically different (analysis of variance).

2. Detecting antibody responses in mice

[00120] BALBc and CD1 mice are divided into six groups of eight. Mice in groups 1-3 receive administration of the vector control, mice in groups 4-6 receive administration of Salmonella typhimurium containing plasmid pVAX10.SCV-1. Both Salmonella typhimurium strains are thawed and administered within 30 min, the working solutions are discarded after use. The treatment dose is 10^8 CFU in 100 µl per administration. The Salmonella strains are administered by oral gavage (*per os, PO*) via a cannula with a volume of 0.1 ml. Regardless of animal groups, each animal receives pre-dose application buffer to neutralize acid in the stomach prior dosing (100 µl / animal / application). This buffer is produced by dissolution of 2.6 g sodium hydrogen carbonate, 1.7 g L-ascorbic acid, 0.2 g lactose monohydrate in 100 ml of drinking water and is applied within 30 min prior application of the Salmonella typhimurium strains. The treatment schedule is as follows:

[00121] The mice in groups 1 (n=8) and 4 (n=8) receive 3 *PO* administrations of respective Salmonella typhimurium at 10^8 CFU every two weeks (Q2WKx3)

[00122] The mice in groups 2 (n=8) and 5 (n=8) receive daily *PO* administrations respective Salmonella typhimurium at 10^8 in CFU every two days for four consecutive times (Q2Dx4).

[00123] The mice in groups 3 (n=8) and 6 (n=8) receive daily *PO* administrations respective Salmonella typhimurium at 10^8 in CFU every two days for four consecutive times (Q2Dx4) and then two boosters every two weeks (Q2WKx2).

[00124] The viability and behavior of the animals is recorded every day, body weights are measured twice a week. Serum is collects on weeks 3, 4, 8, 12, 16, 20, 24 and 28 of study and stored at -20°C until analysis. An autopsy (macroscopic examination of heart,

lungs, liver, spleen, kidneys and gastrointestinal tract) is performed on all terminated animals at the end of the study.

[00125] Briefly, a 96-well EIA plate is coated overnight with 1 microgram per milliliter of N or S protein epitopes or recombinant whole N or S proteins in sodium carbonate buffer (pH 9.5) at 4°C. Next day, plate is washed with 100 millimolar tris-buffered saline/ Tween (TBST) and blocked for 1 hour at 37°C with 3% gelatin. Plate is thoroughly washed with TBST then serum is added to the top row of each plate and 1:1 dilutions prepared down each column with TBST. On each plate, a negative control column is included with no serum. The plate is incubated overnight at 4°C. To develop, plates are washed with TBST and incubated with 1:1000 dilution of Protein G conjugated to alkaline phosphatase (Calbiochem, USA) for 1 hour at 37°C. The OD405 is measured with an ELISA plate reader. Antibody end-point titre is determined as the reciprocal of the dilution required to give 1 standard deviation OD405 above the average OD405 of the negative control.

3. Detecting T cell responses in C57BL6 or BALBc mice

[00126] BALBc and C57BL6 mice are divided into six groups of twelve. Mice in groups 1-3 receive administration of the vector control, mice in groups 4-6 receive administration of Salmonella typhimurium containing plasmid pVAX10.SCV-1. Both Salmonella typhimurium strains are thawed and administered within 30 min, the working solutions are discarded after use. The treatment dose is 10^8 CFU in 100 µl per administration. The Salmonella strains are administered by oral gavage (per os, PO) via a cannula with a volume of 0.1 ml. Regardless of animal groups, each animal receives pre-dose application buffer to neutralize acid in the stomach prior dosing (100 µl / animal / application). This buffer is produced by dissolution of 2.6 g sodium hydrogen carbonate, 1.7 g L-ascorbic acid, 0.2 g lactose monohydrate in 100 ml of drinking water and is applied within 30 min prior application of the Salmonella typhimurium strains. The treatment schedule is as follows:

[00127] The mice in groups 1 (n=12) and 4 (n=12) receive 3 *PO* administrations of respective Salmonella typhimurium at 10^8 CFU every two weeks (Q2WKx3)

[00128] The mice in groups 2 (n=12) and 5 (n=12) receive daily *PO* administrations respective Salmonella typhimurium at 10^8 in CFU every two days for four consecutive times (Q2Dx4).

[00129] The mice in groups 3 (n=12) and 6 (n=12) receive daily *PO* administrations respective Salmonella typhimurium at 10^8 in CFU every two days for four consecutive times (Q2Dx4) and then two boosters every two weeks (Q2WKx2).

[00130] The viability and behavior of the animals is recorded every day, body weights are measured twice a week. One third of the mice in each group (n=4) were euthanized at 14 days, one third (n=4) were euthanized at 28 days, and the remaining one third of the mice (n=4) were euthanized at day 56. At the time of termination spleens and blood samples were collected. Blood was processed for serum, which was stored at -20°C until analysis. Spleens were processed into a single cell suspension. The immunogenicity of the vaccines was evaluated in the splenocyte preparations by IFN-gamma ELISPOT. Briefly, splenocytes were loaded into wells of an ELISPOT plate pre-coated with anti-IFN-gamma (500,000 cells in 0.1 ml). Peptide epitopes from the N or S protein were added to wells in duplicate at 10 micrograms per milliliter. Plates were incubated at 37°C for 18 hours. Next day, plates were developed using AEC kits (Sigma, USA) and individual IFN-gamma secreting cells enumerated using an Immunospot plate reader (Cellular Technologies Ltd, USA). Antibodies were detected in the serum samples by ELISA. Briefly, a 96-well EIA plate was coated overnight with 1 microgram per milliliter of N or S protein epitopes or recombinant whole N or S protein in sodium carbonate buffer (pH 9.5) at 4°C. Next day, plate was washed with 100 millimolar tris-buffered saline/ Tween (TBST) and blocked for 1 hour at 37°C with 3% gelatin. Plate was thoroughly washed with TBST then serum was added to the top row of each plate and 1:1 dilutions prepared down each column with TBST. On each plate, a negative control column was included with no serum. The plate was incubated overnight at 4°C. To develop, plates were washed with TBST and incubated with 1:1000 dilution of Protein G conjugated to alkaline phosphatase (Calbiochem, USA) for 1 hour at 37°C. The OD405 was measured with an ELISA plate reader. Antibody end-point titre was determined as the reciprocal of the dilution required to give 1 standard deviation OD405 above the average OD405 of the negative control.

4. Antigen expression analysis

[00131] Antigen expression analysis is performed by transfecting plasmid pVAX10.SCV-1 into murine 3T3 and human 293T cells. At 24 hours and 48 hours after infection, the cells are harvested and lysed. The obtained whole cell lysates are analyzed by SDS poly-acrylamide gel electrophoresis (SDS-PAGE), followed by Western blotting onto a PVDF membrane. RNA expression will also be confirmed by RT/PCR.

Example 5: Preclinical study - Assessing immune responses elicited by VXM-SCV-3 in healthy mice

[00132] The pVAX10-SCV-3 plasmid (insert SCV-3; SEQ ID NO: 11) encodes SARS-CoV-2 spike protein (SEQ ID NO: 1) with the furin domain removed (amino acid residues 680-683) and the SARS-CoV-2 N protein (SEQ ID NO: 8) (Accession no YP_009724397). The antigens are separated by a 2A self-cleaving peptide sequence (SEQ ID NO: 7) derived from capsid protein precursor of *Thosea asigna* virus (see Figure 3).

[00133] *Salmonella typhimurium* SL7207 vaccines containing pVAX10-SCV-3 were prepared by electroporation. Competent bacteria were incubated on ice with 100-500 ng of plasmid DNA then electroporated in GenePulsar II at 2.5 kiloVolts. Bacteria were incubated in SOC media for 1 hour at 37 degrees Celsius on a shaker plate, then 100 uL were plated on TSB agar plates with 50 ug/mL kanamycin overnight at 37 degrees Celsius. Individual colonies were expanded and frozen in 25% glycerol at -80 degrees Celsius.

[00134] Pathogen free, female BALBc mice, 4-6 weeks of age were purchased from Charles River Laboratories (St Constant, PQ, Canada) and were housed according to institutional guidelines with food and water ad libitum.

[00135] A group of 10 mice was treated with the SL-SCV-3 vaccine. For each treatment mice were pre-treated with 100 microliter dose of administration buffer (310 millimolar sodium bicarbonate, 100 millimolar L-ascorbic acid, 5 millimolar lactose monohydrate) by oral gavage, then received 100 microliter dose of vaccine in administration buffer at 1.5-2x10e9 CFU per milliliter. Mice were treated on days 0, 2, 5, 7, 21, and 35. Mice were bled before the study (pre-immune) then on weeks 2, 4, 6 and 8.

[00136] Vaccine efficacy was assessed by enzyme-linked immunosorbent assay (ELISA), a method that allows the detection of antigen-specific antibody levels in the serum of immunized animals. Briefly, a 96-well EIA plate was coated with antigen SARS-CoV-2 spike protein (ACROBiosystems) overnight at 4 degrees Celsius, blocked with 2% bovine serum albumin for 1 hour at 37 degrees Celsius, then incubated overnight at 4 degrees Celsius with serial dilutions of sera, typically starting at a dilution of 1/200. A secondary reagent (Goat anti-mouse IgG (H+L) Peroxidase, Jackson ImmunoResearch) was then added to each well at a 1/5000 dilution and incubated for one hour at 37 degrees Celsius. Plates were washed thoroughly and 3,3′,5,5′-Tetramethylbenzidine substrate (Life Technologies) was added to the wells for 5-10 minutes, the reaction was stopped by adding 0.16N H2SO4. The absorbance of each well at 450 nanometers was measured using a microtiter plate reader (Cytation5, Biotek). Endpoint titers were calculated as described in Frey A. et al (Journal of Immunological Methods, 1998, 221:35-41). Calculated titers represented the highest dilution at which a statistically significant increase in absorbance is observed in serum samples from immunized mice versus serum samples from naïve, non-immunized control mice.

[00137] Of the 10 mice vaccinated with SL-SCV-3, 2 mice generated antibody responses greater than assay background of 1/400. One mouse achieved peak antibody titer of 1/800 by week 4 and one mouse achieved and maintained peak antibody titer of 1/3200 by week 6 (see Fig. 4). This demonstrates that a salmonella-based SARS-CoV2 vaccine construct targeting the spike protein is able to generate an antigen-specific immune response against the spike protein, i.e. to generate a humoral immune response.

Example 6: Preclinical study - Assessing immune responses elicited by VXM-SCV-30 in healthy mice

[00138] The pVAX10-SCV-30 plasmid (insert SCV-30; SEQ ID NO: 12) encodes SARS-CoV-2 RBD domain of the spike protein (amino acid 319-541 of SEQ ID NO: 1), followed by three repeats of murine C3d (3C3d; SEQ ID NO: 17; KFLNTAKDRNRWEEPDQQLYNVEATSYA) then 2A self-cleaving peptide sequence (SEQ ID

NO: 7) derived from capsid protein precursor of *Thosea asigna* virus, followed by ubiquitin (SEQ ID NO: 9) fused to the SARS-CoV-2 N protein (SEQ ID NO: 8)(Accession no. YP_009724397) (see Fig. 3).

[00139] *Salmonella typhimurium* SL7207 vaccines were prepared with pVAX10-SCV-30 as described in example 5.

[00140] Pathogen free, female BALBc mice, 4-6 weeks of age were purchased from Charles River Laboratories (St Constant, PQ, Canada) and were housed according to institutional guidelines with food and water ad libitum.

[00141] A group of 10 mice was treated with the SL-SCV-30 vaccine. For each treatment mice were pre-treated with 100 microliter dose of administration buffer (310 millimolar sodium bicarbonate, 100 millimolar L-ascorbic acid, 5 millimolar lactose monohydrate) by oral gavage, then received 100 microliter dose of vaccine in administration buffer at 1.5-2x10e9 CFU per milliliter. Mice were treated on days 0, 2, 5, 7, 21, and 35. Mice were bled before the study (pre-immune) then on weeks 3, 4, 6, and 12.

[00142] Serum was analysed for antibodies towards SARS-CoV-2 spike protein as described in Example 5.

[00143] Of the 10 mice vaccinated with SL-SCV-30, one mouse generated antibody responses greater than assay background of 1/400 and reaching 1/3200 by week 3 (see Fig. 5). This demonstrates that a salmonella-based SARS-CoV2 vaccine construct targeting the RBD domain of the spike protein is able to generate an antigen-specific immune response against the spike protein.

Example 7: Preclinical study - Assessing immune responses elicited by VXM-SCV-42 in healthy mice

[00144] The pVAX10-SCV-42 plasmid (insert SCV-42; SEQ ID NO: 13) encodes SARS-CoV-2 S1 domain of the spike protein (amino acid 1-681 of SEQ ID NO: 1), followed by three repeats of murine C3d (SEQ ID NO: 17; 3C3d, SEQ ID NO: 18) then 2A self-cleaving peptide sequence (SEQ ID NO: 7) derived from capsid protein precursor of *Thosea*

asigna virus, followed by ubiquitin (SEQ ID NO: 9) fused to the SARS-CoV-2 N protein (SEQ ID NO: 8) another 2A self-cleaving peptide sequence and SARS-CoV-2 S2 domain of the spike protein (Ser686-Thr1273 of SEQ ID NO: 1) (see Fig. 3).

[00145] *Salmonella typhimurium* SL7207 vaccines were prepared with pVAX10-SCV-42 as described in example 5.

[00146] Pathogen free, female BALBc mice, 4-6 weeks of age were purchased from Charles River Laboratories (St Constant, PQ, Canada) and were housed according to institutional guidelines with food and water ad libitum.

[00147] A group of 10 mice was treated with the SL-SCV-42 vaccine. For each treatment mice were pre-treated with 100 microliter dose of administration buffer (310 millimolar sodium bicarbonate, 100 millimolar L-ascorbic acid, 5 millimolar lactose monohydrate) by oral gavage, then received 100 microliter dose of vaccine in administration buffer at 1.5-2x10e9 CFU per milliliter. Mice were treated on days 0, 2, 5, 7, 21, and 35. Mice were bled before the study (pre-immune) then on weeks 2, 4, 6, and 8.

[00148] Serum was analysed for antibodies towards SARS-CoV-2 spike protein as described in Example 5.

[00149] Of the 10 mice vaccinated with SL-SCV-42, 2 mice generated antibody responses greater than assay background of 1/400 and reaching 1/1600 (see Fig. 6). This demonstrates that a salmonella-based SARS-CoV2 vaccine construct targeting the S1 and/or the S2 subunit of the spike protein is able to generate an antigen-specific immune response against the spike protein.

Example 8: Preclinical study - Assessing immune responses elicited by VXM-SCV-53 in healthy mice

[00150] The pVAX10-SCV-53 plasmid (insert SCV-53; SEQ ID NO: 14; entire plasmid sequence SEQ ID NO: 19) encodes SARS-CoV-2 spike protein (SEQ ID NO: 1) with the furin domain removed (amino acid residues 680-683 deleted) and where the signal domain (Met1-Ser12 of SEQ ID NO: 1) has been replaced with that of invariant chain (Met1-Arg29 of SEQ

ID NO: 15), followed by 2A self-cleaving peptide sequence (SEQ ID NO: 7) derived from capsid protein precursor of *Thosea asigna* virus, followed by ubiquitin (SEQ ID NO: 9) fused to the SARS-CoV-2 N protein (SEQ ID NO: 8). The plasmid also contains the 72 nucleotide SV40 DNA nuclear targeting sequence (DTS) (SEQ ID NO: 16) within a larger SV40 ori enhancer sequence (SEQ ID NO: 20) upstream of the kanamycin resistance gene (see Fig. 3).

[00151] *Salmonella typhimurium* SL7207 vaccines were prepared with pVAX10-SCV-53 as described in example 5.

[00152] Pathogen free, female BALBc mice, 4-6 weeks of age were purchased from Charles River Laboratories (St Constant, PQ, Canada) and were housed according to institutional guidelines with food and water ad libitum.

[00153] A group of 10 mice was treated with the SL-SCV-53 vaccine. For each treatment mice were pre-treated with 100 microliter dose of administration buffer (310 millimolar sodium bicarbonate, 100 millimolar L-ascorbic acid, 5 millimolar lactose monohydrate) by oral gavage, then received 100 microliter dose of vaccine in administration buffer at 1.5-2x10e9 CFU per milliliter. Mice were treated on days 0, 2, 5, 7, 21, and 35. Mice were bled before the study (pre-immune) then on weeks 2, 4, 6, and 8.

[00154] Serum was analysed for antibodies towards SARS-CoV-2 spike protein as described in Example 5.

[00155] Of the 10 mice vaccinated with SL-SCV-53, 3 mice generated antibody responses greater than assay background of 1/400 and reaching 1/800 (see Fig. 7). This demonstrates that a salmonella-based SARS-CoV2 vaccine construct targeting a signal domain modified spike protein is able to generate an antigen-specific immune response against the spike protein.

Example 9: VXM-SCV-X Phase I clinical trial; study design

[00156] The aim of this phase I trial is to examine the safety, tolerability, and immunological responses to VXM-SCV-X. The randomized, placebo-controlled, double blind

dose-escalation study includes 45 subjects. The subjects receive four doses of VXM-SCV-X or placebo on days 1, 3, 5, and 7. Doses from 10^6 CFU up to 10^9 CFU of VXM-SCV-X are evaluated in the study. An independent data safety monitoring board (DSMB) is involved in the dose-escalation decisions. In addition to safety as primary endpoint, the VXM-SCV-1-specific immune reactions are evaluated.

[00157] The objectives are to examine the safety and tolerability, and immunological responses to the investigational anti-SARS-CoV-2 virus vaccine VXM-SCV-X, as well as to identify the maximum tolerated dose (MTD) of VXM-SCV-1. The MTD is defined as the highest dose level at which less than two of up to six patients under VXM-SCV-X treatment experience a dose-limiting toxicity (DLT).

[00158] Primary endpoints for safety and tolerability are adverse events and serious adverse events according to the CTCAE criteria.

[00159] Secondary endpoints, which assess the efficacy of the experimental vaccine to elicit a specific immune response to SARS-CoV-2 S protein, include the number of immune positive patients.

[00160] VXM-SCV-X is manufactured according to Good Manufacturing Practice (GMP) and is given in a buffered solution. The placebo control consisted of isotonic sodium chloride solution.

[00161] The starting dose consists of a solution containing 10^6 colony forming units (CFU) of VXM-SCV-X. This VXM-SCV-X dose was chosen for safety reasons. For comparison, one dose of Typhoral®, the licensed vaccine against typhoid fever, contains $2x10^9$ to $6x10^9$ CFU of Salmonella typhi Ty21a, equivalent to approximately thousand times the VXM-SCV-1 starting dose. The dose is escalated in logarithmic steps, which appears to be justified for a live bacterial vaccine.

[00162] Complying with guidelines for first-in-human trials, the patients of one dose group are treated in cohorts. The first administration of VXM-SCV-X in any dose group is given to one patient. The second cohort of each dose group consists of two patients receiving VXM-SCV-X. This staggered administration with one front-runner, i.e. only one patient receiving VXM-SCV-X first, serves to mitigate the risks.

[00163] A third cohort of patients (three receiving VXM-SCV-X are included in all dose groups.

[00164] The environmental risk inherent to an oral vaccine is the potential of excretion to the environment and subsequent vaccination of people outside the target population. All study patients are confined in the study site for the period during which vaccinations take place plus three additional days. All feces of study patients are collected and incinerated. Body fluids and feces samples are investigated for VXM-SCV-X shedding.

[00165] Hygienic precautions are applied to protect study personnel from accidental uptake. Study personnel are trained specifically for this aspect of the study.

[00166] In addition, specific T-cell activation and antibody formation are measured in this patient setting. A placebo control is included, in order to gain further knowledge on specific safety issues related to the active vaccine vs. the background treatment. In addition, the pooled placebo patients serve as a sound comparator for assessing specific immune activation.

Example 10: VXM-SCV-1 specific T-cell and B cell responses

[00167] Responses to VXM19 are assessed by monitoring the frequencies of SARS-CoV-2 virus S protein specific T-cells in peripheral blood of VXM-SCV-X and placebo treated patients, detected by IFNγ ELISpot, at different time points prior during and post vaccination.

[00168] Firstly, T-cells and peptide pulsed DC are added to wells coated with anti-INFγ antibodies. After a period of incubation, cells are removed with secreted INFγ left binding with the coat antibodies. Then detection antibody is added to detect the bound INFγ, and after a signal amplification, the final yield can be viewed as "color spots" representing single activated and specific T-cells.

[00169] B cell responses are measured by ELISA. Briefly, a 96-well EIA plate is coated overnight with 1 microgram per milliliter of N or S protein epitopes or recombinant whole N or S proteins in sodium carbonate buffer (pH 9.5) at 4°C. Next day, plate is washed with 100 millimolar tris-buffered saline/ Tween (TBST) and blocked for 1 hour at 37°C with 3% gelatin.

Plate is thoroughly washed with TBST then serum is added to the top row of each plate and 1:1 dilutions prepared down each column with TBST. On each plate, a negative control column is included with no serum. The plate is incubated overnight at 4°C. To develop, plates are washed with TBST and incubated with 1:1000 dilution of Protein G conjugated to alkaline phosphatase (Calbiochem, USA) for 1 hour at 37°C. The OD405 is measured with an ELISA plate reader. Antibody end-point titre is determined as the reciprocal of the dilution required to give 1 standard deviation OD405 above the average OD405 of the negative control.

Example 11: Anti-Carrier Immunity

[00170] In order to assess immune responses to the bacterial vehicle, anti-Salmonella typhi IgG and IgM immunoglobulins are detected by ELISA using two commercial assay kits (Salmonella typhi IgG ELISA, Cat. No. ST0936G and Salmonella typhi IgM ELISA, Cat. No. ST084M; Calbiotech. Inc., 10461 Austin Dr, Spring Valley, CA 91978, USA). These assays are qualitative assays. The assays are used as described in the package inserts respectively App. I/I) and as modified as part of the study plan according to the foregoing validation study 580.132.2785.

[00171] Both assays employ the enzyme-linked immunosorbent assay technique. Calibrator, negative control, positive control and samples are analyzed as duplicates. Diluted patient serum (dilution 1:101) is added to wells coated with purified antigen. IgG or IgM specific antibody, if present, bind to the antigen. All unbound materials are washed away and the enzyme conjugate is added to bind to the antibody-antigen complex, if present. Excess enzyme conjugate is washed off and substrate is added. The plate is incubated to allow for hydrolysis of the substrate by the enzyme. The intensity of the color generated is proportional to the amount of IgG or IgM specific antibody in the sample. The intensity of the color is measured using a spectrophotometric microtiter plate reader at 450 nm. The cut off is calculated as follows:

Calibrator OD x Calibrator Factor (CF).

[00172] The antibody index of each determination is determined by dividing the OD value of each sample by cut-off value.

[00173] Antibody index interpretation:

<0.9	No detectable antibody to Salmonella typhi IgG or IgM by ELISA
0.9-1.1	Borderline positive
>1.1	Detectable antibody to Salmonella typhi IgG or IgM by ELISA

Example 12: Vaccination Schedule

[00174] A single dose of VXM19, i.e. from 10^6 to 10^8 CFU is administered orally as 100 ml drinking solution. Vaccination with a single dose each occurs on days 1, 3, 5 and optionally 7. Peak immune response are expected to occur around 10 days after the last vaccination. Boosting may be considered after 2 to 4 weeks or even after 3 to 6 months. Schedule recommendations are derived from vaccine strain Ty21a.

CLAIMS

1. A DNA vaccine comprising a *Salmonella typhi* Ty21a strain comprising a DNA molecule comprising a eukaryotic expression cassette encoding at least a COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof.

2. The DNA vaccine according to claim 1, wherein the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof comprises

   (a) a SARS-CoV-2 full-length S protein;

   (b) a SARS-CoV-2 S protein ectodomain;

   (c) a SARS-CoV-2 S protein subunit S1;

   (d) a SARS-CoV-2 S protein receptor binding domain (RBD); or

   (e) at least 3 immune-dominant epitopes of SARS-CoV-2 S protein.

3. The DNA vaccine according to claim 2, wherein the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein is a SARS-CoV-2 full-length S protein, optionally wherein the SARS-CoV-2 full-length S protein comprises an amino acid sequence of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with SEQ ID NO: 1.

4. The DNA vaccine according to claim 2, wherein the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof comprises the SARS-CoV-2 S protein ectodomain, optionally wherein the SARS-CoV-2 S protein ectodomain comprises an amino acid sequence of amino acid residues 1-1208 of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with amino acid residues 1-1208 of SEQ ID NO: 1.

5. The DNA vaccine according to claim 2, wherein the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof comprises the SARS-CoV-2 S protein subunit S1, optionally wherein the SARS-CoV-2 protein subunit S1 comprises an amino acid sequence of amino acid residues 1-681 of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with amino acid residues 1-681 of SEQ ID NO: 1.

6. The DNA vaccine according to claim 2, wherein the COVID-19 coronavirus (SARS-CoV-2) spike (S) protein or a portion thereof comprises the SARS-CoV-2 S protein receptor binding domain (RBD), optionally wherein the SARS-CoV-2 protein RBD comprises an amino acid sequence of amino acid residues 319-541 of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with amino acid residues 319-541 of SEQ ID NO: 1.

7. The DNA vaccine according to claim 2, wherein the SARS-CoV-2 S protein or a portion thereof is a prefusion-stabilized form of the SARS-CoV-2 full-length S protein or the SARS-CoV-2 S protein ectodomain comprising two stabilizing mutations to proline corresponding to amino acid position K986 and V987 in the amino acid sequence of SEQ ID NO: 1; preferably wherein the SARS-CoV-2 S protein or a portion thereof comprises

   (a) an amino acid sequence of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with SEQ ID NO: 1, comprising two stabilizing mutations K986P and V987P; or

   (b) an amino acid sequence of amino acid residues 1-1208 of SEQ ID NO: 1 or an amino acid sequence having at least 95% sequence identity with amino acid residues 1-1208 of SEQ ID NO: 1, comprising two stabilizing mutations K986P and V987P.

8. The DNA vaccine according to any one of claims 1 to 7, wherein the eukaryotic expression cassette further encodes another SARS-CoV-2 protein or a portion thereof.

9. The DNA vaccine according to claim 8, wherein the other SARS-CoV-2 protein is a SARS-CoV-2 N protein.

10. The DNA vaccine according to any one of claims 1 to 9, further comprising one or more pharmaceutically acceptable excipients.

11. The DNA vaccine according to any one of claims 1 to 10, wherein the vaccine is an oral dosage form.

12. The DNA vaccine according to claim 11, wherein the oral dosage form is an enteric coated capsule, a lyophilized powder or a suspension.

13. The DNA vaccine according to any one of claims 1 to 12 further comprising one or more adjuvants.

14. The DNA vaccine according to any one of claims 1 to 13 for use in the treatment and/or the prevention of coronavirus disease 2019 (COVID-19) or a SARS-CoV-2 infection.

15. The DNA vaccine for use according to claim 14, wherein the DNA vaccine is administered orally.

16. The DNA vaccine for use according to claim 14 or 15, wherein

(a) a single dose of DNA vaccine comprises the *Salmonella typhi* Ty21a strain at about 10^6 to about 10^9 colony forming units (CFU), and/or

(b) the DNA vaccine is to be administered 2 to 4 times in one week for priming, optionally followed by one or more single dose boosting.

17. The DNA vaccine for use according to claim 16, wherein the DNA vaccine is to be administered 2 to 4 times within the first week, followed by one or more single dose boosting each at least 2 weeks later.

Figure 1

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRS
SVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGV
YFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQF
CNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLE
GKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEP
LVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYL
QPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQT
SNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISN
CVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGD
EVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYN
YLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSY
GFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVN
FNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEIL
DITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLT
PTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQ
TQTNSP**RR**A**R**SVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTI
SVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNR
ALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPS
KPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKF
NGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAM
QMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASAL
GKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLD**KV**EAE
VQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLG
QSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPA
ICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGN
CDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDIS
GINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWY
IWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDD
SEPVLKGVKLHYT

Figure 2

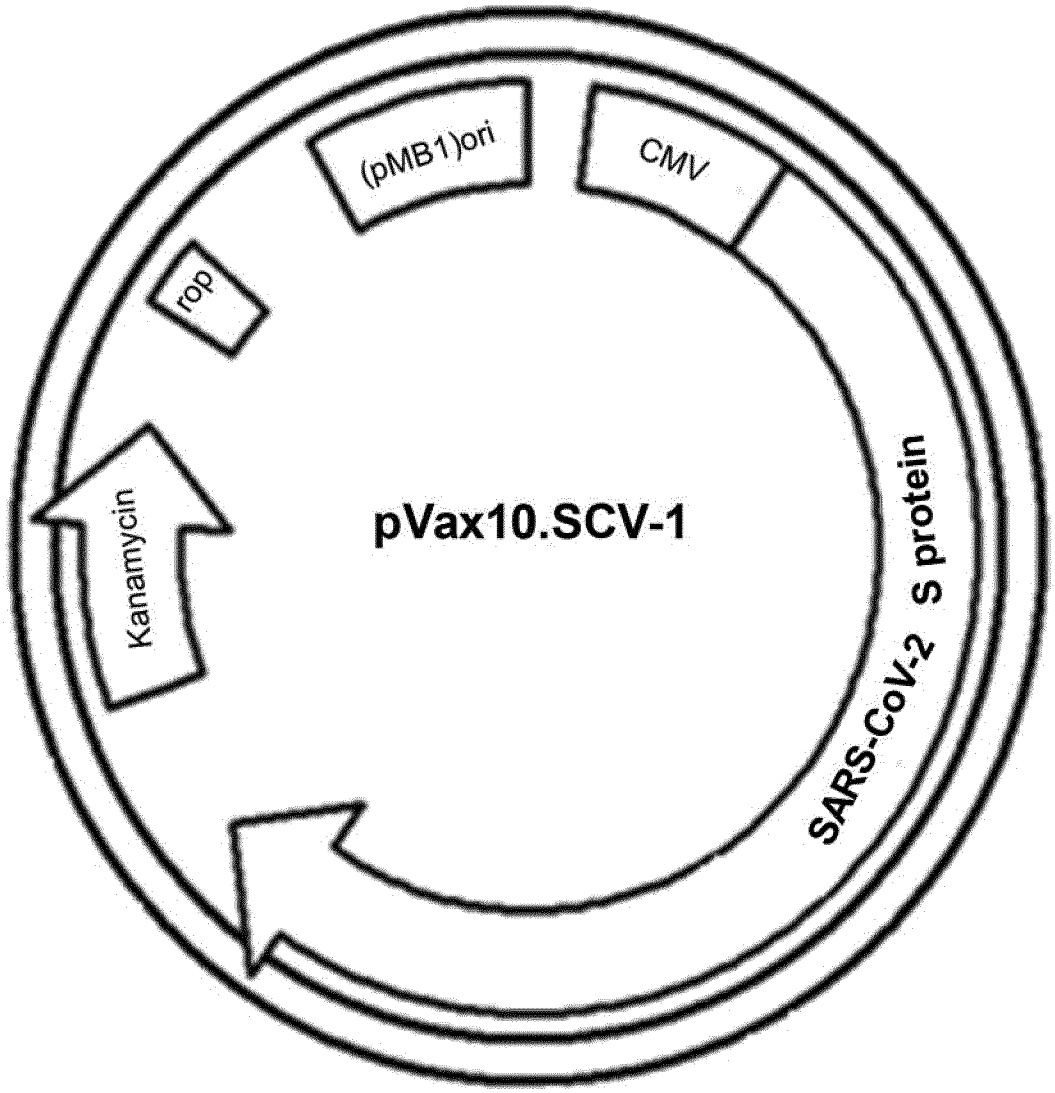


Figure 3

	S FL	S ecto.	S1	RBD	T4 trimer	3C3d	2A	Ubi.	N	S2	SV40 DTS
SCV-1	+										
SCV-2	+					+					
SCV-3	+						+		+		
SCV-4	+						+	+	+		
SCV-5	+					+	+		+		
SCV-6	+					+	+	+	+		
SCV-7		+									
SCV-8		+				+					
SCV-9		+			+						
SCV-10		+			+	+					
SCV-11		+					+		+		
SCV-12		+				+	+		+		
SCV-13		+			+		+		+		
SCV-14		+			+	+	+		+		
SCV-15		+					+	+	+		
SCV-16		+				+	+	+	+		
SCV-17		+			+		+	+	+		
SCV-18		+			+	+	+	+	+		
SCV-19			+								
SCV-20			+			+					
SCV-21			+				+		+		
SCV-22			+			+	+		+		
SCV-23			+				+	+	+		
SCV-24			+			+	+	+	+		
SCV-25				+							
SCV-26				+		+					
SCV-27				+			+		+		
SCV-28				+		+	+		+		
SCV-29				+			+	+	+		
SCV-30				+		+	+	+	+		
SCV-31									+		
SCV-32								+	+		
SCV-33									+	+	
SCV-34								+	+	+	
SCV-35			+				+		+		
SCV-36			+				+	+	+		
SCV-37			+				+		+	+	
SCV-38			+				+	+	+	+	
SCV-39			+			+	+		+		
SCV-40			+			+	+	+	+		
SCV-41			+			+	+		+	+	
SCV-42			+			+	+	+	+	+	
SCV-53	+ *						+	+	+		+

Figure 4:

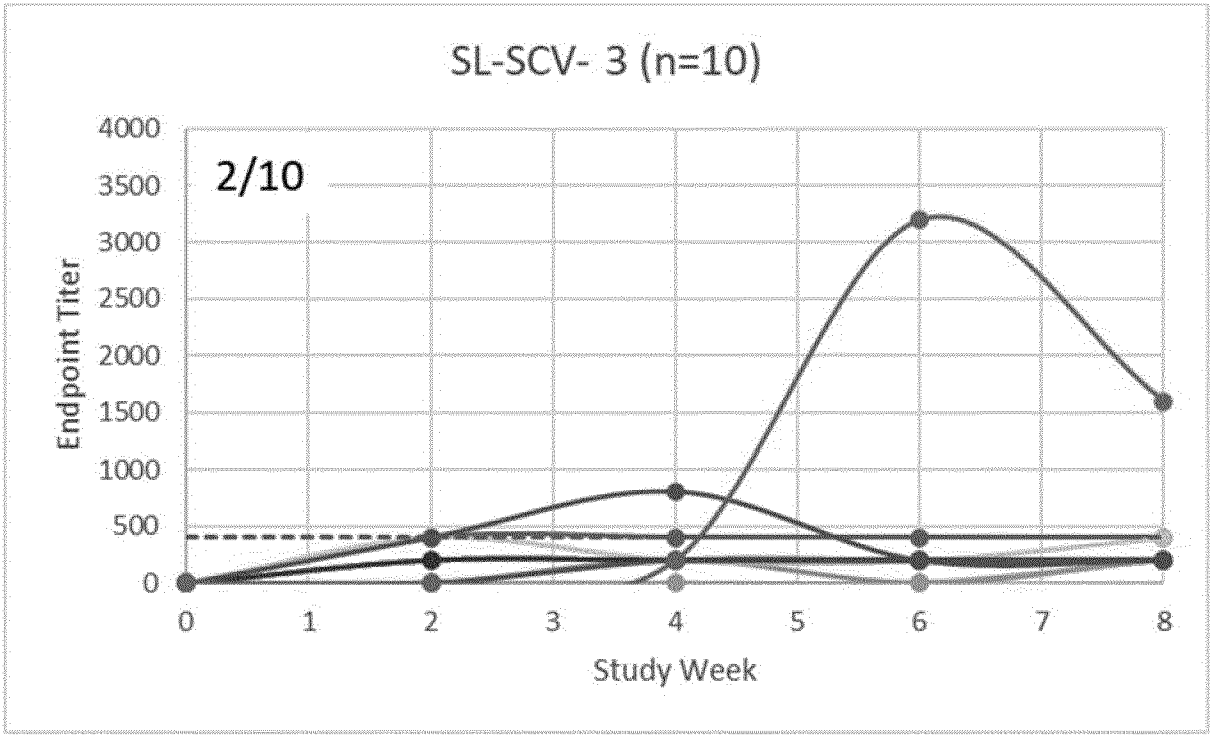


Figure 5:

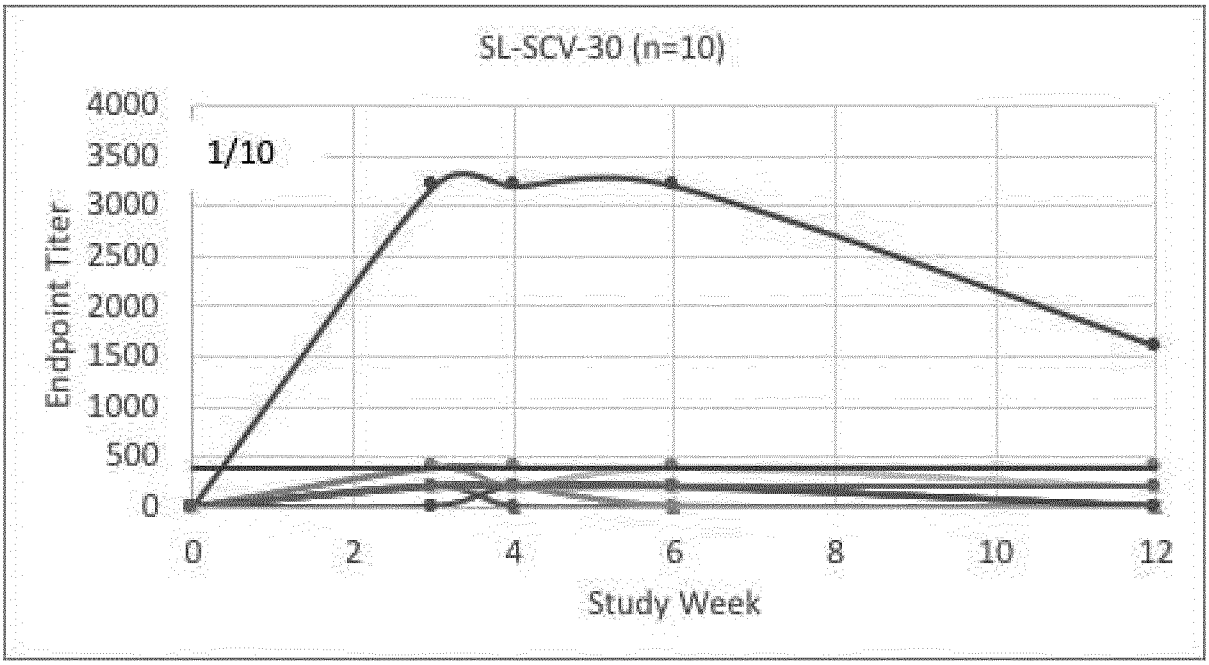


Figure 6:

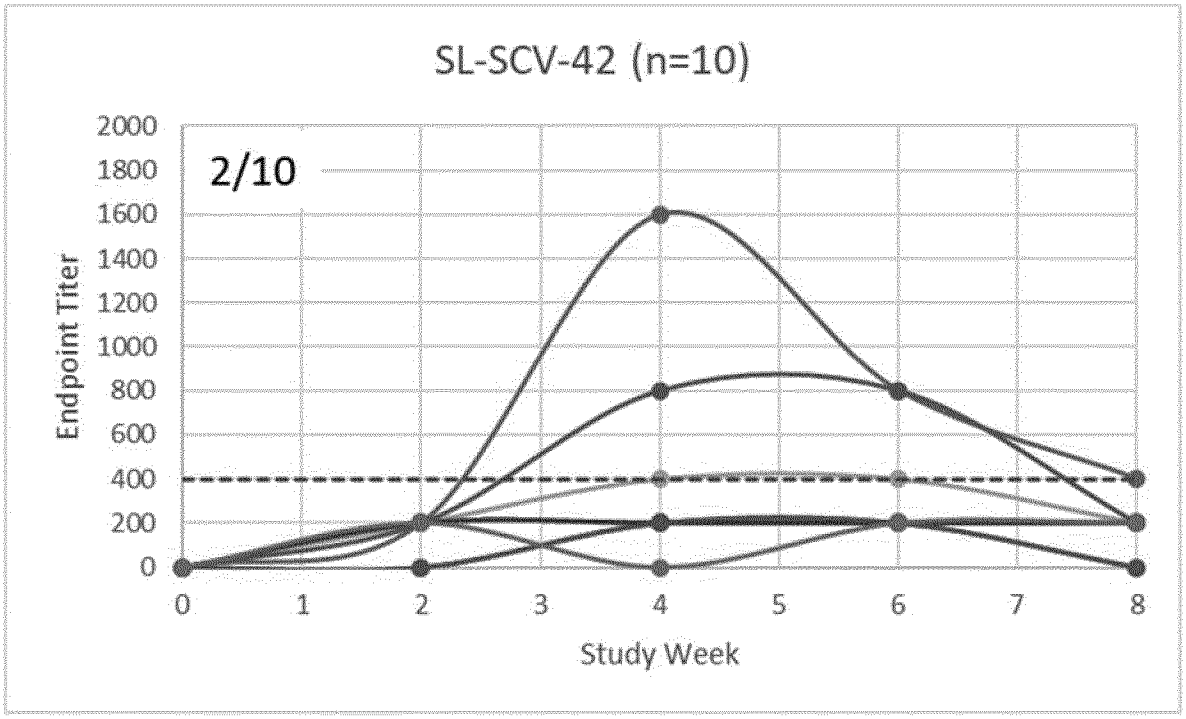
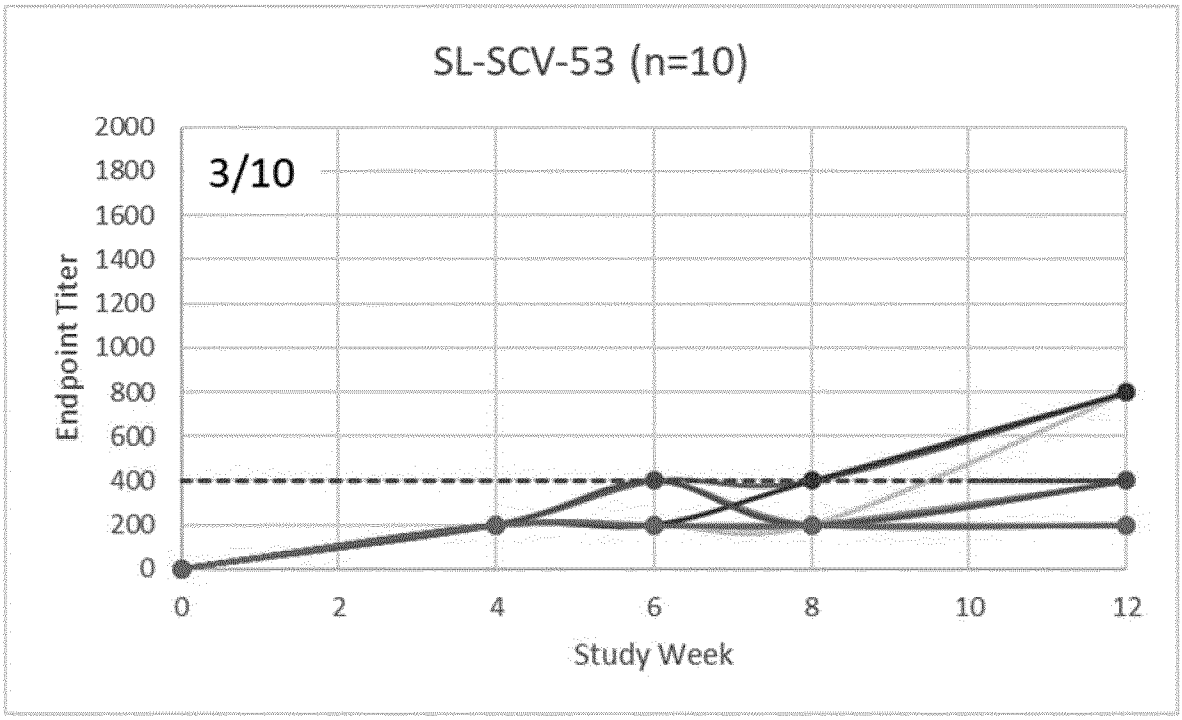


Figure 7:



SEQUENCE LISTING

<110> Vaximm AG

<120> NOVEL SALMONELLA-BASED CORONAVIRUS VACCINE

<130> 116631P855PC

<150> EP20167405.8
<151> 2020-03-31

<160> 20

<170> BiSSAP 1.3.6

<210> 1
<211> 1273
<212> PRT
<213> Coronaviridae

<220>
<223> SARS-CoV-2 S Protein

<400> 1
Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15
Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30
Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45
His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60
Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80
Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95
Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110
Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125
Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140
Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160
Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175
Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190
Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205
Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220
Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr

```
225                 230                 235                 240
Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
                245                 250                 255
Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
            260                 265                 270
Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
        275                 280                 285
Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
    290                 295                 300
Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val
305                 310                 315                 320
Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys
                325                 330                 335
Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
            340                 345                 350
Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
        355                 360                 365
Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
    370                 375                 380
Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
385                 390                 395                 400
Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
                405                 410                 415
Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
            420                 425                 430
Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
        435                 440                 445
Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
    450                 455                 460
Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
465                 470                 475                 480
Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
                485                 490                 495
Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
            500                 505                 510
Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
        515                 520                 525
Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn
    530                 535                 540
Gly Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu
545                 550                 555                 560
Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val
                565                 570                 575
Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe
            580                 585                 590
Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val
        595                 600                 605
Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile
    610                 615                 620
His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly Ser
625                 630                 635                 640
Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val
                645                 650                 655
Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala
```

```
            660                 665                 670
Ser Tyr Gln Thr Gln Thr Asn Ser Pro Arg Arg Ala Arg Ser Val Ala
        675                 680                 685
Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Ala Glu Asn Ser
    690                 695                 700
Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr Asn Phe Thr Ile
705                 710                 715                 720
Ser Val Thr Thr Glu Ile Leu Pro Val Ser Met Thr Lys Thr Ser Val
                725                 730                 735
Asp Cys Thr Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ser Asn Leu
            740                 745                 750
Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Thr
        755                 760                 765
Gly Ile Ala Val Glu Gln Asp Lys Asn Thr Gln Glu Val Phe Ala Gln
    770                 775                 780
Val Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly Phe
785                 790                 795                 800
Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg Ser
                805                 810                 815
Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly
            820                 825                 830
Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala Arg Asp
        835                 840                 845
Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu
    850                 855                 860
Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser Ala Leu Leu Ala Gly
865                 870                 875                 880
Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile
                885                 890                 895
Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr
            900                 905                 910
Gln Asn Val Leu Tyr Glu Asn Gln Lys Leu Ile Ala Asn Gln Phe Asn
        915                 920                 925
Ser Ala Ile Gly Lys Ile Gln Asp Ser Leu Ser Ser Thr Ala Ser Ala
    930                 935                 940
Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn
945                 950                 955                 960
Thr Leu Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val
                965                 970                 975
Leu Asn Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln
            980                 985                 990
Ile Asp Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val
        995                 1000                1005
Thr Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu
    1010                1015                1020
Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val
1025                1030                1035                1040
Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala
                1045                1050                1055
Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu
            1060                1065                1070
Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His
        1075                1080                1085
Phe Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val
```

```
    1090                1095                1100
Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn Thr
1105                1110                1115                1120
Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn Asn Thr
                1125                1130                1135
Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu
            1140                1145                1150
Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly Asp
        1155                1160                1165
Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys Glu Ile Asp
    1170                1175                1180
Arg Leu Asn Glu Val Ala Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu
1185                1190                1195                1200
Gln Glu Leu Gly Lys Tyr Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Ile
                1205                1210                1215
Trp Leu Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile
            1220                1225                1230
Met Leu Cys Cys Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Cys Cys
        1235                1240                1245
Ser Cys Gly Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro Val
    1250                1255                1260
Leu Lys Gly Val Lys Leu His Tyr Thr
1265                1270
```

<210> 2
<211> 3500
<212> DNA
<213> Artificial Sequence

<220>
<223> pVAX10

<400> 2

```
tgggcttttg ctggcctttt gctcacatgt tcttgactct tcgcgatgta cgggccagat       60

atacgcgttg acattgatta ttgactagtt attaatagta atcaattacg gggtcattag      120

ttcatagccc atatatggag ttccgcgtta cataacttac ggtaaatggc ccgcctggct      180

gaccgcccaa cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc      240

caatagggac tttccattga cgtcaatggg tggactattt acggtaaact gcccacttgg      300

cagtacatca agtgtatcat atgccaagta cgcccccctat tgacgtcaat gacggtaaat      360

ggcccgcctg gcattatgcc cagtacatga ccttatggga ctttcctact tggcagtaca      420

tctacgtatt agtcatcgct attaccatgg tgatgcggtt ttggcagtac atcaatgggc      480

gtggatagcg gtttgactca cggggatttc caagtctcca ccccattgac gtcaatggga      540

gtttgttttg gcaccaaaat caacgggact ttccaaaatg tcgtaacaac tccgccccat      600
```

tgacgcaaat gggcggtagg cgtgtacggt gggaggtcta tataagcaga gctctctggc 660

taactagaga acccactgct tactggctta tcgaaattaa tacgactcac tatagggaga 720

cccaagctgg ctagcctcga gtctagaggg cccgtttaaa cccgctgatc agcctcgact 780

gtgccttcta gttgccagcc atctgttgtt tgcccctccc ccgtgccttc cttgaccctg 840

gaaggtgcca ctcccactgt cctttcctaa taaaatgagg aaattgcatc gcattgtctg 900

agtaggtgtc attctattct gggggggtggg gtggggcagg acagcaaggg ggaggattgg 960

gaagacaata gcaggcatgc tggggatgcg gtgggctcta tggcttctac tgggcggttt 1020

tatggacagc aagcgaaccg gaattgccag ctggggcgcc ctctggtaag gttgggaagc 1080

cctgcaaagt aaactggatg gctttctcgc cgccaaggat ctgatggcgc aggggatcaa 1140

gctctgatca agagacagga tgaggatcgt ttcgcatgat tgaacaagat ggattgcacg 1200

caggttctcc ggccgcttgg gtggagaggc tattcggcta tgactgggca caacagacaa 1260

tcggctgctc tgatgccgcc gtgttccggc tgtcagcgca ggggcgcccg gttctttttg 1320

tcaagaccga cctgtccggt gccctgaatg aactgcaaga cgaggcagcg cggctatcgt 1380

ggctggccac gacgggcgtt ccttgcgcag ctgtgctcga cgttgtcact gaagcgggaa 1440

gggactggct gctattgggc gaagtgccgg ggcaggatct cctgtcatct caccttgctc 1500

ctgccgagaa agtatccatc atggctgatg caatgcggcg gctgcatacg cttgatccgg 1560

ctacctgccc attcgaccac caagcgaaac atcgcatcga gcgagcacgt actcggatgg 1620

aagccggtct tgtcgatcag gatgatctgg acgaagagca tcaggggctc gcgccagccg 1680

aactgttcgc caggctcaag gcgagcatgc ccgacggcga ggatctcgtc gtgacccatg 1740

gcgatgcctg cttgccgaat atcatggtgg aaaatggccg cttttctgga ttcatcgact 1800

gtggccggct gggtgtggcg gaccgctatc aggacatagc gttggctacc cgtgatattg 1860

ctgaagagct tggcggcgaa tgggctgacc gcttcctcgt gctttacggt atcgccgctc 1920

ccgattcgca gcgcatcgcc ttctatcgcc ttcttgacga gttcttctga attattaacg 1980

cttacaattt cctgatgcgg tattttctcc ttacgcatct gtgcggtatt tcacaccgca 2040

tacaggtggc acttttcggg gaaatgtgcg cggaacccct atttgtttat ttttctaaat 2100

acattcaaat atgtatccgc tcatgagaca ataaccctga taaatgcttc aataatagca 2160

cgtgctaaaa cttcattttt aatttaaaag gatctaggtg aagatccttt ttgataatct 2220

catgaccaaa atcccttaac gtgagttttc gttccactga gcgtcagacc cccatcagtg 2280

accaaacagg aaaaaaccgc ccttaacatg gcccgcttta tcagaagcca gacattaacg 2340

cttctggaga aactcaacga gctggacgcg gatgaacagg cagacatctg tgaatcgctt 2400

cacgaccacg ctgatgagct ttaccgcagc tgcctcgcgc gtttcggtga tgacggtgaa 2460

aacctctgac acatgcagct cccggagacg gtcacagctt gtctgtaagc ggatgccggg 2520

agcagacaag cccgtcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg cgcagccatg 2580

acccagtcac gtagcgatag cggagtgtat actggcttaa ctatgcggca tcagagcaga 2640

ttgtactgag agtgcaccat atgcggtgtg aaataccgca cagatgcgta aggagaaaat 2700

accgcatcag gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc 2760

tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg 2820

ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg 2880

ccgcgttgct ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac 2940

gctcaagtca gaggtggcga aacccgacag gactataaag ataccaggcg tttcccctg 3000

gaagctccct cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct 3060

ttctcccttc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg 3120

tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct 3180

gcgccttatc cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac 3240

tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt 3300

tcttgaagtg gtggcctaac tacggctaca ctagaaggac agtatttggt atctgcgctc 3360

tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca 3420

ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat 3480

ctcaagaaga tcctttgatc 3500

<210> 3
<211> 15
<212> PRT
<213> Coronaviridae

<220>
<223> Signal peptide S protein

<400> 3
Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys
1   5   10   15

<210> 4
<211> 28
<212> PRT
<213> Homo sapiens

<220>
<223> 28 aa of C3d

<400> 4
Lys Phe Leu Thr Thr Ala Lys Asp Lys Asn Arg Trp Glu Asp Pro Gly
1   5   10   15
Lys Gln Leu Tyr Asn Val Glu Ala Thr Ser Tyr Ala
            20   25

<210> 5
<211> 112
<212> PRT
<213> Artificial Sequence

<220>
<223> 3C3d

<400> 5
Lys Phe Leu Thr Thr Ala Lys Asp Lys Asn Arg Trp Glu Asp Pro Gly
1   5   10   15
Lys Gln Leu Tyr Asn Val Glu Ala Thr Ser Tyr Ala Gly Ser Gly Gly
            20   25   30
Gly Gly Ser Gly Gly Gly Gly Ser Gly Ser Lys Phe Leu Thr Thr Ala
        35   40   45
Lys Asp Lys Asn Arg Trp Glu Asp Pro Gly Lys Gln Leu Tyr Asn Val
    50   55   60
Glu Ala Thr Ser Tyr Ala Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
65   70   75   80
Gly Ser Gly Ser Lys Phe Leu Thr Thr Ala Lys Asp Lys Asn Arg Trp
                85   90   95
Glu Asp Pro Gly Lys Gln Leu Tyr Asn Val Glu Ala Thr Ser Tyr Ala
            100   105   110

<210> 6
<211> 22
<212> PRT
<213> Artificial Sequence

<220>
<223> 2A peptide P2a

<400> 6
Gly Ser Gly Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val
1   5   10   15
Glu Glu Asn Pro Gly Pro
            20

<210> 7
<211> 21
<212> PRT
<213> Artificial Sequence

<220>
<223> 2A peptide T2a

<400> 7
Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu
1   5   10   15
Glu Asn Pro Gly Pro
            20

<210> 8
<211> 419
<212> PRT
<213> Coronaviridae

<220>
<223> N protein YP_009724397.2

<400> 8
Met Ser Asp Asn Gly Pro Gln Asn Gln Arg Asn Ala Pro Arg Ile Thr
1   5   10   15
Phe Gly Gly Pro Ser Asp Ser Thr Gly Ser Asn Gln Asn Gly Glu Arg
            20   25   30
Ser Gly Ala Arg Ser Lys Gln Arg Arg Pro Gln Gly Leu Pro Asn Asn
        35   40   45
Thr Ala Ser Trp Phe Thr Ala Leu Thr Gln His Gly Lys Glu Asp Leu
    50   55   60
Lys Phe Pro Arg Gly Gln Gly Val Pro Ile Asn Thr Asn Ser Ser Pro
65   70   75   80
Asp Asp Gln Ile Gly Tyr Tyr Arg Arg Ala Thr Arg Arg Ile Arg Gly
                85   90   95
Gly Asp Gly Lys Met Lys Asp Leu Ser Pro Arg Trp Tyr Phe Tyr Tyr
            100   105   110
Leu Gly Thr Gly Pro Glu Ala Gly Leu Pro Tyr Gly Ala Asn Lys Asp
        115   120   125
Gly Ile Ile Trp Val Ala Thr Glu Gly Ala Leu Asn Thr Pro Lys Asp
    130   135   140
His Ile Gly Thr Arg Asn Pro Ala Asn Asn Ala Ala Ile Val Leu Gln
145   150   155   160
Leu Pro Gln Gly Thr Thr Leu Pro Lys Gly Phe Tyr Ala Glu Gly Ser
                165   170   175

```
Arg Gly Gly Ser Gln Ala Ser Ser Arg Ser Ser Ser Arg Ser Arg Asn
            180                 185                 190
Ser Ser Arg Asn Ser Thr Pro Gly Ser Ser Arg Gly Thr Ser Pro Ala
        195                 200                 205
Arg Met Ala Gly Asn Gly Gly Asp Ala Ala Leu Ala Leu Leu Leu Leu
    210                 215                 220
Asp Arg Leu Asn Gln Leu Glu Ser Lys Met Ser Gly Lys Gly Gln Gln
225                 230                 235                 240
Gln Gln Gly Gln Thr Val Thr Lys Lys Ser Ala Ala Glu Ala Ser Lys
                245                 250                 255
Lys Pro Arg Gln Lys Arg Thr Ala Thr Lys Ala Tyr Asn Val Thr Gln
            260                 265                 270
Ala Phe Gly Arg Arg Gly Pro Glu Gln Thr Gln Gly Asn Phe Gly Asp
        275                 280                 285
Gln Glu Leu Ile Arg Gln Gly Thr Asp Tyr Lys His Trp Pro Gln Ile
    290                 295                 300
Ala Gln Phe Ala Pro Ser Ala Ser Ala Phe Phe Gly Met Ser Arg Ile
305                 310                 315                 320
Gly Met Glu Val Thr Pro Ser Gly Thr Trp Leu Thr Tyr Thr Gly Ala
                325                 330                 335
Ile Lys Leu Asp Asp Lys Asp Pro Asn Phe Lys Asp Gln Val Ile Leu
            340                 345                 350
Leu Asn Lys His Ile Asp Ala Tyr Lys Thr Phe Pro Pro Thr Glu Pro
        355                 360                 365
Lys Lys Asp Lys Lys Lys Lys Ala Asp Glu Thr Gln Ala Leu Pro Gln
    370                 375                 380
Arg Gln Lys Lys Gln Gln Thr Val Thr Leu Leu Pro Ala Ala Asp Leu
385                 390                 395                 400
Asp Asp Phe Ser Lys Gln Leu Gln Gln Ser Met Ser Ser Ala Asp Ser
                405                 410                 415
Thr Gln Ala
```

<210> 9
<211> 75
<212> PRT
<213> Homo sapiens

<220>
<223> ubiquitin

```
<400> 9
Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
1               5                   10                  15
Val Glu Pro Ser Asp Thr Ile Glu Asn Val Lys Ala Lys Ile Gln Asp
            20                  25                  30
Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
        35                  40                  45
Gln Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile Gln Lys Glu
    50                  55                  60
Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly
65                  70                  75
```

<210> 10
<211> 27
<212> PRT
<213> Artificial Sequence

<220>
<223> T4 Fibritin foldon trimerization domain

<400> 10
Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys
1 5 10 15
Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu
20 25

<210> 11
<211> 5154
<212> DNA
<213> Artificial Sequence

<220>
<223> SCV-3

<400> 11
gctagcgccg ccaccatgtt cgtcttctta gtgcttctgc cccttgtttc atctcagtgc 60

gtaaatctca ccaccagaac ccagctgcct cccgcctaca cgaactcctt tactaggggg 120

gtatattacc ctgacaaggt attcagatcc tcagtactgc attccacaca ggacctcttc 180

ctccctttct tcagcaatgt aacctggttt cacgccattc atgtctccgg gaccaacggc 240

acgaagaggt ttgacaaccc cgtcctaccc tttaatgacg gtgtgtactt tgcgtcgaca 300

gaaaagagta atatcattcg cggctggatt tttggtacaa ccttagattc taagacgcag 360

tcactcctta tagtgaacaa tgccacaaat gttgttatca aggtgtgtga atttcaattc 420

tgcaatgatc ctttcctggg tgtgtactac cataagaata acaaaagttg gatggaatca 480

gaatttagag tttactcttc agctaacaat tgcactttcg agtatgtttc ccagcctttc 540

ctcatggatc tggaggggaa acaagggaat ttcaaaaatc tgcgcgaatt cgtgttcaaa 600

aatatagatg gatatttcaa gatctacagc aaacacacac ctatcaacct tgttcgagat 660

ctgccccagg gattttctgc cctggaacca ctcgtggacc tgcccatcgg tattaacatc 720

actcgattcc agacactact ggccctccac cggagctacc tgacccctgg ggattcttct 780

agtgggtgga cagcaggcgc agctgcttac tacgtgggct atttgcagcc tcggaccttc 840

ttactgaaat acaatgaaaa tggcaccatc acagatgctg tggactgtgc cttagaccct 900

ctctccgaaa ccaaatgtac ccttaagagc ttcacagtcg aaaaagggat ttaccaaacg 960

tcgaacttca gggtgcagcc cactgagagc atcgtcagat ttcccaatat tacgaacctg 1020

tgtcccttcg gagaagtgtt caacgctact cgatttgcct ctgtgtacgc atggaatcgc 1080

aaaaggatct cgaattgcgt agccgactac tccgttctgt ataattccgc atcctttagc 1140

acttttaaat gctacggggt cagccctact aaactgaacg acctctgctt caccaacgtc 1200

tacgctgatt cttttgtcat ccgtggggat gaagtgaggc agatagcccc aggtcagact 1260

ggaaagatcg cggattataa ttacaaactg ccagacgact tcactggctg cgtgatcgcc 1320

tggaacagta acaacctcga cagcaaagtg ggcggaaact acaactattt atatcgcttg 1380

tttagaaagt ccaacttgaa gccttttgaa agagacattt ctactgagat ttaccaggct 1440

gggtcaacac cctgtaacgg cgtggaaggc tttaactgtt acttcccctt acaaagctat 1500

ggcttccagc caaccaacgg agttgggtat cagccctaca gggtcgtcgt gttgagtttc 1560

gagctcttac atgcacctgc caccgtctgc ggccccaaga agtctacgaa cctggtgaag 1620

aacaagtgtg tcaatttcaa ctttaatgga ctaactggga caggcgtgct gacggagtca 1680

aacaagaaat tcctgccatt tcagcagttt ggtcgggaca ttgcagacac aactgacgcg 1740

gtgcgggacc ctcagacgct tgaaatatta gacataacac cttgtagctt tggcggagtg 1800

tctgtcatca ccccgggaac taacaccagc aaccaggtgg cagtcctata tcaggacgtc 1860

aattgcacgg aggttcccgt ggcaatacac gcggaccaac tcacaccgac ttggagagtg 1920

tattccacag gtagcaacgt ctttcaaact cgggctggtt gcctgattgg cgcagagcac 1980

gtgaacaaca gctatgagtg tgacattccc ataggcgcag ggatttgtgc ttcataccag 2040

actcagacta acgcccgtag tgtagccagc cagtccatca ttgcatacac tatgtcactg 2100

ggagcagaga actctgttgc ctacagtaac aatagtattg ctattccgac gaactttacc 2160

atcagtgtga ctaccgagat cctgccagtc tcgatgacca agacctccgt agattgtacc 2220

atgtatatct gcggggacag caccgagtgc tctaatttgt tgctgcagta tgggtccttc 2280

tgtacgcagt tgaatcgggc cttgacaggc atcgctgtcg agcaggataa aaacactcaa 2340

gaggtctttg cacaggttaa gcagatttat aagacacctc ccataaaaga cttcggtggc 2400

tttaacttct cacaaatctt accagaccca agtaagccaa gtaaacgttc tttcatcgag 2460

gacctcctct ttaataaggt gacccttgct gatgcaggtt tcatcaagca gtacggtgac 2520

tgcctgggag acatagcagc ccgggatttg atctgcgcac agaagttcaa cggactaaca 2580

gtgctcccac ctttgctgac ggatgagatg attgcccagt acacatcagc tctcctggcg 2640

ggcaccatta cctcaggatg gacctttgga gcaggagctg cgctccagat cccgtttgcc 2700

atgcagatgg cctataggtt caacggcata ggcgttacac agaacgtgct gtatgagaat 2760

cagaagctca tcgccaacca gttcaattct gccataggaa aaatccagga ttccctttct 2820

tcaactgcct cagccctcgg gaagctacaa gatgtggtca atcagaacgc tcaggcccta 2880

aatactcttg tgaagcagct gtcgtcaaac tttggggcca tcagcagtgt tctgaacgat 2940

atcctgtcac ggctggatcc accagaggcc gaggtccaga ttgaccgact aattaccggc 3000

aggctgcaat ctcttcagac ttacgtcaca cagcagctga tccgagctgc agaaatcaga 3060

gccagtgcaa acctggctgc aacgaaaatg tcagagtgtg tgcttggaca gtcaaaaaga 3120

gtagactttt gtggcaaggg ataccatttg atgtcctttc cacaatccgc tccgcatgga 3180

gttgtgtttt tgcatgttac ctatgtccct gcccaggaaa agaactttac cacagctccc 3240

gccatatgcc acgatgggaa ggcccacttt ccacgcgagg gagtcttcgt gtctaatggt 3300

actcactggt ttgtgaccca acgcaatttt tacgaacctc aaatcattac cacagacaat 3360

acattcgtat caggtaactg cgatgtcgtg attggaatcg tgaataatac cgtttatgac 3420

ccccttcaac cagaactgga ctcatttaag gaagagctcg acaaatattt caaaaaccac 3480

actagcccgg atgtggattt gggcgatata agcggcatca atgctagtgt ggtgaacatc 3540

cagaaggaga ttgaccggct aaatgaagtg gccaagaacc tgaatgagag cttgattgac 3600

ctccaggagc ttggtaaata cgaacaatat atcaagtggc catggtatat ctggctggga 3660

ttcattgccg gcctgatcgc catcgtgatg gtaacaatca tgctgtgctg catgactagc 3720

tgctgctcct gtctgaaggg ctgttgctct tgtggtagtt gttgtaagtt tgatgaagac 3780

gatagcgagc cagttctcaa aggagtgaag ttgcactaca caggatccgg tgaaggccgt 3840

gggtcccttt taacgtgtgg tgacgtagaa gaaaacccag gcccgctcat gagcgacaat 3900

ggaccacaga atcagcgaaa cgcccccagg attacatttg gcggtcccag tgactcgacc 3960

ggcagtaacc agaatgggga gaggtctggg gcgaggtcta agcagaggag gccacaaggc 4020

ctgccaaata atactgcgag ctggttcaca gcactgacac agcatggcaa agaggatctc 4080

aaattcccaa gaggtcaagg cgttccaatc aatactaaca gctctcctga tgatcagatc 4140

ggctactata gacgggccac gcgacgaatt cgcggagggg atggtaagat gaaggatttg 4200

tcaccacgct ggtactttta ttaccttggg acaggacctg aggcagggct tccatacggc 4260

gctaataaag acggcattat ctgggttgcc acagaggggg cactgaatac ccctaaggat 4320

cacatcggca ctagaaaccc tgcgaataac gcagccatag tgctccagtt accacaggga 4380

accacactgc ctaagggttt ctatgctgag gggagcaggg gaggaagtca ggcatcgtcc 4440

agatccagtt ctagatcgcg gaacagcagt cggaattcca ctcctggctc atcaagagga 4500

acctcaccag ccaggatggc agggaacgga ggagacgcag cgctggctct gctcctcctg 4560

gatagactta accaactgga gagcaagatg tccggaaagg gccagcaaca acaaggtcag 4620

accgtgacga agaaatcagc cgccgaagct agtaaaaagc cccggcagaa acggactgct 4680

acaaaggcat acaatgtcac acaagccttc ggtcgcaggg ggccagaaca aacccagggg 4740

aatttcggtg atcaggagct gatacggcag ggaacagact ataagcactg gccgcagatc 4800

gctcagttcg caccctccgc aagcgccttt ttcggcatgt ctcgcatagg aatggaggtg 4860

acgcccagtg gcacatggct aacttataca ggcgccatta aacttgatga caaagacccg 4920

aacttcaagg accaagtcat cttgctgaac aagcatattg atgcttataa gactttccct 4980

cctactgagc cgaaaaagga taaaaaaaag aaagctgacg agacacaggc gcttccccag 5040

cgtcaaaaga aacagcagac agtgacgctg ttgcccgctg ctgatttgga cgatttcagc 5100

aagcaactac agcagtccat gtccagcgct gattccaccc aagcttaact cgag 5154

<210> 12
<211> 2619
<212> DNA
<213> Artificial Sequence

<220>
<223> SCV-30

<400> 12
gctagcgccg ccaccatgtt cgtgtttctg gtgttactcc cattggtgtc tagagtgcag 60

ccgaccgaat ccatcgtgcg cttccctaat atcactaatc tgtgcccctt tggggaagtt 120

tttaatgcca ctaggttcgc atctgtgtat gcatggaacc gcaagcgaat tagcaactgc 180

gtcgctgact attcagtact gtataactct gccagtttct ccactttcaa gtgctatggg 240

gtatccccca ccaagctcaa tgacctttgc tttacaaacg tgtacgccga ttcctttgtg 300

attcgaggag atgaagtgag gcagatagcc ccagggcaga ctggaaagat cgccgactac 360

aattacaaac tgccggatga tttcaccggg tgtgtgattg cttggaatag caataatctg 420

gattctaaag tgggaggcaa ctataattac ctatacagac tgttccggaa gtctaaccta 480

aagccatttg aaagagacat ttcaaccgaa atataccaag cgggctctac cccgtgtaat 540

ggcgttgagg gcttcaattg ctatttccct cttcagtcct acgggttcca gcctacaaat 600

ggtgtcggct accagcccta cagggttgtc gtcttaagtt tcgagctgct gcatgcaccc 660

gccaccgtct gtggcccaaa aaagagtact aacttggtaa aaaacaagtg tgtgaacttt 720

ggatccaaat ttcttaacac cgctaaggat cgtaataggt gggaggaacc agatcagcag 780

ctctacaacg ttgaagccac ctcatacgca ggatcgggtg ggggaggaag cggcggcgga 840

ggcagcggct ctaagtttct gaacacagca aaggacagga atcgttggga ggaacctgac 900

cagcagctct acaacgtcga ggcaaccagt tacgcgggta gtggaggggg cgggtctgga 960

ggcggaggaa gcggttccaa gtttctcaac acagctaaag accgcaaccg gtgggaggag 1020

ccggaccagc aattatataa tgtggaggcc acatcgtatg ccgggagtgg cgagggcagg 1080

ggttcactgc tgacttgtgg agatgtggaa gagaacccag ggccactgat gcagatcttt 1140

gtcaagacac ttacaggaaa aaccattacc ctggaggtgg agccaagtga cacaattgag 1200

aacgttaagg caaaaatcca ggacaaggag ggcatccccc ctgaccaaca acgacttata 1260

tttgcaggta agcaactaga agatggccgc acactttcag actacaacat ccaaaaagag 1320

tctacactcc acctggtctt gagattaaga ggaatgtcag acaacgggcc gcaaaaccag 1380

cggaatgctc ccagaatcac tttcgggggg cctagtgact caaccggtag caaccagaac 1440

ggggagcgtt ccggcgccag gtcaaagcag cgtcggcctc agggcctccc caacaacact 1500

gccagctggt tcactgcttt aacccagcat ggtaaagaag acctgaagtt cccccgaggg 1560

cagggtgtac ctattaacac gaacagctct cctgatgatc agatcggcta ttatcgaaga 1620

gccacaagga gaatacgcgg aggcgacggg aaaatgaaag acctctcccc caggtggtac 1680

ttctactatc tgggcacggg gcctgaggcc gggctgccat acggcgcgaa taaagatgga 1740

atcatctggg tagctacgga gggcgcatta aatactccca aggaccatat tggaaccagg 1800

aatccagcta acaatgcagc tattgttttg cagctccccc aggggacaac tctgcccaag 1860

ggattctatg ctgaaggtag ccgtggaggc agtcaggcgt ctagtcggtc ctctagtaga 1920

agccgaaatt ctagccggaa tagcacccca ggttcctctc ggggcacatc acctgctcgg 1980

atggcaggta acgggggaga tgctgccctg gctctcctgc tgttggatag attgaatcag 2040

ttggagtcta agatgtccgg caagggacaa cagcagcaag gtcagacagt cacgaagaaa 2100

agcgcagcgg aggccagcaa gaaacctcgc cagaagagaa cagccaccaa agcctacaat 2160

gtgacccagg ccttcggccg gcgggggccc gaacagactc aaggaaattt tggtgaccag 2220

gaactgatac gacaggggac agattataag cactggcctc aaatcgctca gttcgctcct 2280

tctgccagtg ccttctttgg tatgagcagg atcggcatgg aagttactcc ttccggcacg 2340

tggcttactt acacaggagc catcaaactt gatgacaaag atccaaactt taaggatcaa 2400

gtcatcttac taaataagca cattgatgct tataaaactt tcccacccac cgaaccaaaa 2460

aaggacaaaa aaaaaaaggc cgacgagacc caagctttgc ctcagcgcca gaagaagcag 2520

cagaccgtga cgctcttgcc agcagcagac ctggacgatt tttcgaaaca gcttcagcag 2580

tccatgtcga gcgcggactc aacacaagca taactcgag 2619

<210> 13
<211> 5787
<212> DNA
<213> Artificial Sequence

<220>
<223> SCV-42

<400> 13
gctagcgccg ccaccatgtt cgtgttcttg gtgctcttac cactggtctc ctcacagtgc 60

gtgaacctga caacacgaac ccaactgcca ccagcataca cgaattcttt tacccgaggt 120

gtgtattatc ctgacaaagt ctttagaagt tcagtcctcc attccactca agatttgttt 180

ctcccgttct tctcgaatgt gacatggttt catgctattc atgtgtcggg gacaaatggc 240

accaagcggt ttgacaatcc tgtcttacct ttcaatgatg gggtttattt cgcatcaaca 300

gaaaagagca acattatacg gggctggata ttcgggacca ccctggattc caaaacccag 360

tctctgctaa ttgtgaataa cgccacaaat gtagttatca aagtgtgtga gttccagttt 420

tgcaacgatc cattcttggg tgtctactat cacaagaaca ataaatcatg gatggagagc 480

gagtttaggg tctattcttc tgcaaacaac tgtacgttcg agtacgtctc ccagccattc 540

ctcatggacc tggagggcaa acaggggaac ttcaaaaatc ttagagaatt cgtcttcaaa 600

aacatcgatg ggtattttaa gatctacagc aaacataccc ccatcaatct cgtgcgggac 660

cttcctcagg gcttttccgc acttgaacct ctcgtcgacc ttcccatcgg aatcaatatt 720

acaaggttcc agacactact cgcccttcac agatcctacc tgacgccagg cgattcctcc 780

tctggctgga cagcgggagc cgctgcctat tatgtgggct atttgcagcc tcgcacattc 840

ctcttgaaat acaacgagaa tgggaccatt acagatgcag tggactgtgc acttgaccct 900

ctgtctgaga caaagtgtac actgaaatcc ttcacggtgg agaagggtat ctaccagact 960

tctaacttcc gggttcagcc cactgagtcc atagtccggt tcccaaacat aactaaccta 1020

tgcccatttg gggaggtttt caatgctacc cgcttcgcct cagtttatgc ttggaacagg 1080

aagagaatta gcaactgtgt ggccgactat agtgtgctgt acaactccgc cagcttctcc 1140

accttcaaat gctatggtgt gtcacctacc aaactgaatg atctgtgctt tacaaatgtt 1200

tacgctgata gttttgtaat cagaggtgac gaagtaagac agattgcccc cggacaaact 1260

gggaagatcg cggattacaa ttacaaactt cccgacgact ttacgggctg cgtgatcgcc 1320

tggaatagca ataatctgga cagcaaggtg ggcgggaatt ataactacct ctaccgcctc 1380

ttcaggaaaa gtaacctgaa gccctttgaa cgtgatataa gcactgagat ataccaagct 1440

ggcagcacgc cttgtaacgg agtggaagga ttcaactgct acttcccttt acaaagttat 1500

ggcttccagc caaccaacgg ggtgggttat cagccataca gggttgttgt gctaagtttt 1560

gagttattac atgcacctgc tactgtctgc ggcccaaaaa agtccaccaa cctggttaaa 1620

aacaaatgtg tcaattttaa tttcaatggc ctaacaggta cgggtgtctt gactgaaagt 1680

aataaaaaat ttcttccatt tcagcagttt gggagagaca tcgcggacac aaccgatgct 1740

gtgagagatc cgcagacact ggagattttg gacatcaccc cgtgtagctt tggcggtgta 1800

tctgtgataa ctcctgggac taatacaagc aaccaggttg ccgtgctgta tcaagacgtg 1860

aactgcactg aggtccctgt ggcaatccat gctgaccagc tgacgcccac ctggagggtt 1920

tactcgaccg gaagtaatgt attccaaact cgggcaggct gcctcatcgg agcagaacac 1980

gtgaacaact cctatgagtg tgatataccc atcggtgcag gaatttgtgc cagctaccag 2040

acacagacaa acagcccagg gagcaagttc ctgaacacgg ccaaggatcg aaaccggtgg 2100

gaggaaccag atcaacaatt gtataacgtt gaggccacgt cttacgccgg tagtggcggg 2160

ggtgggtcag gcggtggagg aagcggatca aagttcctca atactgcgaa ggacaggaac 2220

cgctgggaag agccggacca gcagttgtac aatgtggaag ccacaagcta tgctggaagt 2280

ggtggcggag gctcaggagg agggggatcg gggtctaagt ttctgaacac tgcaaaggac 2340

agaaataggt gggaggagcc tgaccagcaa ctgtacaacg tggaagcaac aagctacgcc 2400

ggctcagggg agggaagagg gagcctcctg acatgtgggg acgtagagga gaaccccggg 2460

ccgctcatgc aaattttcgt aaagactctg acaggcaaga cgatcacact ggaagtcgag 2520

ccatctgaca ccatcgaaaa tgtaaaagcc aaaatacaag acaaagaagg aattccccca 2580

gaccaacagc gactcatctt tgcaggcaag cagctggagg acggacggac cttgtctgac 2640

tacaatattc agaaggagtc tactttgcat ctggttcttc ggctaagagg catgagtgac 2700

aacggcccac aaaaccaaag gaatgccccg aggatcacct tcggcggacc ctccgatagc 2760

acaggcagta accagaatgg ggaaaggagt ggagccaggt caaaacagcg tcggccacag 2820

ggactgccca acaacaccgc ttcctggttc acagcactga cccagcatgg gaaagaggat 2880

ttgaaattcc ctcgtggtca aggtgtcccc attaacacga actcatcccc tgatgaccag 2940

atcggttact acagaagagc aaccaggagg atacgcggcg gagacggcaa aatgaaggat 3000

ctgtctccgc ggtggtattt ctactacctt ggcaccggcc ctgaggctgg actgccttac 3060

ggagcaaaca aagatggaat tatctgggtg gccaccgagg gtgctctcaa cacgcccaag 3120

gaccacatag gtactcgcaa cccagccaat aatgctgcca tcgtgctgca gctgcctcag 3180

ggaacaacac ttcccaaagg attctatgct gagggcagcc gcggaggctc gcaggcctcg 3240

tccaggtcca gttcaaggtc tcgaaattca tcccggaact ctactccagg ttctagtaga 3300

ggaacaagtc cggcaaggat ggccgggaat ggtggagatg ccgctctggc cctcctacta 3360

cttgacagat taaaccagct ggaaagcaaa atgagtggaa agggccagca gcaacagggc 3420

cagacagtga caaagaagtc agctgcggag gcctccaaga agcccaggca gaaacgcact 3480

gctactaaag cttataatgt cacccaagct ttcgggcgca ggggccccga acaaacacag 3540

ggtaattttg gagatcagga actcattcgt cagggcacgg attataaaca ctggccccag 3600

attgcccagt ttgcaccatc tgccagcgcc ttttttggga tgagtcgcat tggaatggaa 3660

gttactcctt ctggcacctg gctgacatat actggagcca ttaagctgga cgacaaggat 3720

cccaatttca aggaccaagt aatactgctc aataagcaca ttgatgccta caaaaccttt 3780

ccgcccactg agcccaagaa agataagaag aagaaagcag atgagaccca ggctctcccc 3840

cagagacaga agaaacagca gacggttacc ctgctgcctg ccgcggacct ggatgacttt 3900

agcaagcaac ttcagcagtc catgtcgtcc gctgatagta ctcaggccgg gtctggggag 3960

gggcggggca gcctgcttac ttgtggtgat gttgaagaga accctggacc ccttagtgtg 4020

gcatcacaat ctattattgc atataccatg tccctcggcg cggagaacag cgtggcctac 4080

agtaacaact ccatagccat tcctaccaac tttaccatct ctgtcaccac agaaattctc 4140

cccgtgtcca tgaccaagac ctcagtggac tgcactatgt acatttgcgg cgattccacc 4200

gagtgctcca acttgcttct gcagtacgga tcgttctgta cacaattgaa tagagcccta 4260

accggcatcg ctgtcgaaca ggataagaac acccaagaag ttttcgccca ggtgaagcag 4320

atttataaga cacccccтat taaggacttc ggaggtttca acttcagcca gatcctgcct 4380

gatccatcca agccttccaa gcggtccttc atcgaggacc ttcttttcaa caaggtcacc 4440

ttagcggacg cagggtttat aaagcaatat ggagattgcc ttggtgatat cgctgcccga 4500

gacctgatct gcgctcagaa atttaatggt ttaacggtgt tgccaccatt attgacagac 4560

gagatgatcg cacagtacac tagtgcgctg ctggctggga caatcacatc tggatggacg 4620

tttggggctg gtgcagccct ccagatccct tttgccatgc aaatggcgta ccgtttcaac 4680

ggcatcggcg tgactcagaa cgtcctatat gagaatcaga agcttattgc taaccagttc 4740

aactccgcta ttggtaagat tcaggacagc ctcagcagca ctgcgtctgc gctgggcaag 4800

ttgcaagacg tggtgaatca gaatgcacag gccttaaaca cactggttaa acaactgtcc 4860

tcgaacttcg gggctatcag ctctgtcctc aatgacatcc tttccagact cgacccaccc 4920

gaggcagaag tccagatcga tcgactgatc acaggcagac tgcagagcct gcagacctat 4980

gtcacccagc agctcatacg agctgctgag atccgagctt ctgcaaacct ggccgcaact 5040

aagatgtctg aatgtgtcct gggtcagagc aagcgggtgg atttttgtgg gaaaggctac 5100

cacctaatgt catttccaca gtctgcccct cacggcgtgg tatttttgca cgtaacctac 5160

gtgccagcgc aagagaagaa ctttaccacc gccccagcaa tctgccatga tggaaaggcc 5220

cacttccccc gtgaaggggt gtttgtatcc aatggaactc actggttcgt gacacagagg 5280

aacttttatg aaccacagat catcacaacc gacaacacct ttgtgtctgg taattgtgac 5340

gtggtaattg ggattgtaaa taacactgtg tatgatcctt tacaaccaga actggattca 5400

tttaaagaag aattggacaa gtactttaag aaccacacta gccccgatgt tgatttagga 5460

gacatctcag gcatcaacgc tagtgtcgta aatattcaga aggagatcga ccgcctgaat 5520

gaagtcgcta aaaatctcaa tgagagtctg attgacttgc aggagctggg caaatacgag 5580

cagtatatca agtggccatg gtacatctgg ttggggttta tagctggtct gatcgctatc 5640

gtgatggtga ctataatgct gtgctgcatg accagctgtt gcagttgcct aaagggatgc 5700

tgttcctgtg gcagctgttg caagtttgat gaagatgatt cagaacccgt ccttaagggg 5760

gttaaattgc actacaccta actcgag 5787

<210> 14
<211> 5427
<212> DNA
<213> Artificial Sequence

<220>
<223> SCV-53

<400> 14
gctagcgccg ccaccatgga tgatcagcgc gatctaattt ctaaccatga acaattgccc 60

attttaggca acagacctag agagccagag cgttgctctc gatcgcagtg tgtcaacctg 120

actacccgca cccaactccc accagcttac acaaatagct ttaccagggg tgtctattac 180

cctgataaag tgtttcgtag ctccgtcctc catagcacac aagatttgtt cctgcccttc 240

ttctcgaatg tgacttggtt ccacgccatc catgtgtcag gaacaaacgg aacgaaaagg 300

tttgacaacc cagtgttgcc cttcaatgac ggcgtttact ttgcgagtac cgagaagagc 360

aacattatcc gaggctggat cttcgggaca actctcgaca gcaagactca gtcgttgctc 420

attgttaaca atgcaaccaa cgtggtcatc aaggtgtgtg agttccagtt ctgtaatgat 480

ccgtttctgg gggtttacta ccacaaaaat aataaatctt ggatggagtc agagttccgg 540

gtgtattctt ctgcaaataa ctgcacattt gaatatgtta gtcagccttt cctcatggac 600

ttggaaggca agcaggggaa cttcaaaaat ctcagagagt ttgtgtttaa gaatattgat 660

ggttacttca agatctactc caaacacact cctataaacc ttgtgcggga tctacctcaa 720

ggcttctcag cactggagcc tttagttgac ctccccattg gcattaacat tactaggttt 780

caaacacttc tggccctaca ccggagctac cttacccctg gtgattcctc aagtggatgg 840

acagctgggg cagcagctta ctatgtggga tacctgcagc caagaacttt cttgttgaag 900

tacaatgaga acggaaccat cacggatgcc gtggattgtg cattggatcc cttgagtgaa 960

accaaatgta cactcaagtc tttcacagtt gaaaaaggaa tataccaaac ctccaacttc 1020

cgagtgcagc caacagaatc cattgttcgg ttccccaata tcacgaatct ctgcccattt 1080

ggggaggttt ttaatgccac aaggtttgcc tccgtgtatg cctggaacag gaagaggatc 1140

tctaattgtg tagctgacta ctcagtccta tacaactcgg cctctttttc cacttttaag 1200

tgctacggcg ttagtcctac caaacttaac gacctgtgtt ttaccaatgt ttacgcagac 1260

agctttgtga tccgggggga cgaggttagg cagatagctc ctggtcaaac aggtaagatc 1320

gccgattaca actataaact gcctgacgac ttcacaggct gtgtgatcgc ctggaattcg 1380

aacaacctgg attccaaggt tggcggcaat tacaactacc tgtaccgtct cttcagaaaa 1440

agcaatctca agcccttga gagggatata agcactgaga tctaccaggc cggtagcact 1500

ccctgcaacg gcgtggaagg attcaactgc tattttcctt tgcaatccta tggttttcag 1560

ccgactaatg gagtgggcta ccaaccctat agggtggttg tcttatcgtt tgagctccta 1620

catgctccag ccacagtgtg cgggccaaag aagtccacca accttgtgaa gaacaaatgc 1680

gtgaacttca acttcaacgg gctcactgga actggggtgc tgacagagtc caacaaaaag 1740

tttctaccct tccaacaatt cggcagggac attgccgata caactgacgc agttcgcgac 1800

cctcagacac ttgagatctt agatataacg ccctgctcct ttggaggcgt ctccgttatt 1860

acccctggca ccaacacaag taaccaggtg gcggtgctct accaggatgt gaactgtaca 1920

gaagtacctg tggctatcca cgctgaccaa ctgaccccaa cctggagagt gtacagcacg 1980

ggatccaatg tcttccagac aagagcagga tgcttgatag gcgccgagca cgtcaacaac 2040

agttatgagt gtgatatccc catcggagct ggcatctgcg catcatatca gacccagacg 2100

aatgcccgtt ctgtagccag ccagtctatc attgcctata ccatgtcact gggagcagaa 2160

aactctgtgg catacagtaa caacagcata gccattccta ccaactttac aatcagtgtc 2220

accaccgaaa tcctgccggt gtccatgact aagaccagcg tggactgtac tatgtacatt 2280

tgcggcgact ctactgagtg ctcaaatctg ctgctacagt atggctcatt ttgcacgcag 2340

ctaaaccggg ccttaactgg cattgctgtg gagcaggaca agaatactca agaggtgttc 2400

gcccaagtta agcagatcta caaaacaccc cccatcaaag actttggtgg ctttaacttc 2460

tctcaaatac ttcccgatcc tagcaagccg tctaagaggt ctttcatcga agatcttttg 2520

tttaacaaag tcaccctggc agatgccggt tttattaagc agtatgggga ctgcctggga 2580

gacatagctg cgagagacct gatctgtgct cagaagttca atggcctgac tgtactccct 2640

ccattgctaa ccgatgagat gatagcacag tacacttctg ccctgctggc tggcacaatc 2700

acctcaggct ggaccttcgg agccggtgcc gccttgcaga tcccattcgc aatgcagatg 2760

gcttaccgat ttaatgggat cggagttacc cagaacgtcc tgtatgagaa ccaaaaattg 2820

attgctaatc agttcaacag cgccatcggc aagatccagg atagcttgtc aagtacagcc 2880

tccgctctgg gaaagctgca ggacgtcgtg aatcagaatg ctcaggcact taacacactc 2940

gtgaaacagc tcagttctaa tttcggagct atttccagtg tccttaatga catcctgtca 3000

cgcctggatc cccctgaagc tgaagtgcag atcgaccgat taattaccgg gcgtctgcag 3060

tctcttcaga cttatgtcac tcagcaatta atcagggcgg ccgaaattcg ggcatccgcc 3120

aacttagcgg ccacaaagat gtcagaatgt gtactgggcc aaagtaaaag agtagatttt 3180

tgtggaaaag gatatcatct gatgtctttc cctcagagcg cccctcatgg ggtcgtattc 3240

ttacacgtga cctatgtccc agcccaggag aagaacttca cgacagctcc cgcgatctgt 3300

catgacggga aagcgcattt ccctcgggag ggagtatttg tttcaaacgg aactcactgg 3360

tttgtcacac agaggaactt ttatgaacct cagataatta caacggataa cacctttgtc 3420

tccggaaatt gtgacgtagt gataggtatt gtaaataata cagtctatga cccattacag 3480

cctgagctcg acagtttcaa agaagaactg gacaaatact tcaagaacca caccagtcca 3540

gatgtggatc tgggagacat ctccggcatc aatgcttctg tggtgaacat ccagaaggaa 3600

attgaccgct tgaatgaagt cgccaaaaac ctgaacgaaa gcttaattga tctgcaggaa 3660

ctggggaaat atgagcagta tataaagtgg ccctggtaca tctggctggg attcattgct 3720

ggtctgatcg cgattgtgat ggtgacgatc atgctgtgtt gcatgacatc ctgttgttct 3780

tgtctgaaag gctgctgctc ctgtggtagc tgctgcaaat ttgatgaaga cgatagcgag 3840

ccggtcctga agggtgttaa gctgcattat actgggagcg gagagggtag ggggtcgttg 3900

ctgacctgcg gagatgttga ggagaatcca ggccccatgc agatctttgt gaaaaccctg 3960

accggtaaga ccatcactct ggaggtggag ccctcggaca ccattgaaaa tgtaaaggcc 4020

aagattcagg acaaggaggg cattccacct gaccagcagc gacttatttt tgcagggaag 4080

cagctggagg atgggcgcac tctgtcggat tataacatac aaaaggagag tacactccac 4140

ctggtgctcc ggctgcgagc catgagcgac aatggacccc aaaaccagag aaacgcaccc 4200

cggatcacat tcggagggcc gtcagattca accggctcca atcaaaacgg agagcggtcc 4260

ggcgcgagaa gcaagcaaag aaggccccag ggtctgccaa acaacactgc atcctggttc 4320

accgcactta ctcagcacgg caaagaagac cttaagttcc ctaggggcca gggggtacct 4380

atcaacacaa actctagccc cgacgaccag ataggctatt atcgccgcgc cactcggcgg 4440

attagaggag gtgatgggaa aatgaaggac ctgtcgccaa gatggtattt ctactacctc 4500

ggaacagggc cagaggcagg cctcccctac ggggccaata aggatggaat tatatgggtg 4560

gctacagagg gtgctctcaa tactcccaaa gaccacatag ggactcggaa tcccgccaac 4620

aatgcagcca tcgtcctgca actacctcaa gggactaccc tgccaaaagg cttctatgct 4680

gagggcagcc gaggaggctc tcaagccagt tcccgttcaa gcagccgcag cagaaattca 4740

agcagaaatt ctacaccggg ttcttcacgt ggtaccagcc cggctaggat ggcgggcaac 4800

ggtggggacg ctgcattggc tcttttactt ttggacagat tgaaccagct ggaatccaaa 4860

atgtctggaa aagggcagca acagcaggga cagactgtga ccaagaaaag tgccgcagag 4920

gccagcaaga aacctcgcca gaagagaacc gctacaaaag cctacaacgt cacccaggca 4980

tttggccggc gcgggccaga acagacacaa gggaattttg gggatcagga gctgataaga 5040

cagggcacag actacaagca ctggccacag atcgctcagt ttgctccaag tgctagtgcc 5100

ttctttggta tgtcccggat agggatggaa gtgactcctt ctggaacatg gctcacctac 5160

accggcgcga ttaaactgga tgacaaggat ccaaatttca aggaccaggt cattctccta 5220

aataagcata tcgatgctta taagacgttc ccacccacag agcccaaaaa agataagaag 5280

aaaaaggcag acgaaaccca agcactgccg cagaggcaga agaagcagca gacagtgacg 5340

ctgctgccag ctgcagacct ggatgacttc tccaaacagc tccagcagtc tatgagttca 5400

gcggattcta ctcaggctta actcgag 5427

<210> 15
<211> 140
<212> PRT
<213> Mus musculus

<220>

<223> EDL09808 (invariant chain)

<400> 15
Met Asp Asp Gln Arg Asp Leu Ile Ser Asn His Glu Gln Leu Pro Ile
1   5   10   15
Leu Gly Asn Arg Pro Arg Glu Pro Glu Arg Cys Ser Arg Gly Ala Leu
            20   25   30
Tyr Thr Gly Val Ser Val Leu Val Ala Leu Leu Leu Ala Gly Gln Ala
        35   40   45
Thr Thr Ala Tyr Phe Leu Tyr Gln Gln Gln Gly Arg Leu Asp Lys Leu
    50   55   60
Thr Ile Thr Ser Gln Asn Leu Gln Leu Glu Ser Leu Arg Met Lys Leu
65   70   75   80
Pro Lys Ser Ala Lys Pro Val Ser Gln Met Arg Met Ala Thr Pro Leu
                85   90   95
Leu Met Arg Pro Met Ser Met Asp Asn Met Leu Leu Gly Pro Val Lys
            100   105   110
Asn Val Thr Lys Tyr Gly Asn Met Thr Gln Asp His Val Met His Leu
        115   120   125
Leu Thr Val Ser Leu Asp Pro Trp Ser Thr Arg Ser
    130   135   140

<210> 16
<211> 72
<212> DNA
<213> Artificial Sequence

<220>
<223> SV40 DNA nuclear targeting sequence

<400> 16
ggtgtggaaa gtccccaggc tccccagcag gcagaagtat gcaaagcatg catctcaatt   60

agtcagcaac ca   72

<210> 17
<211> 28
<212> PRT
<213> Mus musculus

<220>
<223> C3d (derived from P01027.3)

<400> 17
Lys Phe Leu Asn Thr Ala Lys Asp Arg Asn Arg Trp Glu Glu Pro Asp
1   5   10   15
Gln Gln Leu Tyr Asn Val Glu Ala Thr Ser Tyr Ala
            20   25

<210> 18
<211> 112

<212> PRT
<213> Mus musculus

<220>
<223> 3C3d (P01027.3)

<400> 18
Lys Phe Leu Asn Thr Ala Lys Asp Arg Asn Arg Trp Glu Glu Pro Asp
1   5   10   15
Gln Gln Leu Tyr Asn Val Glu Ala Thr Ser Tyr Ala Gly Ser Gly Gly
            20   25   30
Gly Gly Ser Gly Gly Gly Gly Ser Gly Ser Lys Phe Leu Asn Thr Ala
        35   40   45
Lys Asp Arg Asn Arg Trp Glu Glu Pro Asp Gln Gln Leu Tyr Asn Val
    50   55   60
Glu Ala Thr Ser Tyr Ala Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
65   70   75   80
Gly Ser Gly Ser Lys Phe Leu Asn Thr Ala Lys Asp Arg Asn Arg Trp
                85   90   95
Glu Glu Pro Asp Gln Gln Leu Tyr Asn Val Glu Ala Thr Ser Tyr Ala
            100   105   110

<210> 19
<211> 9284
<212> DNA
<213> Artificial Sequence

<220>
<223> pVAX10-SCV-53

<400> 19
tgggcttttg ctggcctttt gctcacatgt tcttgactct tcgcgatgta cgggccagat 60

atacgcgttg acattgatta ttgactagtt attaatagta atcaattacg gggtcattag 120

ttcatagccc atatatggag ttccgcgtta cataacttac ggtaaatggc ccgcctggct 180

gaccgcccaa cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc 240

caatagggac tttccattga cgtcaatggg tggactattt acggtaaact gcccacttgg 300

cagtacatca agtgtatcat atgccaagta cgcccccctat tgacgtcaat gacggtaaat 360

ggcccgcctg gcattatgcc cagtacatga ccttatggga ctttcctact tggcagtaca 420

tctacgtatt agtcatcgct attaccatgg tgatgcggtt ttggcagtac atcaatgggc 480

gtggatagcg gtttgactca cggggatttc caagtctcca ccccattgac gtcaatggga 540

gtttgttttg gcaccaaaat caacgggact ttccaaaatg tcgtaacaac tccgccccat 600

tgacgcaaat gggcggtagg cgtgtacggt gggaggtcta tataagcaga gctctctggc 660

taactagaga acccactgct tactggctta tcgaaattaa tacgactcac tatagggaga 720

cccaagctgg ctagcgccgc caccatggat gaccagcgcg atctgattag caaccatgaa 780

cagctgccaa ttctgggcaa ccgcccacgc gaacccgaac gctgcagccg cagccagtgc 840

gtgaacctga ccacccgcac ccagctgcca cccgcttata ccaacagctt tacccgcggc 900

gtgtactatc ctgataaagt gtttcgcagc tcagtcctgc atagcaccca ggatctgttt 960

ctgccttttt ttagcaacgt gacctggttt catgcaattc atgtgagcgg caccaacggc 1020

actaaacgct ttgataaccc agtgctgcca ttcaacgatg gcgtgtattt tgccagcacc 1080

gaaaaatcaa acattattcg cggctggatt tttggcacca ccctggatag caaaacccag 1140

agcctgctga ttgtgaacaa cgcaaccaac gtggtgatca aagtgtgcga atttcagttt 1200

tgcaacgatc cctttctggg cgtgtactat cacaaaaaca acaaaagctg gatggaaagc 1260

gaatttcgcg tgtatagcag cgcaaacaac tgcacctttg agtatgtgtc ccagccattt 1320

ctgatggacc tggaaggcaa gcaaggaaac tttaaaaacc tgcgcgaatt tgtgtttaaa 1380

aacattgatg gctattttaa aatttattcc aaacataccc ccattaacct ggtgcgcgat 1440

ctgccccaag gctttagcgc cctggaacct ctggtggatc tgcctattgg tattaacatt 1500

acccgctttc agaccctgct ggctctgcat cgcagctatc tgacccctgg cgatagcagc 1560

agcggctgga ccgccggcgc agcagcttat tatgtgggct atctgcaacc ccgcaccttt 1620

ctgctgaaat acaacgaaaa cggcaccatt accgatgccg tggattgcgc actggatccc 1680

ctgagcgaaa ccaaatgcac cctgaaaagc tttaccgtgg agaaaggcat ttatcagacc 1740

tccaactttc gcgtgcagcc caccgaaagc attgtgcgct ttcctaacat taccaacctg 1800

tgccctttg gcgaagtgtt taacgcaacc cgctttgcca gcgtgtatgc ttggaaccgc 1860

aaacgcatta gcaactgcgt ggctgattac agcgtgctgt ataacagcgc cagctttagc 1920

acctttaaat gctatggcgt gagcccaacc aagctgaacg atctgtgctt caccaacgtg 1980

tatgcagata gttttgtgat tcgcggcgat gaagtgcgcc agattgctcc tggccagacc 2040

ggtaaaattg ctgattataa ctataaactg ccagatgatt ttaccggctg cgtgattgca 2100

tggaacagca acaacctgga tagcaaagtg ggcggtaact ataactatct gtatcgcctg 2160

tttcgcaaaa gcaacctgaa acctttttgaa cgcgatatta gcaccgaaat ttatcaggca 2220

ggtagcaccc catgcaacgg cgtggaaggc tttaactgct attttcctct gcagagctat 2280

ggctttcaac caaccaacgg ggtgggctat caaccatatc gcgtggtggt gctgtctttt 2340

gaactgctgc atgctccagc taccgtgtgc ggccccaaaa aaagcaccaa cctggtgaaa 2400

aacaaatgcg tgaactttaa ctttaacggc ctgacaggca ccggcgtgct gaccgaatca 2460

aacaaaaaat tcctgccctt tcagcagttt ggccgcgata ttgctgatac aaccgatgca 2520

gtgcgcgatc ctcagaccct ggaaattctg gatattaccc cctgcagctt tggcggcgtg 2580

agcgtgataa cccctggcac caacaccagc aaccaggtgg ctgtgctgta tcagggcgtg 2640

aactgcaccg aagtgcccgt ggccattcat gctgatcaac tgacccccac ctggcgagtg 2700

tatagcacag gcagcaacgt gtttcagacc cgcgccggct gcctgattgg cgccgaacat 2760

gtgaacaaca gctatgagtg cgatattcct attggcgctg gcatttgcgc aagctatcag 2820

acccagacca acgcccgcag cgtggccagc cagagcatta ttgcctatac catgagcctg 2880

ggcgcagaaa acagcgtggc ttatagcaac aacagcattg ctataccaac caacttcacc 2940

attagcgtga ccaccgaaat tctgcccgtg agcatgacca aaaccagcgt ggactgcact 3000

atgtatattt gcggcgatag caccgaatgc agcaacctgc tgctgcagta tggcagcttt 3060

tgcacccagc tgaaccgcgc cctgactggc attgcagtgg aacaggataa aaacacccaa 3120

gaagtgtttg cccaggtgaa acagatttac aaaaccccccc ctattaaaga ttttggcggc 3180

ttcaacttta gccagatact gccagatcct agcaaaccta gcaaacgcag ctttatcgaa 3240

gatctgctgt ttaacaaagt gaccctggct gatgctggct ttattaaaca gtatggtgat 3300

tgcctgggcg atattgccgc ccgcgatctg atttgcgctc agaaatttaa cggcctgacc 3360

gtgctgcccc ctctgctgac cgatgaaatg attgcacaat acaccagtgc tctgctggct 3420

ggaaccatta cctcaggctg gacttttggc gcaggtgccg ccctgcagat tccatttgcc 3480

atgcagatgg cctatcgctt taacggcatt ggcgtgaccc agaacgtgct gtatgaaaac 3540

cagaaactga ttgccaacca gtttaacagc gctattggca aaatccagga tagcctgagc 3600

agcaccgcca gcgcactggg caaactgcag gatgtggtga accagaacgc tcaggcactg 3660

aacaccctgg tgaaacagct gagcagcaac tttggcgcca ttagcagcgt gctgaacgat 3720

attctgagcc gcctggaccc tccagaagct gaagtgcaga ttgatcgcct gattacaggc 3780

cgcctgcaga gcctgcagac ttatgtgacc cagcagctga ttcgcgccgc tgaaatccgc 3840

gcaagcgcaa acctggctgc caccaaaatg agcgaatgcg tgctgggcca gagcaaaagg 3900

gtggattttt gtggcaaagg ctatcacctg atgagctttc cccagagcgc tccacatggc 3960

gtggtgtttc tgcacgtgac ctatgtgccc gctcaggaaa aaaactttac cacagcccca 4020

gcaatttgcc atgatggcaa agctcatttt ccccgcgaag gcgtgtttgt gagcaacggc 4080

acccattggt ttgtgaccca gcgcaacttt tatgaacctc agattattac caccgataac 4140

acctttgtga gcggcaactg cgatgtggtg attggcattg tgaacaacac cgtgtatgat 4200

cccctgcagc ccgaactgga tagctttaaa gaagagctgg ataaatattt caaaaaccat 4260

accagccccg atgtggacct gggcgatatt agcggcatta acgccagcgt ggtgaacatt 4320

cagaaagaaa ttgatcgcct gaacgaagtg gctaaaaacc tgaacgaaag cctgattgat 4380

ctgcaagaac tggggaaata tgaacagtat attaaatggc cctggtatat ttggctgggc 4440

tttattgctg gcctgattgc cattgtgatg gtgaccatta tgctgtgctg catgaccagc 4500

tgctgtagct gcctgaaagg gtgctgcagc tgcggcagct gctgcaaatt tgacgaagat 4560

gatagcgaac ctgtgctgaa aggcgtcaaa ctgcattata ccggcagcgg cgaaggaagg 4620

gggagcctgc tgacctgtgg cgatgtggaa gagaaccctg gcccactgat gcagattttt 4680

gtgaaaaccc tgacaggcaa gaccatcacc ctggaagtgg aacctagcga taccattgaa 4740

aacgtgaagg caaaaatcca ggataaagaa gggattcctc agaccaaca gcgcctgatt 4800

tttgccggca aacaactgga agatggccgc acgctgagcg attataacat tcagaaggaa 4860

agcaccctgc atctggtgct gcgcctgcgc ggcgtgatga gcgataacgg ccctcagaac 4920

cagcgcaacg caccacgcat tacctttggc ggtccatcag atagcactgg cagcaaccag 4980

aacggcgaac gcagcggagc ccgcagcaaa caacgccgcc cacagggcct gcctaacaac 5040

accgcaagct ggtttaccgc cctgacccag cacggcaaag aagatctgaa atttccccgc 5100

ggccagggcg tgccaattaa caccaacagc agcccagatg atcagattgg ctattatcgc 5160

agggccaccc gccgcattcg cggaggcgat ggcaaaatga aagatctgag cccacgctgg 5220

tacttttatt atctgggaac cggccccgaa gccggtcttc catatggcgc caacaaagat 5280

ggcatcattt gggtggcaac cgaaggagct ctgaacaccc caaaagatca tattggcacc 5340

cgcaaccccg caaacaacgc tgcaattgtg ctgcagcttc ctcagggcac caccctgcca 5400

aaaggctttt atgccgaagg cagccgcggc ggcagccagg caagctctag gagcagcagc 5460

cgcagccgca acagcagccg gaacagcacc cctggaagca gcaggggcac cagcccagcc 5520

aggatggcag gtaacggcgg cgatgctgcc ctggctctgc tgctgctgga tcgcctgaac 5580

cagctggaaa gcaaaatgag cggcaaaggc cagcagcagc agggccagac cgtgaccaag 5640

aaaagcgcag ccgaagccag caaaaaacca cgccagaaaa gaaccgccac caaagcctat 5700

aacgtgacgc aggcttttgg ccgccgtggc cccgaacaaa cccagggcaa ctttggcgat 5760

caggaactga ttcgccaggg caccgattat aaacattggc cacagattgc ccagtttgct 5820

cctagcgcca gcgcttttt tggtatgagc cgcattggca tggaagtgac cccctctggc 5880

acctggctga cctataccgg cgctattaaa ctggatgata aagatccaaa ctttaaagac 5940

caggtgattc tgctgaacaa acatattgat gcctataaaa cctttccccc aaccgaaccc 6000

aaaaaagata aaaagaaaaa agccgatgaa acccaggccc tgccacagcg ccagaaaaaa 6060

cagcaaaccg tgaccctgct gcctgcagcc gatctggatg atttcagcaa acagctgcaa 6120

cagagcatga gcagcgcaga tagcacgcag gcttaactcg agtctagagg gcccgtttaa 6180

acccgctgat cagcctcgac tgtgccttct agttgccagc catctgttgt ttgcccctcc 6240

cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tcctttccta ataaaatgag 6300

gaaattgcat cgcattgtct gagtaggtgt cattctattc tgggggggtgg ggtggggcag 6360

gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct 6420

atggcttcta ctgggcggtt ttatggacag caagcgaacc ggaattgcca gctggggcgc 6480

cctctggtaa ggttgggaag ccctgcaaag ttaacaaatg agctgattta acaaaaattt 6540

aacgcgaatt aattctgtgg aatgtgtgtc agttagggtg tggaaagtcc ccaggctccc 6600

cagcaggcag aagtatgcaa agcatgcatc tcaattagtc agcaaccagg tgtggaaagt 6660

ccccaggctc cccagcaggc agaagtatgc aaagcatgca tctcaattag tcagcaacca 6720

tagtcccgcc cctaactccg cccatcccgc ccctaactcc gcccagttcc gcccattctc 6780

cgccccatgg ctgactaatt ttttttattt atgcagaggc cgaggccgcc tctgcctctg 6840

agctattcca gaagtagtga ggaggctttt ttggaggcct aggcttttgc aaaaagctcc 6900

cgggagcttg tatatccatt ttcggatctg atcaagagac aggatgagga tcgtttcgca 6960

tgattgaaca agatggattg cacgcaggtt ctccggccgc ttgggtggag aggctattcg 7020

gctatgactg ggcacaacag acaatcggct gctctgatgc cgccgtgttc cggctgtcag 7080

cgcaggggcg cccggttctt tttgtcaaga ccgacctgtc cggtgccctg aatgaactgc 7140

aagacgaggc agcgcggcta tcgtggctgg ccacgacggg cgttccttgc gcagctgtgc 7200

tcgacgttgt cactgaagcg ggaagggact ggctgctatt gggcgaagtg ccggggcagg 7260

atctcctgtc atctcacctt gctcctgccg agaaagtatc catcatggct gatgcaatgc 7320

ggcggctgca tacgcttgat ccggctacct gcccattcga ccaccaagcg aaacatcgca 7380

tcgagcgagc acgtactcgg atggaagccg gtcttgtcga tcaggatgat ctggacgaag 7440

agcatcaggg gctcgcgcca gccgaactgt tcgccaggct caaggcgagc atgcccgacg 7500

gcgaggatct cgtcgtgacc catggcgatg cctgcttgcc gaatatcatg gtggaaaatg 7560

gccgcttttc tggattcatc gactgtggcc ggctgggtgt ggcggaccgc tatcaggaca 7620

tagcgttggc tacccgtgat attgctgaag agcttggcgg cgaatgggct gaccgcttcc 7680

tcgtgcttta cggtatcgcc gctcccgatt cgcagcgcat cgccttctat cgccttcttg 7740

acgagttctt ctgaattatt aacgcttaca atttcctgat gcggtatttt ctccttacgc 7800

atctgtgcgg tatttcacac cgcatacagg tggcactttt cggggaaatg tgcgcggaac 7860

ccctatttgt ttatttttct aaatacattc aaatatgtat ccgctcatga gacaataacc 7920

ctgataaatg cttcaataat agcacgtgct aaaacttcat ttttaattta aaaggatcta 7980

ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca 8040

ctgagcgtca gacccccatc agtgaccaaa caggaaaaaa ccgcccttaa catggcccgc 8100

tttatcagaa gccagacatt aacgcttctg gagaaactca acgagctgga cgcggatgaa 8160

caggcagaca tctgtgaatc gcttcacgac cacgctgatg agctttaccg cagctgcctc 8220

gcgcgtttcg gtgatgacgg tgaaaacctc tgacacatgc agctcccgga gacggtcaca 8280

gcttgtctgt aagcggatgc cgggagcaga caagcccgtc agggcgcgtc agcgggtgtt 8340

ggcgggtgtc ggggcgcagc catgacccag tcacgtagcg atagcggagt gtatactggc 8400

ttaactatgc ggcatcagag cagattgtac tgagagtgca ccatatgcgg tgtgaaatac 8460

cgcacagatg cgtaaggaga aaataccgca tcaggcgctc ttccgcttcc tcgctcactg 8520

actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc agctcactca aaggcggtaa 8580

tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc 8640

aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg ctccgccccc 8700

ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg acaggactat 8760

aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc 8820

cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct 8880

cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg 8940

aacccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc 9000

cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga 9060

ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa 9120

ggacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa agagttgtta 9180

gctcttgatc cggcaaacaa accaccgctg gtagcggtgg ttttttgtt tgcaagcagc 9240

agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatc 9284

<210> 20
<211> 281
<212> DNA
<213> Artificial Sequence

<220>
<223> SV40 ori enhancer sequence

<400> 20
tgtgtcagtt agggtgtgga aagtccccag gctccccagc aggcagaagt atgcaaagca 60

tgcatctcaa ttagtcagca accaggtgtg gaaagtcccc aggctcccca gcaggcagaa 120

gtatgcaaag catgcatctc aattagtcag caaccatagt cccgccccta actccgccca 180

tcccgccccc aactccgccc agttccgccc attctccgcc ccatggctga ctaatttttt 240

ttatttatgc agaggccgag gccgcctctg cctctgagct a 281