

SARS-CoV-2 (2019-nCoV) vaccine

Rapid Deployment Vaccine Collaborative (RaDVaC)

Preston Estep, Don Wang, Alex Hoekstra, Ranjan Ahuja, George Church, Brian Delaney, and over 100 others

Table of Contents

[Table of Contents](#)

[TERMS OF USE/INFORMED CONSENT \(updated 2020-08-17\)](#)

[OVERALL PROJECT GOAL \(updated 2021-02-28\)](#)

[BACKGROUND \(updated 2021-03-07\)](#)

[VACCINE BACKGROUND AND DESIGN \(updated 2020-12-07\)](#)

[POSSIBLE MECHANISMS OF VACCINE-ENHANCED DISEASE, VACCINE-INDUCED AUTOIMMUNITY, AND MITIGATION STRATEGIES \(updated 2020-10-24\)](#)

[KEY TECHNICAL FEATURES AND SPECIFICATIONS \(Updated 2021-03-06\)](#)

[GOALS](#)

[MATERIALS AND METHODS \(updated 2021-03-07\)](#)

[Selection and purchase of the vaccine ingredients](#)

[Preferred equipment and materials \(updated 2020-11-04\)](#)

[Selection of epitopes \(updated 2021-03-06\)](#)

[Peptides and vaccine composition \(last updated 2021-03-05\)](#)

[Peptide synthesis general guidelines](#)

[Gen 10](#)

[Gen 9.1 Optional peptides](#)

[Spike or S protein amino acid sequences represented by epitope peptides](#)

[Nucleocapsid or N protein](#)

[B-cell epitope peptides](#)

[Nucleocapsid protein epitopes](#)

[T-cell epitope peptides](#)

[SUMMARY OF THE PROTECTIVE STRATEGY OF THE GEN 10 VACCINE DESIGN \(updated 2021-03-07\)](#)

[PREPARATION OF CHITOSAN NANOPARTICLE VACCINE \(updated 2021-03-07\)](#)

[ADMINISTRATION OF THE VACCINE \(updated 2021-03-07\)](#)

[ASSESSMENTS OF IMMUNE RESPONSE \(updated 2021-03-07\)](#)

[Samples to be collected and tested](#)

[Antibodies and B-cell immune response](#)

[B-cell epitopes and ELISA antigen targets](#)

[Nucleocapsid protein B-cell epitopes / antigen targets](#)

[Back to top](#)

[T-cell immune response](#)

[SUPPLEMENTAL MATERIALS](#)

[SELECTED PRIOR VACCINE GENERATIONS AND EXPERIMENTAL VARIATIONS](#)

[RETIRED EPITOPE PEPTIDES](#)

[ADDITIONAL SEQUENCES](#)

[Spike protein sequence](#)

[Nucleocapsid protein sequence](#)

TERMS OF USE/INFORMED CONSENT (updated 2020-08-17)

This document describes the rationale, design, formulation, and self-administration of a vaccine for SARS-CoV-2. By using this information you agree to the following: 1) you are a consenting adult (in the USA, at least 18 years of age); and 2) you take full responsibility for your use of RaDVaC information and vaccine or material—including redistribution, modification, vaccine formulation, production and administration.

A foundational principle of the vaccine development and deployment strategy of RaDVaC is rapid iteration and testing of vaccine designs, based on newly published information from the forefront of biomedical research. This agile approach has the potential to produce better vaccines much more quickly than traditional approaches. However, such rapid design improvements currently are not compatible with established clinical trial requirements of fixed vaccine design, and Institutional Review Board (IRB) review and approval (due to shifting vaccine design and continually updating self-experimentation and testing protocols). Therefore, the information presented within has not been approved by an IRB, and you understand and agree that any implementation of this information constitutes self-experimentation.

The purpose of this open-source vaccine effort is to reduce risk of harm from SARS-CoV-2. In addition to providing the results of our own research and experimentation, we hope to motivate others to build on our work, and to pursue diverse evidence-based approaches. Given the immense complexity and variability of individual human biology, it is not possible to predict all potential physiological responses to any vaccine. But as is true for most or all vaccines used on a large scale, there is a tradeoff between a larger known risk from the disease the vaccine is designed to prevent or mitigate, and the smaller risk introduced by the vaccine itself. Any vaccine poses risks, and, if used in enough people, will cause some degree of harm. Furthermore, certain harm, such as allergic and possibly anaphylactic response, will be readily seen and measured, whereas benefit is more difficult and takes longer to assess. This vaccine is no different; and because quality of delivery is highly dependent on the meticulousness of individual end users, it might pose unique risks not posed by typical commercial vaccines.

[Back to top](#)

Does Not Constitute or Substitute for Medical Advice

Information presented here is ongoing research, and is not intended as a substitute for medical advice. RaDVaC is not responsible for the decision to administer, or to receive administration of, any vaccine.

No Promises or Guarantees of Efficacy

Vaccines are often received with the false hope of efficacy, without testing to determine the degree of individual immune response. For example, influenza killed about 100,000 people in the U.S. between late 2016 and early 2018. Yet, the influenza vaccines available in that period were substantially less than 50% effective against H3N2, the flu strain mainly responsible for the death toll. Many who died were vaccinated but not protected from the virus, and testing for vaccine-induced immunity was essentially non-existent. Vaccine-induced immunity can be more challenging to assess than immunity due to viral infection, and such is probably the case for the nasal vaccine described here. Because this work is a research undertaking, no expectation is given regarding any degree of efficacy in granting protection against SARS-CoV-2. On the contrary, see possible risks and uncertain benefits below.

Preventive, Not Therapeutic

Even if this vaccine works as intended, it will not help someone who has been infected. It will only work as a preventive measure taken weeks in advance of virus exposure.

Not a Clinical Trial

Any use of the information in this document, or provided in correspondence with us does not constitute 'partnership' or 'recruitment' for an organized trial in any way. The supported conjecture here is given as a starting point for additional individual or organized, sponsored research. As of public release of this information, no organized clinical trials have been performed to test this vaccine.

Not Approved or Reviewed by the FDA

The information, procedures, and conclusions presented here have not been approved, or reviewed by the FDA, or any other regulatory body.

Possible Risks and Uncertain Benefits

- Immediate allergic or other serious reaction
- Unforeseen long-term effects

- Instillation/administration of the vaccine in an inappropriate way or in an infected area might increase the risk of infection by enhancing viral entry into the body
- Benefits are uncertain. There are extensive published histories of the materials and procedures described in this document, but every novel vaccine should be considered experimental, with the possibility there will be no benefit.
- Even if there are signs of immune response there is no guarantee this response is indicative of protection from SARS-CoV-2 infection, or if protection is achieved, how long it will last.
- Even if the vaccine confers protection from the virus, certain methods for assessing protection--such as measuring antibodies due to previous infection--might not capture vaccine-induced immunity. In such cases, infection, and thus a positive test result, are unlikely to occur. Immunity passports and other privileges given to convalescents might be difficult to obtain for those with vaccine-induced immunity.
- Use of this vaccine may change the efficacy of any future vaccines you may take or that are administered to you, in unknown ways

No Offer of Service and No Access to Materials

All information in this document or via correspondence is made available as open science. We do not sell equipment or services, e.g. we will not prepare this formulation or any other on behalf of a requestor. All vaccine materials at our disposal are kept in a secured laboratory. We cannot and will not keep materials outside the lab for easy access or distribution. We will not provide laboratory or equipment access for vaccine production.

Probability and Coincidence

Vaccines have become associated with negative outcomes that do not result from the vaccine, but that occur soon after administration. Increasing use of any vaccine will increase the probability of recipients experiencing an unconnected negative outcome. By using the information presented here, you acknowledge the increasing likelihood of such coincidences, and assume full responsibility for any use of the information herein and for real or perceived negative outcomes, irrespective of the cause.

Self-Experimentation

By utilizing the information and scientific opinions in this document, and in any correspondence with us, you acknowledge and agree that any use to develop and self-administer a substance is an act of self-experimentation. Laws and regulations on such actions may differ in your national or state jurisdiction, and it is your responsibility to be informed about them.

Additionally, the type of work described in this document requires certain equipment and level of skill with laboratory techniques. You agree to assume full responsibility for acquiring proper equipment, knowledge, and training, and for attempting to formulate or administer vaccine.

Licenses

All contents of the RaDVaC website and white paper are covered by a Creative Commons Attribution 4.0 License (CC BY 4.0), and by an Open COVID License – Patent 1.1 (OCL-P v1.1). You are free to Share (copy and redistribute the material in any medium or format), and Adapt (fork, alter, and build upon) the material for any purpose. The only requirement is Attribution.

OVERALL PROJECT GOAL (updated 2021-02-28)

The ultimate goal of the RaDVaC project is to create a modular vaccine platform for extremely rapid deployment at the beginning of a serious outbreak of a pathogen for which no good vaccine exists. The founding philosophy of RaDVaC is that the best way to accomplish this goal is to share freely and openly information on vaccines and testing throughout a global network of researchers engaged in agile R&D. For many, this will include highly localized vaccine production. While the SARS-CoV-2 pandemic has been the catalyst for initiation of the project, it will not be the last serious outbreak; the establishment of a solid scientific and technical foundation for rapid vaccine deployment is a long-term endeavor that will not end as the SARS-CoV-2 pandemic winds down.

BACKGROUND (updated 2021-03-07)

The SARS-CoV-2 virus (a.k.a. 2019-nCoV; disease: COVID-19) is responsible for a worldwide pandemic far beyond the scope of any other public health crisis in living memory. As of late February 2021 the pandemic has caused over 2.5 million deaths worldwide, but the impact extends far beyond the death toll. The best data at the time show that high proportions--in some studies, well over half--of convalescents have significant heart, lung, and neurological damage months after recovery^{1,2,3,4}.

Most leading experts have said reliable testing will be helpful in managing the pandemic and monitoring individual exposure, but that the ultimate goal is an effective vaccine. Unfortunately, a vaccine for the virus was not available outside of relatively small clinical trials until December 2020, during which time about 1.7 million deaths were attributable to COVID-19, globally⁵. In the U.S. as with other high-income nations, due to limitations in production capacity and distribution, commercial vaccines will not be available to many or most people until mid 2021, if all goes well⁶. Dozens of low income nations will likely not have access to commercial vaccines until at least 2022[Citation]. These significant delays not only lead to continued deaths and chronic illness but also allow additional viral strains and escape mutants to develop, compounding the public health and scientific challenges.

¹ <https://jamanetwork.com/journals/jamacardiology/fullarticle/2768916>

² <https://www.sciencemag.org/news/2020/07/brain-fog-heart-damage-covid-19-s-lingering-problems-alarm-scientists>

³ <https://news.berkeley.edu/2020/07/08/from-lung-scarring-to-heart-damage-covid-19-may-leave-lingering-marks/>

⁴ <https://www.statnews.com/2020/07/27/covid19-concerns-about-lasting-heart-damage/>

⁵ <https://coronavirus.jhu.edu/map.html>

⁶ [https://www.technologyreview.com/s/615331/a-coronavirus-vaccine-\[...\]least-18-monthsif-it-works-at-all/](https://www.technologyreview.com/s/615331/a-coronavirus-vaccine-[...]least-18-monthsif-it-works-at-all/)

Vaccines are designed to stimulate an immune response, ideally one that provides long-lasting protection against infection by a pathogen. In a pandemic situation, ideally the first choice one must make in the decision to use a vaccine is which of the two possible ways one will be exposed to immunogenic elements of a pathogen: by infection, or by vaccination. Pathogens like viruses possess many complex mechanisms for taking over the body's processes, evading immune responses, and ultimately compromising health and causing extreme suffering and possibly death. We believe that a vaccine is a far better choice; unfortunately, that option has not been available to many people who have become infected. This document describes the rationale, design, formulation, and intranasal self-administration of a peptide-based vaccine against SARS-CoV-2. Starting in March, 2020, our group has followed these protocols, which are based on prior publications and a rapidly growing preprint literature on SARS-CoV-2, to produce and self-administer multiple generations of incrementally improved vaccines.

There is substantial published information on all aspects of vaccine production, testing, and delivery, some of which is specific to the SARS-CoV-2 virus. This published information is sufficient to create a vaccine and tests for correlates of protection on a timeline much shorter than that given for commercial vaccines.

Vaccines are the safest of all therapeutic classes. Because of their safety and efficacy, they have by far the highest probability of success in the overall course of clinical trials (over 40%, which is about twice the rate of the next highest class).⁷ Because of this and the impact of the pandemic, commercial trials have been greatly accelerated.

Nevertheless, commercial vaccines must be designed to enable large scale production and deployment, and they require regulatory approval for sale, both of which greatly constrain and slow progress. Critically, certain features of commercial vaccines aren't required for research-level production and testing; therefore, a much shorter timeline is possible for a smaller scale, self-administered vaccine. Herein we describe the formulation of a vaccine using proven, inexpensive, and mostly off-the-shelf components. The open-source SARS-CoV-2 vaccine described can be produced quickly and inexpensively in a wide range of lab and physician office settings. Our group, the Rapid Deployment Vaccine Collaborative (RaDVaC), has produced and first administered the vaccine to ourselves in April 2020. As of late February, hundreds of us have self-administered multiple generations of progressively updated vaccines.

This will be a versioned living document, embodying and enabling incremental improvements in vaccine design and testing.

⁷ <https://projectalpha.mit.edu/pos/>

VACCINE BACKGROUND AND DESIGN (updated 2020-12-07)

There are many advances in vaccine research and technology development that have not appeared in a commercial product, in part because commercial vaccine design and production is constrained by different factors than a small-scale research vaccine. Key features for a RaDVaC open-source vaccine are high safety, low cost, and ease of production and administration. Another key feature of the RaDVaC vaccine is rapid iteration of updated designs as new discoveries about the biology and clinical manifestations of SARS-CoV-2 are published in the biomedical literature and on preprint servers. This feature is critical to the RaDVaC mission of rapid development and deployment of effective vaccines. We address all of these features in this document, and consider subcategories, such as near-term and long-term safety.

Key issues and some differentiating factors between commercial vaccines and the RaDVaC research vaccine platform:

- **Rapidly responsive design.** A key feature of our rapid deployment strategy is the flexibility of the vaccine platform, allowing very fast and efficient updating of vaccine designs in response to recent research. On multiple occasions we have updated vaccine designs with important information published in the biomedical literature only weeks or days earlier. One important discovery shows that T-cell responses are found in patients almost two decades after recovery from SARS-CoV-1 (a very closely related coronavirus)⁸. The portions of the virus responsible for these very long-lived T-cell responses (CD4 and CD8) are in highly conserved parts of the virus. They are not present in the Spike protein, which is not highly conserved, and which is the primary target of leading vaccines in the race to commercialization. A second publication reveals the identities of 29 experimentally verified, immunodominant CD8 T-cell epitopes of SARS-CoV-2⁹. This publication confirms several important findings of a prior preprint¹⁰, and some of these epitopes are shared among the small group of long-lived epitopes from SARS-CoV-1, none of which resides in the Spike protein. Only 3 of 29 epitopes lie within the Spike protein, and only one lies within the RBD region of Spike. Another subsequent publication experimentally confirmed and extended these findings, also showing that Spike epitopes are not among the top 10 immunodominant CD8 T-cell epitopes¹¹. These findings suggest

⁸ <https://www.nature.com/articles/s41586-020-2550-z>

⁹ <https://www.medrxiv.org/content/10.1101/2020.07.24.20161653v2.article-info>

¹⁰ <https://www.biorxiv.org/content/10.1101/2020.06.05.134551v1.full.pdf>

¹¹ <https://www.medrxiv.org/content/10.1101/2020.07.31.20165647v2.article-info>

that the Spike protein might not provide extensive and robust T-cell protection, and that RBD will provide even less protection for many people, especially the large numbers of people with MHC haplotypes not represented by these antigens. Consistent with this expectation, early data on leading vaccines in the race to large-scale commercial deployment suggest that they might not confer sterilizing immunity (the immune system is able to stop a pathogen from entering the body, or if it gains entry, from replicating within the body). Again, the RaDVaC strategy allows for frequent updating of vaccine designs based on ongoing research throughout the world.

- **Near-term safety.** Near-term safety of the intranasal vaccine described here should be excellent when prepared appropriately. Intranasal vaccines have a long history of demonstrated safety. The vaccine formulations described here can also be used for other delivery routes, including inhalation, oral, or parenteral delivery (injection), but these, especially the latter, should be attempted only by skilled practitioners. In general, a vaccine is more safely used for inhalation into the lungs¹², oral, or intranasal delivery, than for injection. This document is focused on intranasal delivery, but experts will know how to adapt the information herein to another delivery route. For any delivery mode, sourcing of high-quality materials and meticulous preparation are essential for maintaining vaccine safety.
- **Long-term safety.** The long-term safety of any vaccine is currently difficult or impossible to predict. Even widely-deployed commercial vaccines have resulted in serious and unforeseen complications. Vaccines that have shown serious side effects are injectable, whole virus or subunit formulations. There are at least three potentially serious complications that might arise in the long term: tolerance, vaccine-enhanced disease (VED), and adjuvant triggered immune or neurological complications.
 - **Immune tolerance** is a term meaning diminished immunity resulting from exposure to an antigen. This attenuated immune response is commonly observed in food antigens, in “self” antigens of one’s own body, and in antigens from commensal organisms. In general, it is thought that extremely large and/or frequent exposures and oral doses lead to tolerance.
 - **Vaccine-enhanced disease (VED).** A small number of injected vaccines have led to enhancement of disease, meaning that infectivity is enhanced, or the disease is made more serious in people who have been vaccinated, relative to unvaccinated controls. Possible mechanisms underlying VED, such as antibody-dependent enhancement (ADE) are explored at length in the section below, together with various vaccine design features and strategies for reducing the likelihood of VED.
 - **Adjuvant hyperstimulation or toxicity.** Adjuvants help stimulate a robust immune response to a vaccine; however, certain adjuvants have caused hyperstimulation

¹² <https://www.sciencedirect.com/science/article/pii/S1549963415000313>

and other serious side effects. For example, alum produces a robust Th2 immune response, but an unbalanced ratio of Th2:Th1. A Th2 polarized response, and alum in particular, have been implicated in immunopathology, including ADE. Adjuvants can also be toxic. As one example, the intranasal use of a detoxified mutant form of *Escherichia coli* Heat Labile Toxin has resulted in transient Bell's palsy, or facial nerve paralysis.¹³ Specific adjuvants have also been suggested to play a role in the development of autoimmunity. For example, the influenza vaccine Pandemrix appears to have increased the incidence of narcolepsy in Scandinavian countries¹⁴. The leading explanation is that the vaccine triggered a CD4-mediated autoimmune reaction against sleep-regulating neurons (although the vaccine-attributable risk is very low, with 1 case in 18,400 vaccinations, or about 0.005%). One suggested contributing factor is the strong squalene-based adjuvant. The primary reason for using strong adjuvants in commercial vaccines is to trigger a robust immune response with a single administration, avoiding a booster schedule.

- **Stability.** Stability is a key determinant of a commercial vaccine. Formulations that are both safe and effective in a research setting, but with limited shelf life, are generally excluded from commercial products. We have discovered that there are formulations for vaccines that are extremely simple, safe, and effective, but have only short-term stability (on the order of weeks). For example, chitosan gel nanoparticles have been shown to be effective and extremely simple to formulate, but short shelf life has contributed to their limited use in commercial vaccines.
- **Intranasal delivery.** Intranasal vaccines have advantages over other delivery approaches, including the most common delivery modality, parenteral/injection. Intranasal delivery has been demonstrated to be very safe, with mild side effects typically equal to those seen in placebo-treated subjects. Importantly, it can elicit not only systemic immunity, but also mucosal immunity at the point of infection for respiratory viruses.¹⁵ Commercial intranasal influenza vaccine is available, and relative to parenteral (injected) administration shows equal efficacy for systemic immunity, but greater efficacy for mucosal immunity at sites of entry (nose, lungs) for respiratory viruses.¹⁶ This is critically important for SARS-CoV-2 because early studies show that a majority of infections begin in the nasal cavity¹⁷. As of March 2021, most or all commercial SARS-CoV-2 vaccines in development are designed to be injected (parenteral), a route which is unlikely to provide mucosal immunity to infection.

¹³ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2737308/>

¹⁴ <https://pubmed.ncbi.nlm.nih.gov/28847694/>

¹⁵ <https://pdfs.semanticscholar.org/6e55/db027b08b367eaac55ec54a730f4c99061ec.pdf>

¹⁶ <https://www.tandfonline.com/doi/abs/10.1586/erv.12.31>

¹⁷ [https://www.cell.com/cell/pdf/S0092-8674\(20\)30675-9.pdf](https://www.cell.com/cell/pdf/S0092-8674(20)30675-9.pdf)

- **Prime-Boost.** Intranasal delivery can be as effective as injection, but to achieve this level of immune response and protection--especially with milder adjuvants--multiple doses are often required. The initial dose is the prime and subsequent doses are given to boost or increase the immune response. Prior exposure to a pathogen or closely related one has a similar effect to a prime dose of vaccine. The only commonly used intranasal vaccine is for influenza. Since essentially everyone is naturally exposed to influenza, the nasal vaccination with attenuated virus is in essence a booster. Requirement of multiple boosters has limited the commercial production of intranasal (and other mildly to moderately immunostimulating) vaccines for which prior ubiquitous exposure is unlikely. An important corollary to a multiple-dose design is titration of adequate of immune response for each individual. Whereas the immunogenicity-safety profile of commercial formulations must be standardized to an "average" level--overdosing some and underdosing others--with accompanying testing the RaDVaC approach should allow for more precise dosing.
- **Efficacy.** Intranasal delivery of chitosan-based vaccines have shown mild side effects and high levels of efficacy of both mucosal and systemic immunity, when delivered in a prime-boost regimen (in both animal models and human trials; see one review by Smith et al.¹⁸).

Given the above key points, it is clear that a rapid-deployment vaccine that is both safe and effective is likely to have multiple attributes that differentiate it from a commercial vaccine. And it becomes clear why the formulation described here has not been used in a commercial product; it is not lack of safety or efficacy, but other factors related to stability, scalability, patient compliance, profitability, and regulatory complexities of commercial vaccines. Commercial vaccines are designed to produce high levels of immunity with a single dose. A prime and multi-boost schedule is used routinely in the research setting, but concerns about compliance (failure to boost) limit commercial deployment.

For multiple reasons, including those given above, intranasal delivery is difficult to achieve for a commercial vaccine; however, for a research vaccine, it is not only relatively easy, in certain cases and for some pathogen classes like respiratory viruses, it is the preferred mode of delivery. Relative to injection, there is no risk of needle injury or transmission of blood-borne infection. Immunization by the intranasal route not only prevents viral infection through the nasal membranes, but also efficiently stimulates a strong mucosal immune response in the lungs and upper respiratory tract. As one example, Gai and colleagues have shown that a SARS vaccine delivered intranasally elicits a robust mucosal immune response, protecting against initial infection, whereas the same vaccine delivered by injection does not.¹⁹ This difference is important

¹⁸ <https://www.tandfonline.com/doi/full/10.4161/hv.27449>

¹⁹ https://scholar.google.com/scholar?cluster=13722847195977601349&hl=en&as_sdt=0,22

because the area of the mucosal surfaces (nasal, lungs, gastrointestinal, urogenital, etc.) is very large, about 200 times the surface area of the skin, and about 70% of pathogens enter through these routes.²⁰ Compliance is also very high for a single dose because intranasal delivery doesn't involve needles or cause pain. High safety and ease of administration are expected to contribute to high rates of immunization. For a recent review of research on nasal nanovaccines, see Bernocchi et al²¹ (Table 1).

Synthetic peptide-based vaccines have advantages over the most widely used vaccine designs based on attenuated viruses or even full-length Open Reading Frames (ORFs) of key epitope proteins. In commercial vaccines, there is a trend toward highly defined, minimal designs. According to Pompano and colleagues "Vaccine design is moving away from using whole pathogens in favor of selecting only the most protective antigens for immunization with a suitable adjuvant. The most tailored of these sub-unit vaccines utilize specific B cell or CD8+ T cell epitopes, such as the short peptides and carbohydrates that have been used to elicit antibody or killing responses against malaria pathogens, bacterial infections, or tumors. In addition to the desired target epitope(s), these vaccines also require one or more CD4+ T cell epitopes to engage CD4+ helper T cells."²²

Such compact and efficient designs are not easily achieved, and require substantial data on the immunogenicity of all epitopes of a pathogen, but such designs have the advantage of limiting pathogenic effects of certain viral antigens. Coronavirus Spike and Nucleocapsid full-length proteins have been associated with ADE in animals and human cellular studies. Yasui and colleagues showed that Nucleocapsid vaccination does not provide protective immunity, yet enhances immunopathology.²³ Vaccination with certain epitopes of Spike protein do confer protection, but use of the full length Spike might be problematic. For example, from Tais and colleagues: "... full-length S protein should be used with caution. Kam *et al.* reported that although a recombinant, trimeric SARS-CoV S protein vaccine elicited a protective immune response in mice the anti-S antibodies also mediated antibody-dependent enhancement of viral entry into human B-cells *in vitro*. In another study, ferrets vaccinated with SARS-CoV full-length S protein expressed by a recombinant modified vaccinia Ankara grown in BHK21 and Vero E6 cells show enhanced virulence of hepatitis induced by SARS-CoV. Furthermore, the use of a SARS S protein vaccine may lead to enhanced disease and immunopathology instead of protection as seen for feline coronavirus, feline infectious peritonitis virus. Given these concerns, the use of a SARS vaccine strategy in which the full-length S protein is used may not be optimal for humans. Hence, the best approach would probably be to use small S protein epitopes that are major

²⁰ <https://pdfs.semanticscholar.org/6e55/db027b08b367eaac55ec54a730f4c99061ec.pdf>

²¹ <https://www.ncbi.nlm.nih.gov/pubmed/28698066>

²² <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4227912/>

²³ <https://www.jimmunol.org/content/181/9/6337>

neutralization determinants.”²⁴ Vaccines against other pathogens that similarly use full-length constructs have also elicited enhancement of viral sensitivity or disease. The vaccine design described here is based on such B-cell peptide epitopes of the S or spike protein, predicted in the literature to be major neutralization determinants, as well as predicted and experimentally tested effective T-cell epitopes. Several are combination B and T-cell epitopes.

Synthetic peptide synthesis provides freedom to design epitopes of sufficient length for immunogenic stimulation, but predicted not to trigger these serious side effects. This approach also allows the use of multiple epitope peptides, either concatemerized in a single, multimeric linear peptide, or as a collection of individual peptides. The approach taken here uses B-cell epitope peptides of S protein sequences, and T-cell epitopes from S and other proteins. Most or all of the B-cell epitope amino acid sequences contain predicted T-cell epitopes. We have employed many of these combination epitopes in order to keep the design compact and relatively simple and inexpensive to produce. Later versions of epitopes were selected based on data for neutralizing antibodies and T-cell responses from previously SARS-CoV-2 infected convalescent patients.

Choice of adjuvant is important for safe enhancement of immune response. Many adjuvants have been compared for their ability to elicit various aspects of immune response. These include alum, chitosan, inactive cholera toxin (CT), CpG DNA, monophosphoryl lipid A (MPL), poly (I:C), imiquimod and related compounds, and *E. coli* Heat Labile Toxin. Alum is an aluminum salt (Aluminum potassium sulfate, $\text{AlK}(\text{SO}_4)_2$) and is the most widely used adjuvant in commercial vaccines. It has been used for nearly a century as an effective adjuvant. In direct comparisons, higher doses of alum are superior to most other adjuvants, including in intranasally delivered vaccines. Alum has been found to be equal or superior to most others for eliciting systemic immunity (in IgG1, IgG2a, IgG2b), and superior for eliciting secretory IgA (slgA) mucosal immunity and protection against virus challenge in animal studies.²⁵ Poly (I:C) can act as an additional adjuvant; however, it shows weak activity relative to alum, and at higher doses it has triggered severe reactions and autoimmunity in animal models.²⁶

Chitosan-based therapeutics have been developed for many biomedical applications. Chitosan is a partly or completely deacetylated derivative of chitin, the linear polysaccharide found in the shells of crustaceans like shrimp, exoskeletons of insects, and in fungi. Chitosan nanoparticles were first developed in the mid 1990s by groups in Japan and Spain for drug delivery^{27, 28}. For vaccine applications, chitosan acts as both delivery nanoparticle and adjuvant and has been

²⁴ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2756483/>

²⁵ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3868302/>

²⁶ <https://www.ncbi.nlm.nih.gov/pubmed/?term=32122615>

²⁷ <https://pubmed.ncbi.nlm.nih.gov/8383199/>

²⁸ [https://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1097-4628\(19970103\)63:1%3C125::AID-APP13%3E3.0.CO;2-4](https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-4628(19970103)63:1%3C125::AID-APP13%3E3.0.CO;2-4)

formulated for intranasal administration, parenteral injection, oral and sublingual delivery, and more. It has been used widely in animal experiments, and it has been used safely in human clinical and preclinical trials (see Smith et al. review of trials by El-Kamary et al. and Atmar et al.²⁹) (some trials used the commercial product Viscogel)^{30, 31} (also see selected human trials reviewed by Guro Gafvelin and Hans Grönlund, in *Molecular Vaccines: From Prophylaxis to Therapy - Volume 2*, Edited by Matthias Giese; Chapter 39, pp. 624-629; Springer. See Table 39.1). The use of chitosan in intranasal delivery of an influenza vaccine to healthy human volunteers produces systemic (IgG) protection, though less than parenteral vaccine with alum adjuvant (although 15 microgram dosage is low)³²; plus, it also induces mucosal immune response, unlike parenteral vaccine³³. Chitosan is a cationic polymer typically insoluble at physiological pH, which has led to the development of a range of more soluble derivatives, including trimethyl chitosan (TMC) and quaternary or quaternized chitosan, including N-(2-hydroxy)propyl-3-trimethyl ammonium chitosan chloride (a.k.a. 2-hydroxypropyltrimethyl ammonium chloride chitosans (HTCC or HACC))(for review see Wang et al.³⁴). Cationic chitosan and its derivatives generally are all minimally immunostimulatory and other than physiological pH solubility, their preparation methods are quite similar.

Chitosan as an adjuvant has been compared to alum and to other adjuvants, alone and as part of an adjuvant cocktail, in injectable and intranasal forms. Chitosan alone is a potent mucosal and systemic adjuvant, and it is synergistic with alum and other adjuvants^{35,36}. Importantly, intranasal delivery of chitosan-based vaccines has the ability to elicit both mucosal and systemic immunity.

²⁹ <https://www.tandfonline.com/doi/full/10.4161/hv.27449>

³⁰ <https://cordis.europa.eu/project/id/261954/reporting>

³¹ <https://pubmed.ncbi.nlm.nih.gov/25218298/>

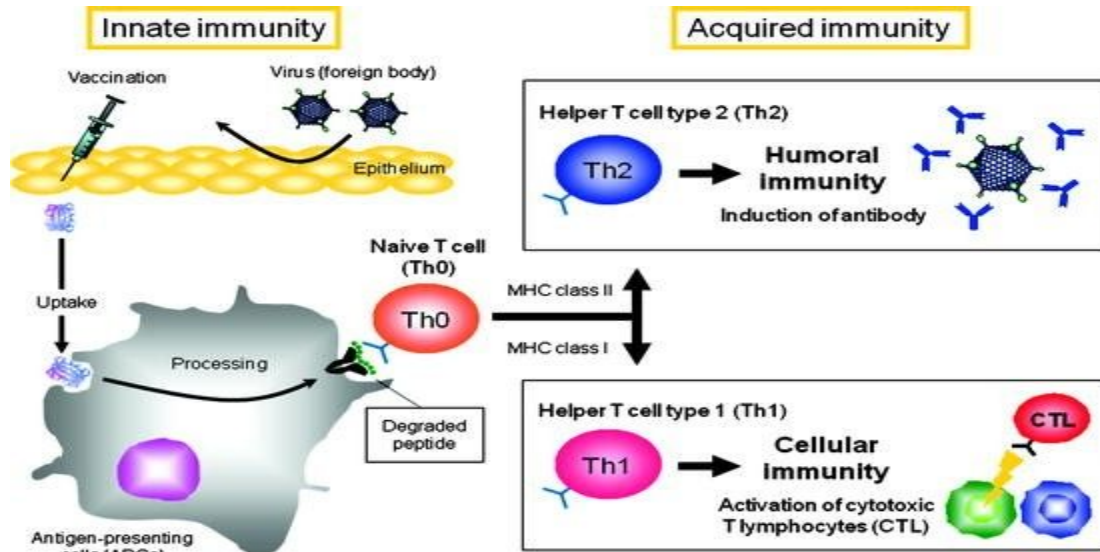
³² <https://pubmed.ncbi.nlm.nih.gov/15916838/>

³³ <https://www.sciencedirect.com/science/article/abs/pii/S0169409X01001715>

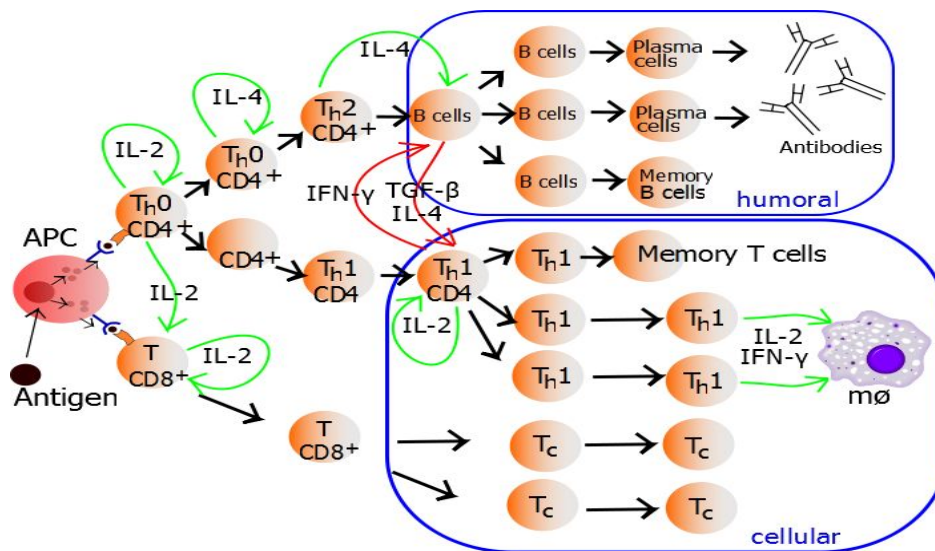
³⁴ <https://www.mdpi.com/1422-0067/21/2/487>

³⁵ <https://pubmed.ncbi.nlm.nih.gov/23933339/>

³⁶ <https://pubmed.ncbi.nlm.nih.gov/17644205>



Schematic showing pathways for Th1 and Th2 acquired immunity (a.k.a. adaptive immunity), and roles of MHC class I and class II. Figure from Akagi, Baba and Akashi: (https://link.springer.com/chapter/10.1007/12_2011_150).



Th1/Th2 Model for helper T-cells. An antigen is ingested and processed by an Antigen Presenting Cell (APC). It presents fragments from the antigen to T-cells. The upper, Th0, is a T helper cell. The fragment is presented to it by MHC2. IFN-γ, interferon γ; TGF-β, transforming growth factor β; mφ, macrophage; IL-2, interleukin 2; IL-4, interleukin 4. From Wikipedia, https://en.wikipedia.org/wiki/T_helper_cell.

Chitosan alone elicits both Th1 and Th2 responses, whereas alum triggers strong Th2 and very weak Th1 responses (see Moran et al. for review³⁷). In general, the Th1 response plays a dominant role in adaptive immunity (a.k.a. acquired immunity) to viral infections, and is responsible for cellular immunity and killing of infected cells by cytotoxic T lymphocytes (CTL), whereas Th2 is humoral and promotes antibody production. Ideally, both Th1 and Th2 processes are elicited by appropriate antigens and adjuvants, and which to target by vaccination are likely to be pathogen specific. Spike protein stimulation of peripheral blood mononuclear cells (PBMC) of SARS-CoV-2 patients results in secretion of high levels of primarily Th1 cytokines, and elevated but much lower levels of Th2 cytokines³⁸. This suggests that SARS-CoV-2 infection provokes a primarily Th1 T-cell response, and a protective vaccine should similarly produce a primarily Th1-based prophylactic response.

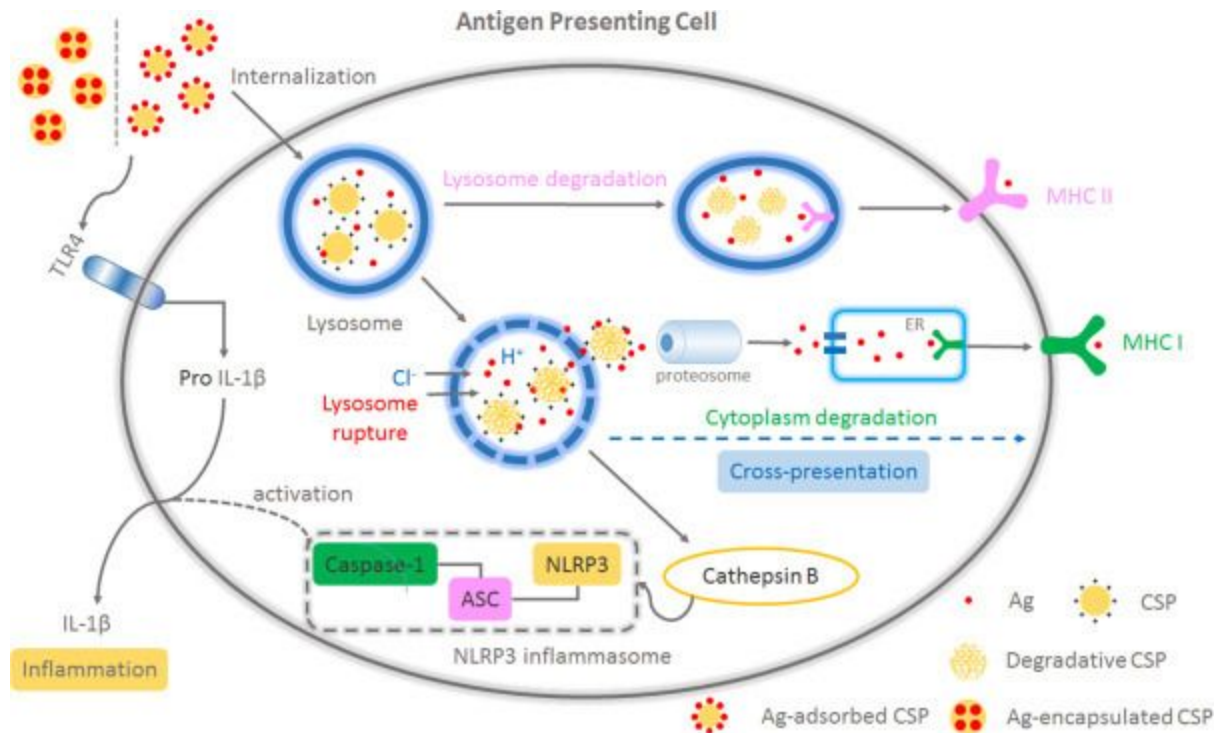
Antibody-dependent enhancement of disease and immunopathology appears to be mediated primarily by a Th2 response³⁹. Extensive research on Respiratory Syncytial Virus, SARS, and other respiratory viruses shows that ADE immunopathology will be more likely with the use of alum or other Th2-boosting adjuvant.⁴⁰ From Honda-Okubo et al: although an anti-SARS vaccine “formulated with alum protected against mortality, these mice developed severe lung eosinophilia at day 6 postchallenge, reminiscent of the lung pathology induced by alum-adjuvanted RSV vaccines and indicating alum's general lack of suitability as a coronavirus vaccine adjuvant.” These concerns make alum alone a poor adjuvant choice, and chitosan a preferred choice. However, seroconversion in response to vaccination becomes more difficult with age; thus, older people might benefit from the addition of additional potent adjuvants, including those that promote some degree of Th2 response. This is an important consideration for SARS-CoV-2, since older people are most susceptible.

³⁷ <https://pubmed.ncbi.nlm.nih.gov/30195140>

³⁸ <https://www.medrxiv.org/content/10.1101/2020.04.11.20062349v1>

³⁹ <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0035421>

⁴⁰ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4337527/>



Model of chitosan nanoparticle (CSP)-induced NLRP3 inflammasome activation and cross-presentation: Ag together with CSP (via adsorption or encapsulation) was internalized by APCs. The protonation of the amino groups (“proton sponge effect”) leads to an extensive inflow of ions and water into the lysosome, which causes the osmotic swelling and deconstruction of the lysosome. The entrapped components (CSP and Ag) are released and finally presented onto MHC I, by cytoplasm degradation (with proteasome and ER involved); After the rupture, lysosome enzymes, cathepsin B, was also leaked into the cytoplasm, which induced the assembly and activation of NLRP3 complex. The capacity of TLR4 stimulation of CSP also played an important role in the intracellular synthesis of pro IL-1 β and triggered the secretion of IL-1 β and inflammatory responses, together with NLRP3 activation. (Ag: antigen; CSP: chitosan particle.) Figure and caption from Xia et al., <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7185844/>.

The amount of vaccine delivered is important. Too little antigen is unlikely to elicit a strong immune response, but too much might attenuate response. Delivery route is also important, with oral delivery being the route most associated with tolerance. Furthermore, high doses of mucosal antigens appear to be required in humans (50+mg) for inducing systemic tolerance, whereas low doses do not appear to induce systemic tolerance.⁴¹

⁴¹ <https://academic.oup.com/intimm/article/26/9/517/2950826>

One challenge for intranasal vaccines is verifying efficacy. Many published reports show that nasally delivered vaccines have high efficacy for prevention of infection, typically equivalent or superior to injected vaccines for this most important measure. However, efficacy is not as easily measured or predicted by traditional measures, such as anti-virus or anti-epitope antibody titer (e.g. as measured by ELISA) in serum. This is in part because blood is more quantifiable than mucosal secretions, and in part because the primary means of conferring immunity is through mucosal stimulation and response. Measuring efficacy against SARS-CoV-2 is also difficult because the B-cell/antibody response to coronaviruses is highly variable and provides uncertain immune protection. Sustained immune protection appears to be mediated in large part by the T-cell response, which is much more difficult to measure than the antibody response. In certain cases, nasal vaccination can impart mucosal immunity, but not systemic immunity. In such cases, resistance to viral infection due to mucosal immunity will not be predicted accurately by a negative result for IgG systemic antibody. Antibody can be measured in nasal wash, however this is less quantitative and reliable than measurement in serum. Because of such issues, alternative approaches to predicting overall immunity have been developed and published. These include systems biology approaches of blood-based profiling of immune proteins and transcriptomics.

Summary of key general features of an open-source, citizen-science vaccine:

- **Safe ingredients.** Long history of published results.
- **Ease and robustness of production.** Simple to make but with the possible tradeoff of short stability. Nanoparticles of chitosan or its derivatives form spontaneously and reproducibly, over a wide range of conditions and ingredient concentrations.
- **Intranasal delivery.** The preferred delivery mode for respiratory and nasal viruses. Highest safety profile. Potentially stimulates both protective mucosal immunity at the site of infection, and systemic immunity.
- **Peptide epitope antigens.** Short peptides of linear epitopes are easiest to procure or produce, and are predicted to function best without specific structural constraints. These can be mined from a rich literature, or produced de novo. Antigens for incorporation into a vaccine can be produced synthetically or as recombinant expression proteins.
- **Booster schedule.** Allows use of intranasal delivery and a lower dose or milder adjuvant, yet has the potential to produce immune responses equivalent to a single dose of hyper-stimulatory adjuvants.

POSSIBLE MECHANISMS OF VACCINE-ENHANCED DISEASE, VACCINE-INDUCED AUTOIMMUNITY, AND MITIGATION STRATEGIES (updated 2020-10-24)

As mentioned briefly in the section above, a small number of injected vaccines have led to vaccine-enhanced disease (VED), meaning that disease is made more serious in people or experimental animals that have been vaccinated, relative to unvaccinated controls. This has occurred in response to vaccines for respiratory syncytial virus (RSV), dengue, Zika, and others. Some of these cases involve experimental vaccines, although RSV and dengue were deployed in healthy human populations. Although there isn't a single accepted or proven mechanism known to cause or contribute to VED, there are historical vaccine design features and adjuvant factors that are suspected contributors. The RSV vaccine that caused VED in the 1960s was formalin inactivated whole virus, adjuvanted with alum. This vaccine triggered a weak initial immune response, and theoretical explanations for how it caused VED have focused on the chemical alteration of antigenic structures caused by formalin and on the Th2 bias induced by alum. Recent studies have shown that, in addition to Th2 bias, RSV VED is driven by a few factors including deficiencies in Toll-like receptor activation and CD8 response, and hyperactivation of a distinct subset of CD4 T-cells.^{42, 43, 44} Mitigation of any of these reduces or prevents VED. In the case of dengue virus, a second infection with a distinct dengue serotype often causes more severe disease than the first infection. Use of the dengue vaccine Dengvaxia similarly results in more severe disease than in unvaccinated controls, suggesting that vaccination sets up a serious subsequent infection by similar mechanisms to a prior dengue infection.

One proposed mechanism for VED in dengue is antibody-dependent enhancement (ADE), in which antibodies of the systemic immune system act in "Trojan horse" fashion to facilitate viral entry into certain host immune cells. (for a review of ADE specific to SARS-CoV-2 see Eroshenko et al.⁴⁵). Waning or weak antibody response, including a substantial non-neutralizing antibody component to this response, has been shown to facilitate viral entry into macrophages and monocytes. There are multiple possible ways this might occur. In Dengue infection, there are multiple viral serotypes, which are distinguished by substantial protein sequence divergence. Amino acid differences in key epitopes can alter binding of antibodies, and the result is that antibodies that are neutralizing for one serotype bind antigens of another serotype, but in a manner that is weakly or non-neutralizing. One possible contributing factor in this process is a phenomenon referred to as "original antigenic sin." Exposure to an initial antigen produces

⁴² <https://www.nature.com/articles/nm.1894>

⁴³ <https://pubmed.ncbi.nlm.nih.gov/25769044/>

⁴⁴ <https://pubmed.ncbi.nlm.nih.gov/17911628/>

⁴⁵ <https://www.nature.com/articles/s41587-020-0577-1>

memory immune cells that recognize the amino acid sequence of this antigen, and upon subsequent exposure to a related but slightly different antigen, there is preferential expansion of memory cells that recognize the original antigen rather than cells that recognize and bind more tightly to the second antigen.⁴⁶ Thus, upon second infection, the antibody response is directed mostly at epitopes of the previous infection, and less robustly against the current infection. Original antigenic sin is not limited to B-cell epitopes, and essentially the same phenomenon has been reported for T-cell epitopes and responses.⁴⁷

Other aspects of the T-cell response also have been implicated in VED. In particular, a robust CD8 response is a key immune correlate of mild or asymptomatic infection for a number of respiratory viruses, including dengue, influenza, SARS-CoV-1, and SARS-CoV-2. Multivalent whole-virus dengue vaccines that provide robust and broad protection against dengue infection produce a robust CD8 response.⁴⁸ Notably, the CD8 epitopes that produce the strongest protection against all dengue serotypes are derived from conserved non-structural proteins, rather than the viral outer surface proteins, which have much greater amino acid sequence variation.⁴⁹ This appears to be a common feature of respiratory viruses, including influenza and SARS-CoV-2.^{50, 51}

In contrast, Dengvaxia is based on a hybrid virus that uses the non-structural backbone of the related yellow fever virus, together with the outer structural proteins from dengue. Dengvaxia elicits a strong initial antibody response, but this strong response wanes relatively quickly. Since Dengvaxia lacks the immunodominant CD8 epitopes of the non-structural proteins, the cytotoxic response is directed toward the weaker dengue surface epitopes, and the immunodominant non-structural proteins of the yellow fever virus. Therefore, not only is the CD8 response specific to dengue epitopes relatively weak, but a possible exacerbating factor is that the primary CD8 response triggered by Dengvaxia is directed away from dengue by immunodominant yellow fever epitopes. To summarize one model of Dengvaxia-induced VED favored by some leading immunologists: upon waning antibody response (and possible ADE), compromised cellular immunity is weak, and upon subsequent infection T-cells mount a pathological cellular immune response, possibly exacerbated by original antigenic sin, driven by the highly variable surface T-cell epitopes of the dengue serotype of the subsequent infection.

It is important to reiterate that immunopathological responses are not limited to vaccines, and are known to be mediated by viral exposures, as occurs in primary dengue virus infection followed by

⁴⁶ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7291596/>

⁴⁷ <https://pubmed.ncbi.nlm.nih.gov/9697771/>

⁴⁸ <https://jvi.asm.org/content/89/1/120.full>

⁴⁹ <https://www.frontiersin.org/articles/10.3389/fimmu.2019.02125/full>

⁵⁰ <https://pubmed.ncbi.nlm.nih.gov/18802496/>

⁵¹ <https://www.nature.com/articles/s41590-020-0782-6>

second infection by a distinct dengue serotype. However, as seen in RSV and dengue, vaccine designs can contribute to the likelihood of immunopathology. One early report on SARS-CoV-2 suggests that neutralizing IgG antibodies for the SARS-CoV-2 receptor binding domain (RBD) do not exhibit such enhancement; and as mentioned above, some experts suggest that ADE is not a clear concern in the development of vaccines against this virus^{52,53}. However, other experts are less confident that SARS-CoV-2 vaccine-induced ADE or other VED will not occur, and we take seriously even a low risk.

Given the immunopathology associated with Dengvaxia, we are especially concerned about the use of the same general design features that are suspected to have caused or contributed to VED. In particular, if SARS-CoV-2 immunodominant MHC class I/CD8 epitopes are concentrated on non-surface proteins, use of SARS-CoV-2 outer surface proteins alone might produce a weak CD8 response; and if this feature is combined with the backbone of another virus carrying potentially immunodominant CD8 epitopes, such a vaccine fits a general risk profile similar to Dengvaxia. As of late September, the most extensive quantitative T-cell epitope mapping data have been published by Adaptive Biotechnologies and made public in the ImmuneCODE database⁵⁴. The mapping of SARS-CoV-2 epitopes shows clear CD4 immunodominance, and also some degree of CD8 immunodominance. It is notable and concerning that the majority of immunodominant MHC class II/CD4 epitopes are found in Spike and other surface proteins (although these class II data do not include Orf1ab), while the majority of immunodominant MHC class I/CD8 epitopes (7 of the top 8; 12 of the top 15) of the entire proteome are found in non-structural proteins, analogous to the distribution found in dengue.

Vaccine-induced autoimmunity (VIA) is another immunopathological complication of vaccination. It is a rare but worrisome side effect. Over the past decades, several vaccines have been implicated as autoimmune triggers⁵⁵. Four possible factors have been proposed in explanations for mechanisms of VIA: molecular mimicry and cross-reactivity to host antigens; unbalanced or hyperstimulation of immunity by specific adjuvants; host genetics; and concurrent vaccination and infection. A combination of molecular mimicry of host antigens and certain potent adjuvants is especially concerning, and has been suggested as a leading explanation for VIA in narcolepsy triggered by Pandemrix, although specific Scandinavian host genetics appear to play a role. Other serious autoimmune diseases believed to be triggered by vaccines are Guillain-Barré syndrome, Multiple Sclerosis, and lupus. While these side effects are quite serious, they are also quite rare. The incidence of narcolepsy increased over 10-fold in those vaccinated with Pandemrix, but the vaccine-attributable rate was only 1 in 18,400⁵⁶. In each case, molecular

⁵² <https://www.biorxiv.org/content/10.1101/2020.04.10.036418v1>

⁵³ <https://www.nature.com/articles/s41587-020-0577-1>

⁵⁴ <https://www.researchsquare.com/article/rs-51964/v1>

⁵⁵ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6078966/>

⁵⁶ <https://pubmed.ncbi.nlm.nih.gov/28847694/>

mimicry is a suspected primary contributor, and viral infection can trigger these same autoimmune reactions, sometimes even more frequently than vaccines. Therefore, as with other decisions regarding vaccination during a global pandemic, if certain antigens pose an increased risk of autoimmunity, one is likely to be exposed to antigens either by viral infection, or through vaccination.

In the list below, we summarize possible mechanisms of VED in precedent cases that might contribute to VED immunopathology in response to vaccination against SARS-CoV-2, along with some of the mitigation strategies we have used to address each of these in our vaccine designs.

1. Th2 and/or Th17 immune hyperpolarization, possibly accompanied by an attenuated Th1 response have been proposed to be leading risk factors for VED. It has been suggested that use of alum as an adjuvant creates such a Th2 bias and inhibition of Th1, while the use of other adjuvants, such as chitosan and toll-like receptor (TLR) agonists shift the balance in favor of Th1.
 - a. In many published studies of chitosan, our chosen delivery and adjuvant vehicle, intranasal delivery elicits a balanced Th1/Th2 immune response.
 - b. TLR agonists have been used together with chitosan, including in intranasal formulations to further enhance the Th1 response. In order to simplify vaccine production and to create a higher margin of short-term safety, we have not included additional adjuvants in our designs; however, the inclusion of such secondary adjuvants is relatively straightforward, and might be used to further increase the Th1:Th2 ratio.
2. ADE of viral infection is primarily associated with IgG and the Fc γ receptor, and much less often with IgA and IgM, and their Fc receptors.⁵⁷ Secretory IgA (sIgA) of mucosa has been shown to mediate ADE of bacterial infection but has not been demonstrated for viral infection.⁵⁸
 - a. Secretory IgA is the dominant antibody isotype at mucosal surfaces. It remains unknown if ADE of viral infection can occur through sIgA, but it has not been reported. Nevertheless, a robust mucosal immune response should greatly reduce or prevent IgG-mediated ADE in the systemic response by reducing or abrogating initial infection.
3. Certain B-cell epitopes appear to be more likely to give rise to ADE, including on the Spike protein of SARS-CoV-1, resulting in concern about the use of full-length Spike protein in SARS-CoV-2 vaccines.⁵⁹

⁵⁷ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7119964/>

⁵⁸ <https://pubmed.ncbi.nlm.nih.gov/12967664/>

⁵⁹ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2756483/>

- a. Spike is the primary target for generating antibody-based protective immunity. It is a very large protein (1273 amino acids) with many antibody binding epitope sites, but only a small number of prospective epitopes are potentially neutralizing. This has been proposed to be a potential contributing factor to ADE.
 - b. Selected target epitopes included in the vaccine represent only a fraction of epitopes, including those shown to be bound by neutralizing antibodies. In the current vaccine design, only about 5% of the Spike protein sequence is represented as B-cell epitopes. If original antigenic sin occurs, it should only apply to these epitopes, and not affect other neutralizing epitopes of the virus or subsequent vaccines. This allows epitopes of this subsequent exposure to trigger activation, maturation and expansion of naive B-cells into plasma cells that secrete antibodies specific to these neutralizing epitopes of the second infecting viral serotype. There are multiple Spike protein conformational B-cell epitopes bound by neutralizing antibodies, and peptide vaccines such as those described in this white paper series do not include many of these portions of Spike.
 - c. Incorporation into a vaccine of only a small portion of all viral antigens also reduces the likelihood that a viral antigen will trigger VIA through molecular mimicry and immune cross-reactivity to a host antigen. Again, during a pandemic, if certain antigens of a pathogen pose an increased risk of autoimmunity, one is likely to either be exposed to antigens by viral infection, or through vaccination. Selected epitope peptide sequences can be compared to the human proteome (e.g. by BLAST) to further reduce cross-reactivity to host antigens. We perform a BLAST for every selected epitope peptide to minimize the potential for high-scoring matches to potentially cross-reactive host antigens.
4. Waning or weak T-cell response, possibly driven by T-cell cross-reactivity to divergent epitopes derived from prior exposures to different viral serotypes. One model for how this might occur is through a T-cell original antigenic sin, analogous to the phenomenon described for antibodies.⁶⁰
 - a. Again, it is impossible to completely avoid this since it is not restricted to vaccines, but also applies to viral infection. Nevertheless, selection of a few conserved epitopes should reduce the likelihood that a SARS-CoV-2 mutant will arise that results in this phenomenon. Following the same logic described for B-cell epitopes, if original antigenic sin occurs it can be limited to only those epitopes supplied in the vaccine. In the current vaccine design, less than 1% of the SARS-CoV-2 proteome (entire protein sequence) is represented as T-cell epitopes, and only about one-quarter to one-third of empirically verified immunodominant epitopes.

⁶⁰ <https://www.jimmunol.org/content/176/6/3821.long>

5. Similar to SARS-CoV-2, severe dengue infection is CD8 deficient, while mild infection is accompanied by a robust CD8 response. Multivalent dengue vaccines that elicit protective immunity include immunodominant MHC class I/CD8 epitopes from non-structural proteins. The Dengvaxia design employs a hybrid virus that uses the non-structural backbone of the related yellow fever virus, together with the outer structural proteins from dengue. Immunodominant MHC class I/CD8 epitopes are concentrated in the non-structural proteins of dengue, SARS-CoV-2, and other respiratory viruses, while immunodominant MHC class II/CD4 epitopes are concentrated in the outer structural proteins. Dengvaxia elicits a strong initial, but rapidly waning antibody response, and a weaker and possibly immunopathological CD8 T-cell response in response to subsequent infection.
 - a. RaDVaC vaccine designs incorporate a few immunodominant MHC class I/CD8 epitopes from non-surface and non-structural proteins. Some of these epitopes are highly conserved, and the most common HLA alleles are covered redundantly by at least 2 separate epitopes. We have not included all immunodominant epitopes for CD8 (or for B-cells or HLA class II/CD4), which should prevent original antigenic sin from compromising a future robust immune response to a cross-reactive serotype.
6. VED might arise through vaccine design-induced distortion of viral epitopes, as has been proposed to occur by formalin treatment of RSV.
 - a. Our preferred epitope type is synthetic peptides. Such peptides are chemically well defined and have a high degree of structural integrity.

KEY TECHNICAL FEATURES AND SPECIFICATIONS (Updated 2021-03-06)

Certain technical features of the RaDVaC strategy follow from the above considerations, and are key to an open-source, citizen-science vaccine.

1. **Intranasal delivery** is probably the safest choice of administration routes. For a respiratory virus, intranasal delivery has the advantage of eliciting a mucosal immune response at sites of viral entry. Intranasal delivery can also elicit systemic immunity, although often not as robust as parenteral delivery after a single dose.
2. **Chitosan nanoparticle and adjuvant(s)**. Successful vaccines depend on successful delivery and on immunogenic stimulation, often by the use of adjuvants that promote appropriate stimulation of the immune system. Many adjuvants are proven and available, with various strengths and weaknesses. A leading nanoparticle combined nasal delivery and adjuvant system is based on chitosan and sodium triphosphate (STP) a.k.a.

tripolyphosphate (TPP)⁶¹, which is the present top choice for this intranasal vaccine. Chitosan nanoparticles of various sizes can be created with simple tunable parameters.⁶²,⁶³ The virus has an average size of about 100 nm to 150 nm.⁶⁴ Nanoparticle size ideally will be tuned to be between 100 nm and 200 nm. As described above, chitosan has been shown to be safe and well tolerated, and intranasal delivery elicits both mucosal and systemic immune responses. Chitosan is available in various molecular weights, deacetylation ranges, and forms such as pure chitosan, trimethyl chitosan, and quaternary or quaternized chitosan, including hydroxypropyl-trimethyl ammonium chitosan chloride (HTCC or HACC). It is important to obtain high purity chitosan (including a certificate of analysis). Low-grade chitosan might not work well, or might be contaminated with allergenic shellfish proteins. Multiple forms, molecular weights and suppliers have been tested and published. Very low molecular weight chitosans (e.g. 1 to 20 kilodaltons; sometimes referred to as oligochitosan) have the advantage of being water soluble, yet the lower the molecular weight, the less stable the chitosan gel. Unmodified higher molecular weight (100+ kDa) chitosans are only soluble at pH values under their pKa values of 6.2 to 6.5 because their free amino groups are protonated. Deprotonation at physiological pH causes insolubility and loss of integrity of nanoparticles formed by ionic crosslinking. Trimethyl chitosan and HACC do not suffer from this problem, because they possess amino groups that are protonated at physiological pH. Therefore, after much experimentation we currently use HACC as our primary chitosan derivative, although some RaDVaC collaborators (aka Radvackers) are experimenting with trimethyl chitosan. Throughout this document, the general term chitosan can mean unmodified chitosan along with soluble derivatives, such as HACC and trimethyl chitosan.

- a. Possible additional adjuvants/immunostimulants. Chitosan is a Th1/Th2 immune triggering, self-adjuvanting polysaccharide; therefore, additional adjuvants or immunostimulants are not strictly necessary, and none will be incorporated into the initial formulation. However, if chitosan alone produces insufficient immune stimulation, we are considering experimenting with the incorporation of other adjuvants that enhance a Th1 or specific T-cell targeted response. Good candidates are alum (Th2 biased), imiquimod (TLR7 agonist, Th1 biased), all-trans retinoic acid (ATRA)⁶⁵, poly (I:C) derivatives (TLR3 agonist, Th1 biased), compound

⁶¹ <https://www.frontiersin.org/articles/10.3389/fimmu.2018.00934/full>

⁶² <https://www.nature.com/articles/s41598-018-23064-4>

⁶³ <https://www.ncbi.nlm.nih.gov/pubmed/28578269>

⁶⁴ <https://www.nejm.org/doi/full/10.1056/NEJMoa2001017>

⁶⁵ <https://jvi.asm.org/content/85/16/8316>

48/80 (mast cell agonist with balanced Th1/Th2/Th17 response)⁶⁶, and other TLR⁶⁷ (mostly Th1 biased) and mast cell⁶⁸ agonists .

3. **Peptide antigen.** A peptide-based vaccine is the method of choice, based on and bearing in mind the following considerations:
 - a. **Multimeric.** Multimeric/multi-epitope vaccines have been shown to be effective. The approach taken here uses multiple peptides, each of which will carry one or more B and/or T-cell epitopes.
 - b. **Fast and flexible.** Synthetic peptides are inexpensive, and can be made to order quickly by many commercial peptide producers. They can be synthesized with many possible chemical modifications that have been reported to increase immunogenicity.
 - c. **Epitope conformation and presentation.** Gen 10+ designs feature surface display of B-cell epitope peptides for enhanced antibody production. Early versions of the RaDVaC vaccine contained simple linear epitopes, with no special regard to conformation or presentation. Some of the B-cell epitope peptides from Gen 3 onward were conformationally constrained with disulfide bonds. Ideally, 3D structures of viral proteins should be imaged and linear epitopes should be chosen that either don't require special conformational constraints, or attempts should be made to achieve natural structural conformation of the epitope peptide.
 - d. **Non-peptide antigens.** Other antigens are possible and can be delivered by chitosan intranasally, including proteins, DNA or RNA.
4. **Epitope selection** is critical. B-cell and T-cell epitopes have been selected and published by others. Multiple epitopes of both types should be selected, preferably in important functional and/or conserved regions of the virus. This gives a higher probability that at least one epitope will successfully stimulate immunity, especially given variability in individual immune responses, rather than relying on a single epitope.
 - a. B-cell epitopes. Studies of antibody mapping from convalescent patients help identify portions of the virus available for antibody binding to B-cell epitopes, and some of these antibodies have been shown to neutralize viral infection in cellular studies or specially designed neutralization assays. Certain B-cell epitopes that score highly by common machine learning predictions have not been reported at high frequency in convalescent sera. This might be due to linear peptide-based approaches used to map epitopes, the extensive glycan shield around the virus, or other complicating factors. Whatever the explanation, we regard the convalescent antibody data as superior to purely computational predictions.

⁶⁶ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2743390/>

⁶⁷ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3000864/>

⁶⁸ <https://www.nature.com/articles/nm1757>

- b. T-cell epitopes. Empirical methods similar to but more complex than those used for B-cell epitope selection have been used for selecting superior T-cell epitopes.
- 5. **Testing.** Success of mucosal and systemic humoral immunogenic stimulation will be assessed by testing antibody titers of nasal wash, saliva, and serum. Tests of cellular immunity (T-cell) will also be performed, although as of mid Feb 2021 there are no commercial tests yet available for these kinds of tests, which require substantial laboratory infrastructure. Testing of immunity should be performed using standard assays and reagents, as well as the use of newer technologies, such as transcriptome profiling of peripheral blood mononuclear cells (PBMCs).

GOALS

1. Design and incrementally improve a vaccine made with commercially available materials, and with the following properties:
 - a. All ingredients have undergone extensive testing for efficacy and safety in animal models, and preferably in human trials. Delivery and adjuvant materials and technologies must have decades of proven safety.
 - b. Use of best practices to achieve high probability of efficacy.
 - c. Mode of administration must be safe and proven. Intranasal delivery is preferred for safety and for establishment of both mucosal and systemic immunity. Inhalation and injection are possible delivery modes for this vaccine but should only be pursued by those with sufficient training and experience.
 - d. Vaccine should be extremely easy to make (including sourcing of ingredients) and administer.
2. Produce vaccine for initial testing, and self-administration.
3. Produce and administer updated generations of vaccines as new information becomes available.
4. Design and implement testing for immune responses to the vaccine. Tests will be performed with published methods, and new methods should be developed to predict efficacy. Ideally, biological samples such as blood and nasal wash are collected before initial vaccination, and at various time points.
5. Openly share and disseminate information, enabling many others to produce the vaccine. The only custom components are the epitope peptides. By themselves, these synthetic peptides are essential components not only of vaccine, but also of standard types of testing (ELISA, ELISpot, etc.), and thus are not drugs, and might be distributed to facilitate progress in multiple kinds of research.
6. Maintain ongoing long-term self-monitoring of health and symptoms to collect empirical evidence of degree of efficacy and possible unforeseen effects.

7. Create an online forum for the exchange of data, ideas and best practices with other individuals and groups. Over time, results can be aggregated for broader analysis.

MATERIALS AND METHODS (updated 2021-03-07)

Selection and purchase of the vaccine ingredients

All materials and ingredients are commercially available. However, some vendor deliveries might be interrupted by public health measures. The vaccine is very simple and consists of five ingredients: epitope peptide(s), chitosan, sodium triphosphate, water, and sodium chloride.

- Epitope peptides. Multiple vendor options are available for synthetic peptides. Peptides or concatemers of the selected epitopes can also be produced by other means, such as recombinant plasmid-based expression and purification (e.g. *E. coli* His tag). Options should be considered and tested, as supply channels might be disrupted by the pandemic.
- Chitosan derivatives. Chitosan is available in various molecular weights, deacetylation ranges, and forms such as pure chitosan, trimethyl chitosan, and quaternary or quaternized chitosan, including N-(2-hydroxy)propyl-3-trimethyl ammonium chitosan chloride (HTCC or HACC, also known as N-HTCC or N-HACC). (Note that O-HTCC or O-HACC are not the same as N-HACC, and degrees of substitution and therefore molecular weights will differ, although they likely will work similarly.)⁶⁹ It is important to obtain high purity chitosan (including a certificate of analysis). Low-grade chitosan might not work well, or might be contaminated with allergenic shellfish proteins. Because unmodified chitosan is water insoluble at physiological pH, undermining nanoparticle integrity *in vivo*, we currently use HACC as our primary chitosan derivative (although some Radvackers are experimenting with trimethyl chitosan). Acceptable ranges: medium to high molecular weight range of 100-400 kDa. Deacetylation = 75% to 90%. Depending on the source, some initial experimentation is likely necessary to test various behaviors of the stock solution. If you only have access to unmodified chitosan, pre-Gen 10 protocols can be used⁷⁰; however, soluble chitosan derivatives will give superior results.
- Sodium triphosphate, Na₅P₃O₁₀. Multiple suppliers.
- Sodium chloride (NaCl). 5M stock solution.
- dH₂O, deionized water. Widely available.

⁶⁹ <https://onlinelibrary.wiley.com/doi/pdf/10.1002/app.26038>

⁷⁰ <https://radvac.org/wp-content/uploads/2020/12/White-Paper-SARS-CoV-2-vaccine-ver-3-3-1.pdf>

- OPTIONAL (If you are using unmodified chitosan instead of soluble chitosan derivative): Distilled white vinegar (or acetic acid CH_3COOH). White vinegar is typically about 5% acetic acid.
- OPTIONAL: Granular activated charcoal for circularizing optional peptides.
- OPTIONAL: Sodium bicarbonate NaHCO_3 , or other mild base for adjusting the pH.

Preferred equipment and materials (updated 2020-11-04)

Certain equipment is required for efficient vaccine production and administration. Vaccine can be produced without specialized laboratory equipment but the process is more laborious and the results are likely to be more variable.

- Pipettes: 1000 microliter, 200 microliter
- Sterile filtered pipette tips: 1000 microliter, 200 microliter
- Scale: gram scale accurate to 0.1 grams, or jeweler's scale for very small amounts
- Clean spatula for dry reagents
- Small beakers for mixing and stirring. 10 ml to 25 ml beaker for stirring
- Small bottles for stock solutions; 100 ml to 250 ml
- 15 ml conical tubes
- 50 ml conical tubes
- 1.5 to 2 ml microcentrifuge tubes
- Stirring apparatus: magnetic stir plate and small stir bar to fit in beaker
- OPTIONAL but preferable: pH strips, range at least 2.0 to 9.0
- Nasal spray apparatus. These are available through multiple vendors. We selected small nasal spray bottles (about 5 milliliter) with a pump-top apparatus. A tube stem about 4 cm long and approximately 1 mm inner diameter extends from the bottom of the apparatus, which can be placed into the bottom of a 2-ml tube containing vaccine. Our spray apparatus delivers about 100 microliters per pump.

Selection of epitopes (updated 2021-03-06)

Epitopes are a critical element of a vaccine because they are specific to each pathogen. An epitope is the portion of an antigen (typically a protein) that interacts with the immune system, such as through binding of antibodies or T-cell receptors. Multiple data types can be used in the selection of epitopes. This section describes general guidelines and issues for consideration in epitope selection and design.

- **Epitopes for B-cell and T-cell responses can be quite different**, and there are critical differentiating features. Here are key considerations in the selection and design of B-cell and T-cell epitopes.
 - T-cell epitopes are presented as short linear peptide antigens, but B-cell epitopes (which can be linear or conformational) are presented in their native conformation.
 - The B-cell response is relatively easily measured by antibodies, but the T-cell response is more difficult to measure and more rarely performed. Antibody testing is a large-scale commercial activity (using lateral flow, ELISA, etc.) but T-cell testing is generally a research activity (using ELISpot, etc.). Therefore, T-cell epitopes are often based on computational prediction. This was true of early T-cell epitope selections in the biomedical literature and preprint servers. However, beginning in mid July, 2020, substantial empirical data have been published on T-cell epitopes.
 - B-cell or antibody binding epitopes are often mapped into two general classes: linear and conformational (or structural). Linear epitopes (single peptide chain of nearby amino acids) are typically mapped by binding antibody-containing serum from convalescent patients to large collections or arrays of short epitope peptides collectively representing viral proteins. Conformational epitopes often consist of amino acids that are far apart on a protein sequence but in close proximity in the folded protein. Linear epitopes are easier to map than conformational epitopes. However, conformational epitopes can be critically important.
 - There are 2 classes of T-cell epitopes: MHC Class I and MHC Class II (for a thorough review, see Rock et al.⁷¹). Core motif binding of both Class I and Class II is about 9 amino acids but both are variable. The Class I binding site has closed ends, and epitopes range from 8 to 15 aa, but infrequently diverge from 9 aa. Class II binding is open-ended, and natural epitopes extend longer than the 9 aa core, and range from 13 to 25 aa⁷². Many computational predictive algorithms are available, and a large number of preprints provide Class I and Class II predictions for inclusion in a vaccine. If empirical data are available for protective T-cell

⁷¹ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5159193/>

⁷² <https://pubmed.ncbi.nlm.nih.gov/1380674/>

epitopes (e.g. Le Bert et al, Adaptive Biotechnologies, Peng et al⁷³, and Zhang et al⁷⁴), these data should guide T-cell epitope selection.

- Immune responses (both B and T-cell) can vary substantially with different ancestry/ethnicity. We have used predictions and empirical data for wide coverage of immune responses (our group is highly ethnically diverse). For specific geographies and ancestries, local immune response information should be taken into account.
- Although there can be differences between T and B-cell epitopes, they can overlap. If possible, peptides that contain experimentally validated epitopes for both B and T-cells are ideal.
- **Immunodominance and limited effective epitope space.** Immunodominance is the phenomenon of antibody and/or T-cell responses to a small number of epitopes, despite a large number of possible epitopes within the virus. Findings to date suggest both B-cell and T-cell responses to SARS-CoV-2 display pronounced immunodominance. Some key findings about the immunodominant landscape of SARS-CoV-2:
 - A publication by Snyder and colleagues shows a few epitope peptides are responsible for over 50% of CD8 T-cell receptor binding events⁷⁵.
 - The immunodominance landscapes of CoV-1 and CoV-2 are:
 - Very different between CoV-1 and CoV-2 Spike proteins, possibly largely due to glycosylation differences.⁷⁶
 - Highly similar for CoV-1 and CoV-2 Nucleocapsid proteins^{77, 78}.
 - One early but thorough preprint by Zhang and colleagues suggests immunodominance is a feature of the B-cell immune response to SARS-CoV-2⁷⁹.
 - Farrera and colleagues reported a set of immunodominant B-cell epitopes similar to those reported by Zhang, but in a study of Swiss convalescent patients⁸⁰.
 - Most of the epitopes reported by Zhang et al and Farrera et al were subsequently confirmed by Li and colleagues in the first large-scale linear B-cell epitope mapping of over 1,000 convalescent patients in Wuhan, China⁸¹.
 - The emerging consensus suggests that there are about half a dozen dominant B-cell epitopes, which are the same in European and Chinese patients. The

⁷³ <https://www.nature.com/articles/s41590-020-0782-6>

⁷⁴ <https://pubmed.ncbi.nlm.nih.gov/32668444/>

⁷⁵ <https://www.medrxiv.org/content/10.1101/2020.07.31.20165647v2>

⁷⁶ <https://pubs.acs.org/doi/full/10.1021/acsinfecdis.6b00006>

⁷⁷ <https://jcm.asm.org/content/42/11/5309.full>

⁷⁸ <https://www.biorxiv.org/content/10.1101/2020.03.26.994756v1>

⁷⁹ <https://www.biorxiv.org/content/10.1101/2020.04.23.056853v2.article-info>

⁸⁰ <https://www.medrxiv.org/content/10.1101/2020.06.15.20131391v1>

⁸¹ <https://www.medrxiv.org/content/10.1101/2020.07.13.20152587v1.full.pdf>

majority are found in key functional portions of the Spike protein, including the following (amino acid coordinates shown are approximate):

- Spike 450-500; ACE2 binding residues of the RBD (Zhang et al); low degree of conservation; probably moderate to high mutant escape potential
 - Spike 655-685; immediately adjacent to the S1/S2 protease cleavage sites; high degree of conservation in the antibody-binding N terminal portion of the epitope; likely low to moderate mutant escape potential. Cleavage of this site greatly increases efficiency of host cell infection⁸².
 - Spike 805-820; spanning the S2' protease cleavage site (815-816); high degree of conservation; likely low mutant escape potential.
 - Spike 1145-1160; adjacent to the heptad repeat 2 (HR2) region, which is critical for fusion of the viral and host cell membranes of the SARS virus⁸³; very high degree of conservation; likely extremely low mutant escape potential.
- These immunodominance findings are supportive of the peptide vaccine approach, since only a limited portion of the virus needs to be represented in the vaccine for it to be effective; plus, potentially counterproductive portions of the virus (such as those that might contribute to ADE) are not present in the vaccine.
- **The critical importance of MHC Class I/CD8 T-cell epitopes and immunity.** Antibodies get a lot of press attention (in large part because simple commercial antibody tests are widely available but T-cell tests are more complex and must be performed on isolated cells, and therefore are done in specialized research labs) but the most protective immune response consists of robust humoral (B-cell mediated) and cellular (T-cell mediated) responses. Early indications suggest T-cell responses might be key to both minimizing severity of illness and to longevity of immunity, but the specific type(s) of T-cell response that might be most protective is a critical yet still unresolved issue. Here are a few key findings about the importance of T-cell immunity to SARS-CoV-2 infection:
 - Mild COVID-19 is seen in convalescents with robust T-cell responses, including those who do not have detectable antibodies (are seronegative)⁸⁴
 - Severe COVID-19 is accompanied by higher antibody levels than mild cases, but with unbalanced or impaired T-cell response, especially the CD8 response⁸⁵. This study by Peng et al identified 6 immunodominant CD4 epitopes (3 of which are located on the Spike protein), a similarly small number of immunodominant CD8

⁸² <https://www.pnas.org/content/117/21/11727>

⁸³ <https://www.thelancet.com/journals/lancet/article/PIIS0140673604157887/fulltext>

⁸⁴ <https://www.biorxiv.org/content/10.1101/2020.06.29.174888v1>

⁸⁵ <https://www.biorxiv.org/content/10.1101/2020.06.05.134551v1.full.pdf>

epitopes, and they showed that a higher ratio of CD8/CD4 response is found in those with milder disease versus those with severe disease.

- The effect of T-cell epitopes can vary depending on the protein of origin. Peng and colleagues also report that T-cell epitopes from Spike, Membrane and ORF3a proteins were associated with more severe disease. Higher frequencies of multi-cytokine producing Membrane and Nucleocapsid-specific CD8 T-cells (relative to Spike-specific CD8 T-cells) were observed in patients with mild cases. Because of these findings, Peng et al conclude that “The identification of T cell specificity and functionality associated with milder disease, highlights the potential importance of including non-spike proteins within future COVID-19 vaccine design.” Many subsequent publications have further strengthened this conclusion. Nevertheless, the CD4 epitopes included in all RaDVaC designs to date do not include the 3 immunodominant CD4 Spike protein epitopes identified by Peng et al, and we have no plan to include these in future designs until this concern about CD4 hyperstimulation is resolved.
- Peng et al also used single cell analysis of bronchoalveolar lavage fluid from COVID-19 patients and found substantial clonal expansions of CD8 T cells in mild but not severe patients, further reinforcing evidence that a robust CD8 response protects against severe disease.
- T-cell exhaustion or deficiency is correlated with disease severity (low T-cell response with more severe disease)^{86, 87, 88, 89}
- Population MHC class I binding (and CD8 response) is predictive of mortality⁹⁰.
- Vaccine-enhanced disease involves a Th2-polarized response that is CD8 T-cell deficient⁹¹. A robust CD8 response is observed in people and animal models who do not develop vaccine-enhanced disease.
- Many reports have shown rapid declines of anti-SARS-CoV-2 antibodies in convalescents. It remains unclear how this decline impacts protection from reinfection or severity of disease, but initial results also show that T-cell immunity might be very long-lived. One report by Le Bert and colleagues shows that patients who recovered in 2003 from SARS-CoV-1 infection harbor SARS-CoV-2 cross-reactive T-cells 17 years later⁹².

⁸⁶ <https://www.nature.com/articles/s41423-020-0401-3>

⁸⁷ <https://www.frontiersin.org/articles/10.3389/fimmu.2020.00827/full>

⁸⁸ <https://www.medrxiv.org/content/10.1101/2020.06.16.20130914v1>

⁸⁹ [https://www.cell.com/immunity/fulltext/S1074-7613\(20\)30333-2](https://www.cell.com/immunity/fulltext/S1074-7613(20)30333-2)

⁹⁰ <https://www.medrxiv.org/content/10.1101/2020.05.08.20095430v3.full>

⁹¹ <https://link.springer.com/content/pdf/10.1007/s12026-007-0071-6.pdf>

⁹² <https://pubmed.ncbi.nlm.nih.gov/32668444/>

- Optimal T-cell epitopes vary by ethnicity and geography (MHC/HLA allele frequencies can be calculated using various tools⁹³. (MHC in humans is also called Human Leukocyte Antigen, or HLA). Epitopes providing extensive MHC coverage have been reported for many pathogens, including SARS-CoV-2 (for example, as published by Ferretti et al⁹⁴ and by Liu and colleagues⁹⁵). Such publications are useful for selection of T-cell epitope peptides for use in specific geographical regions and ethnicities. The publication by Ferretti and colleagues provides empirical validation of CD8 T-cell epitopes, broad coverage of the most common MHC haplotypes, and largely confirms the empirical findings of Le Bert and of Snyder et al⁹⁶. The study by Liu is computational and theoretically provides more complete haplotype coverage; however, only one of the SARS-CoV-1 cross-reactive T-cell epitopes described by Le Bert et al is present in this set of predicted epitopes, and it is predicted by Liu et al to be optimal for MHC1 and MHC2 haplotypes found mainly in Central Africa, a population not represented in the Singapore convalescents (mainly East Asian and European ancestry). And the top immunodominant MHC Class I / CD8 epitopes empirically verified by Snyder and colleagues are not represented in the computationally predicted set of Liu. Therefore, such computational predictions should be used with the caveat that they are not always consistent with empirical data.
- **Peptide epitopes** are preferred because they have the following advantages over other epitopes:
 - Well defined
 - Potentially pure
 - Enrichment of antigenic targets; not accompanied by substantial unnecessary and potentially counterproductive/non-neutralizing antigenic targets. Many epitopes are very short and well-represented by peptides (although uncertainty about epitope boundaries and addition of flanking amino acids to improve peptide solubility can warrant the selection of slightly longer sequences).
 - Some critical neutralizing B-cell epitopes can be very small, less than 20 amino acids.
 - The majority of CD4 T-cell epitope peptides are less than 20 amino acids
 - CD8 epitopes are often 9-11 aa.
 - Short production/turnaround time
 - Scalable; we currently use 50 to 100 micrograms total peptide per dose

⁹³ <http://allelefrequencies.net/>

⁹⁴ <https://www.medrxiv.org/content/10.1101/2020.07.24.20161653v2>

⁹⁵ [https://www.cell.com/cell-systems/fulltext/S2405-4712\(20\)30238-6](https://www.cell.com/cell-systems/fulltext/S2405-4712(20)30238-6)

⁹⁶ <https://www.medrxiv.org/content/10.1101/2020.07.31.20165647v2>

- Inexpensive relative to other epitope/antigen choices
- **Empirical evidence** should dominate selection criteria. Here are some best types of evidence:
 - Mapping of epitopes in blood and other samples collected from convalescent patients (ideally stratified by severity of illness). This can be accomplished by a few primary means:
 - 3D structural studies and modeling of neutralizing antibody binding to a viral antigen (e.g. Spike protein)
 - Mapping of linear B-cell epitopes by binding antibodies in convalescent sera to a library of peptides representing viral antigens. A strong signal in a linear epitope mapping study does not guarantee that the epitope peptide in the context of a vaccine will trigger the production of an antibody that binds to this epitope within the context of the virus. However, it is a good indicator that this is at least possible. Peptides can be constrained to approximate native conformation, making it more likely to bind the native epitope.
 - Mapping of T-cell epitopes by stimulating convalescent T-cells with epitope peptides, and measuring their response (e.g. cytokine secretion; ELISpot)
 - Epitope peptides from a peptide vaccine that has shown protection against infection
 - Successful use of epitope peptides in vaccines that elicit antibodies (or serum) effective in virus neutralization assays. B-cell epitopes that allow antibody binding to the virus but don't block viral function might increase risk of antibody-dependent enhancement.
 - Mapped epitopes that are effective in virus neutralization assays (e.g. peptides compete with viral sequences in cellular infection assays).
 - Successful use of epitope peptides in vaccines that elicit T-cell responses, or peptides shown to stimulate T-cells or cytokine production in ELISpot or other T-cell assay in cells from convalescents.
- **Multiple epitopes with low mutant escape potential.** It is important to note that most published neutralizing antibodies target Spike RBD, as do many vaccines in commercial development. However, given the high degree of mutability of the RBD portion of Spike, it is important to identify and select targets outside the RBD because of mutant escape potential. This strategy is summarized by Brouwer and colleagues: "In light of the rapid emergence of escape mutants in the RBD of SARS-CoV-1 and MERS, monoclonal NAb's targeting other epitopes than the RBD are a valuable component of any therapeutic antibody cocktail. Indeed, therapeutic antibody cocktails with a variety of specificities

have been used successfully against Ebola virus disease and are being tested widely in clinical trials for HIV-1. NABs targeting non-RBD epitopes have been identified for SARS-CoV-1 and MERS, supporting the rationale...” for SARS-CoV-2⁹⁷. Three neutralizing epitopes described in the following section, containing or immediately adjacent to the Spike protease recognition sites and heptad repeat 2 (HR2) fusogenic domain are highly conserved among these 3 coronaviruses, suggesting that they have very low mutant escape potential.

- Two of these sites are cleavage sites for host cell proteases. Cleavage of the S1/S2 site by furin is required for subsequent cleavage of the s2' site by TMPRSS2. Cleavage by both is required for efficient fusion of the virus to the host cell membrane, and these sequences are constrained by recognition by these host proteases. Another region of the virus (heptad repeat 2 or HR2) is also required for membrane fusion, and the homologous 50 amino acid region of the SARS-CoV-1 virus is both 100% conserved, and is bound by neutralizing antibodies.
- According to 3D structural model data, relative to the RBD, these epitopes are structurally simple, and the peptides can be short (under 25 amino acids) and are water soluble. These features make them ideal for a peptide vaccine. The two protease sites loop out from the Spike protein surface, which makes them accessible to the required host proteases, but also to antibodies. The HR2 domain is predicted to form a coiled-coil “tether” that links the Spike trimer to the viral membrane. The linear epitope peptide should adopt a coiled coil conformation very similar to that exhibited within the native Spike protein.
- Therefore, rather than focusing on ACE2-binding epitopes in the highly mutation prone RBD to inhibit virus binding to the ACE2 receptor, we targeted these B-cell epitopes in the highly conserved portions of the Spike protein to strategically neutralize proteolytic cleavage and membrane fusion. Furthermore, all three are bound by antibodies present in the sera of large fractions of convalescents, and they produce among the highest signals in linear epitope mapping studies, which are far higher than signals measured for binding to any linear epitope in the RBD.
- The Gen 7+ vaccine formulations still contain epitope peptides from within the RBD, including 2 peptides in the ACE2-binding region. However, these 2 peptides are unlikely to be conformationally correct and elicit antibodies that bind the native Spike conformation. The reason we include these is that they are predicted to be effective CD4 T-cell epitopes, although empirical data from convalescent patients has not yet confirmed animal studies and computational predictions.

⁹⁷ <https://science.sciencemag.org/content/early/2020/06/15/science.abc5902>

- **3D modeling of viral antigens** (especially Spike protein) is very helpful in B-cell epitope selection and peptide design.
 - Surface loop structures are often good antibody binding sites. Unstructured, hydrophilic loops that are stabilized by a minimal number of neighboring amino acid residues make good linear (rather than conformational) epitopes. One prime example in SARS-CoV-2 is Spike 804 to 818, the TMPRSS2 cleavage site (one of a few immunodominant epitopes reported by both Farrera et al and Li et al).
 - Certain amino acid residues are overrepresented and underrepresented in preferred antigens. According to structural studies, these amino acids are overrepresented (from most overrepresented): W, Q, P, K, S, E, N, H, C. And these are underrepresented (from most underrepresented): A, V, I, F, L, Y ⁹⁸.
 - There are multiple excellent 3D modeling packages and online servers. The NCBI supported iCn3d is very good and allows automatic access to Protein Database (PDB) files of viral protein structures⁹⁹. Multiple 3D models of SARS-CoV-2 proteins are available, including Spike in various conformations, and complexed with antibodies and with the ACE2 receptor.
 - Conformational B-cell epitopes can be difficult to produce with peptides, but potentially might be designed as single peptides that join discontinuous portions of virus proteins, or can be produced as multiple peptides that are predicted to physically interact into relatively stable conformations.
- **Physical properties of peptides.** Certain peptides have excellent physical properties such as solubility and stability against proteolysis, and are synthesized easily; others are more problematic. Peptides with marginal properties can be improved by amino acid substitutions, additions, or subtractions. Bear in mind that changes to the viral sequence might substantially alter antigen recognition. It is likely preferable to include peptides that are in critical functional regions of the virus, with minor modifications to improve their use in a vaccine, than to exclude them. It is widely reported throughout the immunology and vaccine literature that both T and B-cell immunity are retained to varying degrees for new mutant strains. It has been reported that T-cell responses to SARS-CoV-2 are present in 34% to 60% of people not exposed to SARS-CoV-2, and are likely due to exposures to distantly related coronaviruses (most epitopes less than 50% aa identity)^{100, 101}. Therefore, conservatively modified epitope peptides should be expected to provide high levels of protection, although possibly not equivalent to the unaltered viral amino acid sequences.

⁹⁸ <https://www.jimmunol.org/content/181/9/6230.long>

⁹⁹ <https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html>

¹⁰⁰ [https://www.cell.com/cell/fulltext/S0092-8674\(20\)30610-3#articleInformation](https://www.cell.com/cell/fulltext/S0092-8674(20)30610-3#articleInformation)

¹⁰¹ <https://www.medrxiv.org/content/10.1101/2020.04.17.20061440v1>

- **Length.** Binding of the S protein to the cellular receptor ACE2 is mediated through the receptor binding domain (RBD). The RBD has been mapped for the related SARS-CoV-1 virus, and for SARS-CoV-2. ACE2 binding residues in RBD have been published widely and can be visualized in Figure 3 of Tarek et al¹⁰². However, multiple studies using short linear peptides for mapping epitopes bound by antibodies in convalescent sera reveal that only low signals are found in peptides within the RBD. This contrasts with several publications showing the isolation of neutralizing antibodies that bind to the RBD. Most of these mapping studies use epitope peptides of 12 to 15 amino acids, while Zhang et al use longer peptides of 20 to 25 amino acids and show that antibodies in convalescent sera show at least moderate binding to multiple RBD epitopes. Therefore, the peptides we have used in our vaccine formulations, especially in the RBD are longer than 20 amino acids, and as long as 40 amino acids.
- **Solubility.** Certain peptides of the high-scoring epitopes are predicted to have low solubility in water. This complicates formulation, since organic solvent or pH adjustment is required to solubilize. When choosing epitopes and peptides, bear in mind aqueous solubility. Solubility calculators are helpful in this process (for example, see ¹⁰³ and ¹⁰⁴). Marginal solubility of peptides can be improved by amino acid substitutions, additions, or subtractions. Acidic (Asp, Glu) and basic (Arg, His, Lys) amino acids aid solubility; combinations of acidic plus basic amino acids can keep the overall peptide charge near neutral.
- **Presentation.** B-cell epitopes presented on the surface of virus-like particles elicit stronger antibody responses than those encapsulated within particles. Both presentation and solubility can be enhanced by the synthesis of a terminal poly-acidic domain (e.g. several aspartic acid or glutamic acid residues), combined with a soluble chitosan derivative^{105, 106}.
- **Stabilization.** Circular peptides are more stable. Synthesis options are available to produce N-terminal to C-terminal cyclization, or to place cysteines in strategic positions on epitope peptides for disulfide formation. Cyclization stabilizes 3D conformation, and increases resistance to exopeptidase degradation. Disulfide formation is quick, easy, and inexpensive by the treatment of peptides with activated charcoal. Cysteines should be designed into epitope peptides at positions that do not change or interfere with normal epitope function, or do so as little as possible. Replacement of amino acids with cysteine is indicated by the

¹⁰² <https://www.medrxiv.org/content/10.1101/2020.04.25.20079426v1>

¹⁰³ <https://pepcalc.com/peptide-solubility-calculator.php>

¹⁰⁴ [https://www.thermofisher.com/us/en/home/life-science/\[...\]/peptide-analyzing-tool.html](https://www.thermofisher.com/us/en/home/life-science/[...]/peptide-analyzing-tool.html)

¹⁰⁵ <https://www.mdpi.com/2076-393X/3/3/730/html>

¹⁰⁶ <https://www.sciencedirect.com/science/article/abs/pii/S0968089619303347?>

lower case letter “c” in certain epitope peptides. Blocking the N terminus of a peptide also has been reported to stabilize it against proteolysis.¹⁰⁷ Peptides might be either N-terminal acetylated or biotinylated.

We have produced several iterations of vaccine, and for each we selected multiple B-cell and T-cell epitopes. Which B-cell epitopes elicit antibody responses can be verified by immunoassay testing, and these epitope peptides can be carried forward in updated versions of vaccines. The use of multiple epitopes also ensures that, if vaccination is successful in eliciting immunity against, say, 2 or 3 epitopes, then a single mutation will not allow the virus to evade immunity. Mutating virus and epitope variants should be monitored (nextstrain.org, bioRxiv, etc) , and new epitope peptides incorporated into evolving designs as needed.¹⁰⁸

Mapping of antibodies from people previously infected but recovered (convalescent) has shown that there are key regions for antibody binding. SARS and MERS are lethal coronaviruses that have been used for comparison and cross-reactivity of convalescent antibodies. Jiang et al used SARS-CoV-2 specific proteome microarrays to show that S and N proteins bind both IgG and IgM, and the C terminal portion of the N protein exhibits the highest convalescent serum antibody binding signal of all tested proteins.¹⁰⁹ Antibodies from convalescent patients have been mapped to short, linear epitope peptides by Poh and colleagues¹¹⁰ (S protein only), Wang et al¹¹¹, and the previously mentioned Zhang et al, Farrera et al, and Li et al.

Neutralizing (or protective) antibodies are preferred to guide B-cell epitope selection. Antibodies to certain functional regions of the virus will neutralize its activity and infectivity. Many antibodies bind to the virus and some will inhibit function but only a subset of all antibodies will truly neutralize the ability of the virus to infect cells and replicate. Ongoing research is helping to identify which portions of the virus are key to neutralizing activity. Neutralizing activity is not observed in antibodies associated with antibody-dependent enhancement of disease (ADE).

High-resolution linear epitope mapping by Wang et al was done with arrays of peptides covering the entire length of ten SARS-CoV-2 proteins. Length of peptides is 15aa, and peptides overlap by 5aa. Convalescent sera IgG and IgM were applied separately to the arrays. Binding events to a given peptide but not to adjacent peptides in the array indicate tight binding primarily to the central 5aa portion. Table 3 shows the 61 bound epitopes. Only five of these span more than one peptide. One in the S protein spans 15aa or three peptides (start = 806). These stretches are highly suggestive of a long linear epitope, whereas shorter epitopes are potentially a portion of a structural epitope. Moreover, the longer sequence provides a larger target. The longer S protein

¹⁰⁷ <https://www.pnas.org/content/pnas/102/2/413.full.pdf>

¹⁰⁸ <https://nextstrain.org/ncov>

¹⁰⁹ <https://www.medrxiv.org/content/10.1101/2020.03.20.20039495v1>

¹¹⁰ <https://www.biorxiv.org/content/10.1101/2020.03.30.015461v1>

¹¹¹ <https://www.biorxiv.org/content/10.1101/2020.03.26.994756v1>

epitope binds both IgG and IgM, and a largely overlapping epitope binds convalescent antibodies in the study by Poh and colleagues, and produces the highest signal in the later and larger-scale linear epitope mapping study by Li and colleagues. The epitope at this position (about Spike 805-820) spans a key proteolytic cleavage site required for viral entry into host cells. We included this epitope in Gen 3+ vaccine formulas, and as updated and progressively rich scientific information has become publicly available, we have added additional B-cell and T-cell epitopes to successive generations of vaccine. Each of these additions is detailed with references in the sections below.

Peptides and vaccine composition (last updated 2021-03-05)

Peptide synthesis general guidelines

Peptides are the most expensive and only custom component of RaDVaC vaccine designs. They can be synthesized at various scales, from individual to population scale with relative ease. The price of a peptide varies considerably between suppliers, and depending on additional requirements such as purity and modification.

- **Purification.** Peptides do not need to be purified, although certain peptides will perform best if purified. In particular, longer peptides, B-cell epitopes, li-key peptides, and those predicted to be difficult to synthesize due to low coupling efficiency are ideally purified. Short peptides predicted to be easy to synthesize (generally due to minimal hydrophobic amino acids) do not require purification. Peptide “purification” is commonly misunderstood but simply means that full-length peptides are separated from partial peptides resulting from truncated synthesis. These truncation products are due to imperfect and sometimes highly variable amino acid coupling efficiencies during synthesis of the growing peptide chain. Since peptides will be chopped up by proteases, purification does not add much value unless predicted coupling efficiencies are low and the resulting fraction of full-length peptide is low. Estimated peptide synthesis coupling efficiencies can be predicted¹¹². Again, for peptides predicted to be difficult to synthesize (especially B cell epitopes and li-key peptides), consider purification.
- **Desalting.** Peptides should have TFA removal and replacement with acetate. This is commonly known as desalting or buffer exchange, and standard grade buffer exchange is relatively inexpensive.
- **Quality.** Peptides ideally should be GMP but high quality non-GMP peptides can be obtained. (GMP is typically quite expensive.)

¹¹²<https://www.thermofisher.com/us/en/home/life-science/protein-biology/peptides-proteins/custom-peptide-synthesis-services/peptide-analyzing-tool.html>

- **Solubility.** Pay close attention to peptide solubility.¹¹³ Water solubility makes vaccine production substantially easier.

In summary, many peptides do not need to be purified, but they should have TFA removed and replaced with acetate, and should be water soluble. Water insoluble peptides require organic solvents, which complicate the overall vaccine formulation. Through experimentation with various peptides, and addition of soluble residues or substitution of hydrophobic residues with soluble ones, we are able to achieve solubility in water while retaining conservative positions of top scoring epitopes represented on these peptides.

Gen 10

RaDVaC generations 1 to 9 vaccine designs consisted largely of improved epitope peptide selections, based on a transition from initial preprint and published computational predictions to a primary focus on empirical data. Gen 10 provides significant improvements over previous generations in essential aspects of vaccine composition, beyond epitope selection. If you don't want to test all aspects of the Gen 10 design, as with all aspects of the RaDVaC project, you are free to experiment with variations on the guidelines provided here. Our network of participating researchers appreciates any information you provide on your experimental variations, and Gen 10 improvements are partly guided by information provided by ongoing experiments and data collection from the RaDVaC network.

The following general improvements have been made to the Gen 10 vaccine design:

- Improved solubility at physiological pH by the use of derivatized chitosan instead of unmodified chitosan.
- Increased T helper activation combined with reduced MHC Class II restriction to more robustly activate cytotoxic T lymphocytes and B-cells for antibody production.
- Surface display of antigens for improved antibody response.
- A smaller set of core peptides (5 peptides) combined with a list of optional peptides, providing greater functionality and improved representation of common MHC Class I alleles.
- An optional epitope sequence that includes an increasingly common variant (N501Y) in the Spike Receptor Binding Motif (RBM). The RaDVaC primary protective strategy remains focused on the more highly conserved epitopes involved in membrane fusion, but we are testing the potential of this epitope sequence to boost the systemic antibody response.
- An optional dendritic cell targeting peptide for delivering T-cell epitopes to dendritic cells, an important cell type in the presentation of T-cell antigens.

¹¹³ <https://pepcalc.com/>

Vaccines prior to Gen 10 were designed simply to allow ease of reproducible production, and conservatively to minimize the likelihood of excess immunostimulation. Given results from hundreds of people who have taken these vaccines, we know they have been extremely safe (no serious adverse events reported), but the intranasally delivered vaccine is likely less effective than possible if antigen presentation had been more effective, and the T helper response had been stronger. We believe this to be the case since many people who have taken the nasal vaccine are testing negative for serum antibodies with commercial and lab ELISA tests, while many who inject the vaccine (subcutaneous or intramuscular) are testing positive (saliva testing appears to be providing evidence of mucosal response among a subset of researchers who have administered the vaccine intranasally). These antibody responses are important indicators that we will work to maximize, and consider together with other evidence of the effectiveness of RaDVaC vaccine designs, including number of people taking vaccine, number of infections, and severity of infections. Encouraged by reports of parenteral response, we remain committed to improving the effectiveness of intranasal delivery, which is safer and will be more widely adopted. Therefore, the Gen 10 design implements multiple changes to address weak systemic antibody responses to nasal administration, but production of the vaccine remains relatively simple—albeit slightly more complex than previously, while maintaining the use of the same minimal set of equipment and materials.

As with Gen 9, the Gen 10 vaccine design consists of a group of core peptides, and a group of optional peptides. Many people desire to make vaccine but have limited resources, so we have made the core peptide group minimal (5 peptides), while retaining essential features and improving others. The optional peptide list provides a range of options for improved geographic and ethnic specificity. Some peptide sequences through Gen 9 are still used in the Gen 10 design, so for those who already have Gen 9 and prior peptides, partial updating can be achieved by addition of peptides new to Gen 10—in particular, the li-key MHC Class II peptides described below. Most pre Gen 10 peptides are usable in Gen 10 as optional or partially redundant additions. The Gen 10 B-cell core epitopes are essentially the same but with additional sequences for improved display on the nanoparticle surface. Much published evidence shows that display of epitopes on the surface of a nanoparticle substantially enhances B-cell activation and antibody production^{114, 115}.

Those familiar with Gen 9 and prior vaccine designs will notice that sequences of B-cell epitopes at approximately Spike 805 and 1145 are shorter in Gen 10. This is the result of a post-Gen 9 publication by Garrett and colleagues providing single amino acid resolution mapping of the

¹¹⁴ <https://www.tandfonline.com/doi/full/10.1080/14760584.2019.1578216?>

¹¹⁵ <https://www.pnas.org/content/pnas/109/4/1080.full.pdf>

boundaries of linear epitope binding¹¹⁶, and another by Ladner and colleagues providing amino acid resolution mapping across coronaviruses to discover the most cross-reactive conserved residues in these SARS-CoV-2 epitopes¹¹⁷. Note that in these situations, flanking sequences might still be important for proper 3D conformation, but these data help define which amino acids are essential for antibody binding, and least likely to mutate away from recognition by neutralizing antibodies. These two B-cell epitopes at approximately Spike 805 and 1145 both bind neutralizing antibodies, and are much more highly conserved than other targets of neutralizing antibodies. Plus, antibodies targeting the HR2 domain of Spike (and the HR2 boundary region at 1145) neutralize a broader range of clinical isolates than antibodies targeting the S1 portion of Spike and the RBD¹¹⁸. These two epitopes are the only two B-cell epitopes in the list of Gen 10 core peptides. Other B-cell epitope peptides are provided in the optional peptides. Two of these are especially notable.

The peptide **Spike N501Y** represents two important developments. First, it represents the first RaDVaC effort to include a mutation from a variant strain of SARS-CoV-2 (the N501Y mutation found in ascendant variants of concern, including B.1.1.7 [UK], 501Y.V2 [South African], and the P.1 variant of the B.1.1.28 lineage [Brazil]). Second, it is a return to the previous RaDVaC focus on the Spike RBD, in particular the ACE2 contact residues of the receptor binding motif (RBM). Gen 8 and prior designs included peptides targeting the RBD but these peptides were long, difficult to synthesize, expensive, and required 3 peptides for achieving a single structural conformation. We were skeptical of the conformational correctness of this complex arrangement; nevertheless, one group has shared positive ELISA results resulting from the intramuscular injection of the Gen 8 vaccine design, which includes these peptides. Their ELISA uses the RBD as a target, and thus it is likely that this RBM construct is eliciting an antibody response. In parallel to their experiments, we have been pursuing a strategy to restore representation of the RBM, in part to address emerging variants. Based on multiple data types, including single amino acid resolution mapping of neutralizing antibody contacts in the RBD/RBM, we have designed a single, relatively easily synthesized peptide that should accurately represent the structure of the portion of the RBM bound by the Regeneron neutralizing antibody REGN10987.

The other B-cell epitope peptide worth noting is **Spike 816-s**. This peptide is designed to present the N terminal end of the fusion peptide that results after TMPRSS2 protease cleavage of the S2' site. This structure mediates the initial stages of contact between the virus and the host cell membrane, leading to fusion of the viral and host cell membranes. The N terminal portion of this sequence is highly conserved among coronaviruses, forming a highly distinctive coil structure

¹¹⁶ <https://www.biorxiv.org/content/10.1101/2020.11.16.385278v1.full>

¹¹⁷ <https://www.sciencedirect.com/science/article/pii/S2666379120302445>

¹¹⁸ <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0050366>

that penetrates the host cell membrane. Single amino acid resolution epitope mapping shows that this region contains key residues in the binding of neutralizing antibodies. Therefore, we hypothesize that the cleaved structure presents a uniquely immunogenic structure that should be recognized by a different set of antibodies than those that recognize the uncleaved structure.

As for T-cell epitopes, three of four standalone MHC Class I/CD8 epitopes are the same as Gen 9 and before (**Orf8 107T**, **Orf1 1636T**, **Orf1 5471T**), while one (**Nuc 359T**) provided redundancy and was removed from the lists of core and optional peptides in Gen 10 (prior generation sequences can be found in previous versions of the white paper). Gen 10 features expanded reporting of the most common MHC Class I alleles in major populations. The list of Gen 10 core peptides provides good HLA/MHC Class I coverage for many geographies and ethnicities, while optional peptides provide substantially expanded coverage. As in previous vaccine generations, the primary selection criterion for each of the peptides in these Gen 10 lists is empirical evidence of function; the MHC classes listed for each peptide represent computationally predicted strong binding alleles for these empirically proven peptide sequences.

To increase T helper activation in Gen 10, the clinically tested li-key technology has been implemented for a few CD4 and combination epitopes¹¹⁹. The li-key portion (the peptide LRMK joined to the N terminus by a 5 to 6 carbon linker) provides enhanced binding to MHC Class II molecules, and thereby increased CD4+ T helper activation¹²⁰. This promotes both CD8/CTL activation and antibody production from B-cell epitopes. Importantly, li-key has been tested safely and successfully in many animal experiments, and in human clinical trials, and a SARS-CoV-2 vaccine design using li-key is currently being tested by a U.S.-China consortium¹²¹. To further enhance antigen presentation and the T helper activation, we also provide the sequence of an optional dendritic cell targeting peptide (DCTP), which has been used in combination with li-key MHC to further boost its immunogenic effects.

The approach we describe for displaying antigens on the nanoparticle surface is an adaptation of two approaches: one that has been used in multiple published studies, and another that has shown promise in a published small-scale pilot study¹²². The former approach involves creating peptides with long poly-anionic termini consisting of a stretch of glutamic or aspartic acid residues, and nanoparticles are formed by mixing the peptides with derivatized cationic chitosan absent another anionic crosslinker¹²³. The second approach, demonstrated in a study by Correia-Pinto et al. was designed to test the hypothesis that antigen display on the surface of

¹¹⁹ <https://www.tandfonline.com/doi/full/10.1517/14712598.6.12.1311>

¹²⁰ <https://pubmed.ncbi.nlm.nih.gov/10781856/>

¹²¹ <https://link.springer.com/article/10.1007/s10549-020-05638-x>

¹²² <https://www.mdpi.com/2076-393X/3/3/730/htm>

¹²³ <https://www.sciencedirect.com/science/article/pii/S174270611830566X>

nanoparticles increases antibody response to B-cell antigens, and their vehicle and antigen of choice of chitosan and synthetic peptides, respectively, provide a good model and protocol for emulsion. The approach is quite simple: synthesizing multiple aspartic acid residues on the N terminus of epitope peptides, and then ionically binding the peptides into the surface of nanoparticles subsequent to particle formation. This publication is especially instructive since their chitosan nanoparticles were produced by sodium triphosphate ionic gelation. However, it is clear that their use of chitosan hydrochloride is not optimal, since chitosan hydrochloride is neutrally charged and water insoluble at physiological pH¹²⁴. Therefore, we have implemented two minor modifications of this approach using a water soluble chitosan derivative, and having the B-cell peptide participate in the ionic gelation formation of the nanoparticle (as described in a later section on vaccine preparation).

Readers of this publication will note that a Pan-DR Epitope peptide (PADRE) was utilized in this antigen surface presentation design. However, we have decided to use the li-key technology instead of the PADRE approach to enhance the T helper response for the following reason. Both PADRE and li-key appear to reduce MHC Class II binding affinity variability due to genetic variation. However, li-key allows the use of pathogen-derived epitopes, and potent MHC Class II epitopes of SARS-CoV-2 have been identified. Including these viral epitopes in a vaccine design instead of a non-viral epitope will provide better T helper response to infection by a closely related virus (e.g. SARS-CoV-2 mutants), since T memory cells will be cross-reactive to the coronavirus epitopes, but not to the PADRE epitope.

To synergize the MHC Class II li-key enhancement with Class I activation, we have designed one core and three optional li-key peptides that provide broad MHC Class I coverage on the same peptide as the li-key and Class II peptide. The li-key combo Class I and II epitope peptide we have selected for the core peptide list is based on the top-scoring immunodominant Class II/CD4 epitope (**Mem 172T-key**) from Adaptive Biotechnologies and from the publication by Peng *et al.* The previous T-cell epitope **Nuc 100-120** has been updated to **Nuc 102T-key**, by slightly shortening it and adding the li-key. This peptide retains broad MHC Class I coverage but might be difficult to synthesize, so it has been moved to the optional peptides list. Similarly, **Nuc 264T-key** is an optional li-key peptide, which is easily synthesized and provides broad MHC Class I coverage not provided by **Mem 172T-key** or **Nuc 102T-key**. We provide another li-key combo MHC Class I and II peptide **Spike 1196-key**, providing very broad coverage of Class I alleles.

To keep the number of core peptides manageable, some of the epitopes with redundant or less common MHC restriction classes have been moved to the list of optional peptides. It is ideal to

¹²⁴ <https://link.springer.com/article/10.1007/s002890050448>

check the expected allele frequencies for MHC Class I for likely class restrictions of your ethnicity/geography, and to include the appropriate epitope peptides (see the Allele Frequency Net Database¹²⁵ and allele frequencies by geographic region¹²⁶). We also include the aforementioned optional dendritic cell targeting peptide (**DCTP**) that has been modified to attach to the surface of the chitosan nanoparticle.¹²⁷

Code key to the lists below:

- **Amino acid sequences and substitutions:** Upper case amino acids letters indicate wild-type sequence. Lower case letters indicate an amino acid substitution.
- **Linkers:** The symbol {-} indicates a 3 to 6 carbon linker, and the symbol {6} indicates a 5 or 6 carbon linker.
- **Surface display:** The letter “s” at the end of a peptide name indicates surface display of an epitope peptide (by extension of the peptide with a poly-aspartic acid sequence). These poly-aspartic acid sequences are OPTIONAL but data suggests that surface display greatly enhances B-cell activation and antibody production.
- **Disulfide circularization:** The letter “c” at the end of a peptide name denotes circularization by disulfide linkage of a pair of cysteines (prior to Gen 10 these were indicated by the suffix “cir” and the # symbol; either order as a disulfide peptide or circularize with activated charcoal). Consistent with the upper/lower case rule on substitutions, Spike 660-cs has two native cysteines for disulfide circularization, which are shown as upper case letter “C”, while Spike 805-cs has two substituted cysteines, which are shown as lower case letter “c”, indicating substitution.
- **Purification:** Some of these peptides will perform substantially better if they are purified; in particular, most B-cell and li-key peptides will perform best if purified. For our most recently synthesized peptides, purification level is indicated. These purification levels are not provided as recommendations, only for data sharing. Note that of the OPTIONAL peptides, our most recent synthesis included a few of these, including the purified B-cell peptides **Spike 816-s** and **Spike N501Y-s**. We have not produced and tested all optional peptides; we only share these designs as prospects for future testing, or as improvements to previous designs.
- **li-key additions:** MHC Class II peptides have four amino acid li-key sequences (LRMK) attached to the N terminus by a 5 to 6 carbon linker, indicated by the {6} symbol.
- **MHC Class I restriction:** The binding preferences for MHC Class I and some MHC Class II peptides are shown for the most common alleles in most populations. Gen 10 Core peptides provide broad but not comprehensive coverage for many geographies and

¹²⁵ <http://www.alleleffrequencies.net/>

¹²⁶ <http://alleleffrequencies.net/top10freqs.asp>

¹²⁷ <https://www.jimmunol.org/content/172/12/7425.long>

ethnicities. The optional MHC Class I peptides provide expanded ethnically specific coverage, while the optional combination MHC Class I/II li-key peptides also provide a wide selection of MHC classes and additional potential T helper cell activation on a single peptide.

Gen 10 Core peptides

1. **Spike 811-s**, dddddd{-}KPSgRSFIEDLLFNKVT-amide (combo B-cell/MHC Class II) (purified ≥85%)
2. **Spike 1147-s**, acetyl-dddddd{-}gSFKEELDKYFKNH-amide (combo B-cell/MHC Class II) (purified ≥90%)
3. **Mem 172T-key**, acetyl-LRMK{6}TSRTLSYYKLGASQRVA-amide (combo Class I/II; A11:01, A33:03, B27:05, B39:01, C06:02, C07:01, C07:02) (purified ≥85%)
4. **Orf8 107T**, DFLEYHDVRVVL (MHC Class I; A02:01, A23:01, A24:02, B39:01, B40:01, A33:03, C04:01, C05:01, C05:09, C06:02, C07:01, C07:02)(Note: this region appears to be deleted in variant B.1.1.7 (UK), and we are researching a possible replacement.)
5. **Orf1 1636T**, HTTDPSFLGRY (MHC Class I; A01:01, A03:02, A11:01, A26:01, A33:03)

Gen 10 Optional peptides

Targeting/effector peptides

1. **DCTP**, acetyl-dddddd{-}rgFYPSYHSTPQRP-amide (dendritic cell targeting peptide) (easy

B-cell peptides

2. **Spike 816-s**, acetyl-SFIEDLLFNKVT{-}dddddddd-amide (B-cell, targeting the N terminal cleaved fusion peptide) possibly difficult synthesis (purified ≥85%)
3. **Spike N501Y-s**, dddddd{-}aQSYGFQPTyGVGYQP{6}NNLDSKVGGNYNYLYa-amide (B-cell, N501Y mutation found in ascendant variants of concern, including B.1.1.7 (UK), 501Y.V2 (South African), and the P.1 variant of the B.1.1.28 lineage (Brazil)); conformational epitope created by linking two discontinuous Spike peptides by the use of a 6 carbon linker. (purified ≥75%) (N501Y mutation creates several MHC Class I sites: A01:01, A11:01, A26:01, A35:01, A35:08, B15:01, B58:01, C07:01, and C07:02)
4. **Spike 21-s**, dddddd{-}RTQLPPAYTNSFTRGVYY
5. **Spike 658-cs**, NSYECDIPIGAGICASYQT{6}QSaIAYT{-}dddddd (combo B-cell/CD4). Linker bypassing Q677 and the S1/S2 cleavage site.

Combination MHC Class I/II li-key peptides

6. **Nuc 264T-key**, acetyl-LRMK{6}ATKAYNVTQAFGRRG-amide (A23:01, A24:02, A33:03, B15:01, B35:01, B35:08, B53:01, C06:02, C07:01, C07:02), easy synthesis
7. **Nuc 102T-key**, acetyl-LRMK{6}KDLSPRWYFYLLGTGPEAGL-amide (A01:01, A23:01, A26:01, B07:02, B08:01, B15, B58:01), potentially difficult synthesis and purification
8. **Nuc 322T-key**, acetyl-LRMK{6}MEVTPSGTWLTYTGAIKLDD-amide (A01:01, A18:01, A26:01, B35:01, B35:08, B40:01, B44:02, B44:03, B53:01), potentially difficult synthesis and purification

9. **Spike 1196-key**, acetyl-LRMK{6}SLIDLQELGKYEYIKWPWYI-amide (A01:01, A03:01, A03:02, A11:01, A23:01, A24:02, B15:01, B18:01, B44:02, B44:03, C06:02, C07:01, C07:02), potentially very difficult synthesis and purification

MHC Class I peptides

10. **Orf1 5470T**, KLSYGIATVR (MHC Class I; A02:01, A03:01, A03:02, A33:03)
11. **Orf1ab 4723T**, RKIFVDGVPFVVSTGYHFRE (MHC Class I; A02:01, A11:01, A23:01, A26:01, A33:03, A35:01, A53:01, B15:01, B35:05, C04:01, C05:01, C05:09)
12. **Orf1ab 5528T** DYGDVVYRGTTTYKL (MHC Class I; A01:01, A03:01, A03:02, A11:01, A23:01, A24:02, A26:01, A33:03, B15:01, B27:05, B35:01, B35:08, B39:01, C6:02, C07:01, C07:02)

The RaDVaC founding group is testing a Gen 10 vaccine including the core peptides plus **DCTP**, and also **Spike 816-s** and **Spike N501Y-s** for expanded antibody and mutational variant potential coverage, and **Nuc 264T-key**, **Nuc 102T-key**, **Orf1ab 4723T**, **Orf1 5471T** and **Orf1ab 5528T** for expanded T-cell potential coverage due to the ethnic diversity of our group.

This section on Gen 9 is provided only as a historical reference and can be skipped

Gen 9

Given the growing number of epitope peptides in recent RaDVaC vaccine designs and questions about a preferred selection of “core” peptides, the Gen 9 design was the first to provide a selective list of core peptides and a separate list of optional peptides. This more compact design reduces the initial cost of peptide synthesis, simplifies production, better balances the epitope ratios (B-cell/CD4/CD8), and increases the relative concentration of each peptide in the vaccine. Several peptides in Gen 8 and previously have been retired for the Gen 9 design (for more detail, see section below entitled SUMMARY OF THE PROTECTIVE STRATEGY OF Gen 7+ VACCINE DESIGNS). Two new additions are broad-coverage HLA class I/CD8 T-cell epitope peptides, which were selected based on data from Adaptive Biotechnologies^{128, 129}. The selection of peptides has changed to emphasize our shift in focus to membrane fusion neutralizing B-cell epitopes, rather than ACE2 interactions with RBD, and to further emphasize the critical importance of MHC class I/CD8 T-cell epitopes. We are using optional peptides in our current vaccine compositions because they likely provide potentially beneficial features, but we think data presently available suggest that the core group will provide most of these in a smaller number of peptides. Listed optional peptides have been used in prior vaccine generations, and might be included to increase certain properties of a vaccine, such as additional CD4

¹²⁸ <https://www.researchsquare.com/article/rs-51964/v1>

¹²⁹ <https://adaptivepublic.blob.core.windows.net/publishedproject-supplements/covid-2020/ImmuneCODE-MIRA-Release002.1.zip>

immunogenicity or HLA allelic coverage. Retired peptides might similarly be considered for inclusion in a vaccine. Previously tested and retired (due to low solubility) or conditional or prospective peptides are listed in Supplemental Materials.

Amino acid substitutions are shown as lower-case letters. Peptide synthesis is most efficient if certain guidelines are followed (see for example¹³⁰). 3D structural models of a few of the B-cell, protein surface-exposed peptides are shown in figures in the relevant sections below. For these we used the imaging website iCn3d¹³¹. Disulfide circularized peptides promote stability and approximate natural conformation. These peptide names end in cir, e.g. “Spike 804-820cir” or c, e.g. “Spike 805-c”. Most of these peptides were circularized by disulfide formation, with exceptions noted below. We did not order these synthesized with disulfide bridge, but performed disulfide formation ourselves by a simple incubation with charcoal as previously described¹³². Certain peptides have substitutions of cysteine for other amino acids, to enable disulfide formation and approximation of conformational correctness. These amino acid substitutions are indicated by the lower case “c”. Certain portions of viral sequence might be represented in more than one peptide. Spike 804-820cir was synthesized first and used in early generations of vaccine, and then Spike 802-823cir was synthesized to add potentially important flanking sequences to this protease recognition loop sequence on the surface of the Spike protein.

In the lists below, primary functions of each peptide are listed next to each peptide name. Most peptides are combination HLA class II (CD4) and B-cell epitopes, or HLA class II (CD4) and HLA class I (CD8) epitopes. HLA allele restrictions are shown for each HLA class I (CD8) epitope. Two of the optional HLA class I (CD8) peptides provide the additional HLA alleles B15 and B58. Bold and underlined portions show CD8 epitopes. Number signs (#) indicate peptides for disulfide formation. A letter T at the end of the sequence name indicates a T-cell epitope.

1. **Spike 802-823cir**, FSQC LPDP SKPSKRSFCEDLLF# (combo CD4/B-cell)
2. **Spike 660**, YECDIPIGAGICASYQTQTN SPRRA# (combo CD4/B-cell)
3. **Spike 1145**, LDSFKEELDKYFKNHTSP (combo CD4/B-cell)
4. **Nuc 100-120**, KMKDL**SPRWYFY**LG TGPEAG (combo CD4/**CD8**; A01, A26, B07, B08, B15, B58)
5. **Nuc 321-345**, G**MEVTPSGTW**LTYTGAIKLDDKDPN (combo CD4/**CD8**; **B40, B44**, C15, C17)
6. **Nuc 359T**, AY**KTFPPTEPK** (combo CD4/**CD8**; A03, A11)
7. **Orf1 1636T**, H**TTDPSFLGRY** (CD8; **A01, A11, A26, A33**)
8. **Orf8 117T**, D**FLEYHDVRV**VL (CD8; **A02, A24, B39, B40**, C05) (Note: should be Orf8 107T)
9. **Orf1 5471T**, KLSYGIATVR (CD8; **A02, A03, A33**)

¹³⁰ [https://www.thermofisher.com/us/en/home/life-science\[...\]/pierce-protein-methods/peptide-design.html](https://www.thermofisher.com/us/en/home/life-science[...]/pierce-protein-methods/peptide-design.html)

¹³¹ <https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html>

¹³² <https://pubmed.ncbi.nlm.nih.gov/9606016/>

Gen 9.1 Optional peptides

1. **Nuc 263T**, ATKAYNVTQAFGRRG (combo CD4/CD8; **B15, B58, C02, C03, C12, C14, C16, C17**)
2. **Spike 21**, RTQLPPAYTNNSFTRGVVYPDK (combo B-cell/CD4/**CD8; C02, C03, C12, C16**)
3. **Nsp7 21T**, RVESSSKLWAQCVQLH (combo CD4/**CD8; A02, B58**)
4. **Spike 442-460**, DSKVGGNYNYLYRLFRKSN (combo B-cell/CD4/**CD8; A24**)

Spike or S protein amino acid sequences represented by epitope peptides

Numbered amino acid sequence for the S protein is shown below. (Sequences without breaks or spaces can be found in the Supplemental Materials section.)

Blue highlight shows the receptor binding domain (RBD), amino acids 319 to 529.¹³³ Red highlight shows amino acids that contact ACE2, the receptor for the virus. Light yellow highlight shows the Heptad Repeat 2 (HR2) domain from aa 1163 to 1213. The S1/S2 protease cleavage site (rrar/sva) is shown in green highlight, with cleavage between Arg 685 and Ser 686. The S2' protease cleavage site (skr/sfi) is shown in magenta highlight, with cleavage between Arg 815 and Ser 816. The transmembrane domain starts at residue 1209, near the C terminus of the protein sequence. Underlined portions of sequence show coverage by epitope peptides.

```

1   mfvflvllpl vssqcvnl tt rtqlppaytn sftrgvvypd kvfrssvlhs
51  tqdlflpffs nvtwfhaihv sgtngtkrfd npvlpfndgv yfasteksni
101 irgwifgttl dsktqsl liv nnatnvvikv cefqfcndpf lgvyvhknnk
151 swmesefrvy ssannctfey vsqpflmdle gkqgnfknlr efvfknidgy
201 fkiyskhtpi nlvrldpqqf saleplvdlp iginitrftt llalhrsylt
251 pgdsssgwta gaaayyvgyl qprtflkyn engtitdavid caldplsetk
301 ctklsftvek giyqtsnfrv qptesivrfp nitnlcpfge vfnatrfasv
351 yawnrkrish cvadysvlyn sasfstfkcy gvsptklndl cftnvyadsf
401 virgdevrqi apgqgkiad ynyklpddft gcviawnsn ldskvggny
451 ylyrlfrksn lkpferdist ei yqagstpc ngvegfn cyf pldsygfapt
501 ngvgvqpyrv vvl sfellha patvcgpkks tnlvknkcvn fnfngltgtg
551 vltesnkkfl pfqqfgrdia dttdavrdpq tleilditpc sfggvsvitp
601 gtntsnqvav lyqdvntev pvaihadqlt ptwrvystgs nvfqtragcl
651 igaehvnnsy ecdipigagi casyqtqtns prarsvasq siiaytmslg
701 aensvaysnn siaiptnfti svtteilpvs mtktsvdctm yicgdstecs
751 nlllqygsfc tqlnraltgi aveqdkntqe vfaqvkqiyk tppikdfggf
801 nfsqilpdps kpskr sfi ed llfnkvltad agfikqygdc lgdiaardli
851 caqkfngltv lppllt demi aqytsallag titsgwtfga gaalqipfam

```

¹³³ <https://jvi.asm.org/content/jvi/early/2020/01/23/JVI.00127-20.full.pdf>

```

901 qmayrfngig vtqnvlyenq klianqfnsa igkiqdslls tasalgklqd
951 vvnqnaqaln tlvkqlssnf gaissvln di lsrl dkveae vqidrlitgr
1001 lqslqtyvtq qliraaeira sanlaatkms ecvlgqskrv dfcggkgyhlm
1051 sfpqsaphgv vflhvtvpa qeknfttapa ichdgkahfp regvfvsngt
1101 hwfvtqrnfy epqiittndt fvsgncdvvi givnntvydp lqpeldsfke
1151 eldkyfknht spdvdlgdis ginasvvniq keidrlneva knlneslidl
1201 qelgkyeqyi kwpwyiwlgf iagliaivmv timlccmtsc csc lkgccsc
1251 gscckfdedd sepv lkgvkl hyt //

```

Nucleocapsid or N protein

Only T-cell epitope peptides were selected from the sequence of the Nucleocapsid protein. Underlined portions of sequence show coverage by epitope peptides.

```

1  msdngpqnr napritfggp sdstgsnqng ersgarskqr rpqglpnnta swftaltqhg
61  kedlkfprgq gvpintnssp ddqigyyrra trrirggdgk mkdlsprwyf vylgtgpeag
121 lpygankdgi iwvategaln tpkdhigtrn pannaaivlq lpqgttlpkg fyaegsrggs
181 qassrssrs rnssrnstpg ssrgtsparm agnggdaala lllldrlnql eskmsgkgqg
241 qggqvtvkks aaeaskkprq krtatkaynv tqafgrrgpe qtqgnfgdqe lirqgtdykh
301 wpqiaqfaps asaffgmsri gmevtpsgtw ltytgaikld dkdpnfkdqv illnkhiday
361 ktfpptepk dkkkkadetq alpqrqkkq tvtl lpaadl ddfskqlqqs mssadstqa
//

```

B-cell epitope peptides

Spike region 439-507

NNLDSKVGGN~~Y~~NYLYRLFRKSNLKPFFERDISTEIIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQP

This region of the Spike protein comprises the majority of the Receptor Binding Motif, and contains many of the ACE2 receptor contact residues.

Spike N501Y-s

Positions 493 to 507 and 439 to 453

dddddd{-}aQSYGFQPTyGVGYQP{6}NNLDSKVGGN~~Y~~NYLYa-amide

IN TESTING, OPTIONAL, Gen 10

MHC Class I: A01:01, A11:01, A26:01, A35:01, A35:08, B15:01, B58:01, C07:01, and C07:02

This is a conformational epitope created by linking two discontinuous Spike peptides by the use of a 6 carbon linker. The epitopes are in reverse order from their linear order in the Spike protein, with the N-terminal peptide start at 493, and the C-terminal peptide start at 439. This epitope contains the N501Y mutation found in ascendant variants, including B.1.1.7 (UK), 501Y.V2 (South African), and the P.1 variant of the B.1.1.28 lineage (Brazil)); L441A substitution to allow easier

synthesis. It includes all but one (R346) of the contact residues of the Regeneron monoclonal antibody REGN10987. Residues bound by REGN10987 (6XDG): R346, N440, L441, K444, V445, G446, N448, Y449, Q498¹³⁴. The Regeneron antibody therapy currently in use to treat COVID-19 is a cocktail of two monoclonal antibodies: REGN10987 and REGN10933, each of which binds a non-overlapping set of epitopes on the Spike protein, at or near the RBD. Mutants at K417N, Y453F (in mink cluster 5) and E484K reduce binding of REGN10933 but do not affect binding of REGN10987¹³⁵. REGN10987 is one of the most resistant antibodies to impaired neutralization¹³⁶. This is very likely at least partly due to its target amino acids being largely limited to the Spike Receptor Binding Motif (RBM), the residues that contact the ACE2 receptor. Therefore, we believe this minimal RBM epitope has the potential to induce potent neutralizing antibodies such as REGN 10987, and that the N501Y substitution will either not affect binding, allowing neutralization of virus with essentially any amino acid at position 501, or if Tyr 501 participates in antibody binding, it will enhance neutralization against the many emerging variants that carry this substitution. The N501 substitution creates the peptide sequence QSYGFQPTY that is predicted to bind strongly to multiple allelic variants of MHC Class I receptors including A01:01, A11:01, A26:01, A35:01, A35:08, B15:01, B58:01, C07:01, and C07:02.

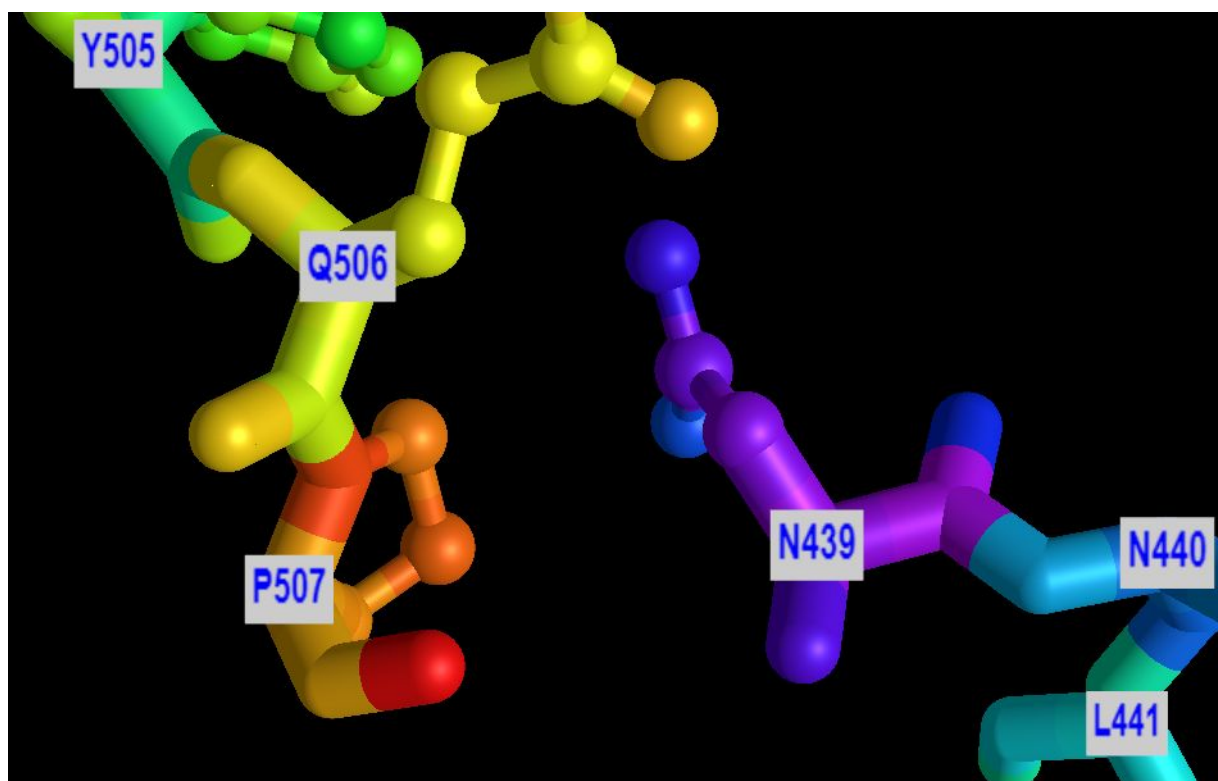


Figure: Linker portion of the peptide Spike N501Y-s, shown in the context of the wild-type Spike protein (PDB ID 6XM4). In the peptide Spike N501Y-s, the C-terminus of Pro507 (shown here in

¹³⁴ <https://science.sciencemag.org/content/369/6506/1010.abstract>

¹³⁵ <https://www.biorxiv.org/content/10.1101/2021.02.18.431897v1>

¹³⁶ <https://www.biorxiv.org/content/10.1101/2021.01.25.428137v3>

red, pointing right) is covalently joined by a flexible carbon linker to the N terminus of Asn439 (shown in dark purple, pointing down).

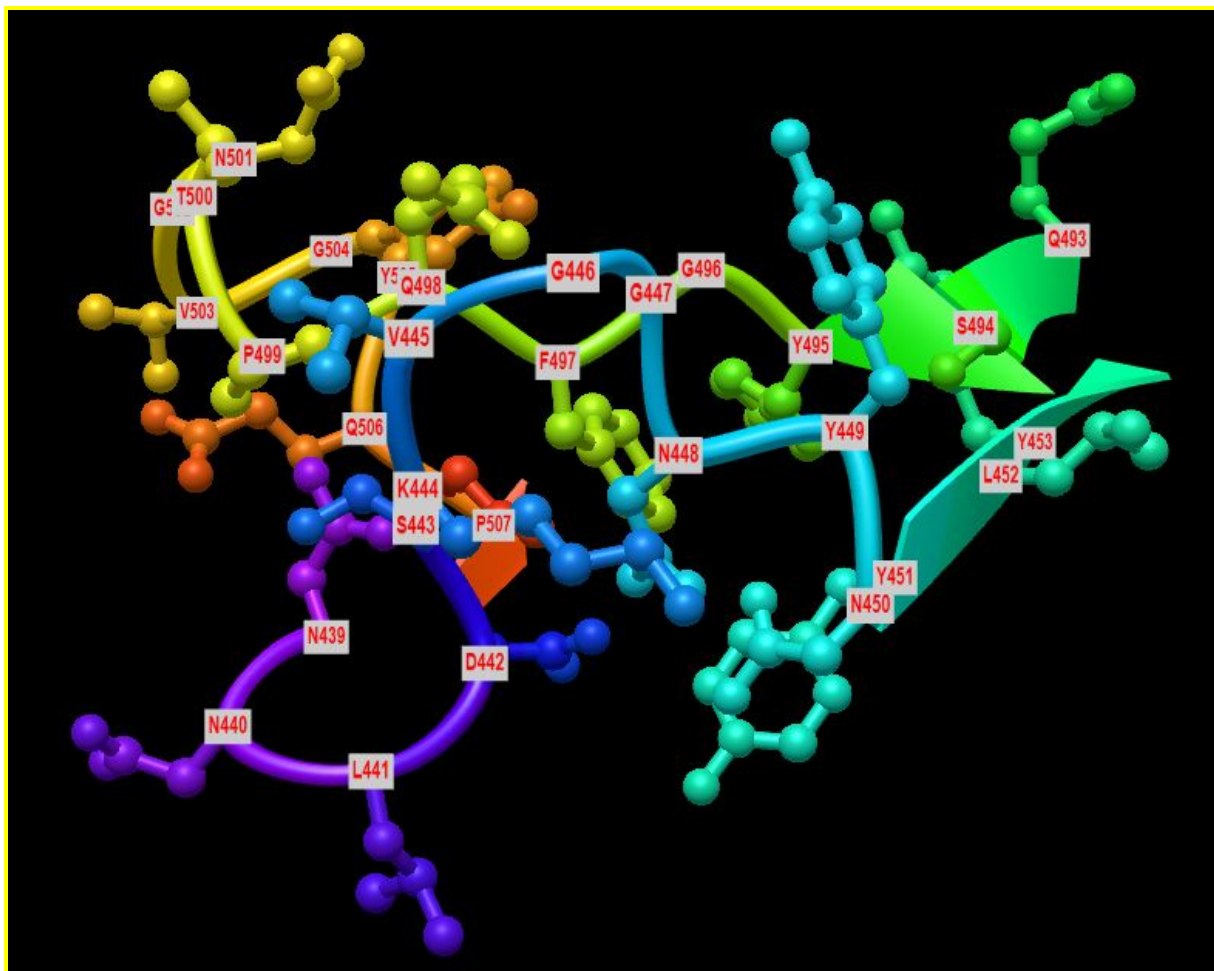


Figure: Portion of the wild-type Spike Receptor Binding Motif (RBM) covered by REGN10987 (from PDB ID 6XM4). REGN10987 contact residues are N440, L441, K444, V445, G446, N448, Y449, Q498. Amino acid N501 is shown at upper left. The **Spike N501Y-s** peptide consists of the amino acids shown and should ideally assume this conformation, except for whatever alteration is due to the N501Y substitution; and, except for features not shown: 6 carbon linker covalently connecting P507 and N439; N501Y; poly-aspartic acid N-terminal segment; and C-terminal amidation. As can be seen, P507 and N439 are on the opposite side of the portion of the RBM from the amino acids contacted by REGN10987; thus, the linker should not perturb or disrupt immune recognition of this structure.

Spike 437-460, a.k.a. Spike1

437 NSNNLDSKVGGNYYLYRLFRKSN

IN TESTING, vaccines Generations 2, 3, 4, 5, 6, 7, 8; OPTIONAL, Gen 9

[Back to top](#)

Spike 442-460442 DSKVGGN^YN^YLYRLFRKSN

Within RBD, seven ACE2 contacting residues. Portion of Spike protein bound by antibodies in SARS-CoV and MERS-CoV.¹³⁷ Li and colleagues demonstrated that a largely overlapping sequence is a linear epitope for a neutralizing antibody¹³⁸. These data suggest the shortened sequence DSKVGGN^YN^YLYRLFRKSN is sufficiently long to be an effective B-cell epitope (see Extended Data Figure 12 of Li et al). In addition to the data of Zhang¹³⁹, this peptide spans two overlapping dominant T-cell epitopes predicted by Grifoni et al¹⁴⁰ and Braun et al¹⁴¹. This epitope also contains the experimentally verified homologous immunodominant T-cell epitope in SARS-CoV-1, Spike 436-443, YNYKYRYL (SARS-CoV-2 Spike 449-456)¹⁴². In this publication by Zhi and colleagues, peptides containing this epitope stimulated the strongest IFN-γ response in the ELISpot assay. Wang et al reported IgG antibodies to a stretch of this peptide. The sequence NYNYLYRLF is a top-scoring cytotoxic T lymphocyte (CTL) epitope predicted by the CTLpred server based on work by Basin and Raghava¹⁴³.

- Predicted T-cell epitope by Grifoni et al.: Class II, NLDSKVGGN^YN^YLYRLFR, 440-457
- Predicted T-cell epitope by Braun et al.: Class II, NLDSKVGGN^YN^YLYRLFR, 440-457

Spike region 658-695

658 NSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYT

This epitope spans the S1/S2 site that is cleaved by the host furin protease, which is required for subsequent cleavage at the S2' site by the host TMPRSS2 protease (Spike region 805-830 epitope). It also includes frequently observed variable amino acid: Q677.

Spike 658-cs

NSYECDIPIGAGICASYQT{-}QSaIAYT{-}dddd

OPTIONAL vaccine Generation 10

We have not yet synthesized or tested this peptide; we share this design as an update to the previous peptide **Spike 660**. **Spike 660** contains a common mutation at Q677 which is addressed in **Spike 658-cs** by the use of a linker that bypasses Q677 and the S1/S2 cleavage site, joining flanking sequences.

¹³⁷ <https://www.biorxiv.org/content/10.1101/2020.02.22.951178v1.full.pdf>¹³⁸ <https://www.biorxiv.org/content/10.1101/2020.09.11.292631v1>¹³⁹ <https://www.biorxiv.org/content/10.1101/2020.04.23.056853v2.article-info>¹⁴⁰ <https://pubmed.ncbi.nlm.nih.gov/32183941>¹⁴¹ <https://www.medrxiv.org/content/10.1101/2020.04.17.20061440v1>¹⁴² <https://www.sciencedirect.com/science/article/pii/S0042682205000620>¹⁴³ <http://crdd.osdd.net/cgibin/ctlpred>

Spike 660

660 YECDIPGAGICASYQTQTNPPRA
 IN TESTING, vaccine Generation 7, 8, 9.

This epitope lies immediately adjacent to the S1/S2 cleavage site (rrar/sva). Immunodominant convalescent antibody-binding epitope described in linear epitope mapping studies by Wang et al, Farrera et al, and Li et al. Neutralizing activity of antibodies that bind the S1/S2 site is hypothesized to be a result of reducing efficient cleavage of this site, preventing virus fusion with the host cell membrane¹⁴⁴.

Spike region 802-830

802 FSQILPDPSKPSKRSFIEDLLFNKVTLADA

Contains the site of the S2' cleavage site of the viral fusion peptide by the host TMPRSS2 protease, which is essential for efficient membrane fusion and host cell infection. The S2' site is cleaved between Arg 815 and Ser 816 (SKR/SFI). An immunodominant epitope bound by convalescent antibodies reported separately by Farrera et al, Poh et al, and Wang et al. Poh and colleagues demonstrated that this epitope is bound by neutralizing antibodies¹⁴⁵. Predicted B-cell epitope by Grifoni et al. Includes a top B-cell epitope prediction from Lon¹⁴⁶ and from Khan¹⁴⁷. Peng and colleagues provide empirical data showing the sequence NFSQILPDPSKPSKR is a CD4 epitope.

Spike 811-s

dddddd{-}KPSgRSFIEDLLFNKT-amide
 IN TESTING, vaccine Generation 10

Minimal epitope peptide containing the site of the S2' cleavage site of the viral fusion peptide, starting at Lys811. For optional enhanced display on the nanoparticle surface, this peptide also features an N-terminal poly-aspartic acid segment attached by a 3 to 6 carbon linker. Single amino acid resolution mapping data show that much of the loop between 805 and 817 is not necessary for typical antibody binding, and that R815 is the most N-terminal amino acid directly involved in antibody binding. Nevertheless, 3D models show that certain amino acids participate in stabilizing interactions with the antibody binding portion of the peptide. In particular, Lys811 forms a stabilizing salt bridge with Asp820. Lys 814 plays no role in the stabilization of the antibody binding region, and to prevent it from binding to the poly-aspartic N terminal portion of the peptide, it was replaced by glycine (K814G).

¹⁴⁴ <https://www.sciencedirect.com/science/article/pii/S1097276520302641>

¹⁴⁵ <https://www.nature.com/articles/s41467-020-16638-2>

¹⁴⁶ <https://www.biorxiv.org/content/10.1101/2020.04.03.022723v1>

¹⁴⁷ <https://www.biorxiv.org/content/10.1101/2020.05.03.074930v1>

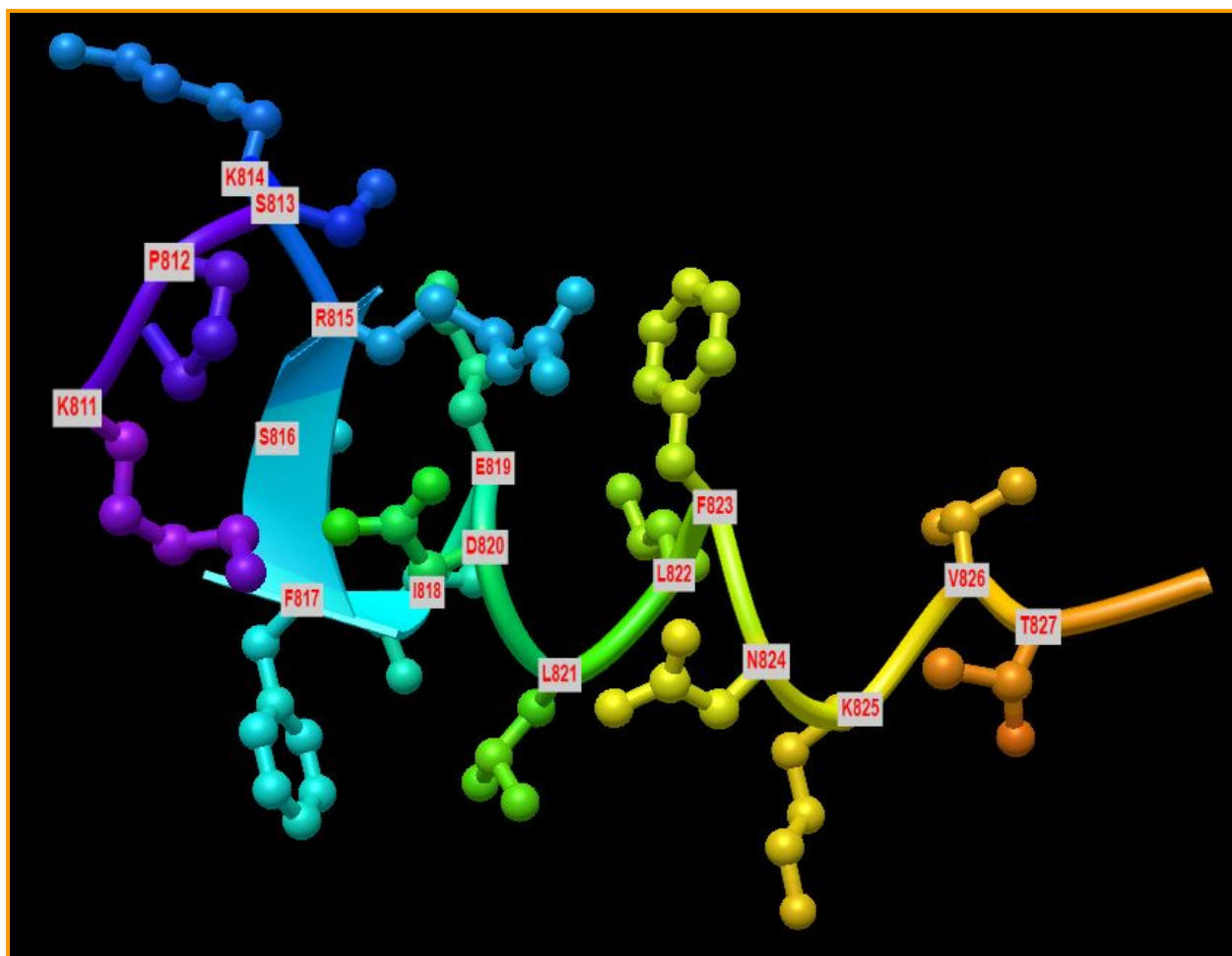


Figure: Structure of Spike 811-827 (from PDB ID 6XM4). This structure shows the uncleaved S2' site and fusion peptide. Note the positions of K811 and R815, and their interactions with E819, D820, and F823. Cleavage of the S2' site separates this structure between R815 and S816, exposing the fusion peptide.

Spike 816-s

acetyl-SFIEDLLFNKVT{-}dddddddd-amide

IN TESTING, vaccine Generation 10

This peptide is designed to present the N terminal end of the fusion peptide that results after TMPRSS2 protease cleavage of the S2' site. This structure mediates the initial stages of contact between the virus and the host cell membrane, leading to fusion of the viral and host cell membranes. The N terminal portion of this sequence is highly conserved among coronaviruses, forming a highly distinctive coil structure that penetrates the host cell membrane. Single amino acid resolution epitope mapping shows that this region contains key residues in the binding of neutralizing antibodies. Therefore, we hypothesize that the cleaved structure presents a uniquely immunogenic structure that should be bound by antibodies with different affinities than those that recognize the uncleaved structure.

[Back to top](#)

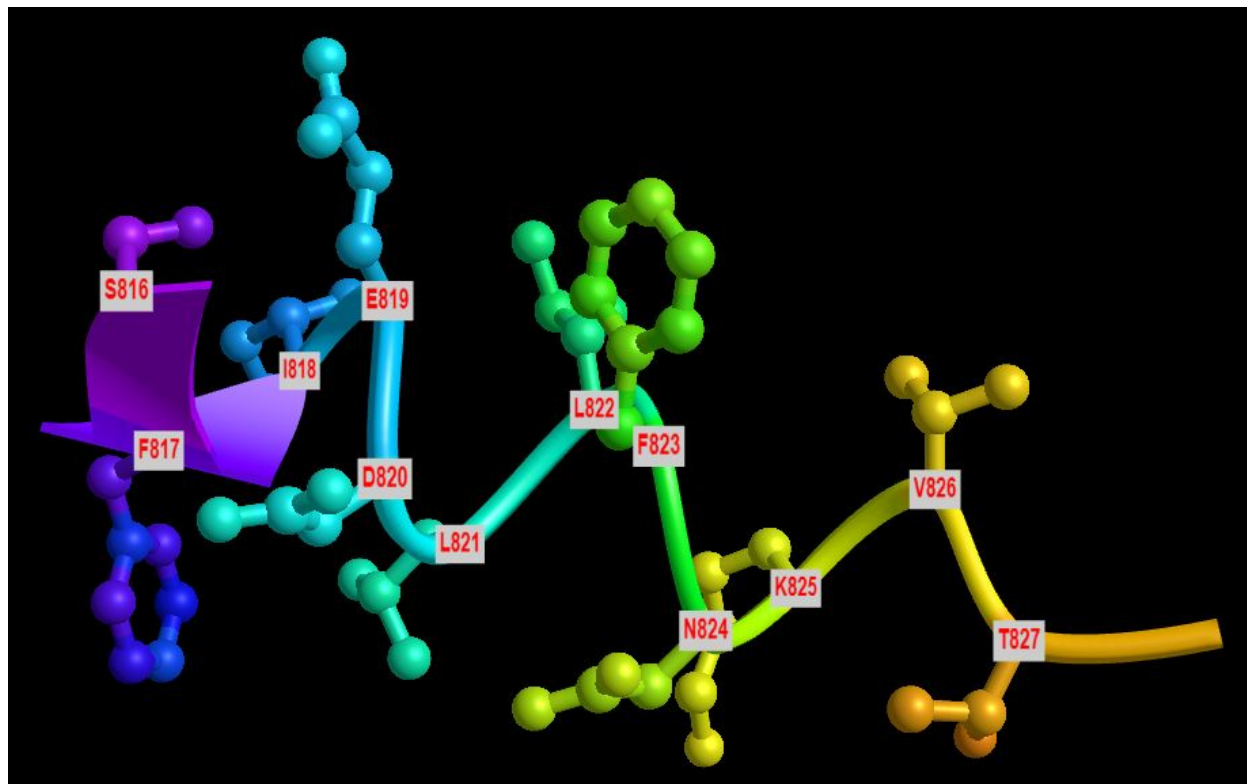


Figure: Structure of Spike 816-827 showing the N terminus of the S2' cleavage site and the fusion peptide (from PDB ID 6XM4). This represents the amino acids of the fusion peptide sequence in native conformation. Following cleavage the structure might undergo a substantial rearrangement upon separation from the rest of the Spike protein, resulting in a different structure from the one depicted here.

Spike 802-823cir

FSQCLPDPSKPSKRSEEDLLF (Cys4, Cys17 disulfide)

IN TESTING, vaccine Generations 5, 6, 7, 8, 9

To preserve the loop structure present in the native conformation, we substituted cysteines for amino acids 4 (Ile>Cys) and 17 (Ile>Cys).

Spike 1147-s

dddddd{-}gSFKEELDKYFKNH-amide

IN TESTING, vaccine Generation 10

This peptide provides a shortened version of the previous Spike 1145, and also includes a polyaspartic N terminus for enhanced nanoparticle surface presentation. Most 3D structural models do not include this portion of the Spike protein. However, PDB ID 6XR8 presents a clear model of this structure. This peptide forms an alpha helix, and the three alpha helices of the

[Back to top](#)

Spike trimer form a triple helix bundle. Note that each of the three helices is formed by an identical peptide sequence. The formation of this structure is promoted by interactions of multiple aromatic and hydrophobic amino acids on the interior, surrounded by aromatic amino acids and interhelix salt bridges on the outside of the bundle core. Single amino acid resolution mapping data show the boundary amino acids for antibody binding. Only the most highly conserved, transmembrane proximal two-thirds of the helix participates in binding. Therefore, we have omitted non-essential residues in this epitope peptide.

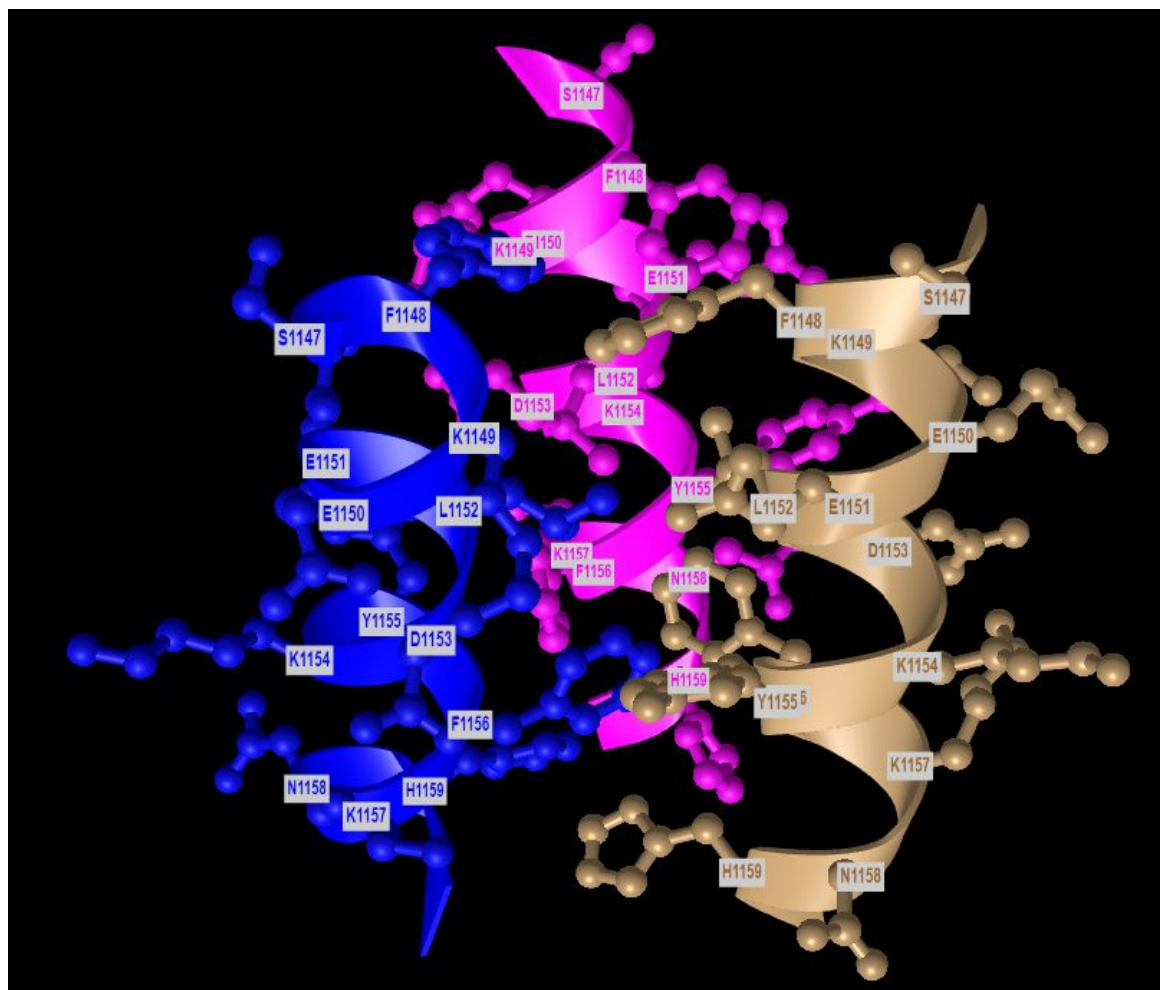


Figure. Spike 1147-1159 from PDB ID 6XR8. This figure shows the triple helix bundle formed by this highly conserved portion of the Spike protein.

Spike 1145

1145 LDSFKEELDKYFKNHTSP

IN TESTING, vaccine Generation 7, 8, 9

Immunodominant convalescent antibody-binding epitope described in linear epitope mapping studies by Farrera et al and Li et al. This region was not covered in the studies by Zhang et al.

[Back to top](#)

This epitope is immediately adjacent to 50 aa heptad repeat 2 (HR2) fusogenic domain, which is essential for fusion of coronavirus and host cell membranes, and is 100% conserved between SARS-CoV-1 and CoV-2 viruses¹⁴⁸. Data by Yi et al suggests that this epitope competes with neutralizing serum from convalescents, indicating that neutralizing antibody binds this sequence

¹⁴⁹.

Spike 21

21 RTQLPPAYTNSFTRGVYYPDK

IN TESTING, vaccine Generation 7, 8; OPTIONAL, Gen 9, 10

Immunodominant convalescent antibody-binding epitope described in linear epitope mapping studies by Wang et al, Zhang et al, and Li et al. An immunodominant neutralizing antibody maps to this region in SARS-CoV-1, and a neutralizing antibody binds this region in SARS-CoV-2¹⁵⁰. A mutation at the neighboring amino acid position 18 (L18F) abrogates binding of a neutralizing antibody¹⁵¹. This portion of the Spike protein is near the N terminus and is sufficiently unstructured as to not appear in currently available 3D structural models of the protein. This fact and the very strong signal in the linear epitope mapping study of Li et al suggest that this epitope peptide might reasonably approximate its native structure in the context of the full-length Spike, and thus is a good candidate for a peptide vaccine.

Nucleocapsid protein epitopes

Nucleocapsid is highly antigenic and produces large signals in ELISA of convalescent sera. Because of concerns about ADE, we are not using Nucleocapsid B-cell epitope peptides. If concerns about ADE risk are resolved, then these selections might be considered for use in a vaccine. However, as described below, certain portions of Nucleocapsid have been shown to be the longest-lasting T-cell epitopes, and we have selected these for eliciting durable T-cell responses.

Nucleocapsid proteins of SARS-CoV-1 and CoV-2 display antibody binding immunodominance, and the portions bound are highly similar, which isn't surprising given the high degree of conservation between these proteins^{152, 153}. The portions of Nucleocapsid that produce high signal in epitope mapping studies (for CoV-1) are amino acids 147-170 and 362-412. CoV-2 mapping by Wang et al produced very similar high binding signals for the C-terminal portion¹⁵⁴.

¹⁴⁸ <https://www.thelancet.com/journals/lancet/article/PIIS0140673604157887/fulltext>

¹⁴⁹ <https://www.tandfonline.com/doi/full/10.1080/22221751.2020.1815591>

¹⁵⁰ <https://science.sciencemag.org/content/369/6504/650>

¹⁵¹ <https://www.biorxiv.org/content/10.1101/2021.01.14.426475v1>

¹⁵² <https://jcm.asm.org/content/42/11/5309.full>

¹⁵³ <https://www.biorxiv.org/content/10.1101/2020.03.26.994756v1>

¹⁵⁴ <https://www.biorxiv.org/content/10.1101/2020.03.26.994756v1>

Mapping of CoV-1 Nucleocapsid protein antibody binding epitopes shows no binding in the region 100-122, and low binding in 321-345, which represent the longer T-cell epitopes included in Gen 7+ vaccine. A CD8 T-cell epitope (at CoV-2 N 359-369) at the border of the C-terminal antibody binding region was selected for inclusion in Gen 8+ vaccine designs.

T-cell epitope peptides

Mem 172T

172 TSRTLSTYYKLGASQRVA

OPTIONAL Generation 9

Mem 172T-key

acetyl-LRMK{6}TSRTLSTYYKLGASQRVA-amide

IN TESTING, vaccine Generation 10

Combination MHC ClassII/CD4 and MHC ClassI/CD8 (A11:01, A33:03, B27:05, B39:01, C06:02, C07:01, C07:02).

The top ranking MHC Class II/CD4 epitope according to data published by Peng and colleagues, as well as the top-ranking MHC Class II/CD4 epitope among Adaptive Biotechnologies' data as of Sept 28, 2020. The peptide KEITVATSRSTYYK and LSYYKLGASQRVAGD both are separately in the top 10 MHC Class II/CD4 epitopes identified by Tarke and colleagues in a whole proteome screen¹⁵⁵. These sequences clearly are among the most potent MHC Class II epitopes in the SARS-CoV-2 proteome. Inclusion of the li-key peptide (LRMK, attached to the N terminus by a 5 to 6 carbon linker) in a vaccine is predicted to substantially increase the MHC Class II/CD4 response.

Nuc 100-120

100 KMKDLSRWYFYLLGTGPEAG

IN TESTING, vaccine Generation 7, 8, 9

Nuc 102T-key

LRMK{6}KDLSRWYFYLLGTGPEAGL

OPTIONAL, vaccine Generation 10

Combination MHC ClassII/CD4 and MHC ClassI/CD8 (A01:01, A23:01, A26:01, B07:02, B08:01, B15, B58:01)

One of the cross-reactive SARS-CoV-2 epitopes to T-cells from SARS-CoV-1 convalescent patients tested by Le Bert and colleagues¹⁵⁶. T-cells from these patients are cross-reactive to homologous epitopes from SARS-CoV-2, which are highly conserved between the two viruses.

¹⁵⁵ <https://www.sciencedirect.com/science/article/pii/S266637912100015X>

¹⁵⁶ <https://pubmed.ncbi.nlm.nih.gov/32668444/>

Also identified by Peng and colleagues as an immunodominant CD4/MHC class II epitope and CD8/MHC class I epitopes. The sequence SPRWYFYLL was also identified experimentally by Ferretti et al as a CD8 epitope with predicted HLA restriction of B07. The second highest ranked CD8 epitope identified by Snyder et al. in their proteome-wide screen, although in subsequent larger public data releases from Adaptive Biotechnologies, its rank dropped to ninth¹⁵⁷.

Nuc 263T

263 ATKAYNVTQAFGRRG

IN TESTING, vaccine Generation 8; OPTIONAL, Gen 9

Nuc 264T-key

acetyl-LRMK{6}ATKAYNVTQAFGRRG

OPTIONAL, vaccine Generation 10

Combination MHC ClassII/CD4 and MHC ClassI/CD8 (A23:01, A24:02, A33:03, B15:01, B35:01, B35:08, B53:01, C06:02, C07:01, C07:02)

The full length epitope ATKAYNVTQAFGRRG was predicted by Fast and Chen as a broad-coverage epitope for MHC-I (74%) and for MHC-II (100%). This sequence largely overlaps a top MHC Class II/CD4 epitope in the public dataset of Adaptive Biotechnologies (PRQKRTATKAYNVTQAFGR), and of Tarke et al. (KRTATKAYNVTQAFG), and one of the cross-reactive SARS-CoV-2 epitopes to T-cells from SARS convalescent patients tested by Le Bert et al (KAYNVTQAFGRRGPE). The sequence KAYNVTQAF is the third highest ranked CD8 epitope identified by Snyder et al, although in subsequent larger public data releases from Adaptive Biotechnologies, its rank dropped to nineteenth. **Nuc 264T-key** provides the li-key version of the same peptide sequence as **Nuc 263T** but fixes a previously incorrect starting position number. Inclusion of the li-key peptide (LRMK, attached to the N terminus by a 5 to 6 carbon linker) in a vaccine is predicted to substantially increase the MHC Class II/CD4 response.

Nuc 321-345

321 GMEVTPSGTWLTYTGAIKLDDKDPN

IN TESTING, vaccine Generation 7, 8, 9

Nuc 322T-key

acetyl-LRMK{6}MEVTPSGTWLTYTGAIKLDD-amide

OPTIONAL, vaccine Generation 10

Combination MHC ClassII/CD4 and MHC ClassI/CD8 (A01:01, A18:01, A26:01, B35:01, B35:08, B40:01, B44:02, B44:03, B53:01)

¹⁵⁷ <https://www.medrxiv.org/content/10.1101/2020.07.31.20165647v2>

One of the cross-reactive SARS-CoV-2 epitopes to T-cells from SARS-CoV-1 convalescent patients tested by Le Bert and colleagues. T-cells from these patients are cross-reactive to homologous epitopes from SARS-CoV-2, which are highly conserved between the two viruses. This sequence contains a predicted dominant T-cell epitope of Grifoni et al (MEVTPSGTWL, HLA restriction B*40:01)¹⁵⁸. MEVTPSGTWL was also verified experimentally by Peng and colleagues as a CD4 epitope and CD8 epitope likely covering HLA type B*40:01. MEVTPSGTWL is also among the top 20 ranked CD8 epitopes identified by Snyder et al. Mateus et al identified the peptide PSGTWLTYTGAIKLD as a CD4 epitope, with common coronavirus cross-reactivity¹⁵⁹. Inclusion of the li-key peptide (LRMK, attached to the N terminus by a 5 to 6 carbon linker) in a vaccine is predicted to substantially increase the MHC Class II/CD4 response.

Spike 1196-key

1196 acetyl-LRMK{6}SLIDLQELGKYEQYIKWPWYI-amide

OPTIONAL Gen 10

Combination MHC ClassII/CD4 and MHC ClassI/CD8 (A01:01, A03:01, A11:01, A23:01, A24:02, B15:01, B44:02, B44:03, C06:02, C07:01, C07:02)

The sequence of this portion of Spike protein lies at the boundary of the viral exterior and includes a few amino acids of the highly conserved transmembrane domain of the protein. It is bound by antibodies in serum from a small fraction of convalescent patients. However, Adaptive Biotechnologies data indicate that it is the most important MHC Class I epitope on the Spike protein, strongly binding a large number of common allelotypes. Given the broad HLA/MHC Class I coverage provided by this peptide it will be an attractive addition to vaccine designs; however, it is predicted to be very difficult to synthesize. According to Class II data published by Adaptive and by Tarke et al, this sequence is only a minor contributor to the MHC Class II response; nevertheless, we have designed it as an optional li-key peptide because of its broad MHC Class I range.

Nuc 359T

359 AYKTFPPTEPK

IN TESTING, vaccine Generation 8, 9 OPTIONAL Gen 9

Class I MHC/HLA restriction: A03:01, A03:02, and A11:01

The sequences KTFPPTEPK / AYKTFPPTEPK are among the top ten ranked CD8 epitopes identified by Snyder et al. The sequence KTFPPTEPK was also identified experimentally by Ferretti et al as a CD8 epitope with predicted HLA restriction of A*03 and A*11. Also identified experimentally by Peng and colleagues as a CD8/MHC class I epitopes likely covering HLA types A*03:01 and A*11:01.

¹⁵⁸ <https://www.sciencedirect.com/science/article/pii/S1931312820301669>

¹⁵⁹ <https://science.sciencemag.org/content/early/2020/08/04/science.abd3871>

Orf1 1636T

1636 HTTDP SFLGRY

IN TESTING, vaccine Generation 8, 9, 10

Class I MHC/HLA restriction: A01:01, A03:02, A11:01, A26:01, A33:03

The sequence HTTDP SFLGRY is the top-ranked MHC class I/CD8 epitope identified by Snyder et al in a proteome-wide screen from Adaptive Biotechnologies, although in subsequent larger public data release from Adaptive, its rank dropped to second. The sequence TTDPSFLGRY was also identified experimentally by Ferretti et al as a CD8 epitope with predicted HLA restriction of A*01. Peng and colleagues didn't include Orf1 in their analysis, explaining the absence of this epitope in their preprint. This epitope sequence is part of the polyprotein Orf1ab; the sub-protein and start position numbering are Nsp3, amino acid 818 (Nsp3 818).

Orf8 107T

107 DFLEYHDVRVVL

IN TESTING, vaccine Generation 9, 10

Class I MHC/HLA restriction: A02:01, A23:01, A24:02, B39:01, B40:01, A33:03, C04:01, C05:01, C05:09, C06:02, C07:01, C07:02

This peptide sequence DFLEYHDVRVVL is among the top twenty MHC Class I/CD8 epitopes in Adaptive Biotechnologies' data. It is highly water soluble, and covers a wide range of common Class I allelotypes.

Orf1 5470T

5470 KLSYGIATVR

IN TESTING, vaccine Generation 9, 10

Class I MHC/HLA restriction: A02:01, A03:01, A03:02, A33:03

This peptide sequence KLSYGIATV is among the top ten MHC Class I/CD8 epitopes in Adaptive Biotechnologies' data and covers Class I MHC/HLA restriction A*02 (KLSYGIATV). With the addition of the flanking C terminal Arg, it is water soluble, and also covers A*03 (KLSYGIATVR) and B*33. This epitope sequence is part of the polyprotein Orf1ab; the sub-protein and start position numbering are Nsp13, amino acid 146 (Nsp13 146).

Nsp7 21T

21 RVESSSKLWAQCVQLH

IN TESTING, vaccine Generation 8; OPTIONAL, Gen 9

Class I MHC/HLA restriction: A02:01, B44:02, B44:03, A58:01

Placed into the list of optional peptides for Gen 9 due to absence from Adaptive Biotechnologies' MHC Class II/CD4 data release in late September 2020 (Orf1ab, which includes Nsp7 was not included in the generation of their data set). It is also absent from the MHC Class II data of Tarke

et al., which does include Orf1ab and Nsp7¹⁶⁰ One of the cross-reactive SARS-CoV-2 epitopes to T-cells from SARS-CoV-1 convalescent patients tested by Le Bert and colleagues. The sequence KLWAQCVQL was also identified experimentally by Ferretti et al as the top-ranked CD8 epitope with predicted HLA restriction of A*02. The same sequence is among the top 20 ranked CD8 epitopes identified by Snyder et al.

Other possible epitopes and peptides, including high solubility synthetic peptides are listed at the end of this document in the Supplemental Materials section.

SUMMARY OF THE PROTECTIVE STRATEGY OF THE GEN 10 VACCINE DESIGN (updated 2021-03-07)

This section describes the evolution of the design strategy of the RaDVaC vaccine to achieve the goal of sterilizing immunity. The B-cell-based protective strategy of the first few generations of vaccine was focused on the Spike RBD interaction with the host ACE2 receptor. However, there are three points that make this strategy suboptimal.

1. The ACE2 contact portion of Spike RBD (the receptor binding motif, RBM) has a low degree of conservation. Structural models show that CoV-1 contacts ACE2 slightly differently from CoV-2, and the related MERS virus interaction is even less conserved. These data suggest that this portion of Spike has higher mutant escape potential than more highly conserved portions of the virus.
2. Mutants with higher transmissibility and potentially higher pathogenicity have arisen in multiple amino acids in and around the RBM. Two in particular have been associated with multiple variants of concern: E484K and N501Y. Both of these mutations appear to alter the interaction of Spike with ACE2, and the binding of certain neutralizing antibodies, but not others.
3. The full RBM is very difficult to produce as synthetic peptides. Even if produced, conformational correctness is a top concern with such a structurally complex portion of Spike, which is likely to be critical for eliciting an appropriate antibody response to this antigen.

Given these concerns, and the fact that early attempts at synthesis of RBD epitope peptides was expensive, inefficient, and complex, RBM peptides were retired in Gen 9. However, we continued to work on solving the problem of creating peptides to represent the RBM, and emerging variants of concern show that mutations in the RBD are important for transmission and pathogenicity. Therefore, for the Gen 10 design, we describe the design of a single peptide that represents key residues on the RBM and includes the N501Y ascendant mutation. We use this simplified

¹⁶⁰ <https://www.sciencedirect.com/science/article/pii/S266637912100015X>

approach to represent a portion of the RBM shown to be bound by neutralizing antibodies, such as REGN 10987, while incorporating this key mutation. But this is not the only or even primary focus of the Gen 10 design. A major focus remains B-cell targets on the Spike protein known to be involved in fusion of the virus and host membranes; in particular, the two targets at about Spike 815 and 1145, with a lesser focus on the less conserved protease cleavage site at about 685. Here is a summary of key reasons for these foci:

1. These portions of the Spike protein are highly to extremely highly conserved. The 50 amino acid Heptad Repeat 2 (HR2) domain is 100% conserved between CoV-1 and Cov-2, as is the immediately flanking sequence at 1145 to 1160. The protease cleavage sites are not quite as conserved but they too are well conserved, and mutant escape potential is constrained because they must be recognized and cleaved by host proteases. The S2' cleavage site and adjoining fusion peptide are highly conserved across all coronaviruses.
2. These 3 targets are required for membrane fusion. Even if one mutates to escape antibody binding, a remaining antibody bound site should inhibit efficient fusion.
3. The two primary target sequences form simple structures that are easily synthesized with relatively short peptides.
4. These 3 targets are strongly bound by antibodies in large fractions of convalescent patients, as described in the linear epitope mapping studies cited in the section above.
5. These 3 targets bind neutralizing antibodies, as described in the studies cited in the section above.

In addition to refocusing the B-cell targeting of the Gen 10 vaccine designs, we maintain our focus on empirically validated T-cell epitopes. Similar to our strategy for B-cell epitope redundancy, we have included multiple T-cell epitopes for common alleles.

PREPARATION OF CHITOSAN NANOPARTICLE VACCINE (updated 2021-03-07)

We have performed many iterations of vaccine preparation, with a range of ingredients and concentrations. The vaccine formula and production protocols below are simpler than earlier versions. Current versions use water soluble derivatized chitosan rather than unmodified chitosan, and this simplifies preparation since unmodified chitosan is only soluble at acidic pH.

Many published protocols exist for the robust production of nanoparticle vaccines from chitosan and its derivatives. Various parameters for creating nanoparticles of specific sizes and characteristics have been studied and optimized. Ideal target nanoparticle size is in the range of about 100 nm to 200 nm, which is in the size range of many viruses, including SARS-CoV-2, although larger nanoparticles have been reported to work well. Key parameters that influence

particle size are chitosan concentration, chitosan to TPP ratio (charge ratio, or mass ratio, CS:TPP), salt concentration, and peptide or protein loading of the nanoparticle (higher load, larger size).¹⁶¹ Chitosan nanoparticle formation is extremely robust to a wide range of ingredient concentrations.

Chitosan is a deacetylated form of chitin, which is found in mushrooms and the shells of crustaceans such as shrimp and crabs (seafood allergies are not allergies to chitin). Deacetylation produces free amino groups, and a high density of positive charges on chitosan. Protonation of the amino groups is pH dependent, and chitosan is generally soluble below pH 6, and insoluble at neutral or physiological pH (it has a pKa in the range of about 6.2 to 6.5). Therefore, many different chitosan derivatives have been developed (for review see Wang et al¹⁶²). Aside from pH/pKa differences, positively charged derivatives behave very similarly to unmodified chitosan; therefore, the following protocols sometimes use the general term chitosan which can include these higher solubility chitosan derivatives.

Many publications have described the production and properties of trimethyl chitosan (TMC), and quaternary or quaternized chitosan, including 2-hydroxypropyl-trimethyl ammonium chloride chitosans (HACC or HTCC). Unfortunately, suppliers of these derivatives are somewhat limited, but HACC appears to be more readily available than TMC. These derivatives provide superior performance relative to unmodified chitosan in vaccine production. To make a chitosan nanoparticle by the extremely common technique of ionic crosslinking, a solution of chitosan is mixed with negatively charged triphosphate (sodium triphosphate, STP a.k.a. tripolyphosphate, TPP) under rapid stirring or vortexing. Ionically bonded gel nanoparticles form spontaneously upon mixing with TPP, including when mixed together with peptide. The methods we use are adapted from published methods for creating peptide-loaded nanoparticles¹⁶³, and protein (ovalbumin) loaded nanoparticles¹⁶⁴, and for the creation of nanoparticles by ionic gelation by mixing a soluble chitosan derivative with poly-anionic peptides as crosslinker^{165, 166}.

Because of its commercial availability and superior properties, we selected HACC for preparation of nanoparticles, which has been used widely in research, including for the production of vaccines that have shown excellent results in animal studies^{167, 168}. The acceptable properties of HACC span a wide range: molecular weight, approximately 100 kDa to 400 kDa; Deacetylation (DA) 70% to 90%; Degree of Substitution (DS) should be as high as possible, ideally >90% (the DS

¹⁶¹ https://scholar.google.com/scholar?cluster=18336291994796435345&hl=en&as_sdt=0,22

¹⁶² <https://www.mdpi.com/1422-0067/21/2/487/html>

¹⁶³ <https://www.frontiersin.org/articles/10.3389/fmicb.2015.00372/full>

¹⁶⁴ <https://www.ncbi.nlm.nih.gov/pubmed/?term=16380189>

¹⁶⁵ <https://www.sciencedirect.com/science/article/pii/S174270611830566X>

¹⁶⁶ <https://www.mdpi.com/2076-393X/3/3/730/html>

¹⁶⁷ <https://www.sciencedirect.com/science/article/abs/pii/S0144861716304623>

¹⁶⁸ <https://www.sciencedirect.com/science/article/pii/S0165247819305863>

of the HACC we use is $\geq 90\%$, derived from Sigma-Aldrich product 417963, DA $\geq 75\%$). Therefore, we assume approximately 70% derivatized chitosan. Some publications report chitosan to TPP w/w ratios; however, degree of deacetylation can influence gelling behavior and nanoparticle size. Therefore, parameters influencing nanoparticle characteristics have also been reported using charge ratios. Chitosan is a polymer composed of D-Glucosamine monomer subunits, and the molecular weight of each monomer is 161.16 Da. The majority of chitosan literature is focused on unmodified chitosan; therefore, we have been guided by the extensive literature on chitosan, plus recent work focused on HACC to guide our vaccine formulations using HACC.

HACC typically is produced by reacting chitosan with 2,3-Epoxypropyltrimethylammonium chloride (a.k.a. Glycidyl trimethyl ammonium chloride), which preserves the single positive charge and increases the molecular weight of each derivatized monomer by 151.63 g/mol. Therefore, for approximately 70% derivatized chitosan, the average mass per monomer is 267.3 g/mol, and the average mass per positive charge is $267.3 \text{ Da} / 0.7 = 381.9$. TPP has a MW of 367.86 Da and a charge of -5. Therefore, the average mass per negative charge is 73.57. An equimolar charge ratio (positive to negative charge) is achieved at a CS:TPP mass ratio of approximately 5.2:1 (compared to the CS:TPP ratio for unmodified chitosan of about 2.8:1).

Nanoparticles are smallest when the chitosan is most dilute. An ideal range of chitosan concentrations is 0.5 mg/ml to 1.5 mg/ml. Given the molecular weight ratio of HACC:CS of about 1.8, it is unsurprising that published protocols for making HACC nanoparticles use higher concentrations. Given the range of key factors tested in previous publications, an ideal range for CS:TPP ratio (w/w) for 80% deacetylated chitosan is about 6:1 to 8:1. Hu et al report that a low CS:TPP ratio increases solution turbidity, probably by creating links between nanoparticles.¹⁶⁹ We have seen this phenomenon with a CS:TPP ratio of less than 4:1. Therefore a reasonable range of weight ratios for HACC:TPP should be approximately 10:1 to 14:1.

Chitosan to epitope peptide ratio (w/w). Ratios of chitosan:peptide have been reported in the literature from 1:1 to 25:1. The 1:1 ratio was achieved using chitosan/TPP as a carrier for insulin¹⁷⁰ and separately for ovalbumin. As the CS:TPP ratio approaches 1:1 loading capacity increases but loading efficiency decreases. Peptide loading of nanoparticles also increases the size of the nanoparticles. A 25:1 chitosan:peptide (w:w) results in approximately 50% increase in particle size, while a 10:1 ratio approximately doubles particle size. NaCl reduces particle size. About 20 to 50 mM NaCl should allow for small particle size, and substantial loading of peptide into the nanoparticle. However increasing NaCl concentration reduces zeta potential of nanoparticles, a measure of stability and resistance to aggregation. A ratio of 1:1 of chitosan to peptide was used by Nevagi and colleagues in self-assembling trimethyl chitosan nanoparticles that required no

¹⁶⁹ https://scholar.google.com/scholar?cluster=10442461531761123020&hl=en&as_sdt=0,22

¹⁷⁰ <https://pubmed.ncbi.nlm.nih.gov/10554100/>

TPP or other additional ionic crosslinker¹⁷¹. This was achieved by synthesizing epitope peptides with additional homopolymeric portions of acidic residues (5, 10, or 15 amino acid polyanions), which supplied sufficient anionic crosslinker for formation of TMC nanoparticles. Both poly-glutamic acid and poly-aspartic acid produced self-adjuvanting vaccine particles, with poly-aspartic acid producing the highest levels of IgG when used in a mouse model. Below, we provide a standard TPP crosslinking protocol for using standard synthetic peptides, and a protocol for hybrid TPP plus polyanion crosslinking using B-cell peptides with extended acidic homopolymers¹⁷².

If the vaccine is being used in outbreak or pandemic conditions, to ensure that the vaccine is free of infectious virus contaminants during administration, solutions should be sterilized prior to mixing. Small volumes of chitosan/peptide and of TPP can be rendered free of infectious virus by heating in capped 15 ml conical tubes to 55° C for at least 15 minutes (longer with larger tubes), allowed to cool to room temperature, then mixed in a small sterile beaker while magnetic stirring rapidly (>500rpm) for a few minutes. Per Amidi et al, 1% w/v Tween 80 can be added to assist solubility for high concentrations of chitosan and/or peptide.

Storage of solutions

All solutions should be stored long term in a freezer at —20°C or below. Peptide and chitosan stock solutions can be divided into smaller aliquots to avoid repeated freeze/thaw cycles.

Stock solutions

Stock solutions only need to be prepared if you are planning to make large quantities of vaccine. If you are planning on making small-scale batches, concentrated stock solutions are probably unnecessary, and you can simply make working solutions. Nevertheless, accurately measuring the very small amounts of required materials is best done by making larger volumes. A vaccine recipe for about 10 ml (about 12 to 15 doses) is provided in a section below, and working solution recipes below allow you to make several of these batches. If you require much more than this, consider making stock solutions. Note that peptide stock and working solutions are the same.

Once prepared, heat stock solutions to 55° C (15 minutes for 50 ml conical tube; 10 minutes for 15 ml conical tube) to ensure that they are free of infectious virus. 5 M NaCl generally does not require heat sterilization but can be treated at the same time as an added safety precaution.

- dH₂O
- 5M NaCl (in dH₂O)
- 10 mg/ml HACC in 100 mM NaCl; 50 ml HACC stock solution recipe:

¹⁷¹ <https://www.sciencedirect.com/science/article/pii/S0968089619303347>

¹⁷² <https://www.mdpi.com/2076-393X/3/3/730/htm>

- a. Add 49 ml dH₂O to a 50ml conical tube
 - b. Add 1 ml 5M NaCl to the tube
 - c. Weigh 500 mg HACC on a jeweler's scale, and add to the tube.
 - d. Cap the tube and shake or vortex occasionally until the chitosan is dissolved. This probably will take at least 30 minutes. Heating helps speed the process.
 - e. Dilute to desired working concentration. To make 50 ml working concentration of 2 mg/ml HACC in 20 mM NaCl, add 10 ml stock solution to a 50 ml conical tube, and then fill the tube to 50 ml with dH₂O
- 5 mg/ml Na₅P₃O₁₀ (sodium triphosphate, a.k.a. tripolyphosphate, TPP) in dH₂O.
 - 1.0 mg/ml to 2.4 mg/ml peptide solution(s). Mix peptides into dH₂O. See protocol below for peptides that require disulfide circularization. A master mix of peptides can be made after each is suspended in water, or peptide powders can be combined into a single tube and resuspended together. For example, multiple peptide solutions of 1.5 mg/ml can be mixed together, and the final concentration of all peptides will be 1.5 mg/ml. If you intend to make vaccine that utilizes T-cell peptides and B-cell epitope peptides in a different manner (for example, use of poly-anionic B-cell peptides as crosslinkers for enhanced surface presentation, as detailed in Protocol 2 below), then separate T-cell and B-cell peptide master mixes can be created.
 - a. Add sufficient water to a conical tube (leaving headspace for mixing)
 - b. Add peptide powder.
 - c. Cap and invert or shake the tube until all of the peptide is dissolved.
 - OPTIONAL: solutions containing additional adjuvants

Working solutions

Working solutions of HACC and TPP can be created from 5x stock solutions by adding 1 part stock solution to 4 parts water. For example, for 50 ml TPP working solution, add 10 ml TPP stock solution to 40 ml dH₂O. Individual working solution recipes are also provided in detail below.

- **HACC:** 2 mg/ml HACC in 20 mM NaCl; 50 ml recipe:
 - a. To a 50 ml conical tube, add NaCl to 20 mM by one of the following methods:
Pipette 200 microliters 5 M NaCl into the tube; or, weigh 59 mg NaCl and add to the tube
 - b. Weigh 100 mg chitosan and add to the tube
 - c. Add dH₂O to 50 ml (about 49.8 ml dH₂O)
 - d. Cap and shake the tube vigorously for about a minute, and then occasionally until all of the chitosan is fully dissolved. This might take up to 30 minutes for larger clumps of chitosan to fully dissolve.
- **Peptide for recipe 1 below:** 1.5 mg/ml peptide solution(s), prepared as above.

- **Ionic crosslinker for recipe 1 below:** 1 mg/ml TPP: recipe for 40 ml:
 - a. Add 40 ml dH₂O to a 50 ml conical tube
 - b. Add 40 mg TPP to the tube
 - c. Cap and shake the tube until all TPP is dissolved.
- **TPP+PAP ionic crosslinker for recipe 2 below:** 2:1 w:w ratio. Current testing range: 0.8 mg/ml to 1.0 mg/ml TPP, 0.4 mg to 0.5 mg/ml PAP: recipe for 10 ml:
 - a. Add 5.0 ml of 1 mg/ml B-cell epitope peptide stock solution to a 15 ml conical tube
 - b. To the same tube, add 3.0 ml dH₂O
 - c. Add 2.0 ml of 5 mg/ml TPP stock solution
 - d. Pipette to mix or cap the tube and gently invert to mix
- **OPTIONAL:** 1M solution of sodium bicarbonate (NaHCO₃, baking soda); recipe for 10 ml:
 - a. To a 15 ml conical tube, add 10 ml dH₂O
 - b. Add 840 mg NaHCO₃ to the tube
 - c. Cap and shake the tube until all NaHCO₃ is dissolved.

Charcoal-assisted circularization of peptides

For peptides that require disulfide circularization. After circularization, charcoal can be filtered out although we find that freezing peptide solutions with charcoal does not appear to cause problems, and the charcoal remains intact in the bottom of the tube upon thawing.

1. Resuspend lyophilized peptide powder in dH₂O at a concentration of 1.5 mg/ml in a tube with headspace and a tight fitting cap.
2. Estimate the amount of peptide in the tube and weigh approximately 1 to 2 equal weights of small (average 0.5 to 1 mm diameter) granular charcoal (e.g. for 10 mg peptide, use 10 to 20 mg charcoal). If the charcoal granules are larger, they can be broken up to increase surface area. Add the charcoal to the peptide solution and tightly cap the tube.
3. Place the tube on agitator (rocker, shaker, rotator, etc). Incubate at room temperature 3 to 5 hours, with mild agitation. Make sure that the charcoal is mixing through the peptide solution, and not simply sitting in the bottom of the tube. This can be done by placing the tube on its side and shaking, or by rocking or inverting the tube on a rotator.

Mixing the vaccine

Protocol 1: Ionic gelation with cationic HACC and anionic TPP as ionic crosslinker - sample protocol to make 10 ml of vaccine (about 12 to 15 doses plus nasal sprayer priming waste)

1. Heat the working solutions to 50 to 60 deg C (20 minutes for 50 ml conical tube; 15 minutes for 15 ml conical tube) to ensure they are virus free.

2. Sterilize the small beaker and stir bar by cleaning with soap and water, and then drying by wiping with an alcohol-soaked kimwipe or paper towel.
3. Place the magnetic stir bar inside the small beaker, and place the beaker onto a magnetic stir plate.
4. Add 8 ml of 1 mg/ml chitosan (HACC) working solution to the beaker. Turn on the stir plate and slowly increase the stir speed to at least 500 rpm.
5. Add 1.0 ml peptide solution (total of 1500 micrograms peptide).
6. OPTIONAL adjuvant solutions should be added at this time.
7. Slowly add 1.0 ml TPP solution one drop at a time, and after the last drop is added continue to stir for at least 1 minute.
8. Test the pH of the vaccine. The pH should be in the range of 3.0 to 4.0. This is a mildly acidic pH, equivalent to very dilute vinegar, and can be used as is. If desired, adjust to pH 4.5 with NaHCO_3 solution. Only a very small amount (about 100 microliters) of 1 M NaHCO_3 should be required per 10 ml vaccine.
9. Aliquot 800 microliters vaccine into sterile capped microfuge or 2 ml tubes. Only about 500 to 600 microliters of this will be used as vaccine; some will be lost due to sprayer priming and other waste.

Protocol 2: Ionic gelation with cationic HACC and poly-anionic peptides (PAP) plus TPP as ionic crosslinkers - sample protocol to make 14 ml of vaccine (about 25 doses plus nasal sprayer priming waste)

This protocol is used for improved surface presentation of B-cell epitopes, based on the aforementioned research by Nevagi and colleagues¹⁷³, and Correia-Pinto et al¹⁷⁴. It employs anionic homopolymeric stretches of amino acids at the termini of B-cell epitope peptides. T-cell epitopes do not need to be surface exposed on the nanoparticle for efficient processing and presentation. Therefore, T-cell peptides do not bear these polyanions and are mixed into the chitosan solution prior to ionic crosslinking, thereby encapsulating them in the nanoparticle. We have designed our B-cell epitope peptides to be synthesized with poly-anionic extensions of 6 to 8 amino acids, which are linked to the epitope sequence by a flexible 3 to 6 carbon linker. We named these poly-anionic peptides or PAPs. These PAPs including linkers can be produced by solid phase synthesis, and should not be more difficult or expensive than standard peptide synthesis.

To crosslink PAPs to the surface of the HACC nanoparticle, the PAPs are mixed together with TPP, and this mixture is added dropwise to the chitosan and T-cell peptide mixture to form gel

¹⁷³ <https://www.sciencedirect.com/science/article/pii/S0968089619303347>

¹⁷⁴ <https://www.mdpi.com/2076-393X/3/3/730/html>

nanoparticles. This results in encapsulation of the PAPs but also should result in binding of PAPs to the surface. Based on the work by Correia-Pinto and their use of a 10:1 or 20:1 w:w ratio of chitosan to PAP, we use this range for our initial experimentation. While their designs provide desirable features, the use of chitosan hydrochloride salt is incompatible with the strategy of retaining ionically bound surface epitopes at physiological pH. Therefore, our protocol is based on HACC, and its positive charge at physiological pH should retain surface-bound PAPs. During dropwise addition of TPP, nanoparticles begin to form at a chitosan:TPP ratio of about 15:1. Although nanoparticle formation for HACC hasn't been studied quite as extensively as for unmodified chitosan, it has been shown to be similar to that of unmodified chitosan, although differences in molecular weight to charge ratio and degree of substitution must be taken into consideration¹⁷⁵. Given these similarities and differences, we expect that nanoparticle formation by dropwise addition of TPP should begin at a HACC:TPP ratio of about 20:1. With the use of HACC, Correia-Pinto's protocol A (mixing the epitope together with TPP) should produce similar results to their protocol B (binding PAPs to the formed nanoparticle), because substantially over half of PAPs should remain in solution and be exposed to the surface after initial formation and ionic condensation of the nanoparticle. This should allow remaining PAPs to undergo equilibrium association with the nanoparticle surface. The range of HACC:TPP ratios we are testing has been guided by prior publications on the formation of HACC nanoparticles by ionotropic gelation with TPP^{176, 177, 178}.

Four ingredients are involved in nanoparticle formation, encapsulation, and ionic crosslinking:

1. HACC, 2-hydroxypropyltrimethyl ammonium chloride chitosans (also called HACC)
2. TPP
3. Crosslinking poly-anionic peptide (PAP) master mix of B-cell epitope peptides; about 1/3 of total peptide mass is poly-anionic
4. Non-crosslinking peptide (NCP) master mix of T-cell epitope peptides

HACC concentration: 2mg/ml

w:w ratios:

Ratio of HACC to TPP = In testing, a range of 10:1 to 12:1

Ratio of HACC to NCP = 5:1

Ratio of HACC to PAP = In testing, a range of 20:1 to 24:1

Ratio of TPP to PAP: 2:1

¹⁷⁵ <https://www.tandfonline.com/doi/abs/10.1080/10601321003742147>

¹⁷⁶ <https://www.mdpi.com/2073-4360/10/11/1226>

¹⁷⁷ <https://www.mdpi.com/1420-3049/26/2/449>

¹⁷⁸ <https://www.tandfonline.com/doi/full/10.3109/10717544.2013.853708>

Premix a solution of 1 mg/ml each TPP and PAP.

1. Heat the working solutions to 50 to 60 deg C (20 minutes for 50 ml conical tube; 15 minutes for 15 ml conical tube) to ensure they are virus free.
2. Sterilize the small beaker and stir bar by cleaning with soap and water, and then drying by wiping with an alcohol-soaked kimwipe or paper towel.
3. Place the magnetic stir bar inside the small beaker, and place the beaker onto a magnetic stir plate.
4. Add 10 ml of 2 mg/ml HACC, 20 ml NaCl working solution to the beaker. Turn on the stir plate and slowly increase the stir speed to at least 500 rpm.
5. Add 2 ml 2.0 mg/ml T-cell peptide solution (total of 4.0 mg peptide) to the beaker.
6. OPTIONAL adjuvant solutions should be added at this time.
7. Add 2.0 ml PAP+TPP solution to the beaker. Slowly add solution one drop at a time, and after the last drop is added continue to stir for at least 5 minutes.
8. OPTIONAL: Test the pH of the vaccine. The pH should be in the range of 6.0 to 7.5. If pH is outside of this range, if desired adjust to pH 7.0 with vinegar or NaHCO₃ solution.
9. Aliquot 500 microliters vaccine into sterile capped microfuge or 2 ml tubes. Only about 400 microliters of this will be used as vaccine; some will be lost due to sprayer priming and other waste.

ADMINISTRATION OF THE VACCINE (updated 2021-03-07)

Dosage amount. Human studies have used as little as 7.5 micrograms inactivated virus or other replication incompetent antigen. Trials by El-Kamary et al. and Atmar et al. use 50 microgram and 100 microgram doses of viral-like particle in chitosan, although these trials use chitosan as a powder, which is a less efficient delivery vehicle than nanoparticles (see review by Smith et al.¹⁷⁹). However, the intranasal inactivated influenza vaccine trial in healthy human adults conducted by Illum and colleagues used 7.5 micrograms and 15 micrograms per dose, with greater than double vaccine effectiveness (VE) at the higher dose¹⁸⁰. We have used 50 micrograms to 100 micrograms total peptide for the initial priming and boost doses. Typical doses have been in the range of 70 micrograms. As of late December 2020, hundreds of self-experimenters have shown good tolerance for this dose. Assuming average coupling efficiencies for our synthetic epitope peptides of 70% to 80%, we expect the minimum peptide dose to be about 50 micrograms full length epitopes. li-key peptide doses in Phase II clinical trials of cancer vaccines have used several doses of 500 micrograms per dose^{181, 182}. With Gen 10, we are trialing a higher dose but

¹⁷⁹ <https://www.tandfonline.com/doi/full/10.4161/hv.27449>

¹⁸⁰ <https://www.sciencedirect.com/science/article/abs/pii/S0169409X01001715>

¹⁸¹ <https://www.sciencedirect.com/science/article/pii/S0923753419356819>

¹⁸² <https://link.springer.com/article/10.1007/s10549-020-05638-x>

still substantially less than this dose. Vaccine preparation protocol 2 allows for administration of about 200 micrograms per dose.

Booster schedule. The chitosan/peptide vaccine is only moderately immunogenic, and has proven to be weakly immunogenic for the stimulation of systemic antibody response. This contributes to short-term safety, but requires higher and more doses for efficacy than injected vaccines with potent adjuvants. A vaccine of this type is most effective if used at least 3 times (as shown by animal studies and human clinical trial data): a priming dose, and administration of 2+ booster doses. People over the age of 60 or with complications that might attenuate immune response, might consider additional booster doses. Booster doses have been taken as soon as a few days after the prior dose, or as long as a few weeks. Ideally, doses are spaced by 1 week or more. Given the seriousness of the pandemic and individual circumstances, even shorter intervals between boosters might be used. This is done rarely even in a research setting, but most research is conducted in support of eventual commercialization, and reflects those constraints. Considering the burden and level of stimulation that a few weeks of CoV-2 infection exerts chronically on the immune system, more and frequent booster doses is not an unreasonable strategy. Nevertheless, there is some evidence of greater efficacy for longer intervals between doses. Dozens of RaDVaC testers have self-administered 4+ doses of progressive generations of vaccine, spaced by 1+ week each. Testing for immune response throughout the booster schedule and a few weeks after is ideal for monitoring building immunity; however, this is not practical for some people. Some have reported after-effects (e.g. pronounced nasal congestion) that increase with later booster doses. This is a positive qualitative sign of immune response, but even positive antibody and other test results should not be relied upon as correlates of protection (a perfect proxy for sterilizing immunity).

Pre-administration. To assess the pre-vaccination state of the immune system, one or more saliva samples, nasal wash or swab, and blood draws are ideal prior to vaccination.

Administration. Vaccination is achieved by nasal administration with the use of small (5ml to 20ml) commercial nasal sprayers. We have tried a few sources of spray bottles. Depression of the sprayer top of those we have tested delivers approximately 100 to 150 microliters of fine mist. This should be tested empirically by spraying into a small tube or beaker, and measuring the volume using a pipette. Dose is adjusted accordingly. Spray should be directed into each nostril. The sprayers we obtained come as a top sprayer unit, and either opaque white plastic, amber, or clear bottle. The sprayer unit stem stretches to the bottom of the bottle. The delivered volumes of vaccine will be too small to be delivered reliably using the bottle. Therefore, we do not use the bottle for vaccination, and instead place the sprayer stem directly into the smaller vaccine vial. The stem of our tested sprayer units are slightly longer than a 2 ml microcentrifuge tube. We

pipette an aliquot of vaccine into the tube, then place the stem of the sprayer into the vaccine for spray administration.

Protocol for vaccine administration

To administer vaccine, it is ideal to wear gloves and have a spray bottle of 70% alcohol (ethanol or isopropanol) to sterilize your gloved hands, vials, and equipment. Select a well-lit work surface that can be sterilized with alcohol.

Required materials:

- Clean and well-lit work surface
 - Gloves (nitrile, latex, etc)
 - 70% alcohol (ethanol or isopropanol), preferably in a spray bottle
 - Small (5ml to 20ml) commercial nasal sprayer bottle
 - Vial of vaccine containing 1 dose plus extra for priming the sprayer (total, 800 microliter)
 - Clean and preferably sterile paper towels or tissues
 - OPTIONAL: saline wash solution and collection tube(s)
 - OPTIONAL: rack or holder for vaccine vial
-
1. Sterilize the work surface with alcohol. Place a paper towel or tissue on the clean work surface, and spray with alcohol. You can place items on the surface and use the towel or tissue to dab or wipe items (e.g. the sprayer stem) while maintaining sterility.
 2. Wipe the vaccine vial with an alcohol wipe. Loosen the cap so that it can be removed by lifting it off, but leave the cap in place, and set the vial upright on the work surface. If you choose to use a tube holder or rack for the vial, make sure it is sterilized.
 3. Blow your nose thoroughly with a tissue or paper towel to clear your nasal passages. Use an alcohol soaked wipe or tissue to clean the outer area of the nose and just inside each nostril. This will help prevent inadvertent contamination of the sprayer tip with infectious virus that might be present on or just inside the nose. We do not recommend cleaning deep into the nasal passage; this is likely to do more harm than good.
 4. Sterilize your spray bottle by placing 2 to 4 ml of 70% alcohol into the spray bottle, screw on the cap, and deploy the sprayer until a fine mist is sprayed.
 5. Remove the spray top, and spray the residual alcohol. While you are doing this **DO NOT TOUCH THE SPRAYER TIP OR STEM** to maintain sterility. After sterilization or the following optional nasal wash, do not return the spray top to the bottle prior to using the vaccine.
 6. **OPTIONAL.** You can use the sterilized spray top apparatus at this point to spray 100 mM NaCl into nostrils for collection of nasal wash samples. Spray saline into nostrils, inhale

slightly, and then evacuate nasal wash into a collection tube. Collect at least 500 microliters.

7. While holding the spray top in one hand, uncap the vaccine vial with the other hand, and place the clean and sterile stem of the sprayer directly into the vaccine vial so that the stem touches the bottom of the vial.
8. Grasp the vaccine vial with one hand, and the spray top with the other hand so that your index finger and middle finger are on the side handles of the sprayer.
9. Deploy the sprayer into the air (not up your nose) until you see a fine mist, in order to clear residual alcohol and/or saline from the sprayer.
10. Insert the spray tip into a nostril. While inhaling through the nose, depress the sprayer (ideally, a total of about 100 to 150 microliters per spray). Sniff in the vaccine. Repeat the dosing, for a total of 200 to 300 microliters into each nostril. The entire procedure can take seconds or minutes but don't let the vaccine drip out of the nose, and don't blow your nose for at least an hour or two.
11. **OPTIONAL.** In later booster doses you can collect nasal washes or saliva after vaccination to assess immune response. At 4 hours post vaccination, collect the sample as described in the previous OPTIONAL nasal wash. Cytokine levels in the sample can be measured as indicators of response, if you have the means to do this.

ASSESSMENTS OF IMMUNE RESPONSE (updated 2021-03-07)

There are multiple important immune responses to an effective vaccine, and, therefore, multiple measures of potential effectiveness. These are not measures of effectiveness, but measurable immune response correlates of protection.

- B-cell/Antibody response
 - Neutralizing vs non-neutralizing
 - Neutralizing titer
 - Avidity/affinity of antibodies
- T-cell response (often measured by cytokine secretion)
 - Cell mediated: Th1/Cytotoxic T Lymphocytes (CTL)/CD8+
 - Humoral: Th2/CD4+

Neutralizing antibodies provide protection against viral infection and replication. They bind to key functional parts of the virus. Non-neutralizing antibodies can be beneficial by marking a virus for destruction, but they do not interfere with virus activities such as infection and replication, and they are also associated with ADE. High antibody titer is generally required for optimal protection. Avidity or affinity of the antibody binding to an antigen is another important measure of an effective antibody. An ideal B-cell response involves high-avidity neutralizing antibodies that bind

to multiple viral epitopes. The sections below address details of measuring each of these key metrics. T-cell responses are often measured by cytokine secretion, or by counting cytokine secreting cells.

Samples to be collected and tested

- Whole blood and serum
- Nasal wash
- Saliva

Baseline samples will be collected, and then post-vaccination samples will be taken at booster timepoints, and also at later dates. Each of the above sample types has been used successfully to detect viral RNA and antibodies.

Collection of samples for immune testing. Ideally, all samples will be collected immediately prior to vaccination, and at various time points after initial and booster vaccinations.

Blood is ideally collected by venipuncture and vacutainer, but can be collected by lancet (finger) and deposited into sterile microcentrifuge tubes, without preservative or clotting chemistry. Antibody assays such as ELISA, typically start at dilutions of 1:100 or 1:200 and thus can be run on a few microliters of serum. About 50% of whole blood is recoverable as serum. Therefore, a few drops of blood (30+ microliters) is sufficient for each individual at each time point, for each test. It is ideal to have enough serum for several tests. Blood and serum can be stored for relatively short periods (weeks) at -20°C until assayed. Whole blood or buffy coat are required for certain assays, such as ELISpot, that use peripheral blood mononuclear cells (PBMCs).

Nasal wash. Method adapted from de Fijter et al ¹⁸³. Use nasal sprayer to deposit 200+ microliters of sterile 100mM saline into one nostril. Collect nasal wash effluent in a clean, labeled tube (a small funnel that fits into the collection tube is very helpful). Collection tube should be free of preservative, or contain non-denaturing preservative such as EDTA. Repeat the procedure with the other nostril. Repeat until you have collected a total of at least 0.5 ml. Don't collect more than 2 ml. Store at -20°C until further processed or assayed. Once all aliquots are collected, vortex with glass beads, then centrifuge the samples at 1000 g for 10 minutes at 4°C to clarify. Pipette the supernatants into fresh microcentrifuge tubes, and either refreeze at -20°C or process immediately.

Saliva. At least 0.5 ml should be deposited into sterile tubes, free of preservative, or containing a non-denaturing preservative such as EDTA. Store at -20°C until assayed. Depending on dilution factor (typically starting at 10 fold or greater), assays can be run on a few microliters of saliva.

¹⁸³ <https://pubmed.ncbi.nlm.nih.gov/8872971>

Antibodies and B-cell immune response

Methods developed for assessing immunity to other respiratory viruses, such as influenza, can be adopted for use. Immunity will be assessed using standard indirect ELISA antibody assays (shown below) of serum, nasal swab or wash, and possibly saliva.

Variations of ELISA allow testing of the following:

- Presence/absence of antibodies that bind a viral specific antigenic target
- Antibody titer (by dilution series)
- Antibody avidity (e.g. by pulse of chaotrope, such as thiocyanate or urea)

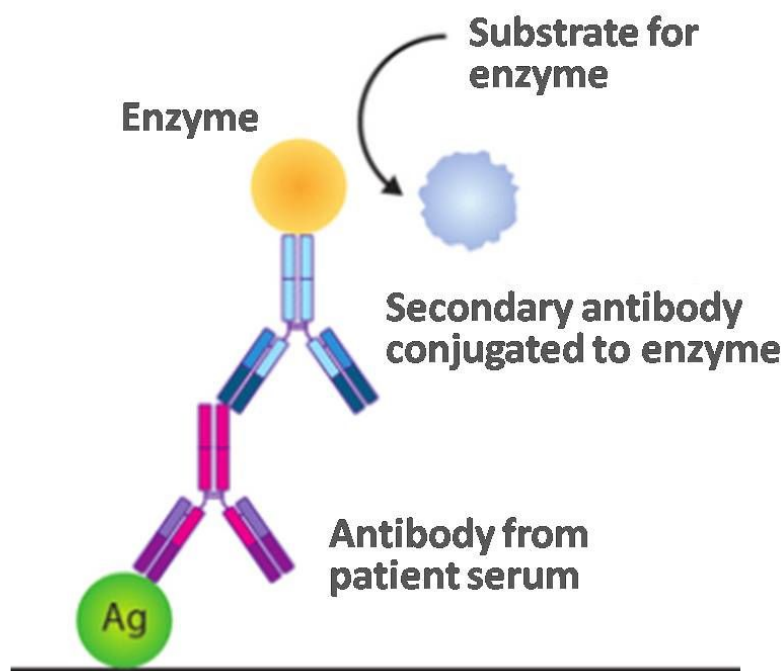
While neutralizing antibodies can't be identified by ELISA, this information can be inferred from published information on which epitopes are bound by antibodies that show neutralizing activity. Thus, if a specific epitope peptide used in a vaccine is bound by an antibody, we can infer that it should be a neutralizing antibody without directly testing it in a neutralization assay.

Initial ELISAs will be performed using recombinant Spike RBD as the target antigen. A few of our B-cell epitope peptides lie within the RBD, and this region is highly represented in immunodominant neutralizing antibodies¹⁸⁴. Followup ELISAs will be performed on various antibody fractions with individual epitope peptides in 96 well plates. Serum will be diluted 1/200, and plated in a standard 2X dilution series. Nasal washes are performed by use of a nasal sprayer and 100 mM saline, and the samples are analyzed by ELISA essentially as by Barria et al.¹⁸⁵ For a review of assessing the effectiveness of nasal vaccines for influenza, see Giancchetti et al¹⁸⁶.

¹⁸⁴ <https://immunology.sciencemag.org/content/5/48/eabc8413>

¹⁸⁵ <https://academic.oup.com/jid/article/207/1/115/874878>

¹⁸⁶ <https://onlinelibrary.wiley.com/doi/full/10.1111/irv.12664>



Schematic of an indirect ELISA assay. Viral antigen (Ag) is attached to a surface (plate well). Patient serum is added to the plate well, washed and blocked. An enzyme-conjugated secondary antibody that binds to human antibodies is added to the plate well. A substrate for the enzyme is added to the plate well, producing a signal if the antigen is bound by antibodies in the patient's serum.

Standard protocols and reagents for testing for SARS-CoV-2 virus exposure have been published.^{187, 188} These protocols and reagents allow testing for immunity either elicited from our vaccine or from virus exposure. Slight modification of the protocols to use our peptides in ELISAs, and for testing of nasal washes, will allow us to test for immunity specific to the epitopes of our vaccine, and for nasal mucosal immunity.

Antibody avidity can be measured by multiple methods, as reviewed by Klasse¹⁸⁹. In our vaccine schema, individual epitope peptides are used to elicit antibodies, and each of these can be used independently to assess avidity of binding antibodies.

B-cell epitopes and ELISA antigen targets

¹⁸⁷ https://labs.icaohn.mssm.edu/krammerlab/wp-content/uploads/2020/03/Expression-and-serology-protocol-3_24_20.pdf

¹⁸⁸ <https://labs.icaohn.mssm.edu/krammerlab/covid-19/>

¹⁸⁹ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4766047/>

Spike protein epitopes / antigen targets

Purified recombinant Spike protein and Spike receptor binding domain (RBD) used as target antigen will allow for testing of viral exposure, and for immunity elicited by our vaccine, but will not be specific for immunity elicited by any specific epitope of our vaccine. Stadlbauer and colleagues use the recombinant RBD protein for testing rather than the full length Spike protein, due to low expression and purification of full-length Spike. We will use RBD as a primary target antigen.

Epitope peptides used in vaccine production can be used in ELISA and other tests for immune response. ELISA testing with recombinant RBD and separately with epitope peptides might help determine if immunity is due to the vaccine, or to virus exposure. ELISA with a protein other than Spike will be more definitive in distinguishing between immunity from the vaccine, and immune response to virus exposure.

Nucleocapsid protein B-cell epitopes / antigen targets

Nucleocapsid is widely used as a highly sensitive antigen target for antibody testing. We ordered a peptide of one of the most antigenic sequences of the highly antigenic Nucleocapsid protein (NP produces large signals in ELISA of convalescent sera). Because of concerns about ADE, we have not yet included Nucleocapsid B-cell epitopes in our vaccine designs. It is possible to use them as negative controls for ELISA. If concerns about ADE risk are resolved, then these selections might be considered for use in a vaccine. We have designed longer peptides, and shorter portions of this peptide.

Nuc 371-399

Longer candidate: DKKKKADETQALPQRQKKQQTVTLPAAD

Long single linear epitope near the C terminus of the Nucleocapsid protein. Substitution of Lys 379 to reduce the positive charge near the N terminus: Lys>Gly. This region has the longest linear epitope discovered by protein microarrays in SARS-CoV-2 (Wang) and SARS-CoV (Zhu)¹⁹⁰. This epitope is split into two overlapping peptides, **Nuc 376-399** and **Nuc 357-382**. **Nuc 376-399** has an additional Glutamate and Alanine at the N terminus to reduce the positive charge and help associate with chitosan.

Substituted **Nuc 371-399**: eaDKKKKADETQALPQRQKKQQTVTLPAAD

Nuc 357-382: IDAYKTFPPTEPKDKKKKADETQAL

Prospective peptide for synthesis

¹⁹⁰ <https://www.pnas.org/content/103/11/4011.long>

Overlaps with **Nuc 376-399**.

Nuc 371-394cir: Acetyl-cDKKKKKADETQALPQRQKKQQTeCL

SYNTHESIZED and ready for use as a negative control in testing

Nucleocapsid antibodies have been implicated in Antibody-directed enhancement. Take notice before selection for use in a vaccine. Useful as a negative control.

Cys1, Cys24 disulfide,

2958 g/mol. 2.96g/ml = 1M. 2.96mg/ml = 1mM 1mg/ml = 338uM

Use of the recombinant RBD as a standard for ELISA will not allow high confidence that an immune response is due to the vaccine rather than virus exposure. Use of envelope or other non-Spike SARS-CoV-2 proteins in ELISA will be more definitive. A positive result in either or both Spike-RBD and Spike1, but negative for non-Spike proteins is suggestive of immune response due to the vaccine.

Positive controls for infection should be obtained. Blood and others samples should be obtained from people who are or have been infected. They should have systemic IgG antibodies to not only Spike protein, but to other proteins as well. They might also have sIgA antibodies in nasal wash and saliva.

T-cell immune response

There are many approaches to testing T-cell immunity specific to the virus. We are considering the following possibilities, and we welcome collaborations with scientists who can contribute to these and other analyses of T-cell responses.

- ELISpot¹⁹¹
- Transcriptomic profiling of PBMC^{192,193}
- Immunome assessment

¹⁹¹ <https://en.wikipedia.org/wiki/ELISpot>

¹⁹² <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3946932/>

¹⁹³ <https://www.ncbi.nlm.nih.gov/pubmed/?term=30205979>

SUPPLEMENTAL MATERIALS

SELECTED PRIOR VACCINE GENERATIONS AND EXPERIMENTAL VARIATIONS

Gen 8

Same peptide composition as Gen 7 vaccine except with the addition of four immunodominant T-cell epitope peptides: Nuc 263T, Orf1 1636T, Nuc 359T, and Nsp7 21T.

1. **Spike 437-460, a.k.a. Spike1**, NSNNLDSKVGGNYNYLYRLFRKSN
2. **Spike 462-476**, KPFERDISTEIIYQAd
3. **Spike 478-502**, kPCNGVEGFNCYFPLQSYGhQPTNG#
4. **Spike 550-574cir**, cgLTESNKKFLPFQQgGRDIADTcD#
5. **Spike 375cir**, cSrDYNSASFSTFKsYGVSPTKcND#
6. **Spike 522cir**, CGPKKSTNLVKNKsVNfNFncd#
7. **Spike 802-823cir**, FSQcLPDPSKPSKRSFcEDLLF#
8. **Spike 21**, RTQLPPAYTNSFTRGVYYPDK
9. **Spike 660**, YECDIPIGAGICASYQTQTNsprRA#
10. **Spike 1145**, LDSFKEELDKYfKNHTSP
11. **Nuc 100-120**, KMKDLSPRWYfYYLGTGPEAG
12. **Nuc 321-345**, GMEVTPSGTWLTYTGAIKLDDKDPN
13. **Nuc 263T**, ATKAYNVTQAFGRRG
14. **Orf1 1636T**, HTTDPSFLGRY
15. **Nuc 359T**, AYKTFPPTEPK
16. **Nsp7 21T**, RVESSSKLWAQCVQLH
17. **Spike 804-820cir**, cILPDPSKPSKRSFcGD* (to be retired as redundant to Spike 802-823cir)
18. **Orf1ab 1544-1564cir**, cFHLdGEVITFDNLKTLLSLREct (computationally predicted T-cell epitope; to be retired due to lack of supportive empirical data)
19. **Spike 462-501**, KPeERDgSTEIIYQAGSTPCNGVEGFNCYFPLQSYGFQPTN (to be retired due to very inefficient synthesis and concerns about conformational correctness)

Gen 8B (basic); Vaccine consisting mostly of peptides added after Gen 6, i.e. peptides new to Gen 7 and Gen 8, but also absent certain peptides. One purpose of this vaccine is to provide a booster for those who received multiple doses prior to Gen 7, in order to avoid excessive stimulation with pre-Gen 7 epitopes. It is also designed to emphasize MHC1/CD8 T-cell epitopes and de-emphasize MHC2/CD4 epitopes, especially those located on the Spike protein. Therefore, multiple peptides bearing Spike MHC II epitopes are not included. Because of synthesis complications, early uses of this vaccine do not include peptide Nsp7 21T. Starting Aug 19, 2020, several RaDVaC members self-administered vaccine containing the peptides listed

below , at about 5 micrograms of each peptide per dose of vaccine. This simpler vaccine design was a forerunner to the design of Gen 9.

1. **Spike 804-820cir**, cILPDPSKPSKRSFc_gD*
2. **Spike 802-823cir**, FSQcLPDPSKPSKRSFcEDLLF*
3. **Spike 21**, RTQLPPAYTNSFTRGVYYPDK
4. **Spike 660**, YECDIPIGAGICASYQTQTNSPRRA*
5. **Spike 1145**, LDSFKEELDKYFKNHTSP
6. **Nuc 100-120**, KMKDLSPRWYFYLLGTGPEAG
7. **Nuc 321-345**, GMEVTPSGTWLTYTGAIKLDDKDPN
8. **Nuc 263T**, ATKAYNVTQAFGRRG
9. **Orf1 1636T**, HTTDPSFLGRY
10. **Nuc 359T**, AYKTFPPTEPK

RETIRED EPITOPE PEPTIDES

Spike 462-501

462 KPFERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTN

Spike 462-501 KP**e**ERD**c**STEIYQAGSTP**c**NGVEGFN**c**YFPLQSYGFQPTN

IN TESTING, vaccine Generation 6, 7, 8.

This 40 aa peptide was ordered on May 9, 2020. Lower case highlighted letters indicate substituted amino acids. Native cysteines are highlighted in yellow, and are to be disulfide bonded (by charcoal treatment). The unmodified peptide was predicted to be marginally soluble, and these substitutions were made to increase solubility. The 40 aa Spike 462-501 peptide failed synthesis twice and was redesigned and made in modified parts as a 15 aa peptide (Spike 462-476) and a 25 aa peptide (Spike 478-502). These peptides arrived May 29. Spike 462-501 arrived June 5. We discussed the failure with the peptide synthesis team and they indicated that coupling efficiency in the amino acids LQSYGF, which are near the C terminus, contributed to the failed syntheses. They flagged Phe 497 as especially difficult. Therefore, the substitution Phe497His was made in Spike 478-502, which is predicted to be a moderately conservative substitution (Tyr was predicted to cause similar coupling inefficiency).

Spike 462-476: KPFERDISTEIQ**A****d**

IN TESTING, vaccine Generations 5, 6, 7, 8.

Spike 478-502: **k**PCNGVEGFNCYFPLQSY**G****h**QPTNG

IN TESTING, vaccine Generations 5, 6, 7, 8.

Spike 462-476 restores the substitutions made to Spike 462-501. The peptide pair Spike 462-476 and Spike 478-502 also contain a paired substitution of amino acids to increase solubility. The substituted amino acids come from the consensus amino acid sequence of SARS-CoV-1 Spike, which has a highly similar 3D structure. The C terminal Gly of Spike 462-476 was replaced with Asp and the N terminal Thr of Spike 478-502 was replaced by Lys. In the SARS-CoV-1 Spike 3D structure, these residues form a salt bridge. The intervening Ser 477 residue was omitted between the C terminus of Spike 462-476 and the N terminus of Spike 478-502, to allow space for the carboxy and amino groups of these peptides.

These peptides contain multiple predicted B-cell and T-cell epitopes and span amino acid sequences of the C terminal portion of 450 and most of 480 from Zhang. Together with Spike 437-460, the total span is about 65 amino acid (a.a.) residues of the ACE2 binding portion of the RBD^{194, 195}, and the highest scoring epitope from Fast and Chen¹⁹⁶. 3D modeling suggests Spike 462-501 (or the peptide pair Spike 462-476 and Spike 478-502) should form a complex with Spike 437-460, with the omission of Gly 461 to allow space for the amino group at the termini of this peptide, and carboxyl group at the C terminus of Spike 437-460. This combined structure contains a linear cluster of amino acids implicated in ACE2 binding of S protein by SARS and by SARS-CoV-2, as shown by Tian et al (Figure 1C)¹⁹⁷, Walls et al (Figure 2C), Park et al (Figure 2)¹⁹⁸, and Sun et al (Figure 1)¹⁹⁹. The nucleotide sequences of multiple SARS-CoV-2 isolates show that the amino acids in this portion of S protein are highly invariant, indicating binding to the receptor depends on the amino acid composition of this epitope.

Solubility of Spike 462-501 is predicted to be marginal and was achieved by substitution: F464E and I468G. These are unlikely to play a role in antibody recognition. However, the sequence containing these substitutions has been identified as a T-cell epitope. Therefore, it would be ideal to restore these substitutions and achieve solubility another way, which was done for Spike 462-476.

- Predicted T-cell epitope by Grifoni et al.: 462-474; DRB1*04:01, KPFERDISTEIYQ
- Predicted T-cell epitope by Braun et al.: 462-474; MHC Class II, KPFERDISTEIYQ
- Peptide Spike 462-476: KPFERDISTEIYQAd

Spike 550-574

¹⁹⁴ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7048180/>

¹⁹⁵ <https://www.ncbi.nlm.nih.gov/pubmed/?term=32155444>

¹⁹⁶ <https://www.biorxiv.org/content/10.1101/2020.02.19.955484v1>

¹⁹⁷ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7048180/>

¹⁹⁸ <https://www.biorxiv.org/content/10.1101/2020.02.22.951178v1.full.pdf>

¹⁹⁹ <https://www.biorxiv.org/content/biorxiv/early/2020/02/20/2020.02.16.951723.full.pdf>

550 GVLTESNKKFLPFQQFGRDIADTTD

Spike 550-574cir: c^gL^gTESNKKFLPFQQ^gGRDIADT^cD

IN TESTING, vaccine Generations 3, 4, 5, 6, 7, 8.

Lower case letters indicate substituted amino acids.

Provided the highest antibody signal in B-cell epitope mapping by Zhang et al. (see Figure 3). Spans one immunodominant epitope reported by Farrera et al²⁰⁰. Also, target of neutralizing convalescent antibodies from Poh et al., and overlapping B-cell predictions from Grifoni et al and Li et al. IgA of two convalescents of SARS-CoV bind the homolog of the N terminal half of this peptide. Disulfide circularized peptide. Cys1, Cys23 disulfide. Cysteines are inserted at positions 550 (G>C) and 573 (T>C), in order to create a disulfide bridge, preserving the looping structure and approximating the geometry of the central portion (TESNKKFLPFQQ) of the peptide shown in the figure below. Substitution of Phe 565 with Gly (F565G). According to structural models, the side chain of Phe 565 projects into the interior of the protein and is therefore unlikely to play an essential role in antibody recognition. It lies in the center of a beta-strand that forms an antiparallel beta-sheet with amino acids 593 to 597. This part of the protein has been omitted and substitution of 551 (V>G) to prevent beta sheet formation between the two remaining beta strands that are separated in the native structure. This portion of the peptide is internal, away from the loop structure predicted to be bound by antibody. However, Braun et al used SARS homology to predict that this Phe is contained within a T-cell epitope.²⁰¹ Therefore, future versions of peptides should restore this Phenylalanine: c^gL^gTESNKKFLPFQQFGRDIADT^cD.

- Predicted T-cell epitope by Braun et al.: MHC Class II, SNKKFLPFQQ^FGRDI, 555-569

Spike 375 One of the immunodominant epitopes identified by Zhang et al. Contains neutralizing residues bound by antibody CR3022. Multiple substitutions made outside the predicted T-cell epitope and antibody binding sites. Cysteine 389 in 3D image below substituted with serine (C389S). Tyrosine 375 and Leucine 397 substituted with Cysteines (Y375C, L397C) to form a disulfide bond for circularization. The native peptide sequence was predicted to be insoluble, so V377R and L378D are designed to provide solubility and preserve the conformation of the structure. Conservative amino acid substitutions are highlighted in green. This epitope also contains the experimentally verified homologous immunodominant T-cell epitope in SARS-CoV-1, Spike 366–374, CYGVSATKL (SARS-CoV-2 Spike 379-387)²⁰².

²⁰⁰ <https://www.medrxiv.org/content/10.1101/2020.06.15.20131391v1>

²⁰¹ <https://www.medrxiv.org/content/10.1101/2020.04.17.20061440v1>

²⁰² <https://www.sciencedirect.com/science/article/pii/S0042682205000620>

- Predicted T-cell epitope by Grifoni et al.: DRB1*07:01, DR8, SASFSTFKCYGVSPTKL, 371-387
- Predicted T-cell epitope by Braun et al.: MHC Class II, SASFSTFKCYGVSPTKL, 371-387

Spike 375cir

375 cSrdYNSASFSTFKsYGVSPTKcND

IN TESTING, vaccine Generations 4, 5, 6, 7, 8.

Spike 522 ATVCGPKKSTNLVKNKCVNFNFNGL

One of the immunodominant epitopes identified by Zhang et al. This epitope also largely overlaps the experimentally verified homologous immunodominant T-cell epitope in SARS-CoV-1, Spike 525-532, VNFNFNGL (SARS-CoV-2 Spike 539-546)²⁰³. Multiple substitutions were made to circularize, to enhance solubility, and to provide a stabilizing salt bridge present in the full Spike protein. Cysteine 525 was retained and was made the N terminal residue. Glycine 545 was substituted by Cysteine to circularize (G545C) by disulfide bridge with C525. Cysteine 538 was substituted with Serine (C538S). Leucine 546 was substituted with Aspartic acid (L546D) to enhance solubility and to stabilize the structure. In the 3D structure, K528 forms a salt bridge with D389, and this C terminal D546 is predicted to substitute for D389.

- Predicted T-cell epitope by Braun et al.: MHC Class II, STNLVKNKCVNFN, 530-544

Spike 522cir

522 CGPKKSTNLVKNKsVNFNFNcd

IN TESTING, vaccine Generations 4, 5, 6, 7, 8

Orf1ab 1543-1576

These peptides are being retired. They were part of early vaccine designs based on strong computational predictions. However, extensive public and private data have failed to support these predictions.

1543 TFHLDGEVITFDNLKTLTSLREVRTIKVFTTVDN

The two highest scoring T-cell epitopes predicted by Fast and Chen are clustered tightly together with a third on the Orf1ab protein. All three are spanned by a predicted epitope included in the vaccine design of Jain and colleagues²⁰⁴.

²⁰³ <https://www.sciencedirect.com/science/article/pii/S0042682205000620>

²⁰⁴ <https://www.biorxiv.org/content/10.1101/2020.03.26.009209v1>

Orf1ab 1544-1564cir: cFHL DGEVITFDNLKTL LSLREct**IN TESTING, vaccine Generations 3, 4, 5, 6, 7, 8. RETIRED**

The highest scoring T-cell epitope predicted by Fast and Chen, and a second overlapping and high scoring T-cell epitope. They predict 74% MHC I coverage for HLDGEVITF (and 100% MHC II coverage), and 83% MHC I coverage for ITFDNLKTL (and 100% MHC II coverage). However, these sequences have not been reported in the empirical CD8 epitope mapping data of Ferretti et al, and Snyder et al (Peng et al did not include Orf1ab in their mapping studies). Therefore, we are considering retiring this peptide as non-essential. Furthermore, this peptide is predicted to be soluble; however, it is not water soluble at room temperature. Therefore, we did not disulfide circularize this peptide because it might further decrease solubility. And because much of the peptide is in the form of solid precipitate, separation from charcoal would not be simple and straightforward.

ADDITIONAL SEQUENCES**Spike protein sequence****without breaks or spaces, for peptide search:**

mfvflvllplvssqcvnltrtqlppaytnsftrgvypdkvfrssvlhstqdlflpffsnvtwfhaih
 vsgtngtkrfdnpvlpfndgvyfasteksniirgwigfttldsktqslivnnavvikvcefqfcnd
 pflgvyyhknnkswmesefrvyssannctfeyvsqpflmdlegkqgnfknlfrefvfkndgyfkiyskh
 tpinlvrdlpqgfsaleplvdlpiginitrfqtlalhrrsyltpgdsssgwtagaayyvgylqprtfl
 lkynengttdavdcaldpsetkctlksftvekgyqtsnfrvqptesivrfpnitnlcpfgevfnat
 rfasvyawnrkrisncvadysvlynsasfstfkcyspstklnldlcfntvyadsfvirgdevrqiapgq
 tgkiadynyklpddftgcviawnsnnldskvggnynlyrlfrksnlkpferdisteiyyagstpcngv
 egfncyfplqsygfgqptngvgyqpyrvvlsfellhapatvcgpkkstnlvknkcvnfngltgtgvl
 tesnkkflpfqqfgrdiadttdavrdpqtileilditpcsfggvsvitpgtntsnqvavlyqdvntevp
 vaihadqltptwrvystgsnvfqtragcligaehvnnsyecdipigagicasyqtqtnsprarsvasq
 siiaytmslgaensvaysnnsiaiptnftisvtteilpvsmtktsvdctmyicgdstecsnlllqygsf
 ctqlnraltgiaveqdkntqevfaqvkkqiyktpikdfggfnfsqilpdpskpskrsfiedllfnkvtl
 adagfikqygdcldgiaardlicaqkfngltvlpplltademiaqytsallagtitsgwtfgagaalqip
 famqmayrfngigvtqnvlyenqklianqfnsaigkiqdslsstasalgklqdvvnqnaqalntlkvql
 ssnfgaissvlnldilsrldkveaevqidrlitgrlqslqtyvtqqlliraaeirasanaatkmsecvlg
 qskrvdfcgkgyhlmsfpqsaphgvvflhvtvypageknfttapaichdgkahfpregvfvsnghwfv
 tqrnfyepqiittntfvsngncdvigivnntvydplqpeldsfkeeldkyfknhtspdvldgdisgin
 asvvniqkeidrlnevaknlneslidlqelgkyeqyikwpwyiwlgfiagliaivmvtimlccmtscs
 clkgccscgscckfdeddsepvlkgvklhyt//

Nucleocapsid protein sequence**without breaks or spaces, for peptide search:**

msdngpqnqrnapritfggpsdstgsnqnqgersgarskqrrpqglpnntaswftaltqhgedlkfprgqgvpinntss
 pddqigyyrratrrirggdgkmkdl sprwyfyytgtgpeaglp ygankdgiwvategalntpkdhigtrnpannaaiv
 lqlpqgttlpkgyfaegsrqgsqassrsssrnssrnstpgssrgtsparmagnggdaalallldrlnqlskmsgk
 gqqqqgqvtvkksaaeaskkprqkrtatkaynvtqafgrrgpeqtqgnfgdqelirggttykhwppqiaqfapsasaffg
 msrigmevtpsgtwltytgaiklldkdnpnfkdqvillnkhidayktfpptepkkdkkkkadetqalpqrqkkqqtvtll

paadlddfskqlqgsmssadstqa