DNASCRIPT

APPLICATION NOTE

SARS-CoV-2 Spike Protein Mutagenesis

Recombinant spike (S) proteins are a critical resource for functional and structural SARS-CoV-2 studies. In this Application Note we demonstrate the use of primers produced with DNA Script's novel Enzymatic DNA Synthesis (EDS) technology in a pragmatic, site-directed mutagenesis strategy for the production of optimized and variant SARS-CoV-2 spike protein expression plasmids. EDS primers supported efficient multi-site mutagenesis and Sanger sequencing confirmation of mutations and cloned inserts. Rapid access to custom oligos enabled predictable project schedules amidst a global synthetic DNA supply bottleneck.

INTRODUCTION

Variants of SARS-CoV-2 have been emerging and circulating around the world since the start of the COVID-19 pandemic.[1-3] Recombinant proteins—particularly those corresponding to variants of spike (S) protein (the primary viral antigen, Figure 1)—are required to study transmissibility and immune escape. Expression plasmids for functional and structural studies may be obtained via a variety of strategies, ranging from the assembly and cloning of PCR products to the production of full-length synthetic constructs. While the latter appears to be the most direct approach, global bottlenecks in commercial oligo supply during the pandemic forced many labs to adopt alternative approaches to advance their critical research.

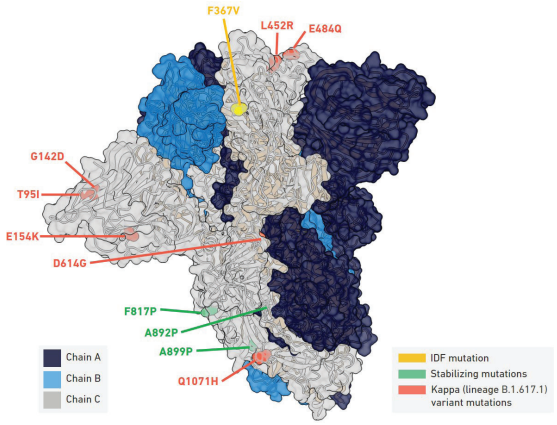

FIGURE 1. CLOSED CONFORMATION STRUCTURE OF THE SARS-COV-2 SPIKE PROTEIN TRIMER with D614G mutation,[4] highlighting amino acid mutations introduced in this study. Not all mutations are shown, as some are not visible or resolved.

Site-directed mutagenesis of previously cloned DNA fragments is a well-established molecular biology technique and offers a flexible and cost-effective approach for producing multiple plasmids for a particular study. All commonly used site-directed mutagenesis methods and kits require oligos for the introduction of specific nucleotide changes at defined positions. With DNA Script's novel Enzymatic DNA Synthesis (EDS) technology[5] and SYNTAX System,[6] oligos can be produced for same- or next-day use in a standard molecular biology lab. This capability enables rapid cycles of plasmid iteration as viral variants and/or project objectives evolve.

This Application Note describes work done to support SARS-CoV-2 research performed at the CNRS Virology lab directed by Félix Rey (Institut Pasteur), at a time when the turnaround time for a new spike variant expression plasmid from external providers exceeded two months. The SYNTAX System was used to synthesize oligos for site-directed mutagenesis of cloned versions of the SARS-CoV-2 S gene, as well as primers for verification of clones by Sanger sequencing. In-house production of custom oligos allowed the development of a pragmatic and iterative, multi site-directed mutagenesis approach. Our workflow (Figure 2) enabled the production of sequence-verified expression plasmids in <2 – 4 weeks. The strategy is broadly applicable to all types of mutagenesis projects for the production of recombinant protein variants. In addition to providing a rapid and cost-effective alternative to full length gene synthesis by commercial suppliers, it also enables full control over project timelines.

INSTITUT PASTEUR

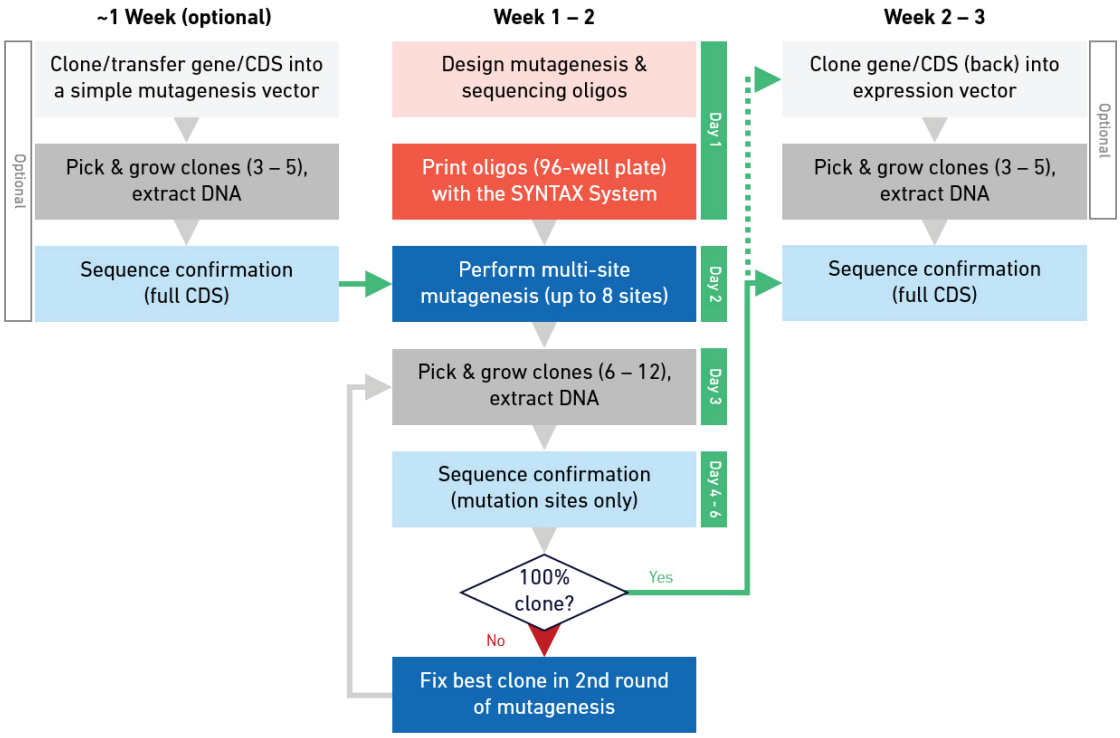

FIGURE 2. ITERATIVE, MULTI SITE-DIRECTED MUTAGENESIS STRATEGY DEVELOPED FOR THE GENERATION OF SARS-COV-2 VARIANT SPIKE PROTEIN EXPRESSION PLASMIDS. Steps in blue require oligos and can be rate-limiting in overall project progress when relying on commercial oligo supply and third-party logistics. With the SYNTAX System, custom, ready-to-use mutagenesis oligos and sequencing primers can be designed and synthesized on demand and in parallel with other workflow steps. This capability enables the production of sequence-verified expression plasmids in <2 – 4 weeks, depending on (i) the vector in which the DNA fragment to be mutagenized is available, and (ii) the complexity of the mutagenesis strategy. Optional subcloning/sequence verification phases are required if the mutagenesis target is not available in a plasmid backbone suitable for multi-site mutagenesis. This step increases the overall time to produce a new expression plasmid, but also increases mutagenesis success rates and greatly reduces the potential for introducing spurious nucleotide changes that may impact expression efficiency (and require fixing in additional rounds of mutagenesis).

EXPERIMENTAL DESIGN

The site-directed mutagenesis strategy with EDS primers developed in this study was used over the course of three months to incorporate 25 amino acid substitutions and 3 deletions across 12 SARS-CoV-2 spike protein expression vectors. These sequence-verified plasmids contained Alpha (lineage B.1.1.7) or Kappa (lineage B1.167.1) spike protein mutations, or amino acid substitutions designed to optimize recombinant protein expression and stability. Only one representative project is described here and is outlined in Figure 3.

The starting point for this project, plasmid pCoVS1_pcDNA3.1(+), was obtained from the Rey lab. It comprises the SARS-CoV-2 spike ectodomain (amino acids 1 – 1,208) with two stabilization mutations (K986P and V987P) fused to a 3'-terminal T4 fibritin trimerization (foldon) motif (spike ectodomain-foldon, or spike EF),[7,8] cloned into the mammalian expression vector pcDNA3.1(+) (Invitrogen/ThermoFisher Scientific).[9] The spike ectodomain was derived from a virus isolate from the Île-de-France (IDF) region of France[10] and is identical to that of the Wuhan Hu-1 reference sequence,[11] except for a phenylalanine (F) residue (instead of valine, V) at amino acid 367.

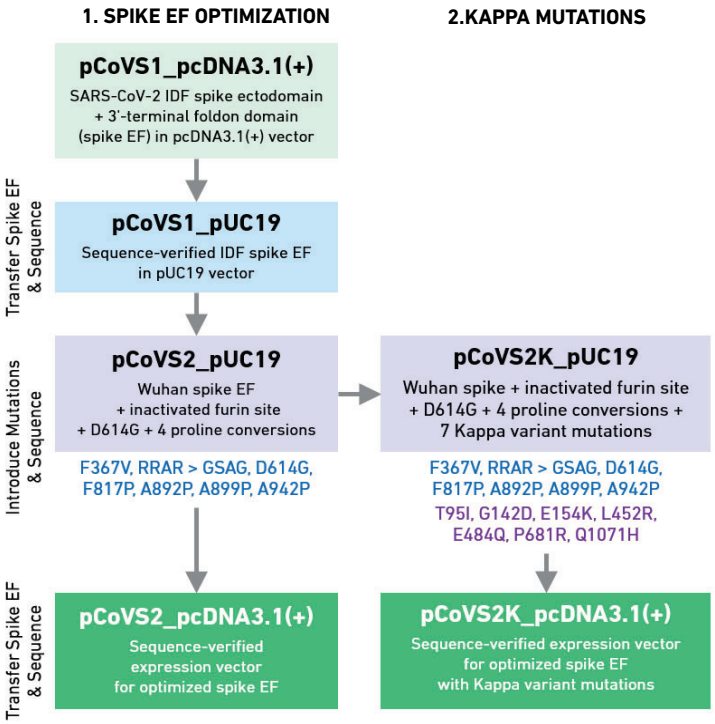

FIGURE 3. PROJECT OVERVIEW.
The cloned spike ectodomain-foldon construct was transferred from the pCoVS1_pcDNA3.1(+) expression plasmid (constructed by the Rey lab) to the smaller, simpler pUC19 backbone to facilitate multi-site mutagenesis. Subcloning, multi-site mutagenesis, and sequencing were performed as described in Materials and Methods. In Phase 1 the spike ectodomain from the Île de France (IDF) isolate was reverted to the reference (Wuhan Hu-1) sequence, and the D614G mutation was introduced. In addition, (i) the furin cleavage site (amino acids 682 – 685, RRAR) was inactivated (by converting it to GSAG), and (ii) four proline residues were introduced in specific locations to optimize recombinant spike protein expression and stability. In the second phase, pCoVS2_pUC19 was used as the starting point for creating a fully optimized Kappa variant spike protein expression vector (pCoVSK2K_pcDNA3.1(+).

Preliminary mutagenesis experiments (not shown) indicated that expression plasmids are challenging templates for site-directed mutagenesis. These are typically large plasmids with complex functional elements (e.g., an origin of replication, promoter(s), and enhancer(s), some of which are comprised of GC-rich regions and repeat sequences) that are at risk of being compromised through successive rounds of manipulation. For this reason, the spike EF was subcloned into the smaller, simpler pUC19 vector (Addgene)[12] to generate less complex templates for site-directed mutagenesis.

In the first phase of the project, four sets of mutations (ten amino acid changes in total) were introduced into the spike EF:

- F367, to revert the cloned IDF spike ectodomain to the Wuhan Hu-1 reference sequence.
- D614G, a spike protein substitution that has became dominant in SARS-CoV-2 variants by increasing the infectivity and stability of virions.[13]
- Amino acids 682 – 685 (RRAR) were converted to > GSAG to inactivate the spike protein's furin cleavage site,[14] thereby improving expression of the recombinant spike protein.
- Amino acids P817, A892, A899, and A942 were converted to prolines (P) to stabilize the protein structure for x-ray crystallography.

Phase 1 yielded a sequence-verified expression plasmid for the optimized spike EF (pCoVS2_pCDNA3.1[+]). In the second phase of the project, the pUC19 equivalent of pCoVS2 was used as the template for introducing seven Kappa (lineage B1.1.67.1) mutations into the optimized spike EF.

A final round of subcloning and sequence verification yielded the second expression vector, pCoVS2K_pcDNA3.1(+), encoding an optimized Kappa variant spike EF.

Full details of the materials and methods used for subcloning, oligo synthesis, multi site-directed mutagenesis, and Sanger sequencing may be found on pp. 7 – 8.

RESULTS AND DISCUSSION

Construction of Optimized Spike EF Expression Plasmid pCoVS2_pcDNA3.1(+)

Seven mutagenesis oligos (see Figure A1 in the Appendix) were included in the initial multi-site mutagenesis reaction to introduce a total of ten mutations. Results are summarized in Table 1. In short:

- Twelve clones were screened by Sanger sequencing (using primers targeting the mutagenesis sites only).
- The F367V mutation was successfully introduced into all twelve clones.
- The more complex RRAR > GSAG (furin site inactivation) mutation was only achieved in two clones.
- Mixed results were obtained with the four proline conversions: the F817P oligo worked in nine of the twelve clones, but spurious insertion or deletion of an adenosine was observed in four of those clones. With the other three oligos (targeting A892P, A899P, and A942P, respectively) success rates ranging from 33% (four of twelve clones) to 91.7% (eleven of twelve clones) were achieved.
- Instead of repeating the original strategy of trying to introduce all ten mutations in one round of mutagenesis, clone #7 was used as the template in a second round of mutagenesis with oligo #6 only to introduce A899P.
- Of the twelve clones screened, ten contained all ten of the Phase 1 mutations.

Sequence confirmation of the full-length spike EF was performed by tiled Sanger sequencing. The optimized spike EF was transferred from pCoVS2_pUC19 back to pcDNA3.1(+) and sequenced again to confirm that no spurious mutations had been introduced in the final expression plasmid.

Introduction of Kappa variant mutations

Results from the two rounds of mutagenesis with Kappa mutation oligos (see Figure A2 in the Appendix) are summarized in the final two columns of Table 1. In short:

- Four of nine clones screened from the first round of multi-site mutagenesis were near-perfect. Three clones each lacked one of the desired mutations; whereas the fourth clone contained a spurious G>A transition resulting in a R577H substitution.
- The missing T95I mutation was added to clone #8 from round one, in a second round of mutagenesis containing oligo #1 only. Five of six clones screened were perfect. The optimized Kappa variant spike EF from clone #4 was cloned back into pcDNA3.1(+) and fully sequenced to yield the final expression plasmid.

Confirmation of Full-Length Spike EF by Tiled Sanger Sequencing

Sequence confirmation of inserts in mutagenesis and expression plasmids was performed after each cloning step, using eleven (five forward and six reverse) sequencing primers (see Appendix). Since the same primers were used in multiple rounds of sequencing, an appreciable quantity of each was ordered from a commercial supplier. However to demonstrate that primers produced by EDS support high-quality Sanger sequencing, a limited quantity of each sequencing primer was also produced using the SYNTAX System. To this end, six replicate wells of each forward or reverse sequencing primer was produced, pooled, and concentrated (by rotatory evaporation) to yield 10 µM working stocks.

An overview of the results from tiled Sanger sequencing of a representative pCoVS2_pUC19 clone is shown in Figure 4. Sequencing read lengths for individual reactions ranged from 1053 – 1244 nt (average of 1181 nt), with quality scores ranging from 37 to 52 (average of 45.8).

Sequencing results obtained with EDS primers were comparable to those achieved with the same primers synthesized with conventional phosphoramidite chemistry by a commercial supplier (not shown).

TABLE 1. SUMMARY OF MUTAGENESIS RESULTS.

	Phase 1: Spike EF optimization		Phase 2: Kappa mutations	
	Round 1	**Round 2**	**Round 1**	**Round 2**
Template	pCoVS1_ pUC19	Clone #7 from round 1	pCoVS2_ pUC19	Clone #8 from round 1
Number of mutagenesis oligos in reaction	7	1	7	1
(Cumulative) number of desired mutations (amino acids)	10 (F367V, RRAR > GSAG, D614G, F817P, A892P, A899P, A942P)	1 (A899P)	7 (T95I, G142D, E154K, L452R, E484Q, P681R, Q1071H)	1 (T95I)
Number of clones screened	12	12	9	6
Minimum number of desired amino acid changes in any clone (this round)	4 (40%)	0	0	0
Maximum number of desired amino acid changes in any clone (this round)	10 (100%)[1]	1 (100%)	7 (100%)[2]	1 (100%)
Perfect clones (% of number screened)	0	10 (83.3%)	0	5 (83.3%)
Clone taken forward	#7	#6 (pCoVS2_pUC19)	#8	#4 (pCoVS2K_pUC19)
Next steps	• Use as template in Phase 1, round 2	• Confirm full-length spike EF sequence • Clone insert back into pcDNA3.1(+) and confirm sequence of optimized spike EF	• Use as template in Phase 2, round 2	• Confirm full-length spike EF sequence • Clone insert back into pcDNA3.1(+) and confirm sequence of optimized Kappa variant spike EF

[1]Clone 7 lacked A899P. Clone 8 had all ten amino mutations, but also contained a spurious dA deletion at the F817P site.
[2]Clones 1 and 8 lacked G142D and clone 9 lacked T95I. Clone 5 had all 7 amino acid mutations, but also contained a spurious R577H mutation.

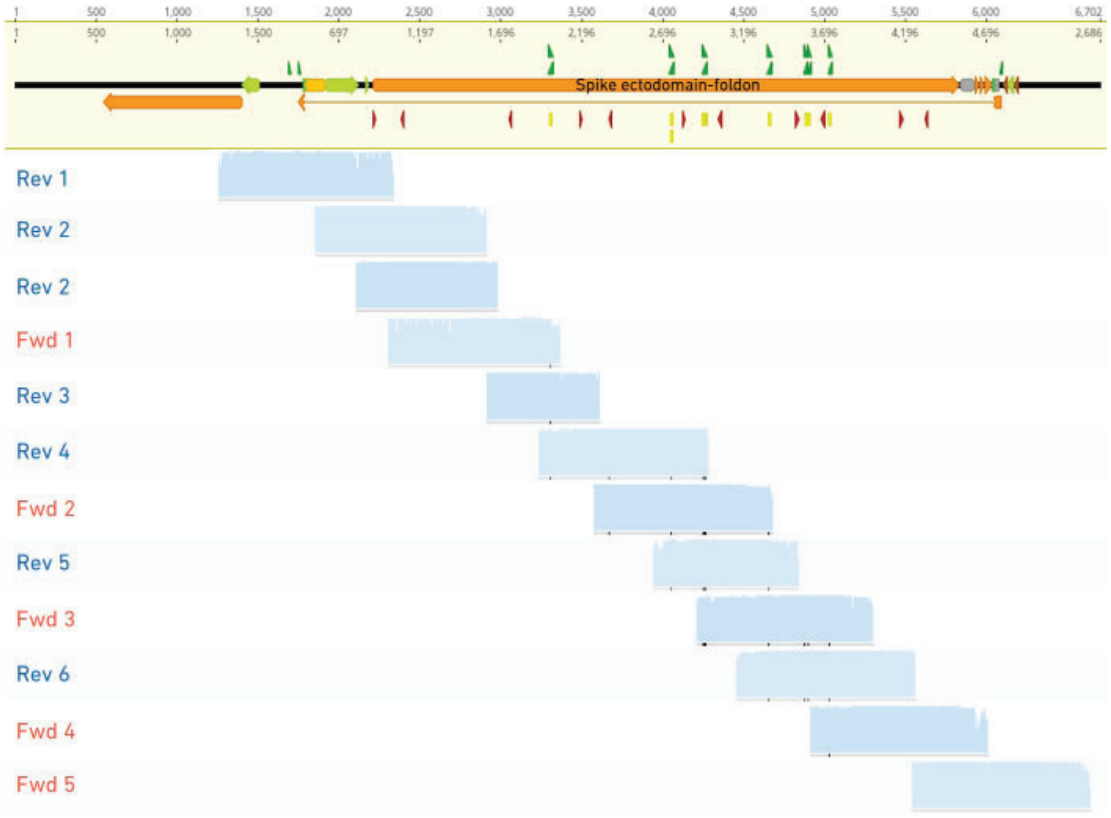

FIGURE 4. SEQUENCE CONFIRMATION OF THE FULL-LENGTH, OPTIMIZED SARS-COV-2 SPIKE ECTODOMAIN-FOLDON FUSION CONSTRUCT (OPTIMIZED) IN A pCoVS2_pUC19 CLONE. Five forward and six reverse sequencing primers (see Appendix) were designed to confirm the sequence of the full-length spike EF in mutagenesis and expression plasmids. Each blue bar represents the chromatogram obtained from a sequencing run performed with the sequencing primer identified on the left, aligned to the plasmid sequence. The theoretical positions of mutations are designated by yellow blocks in the construct map, and by black marks below individual chromatograms. EDS primers yielded high-quality sequencing data. Overlapping reads exceeded 1 kb in length, thus ensuring full coverage of cloned inserts.

CONCLUSION

Access to on-demand oligo synthesis has enabled us to develop a generic, iterative multi site-directed mutagenesis approach for the production of plasmids for recombinant protein expression in a fraction of the time needed to procure full-length synthetic constructs from commercial suppliers. We have successfully used this approach to produce expression plasmids for optimized and variant SARS-CoV-2 spike proteins, thereby supporting critical research during the COVID-19 pandemic.

Oligos produced with DNA Script's EDS technology and benchtop SYNTAX System were shown to enable efficient multi-site mutagenesis (using an established commercial kit) and support verification of mutation sites and full-length inserts by Sanger sequencing. Rapid, in-house DNA synthesis ensured that high-quality, custom oligos were always available when needed for the next step in the project, thereby enabling predictable schedules and steady progress.

ACKNOWLEDGMENTS

This study was performed in collaboration with Prof. Félix Rey, UMR 3569, CNRS Virology Lab, Institut Pasteur, Paris, France. Pablo Guardado Calvo and Marija Backovic from the Rey lab constructed the original SARS-CoV-2 spike CDS plasmids.

MATERIALS AND METHODS

Subcloning of the Spike Protein EF from pCoVS1_pcDNA3.1(+) into pUC19

pCoVS1_pcDNA3.1(+) was digested with restriction enzymes NdeI, XbaI, and SalI, whereas the pUC19 vector was digested with NdeI and SalI. Restriction enzyme digestions were performed for 1 h at 37°C. Digestion products were separated using a 0.8% TBE-agarose gel. The insert (spike ectodomain-foldon fragment) and pUC19 vector backbone were purified using the Monarch® DNA Gel Extraction Kit. DNA fragments were ligated for 5 min at room temperature using the Quick Ligation™ Kit. All restriction enzymes, as well as purification and ligation reagents were from New England Biolabs® and were used according to the manufacturer's recommended protocols.

Four microlitres (4 µL) of the ligation product were chemically transformed into XL-10 Gold® Ultracompetent Cells (Agilent® Technologies), using a 30 sec heat shock at 42°C. Transformed bacterial cells were cultured overnight at 37°C on LB-Agar plates containing 100 µg/mL ampicillin (Fisher Bioreagents). Six individual clones were picked and grown overnight at 37°C in 5 mL of LB broth + 100 µg/mL ampicillin (with shaking at 200 rpm). Plasmid DNA was extracted using the QIAprep® Spin Miniprep Kit (QIAGEN®) according to the manufacturer's instructions.

Confirmation of the Spike EF Sequence in pCoVS1_pUC19

Primers were designed to cover the entire spike CDS with a tiled Sanger sequencing approach. Sequencing primers (17 – 20 nt long) were either ordered from a commercial vendor or synthesized using the SYNTAX System. Primer stocks (10 µM) and plasmid DNA (3 clones per construct) were shipped to a Sanger sequencing service provider (GENEWIZ®). Sequencing data were analyzed with the Geneious™ Prime software (Biomatters, Inc.).

Multi Site-Directed Mutagenesis Strategy for Introducing Optimization Mutations

Mutagenesis oligos (Appendix, Figure A1) were designed using the web-based QuikChange® Primer Design application,[15] according to instructions provided with the QuikChange Multi Site-Directed Mutagenesis Kit (Agilent Technologies). Oligos (27 – 34 nt long) were produced using the SYNTAX System (DNA Script part number 100296 or 100094) and SYNTAX 60-Cycle Kit, 96 oligos (DNA Script part number 100134) with standard synthesis plates according to manufacturer's instructions.

Oligos produced by EDS were diluted in molecular biology-grade water to a final concentration of 25 ng/µL. Mutagenesis reactions were performed using the QuikChange Multi Site-Directed Mutagenesis Kit in a final reaction volume of 25 µL, containing 2.5 µL of 10X QuikChange Multi reaction buffer, 0.5 µL of QuikSolution®, 100 ng of dsDNA template (pCoVS1_pUC19), and 1 µL of QuikChange Multi enzyme blend. Of each mutagenesis oligo, 100 ng was included in reactions performed with 1 – 2 oligos, or 50 ng in reactions performed with 3 – 8 oligos. Mutagenesis reactions were performed using a thermocycler with the following cycling parameters: initial denaturation for 1 min at 95°C, followed by 30 cycles of 95°C for 1 min, 55°C for 1 min, and 65°C for 2 min/kb.

After a 2 min incubation on ice, 1 µL of DpnI enzyme was added to each mutagenesis reaction product. Reactions were incubated for 1 h at 37°C in a thermoshaker to digest the original template DNA. Of each DpnI digestion product, 1 µL was transformed into the XL-10 Gold Ultracompetent Cells provided in the kit according to the manufacturer's instructions. Transformed bacterial cells were cultured in the presence of ampicillin and plasmid DNA was extracted as previously described.

Regions spanning mutagenesis sites were sequenced as previously described, using a subset of the sequencing primers designed for tiled sequencing of the full-length spike EF.

A total of two rounds of multi site-directed mutagenesis and sequencing were performed to obtain a pCoVS2_pUC19 clone with all of the desired mutations and no spurious nucleotide changes.

Generation and Sequence Confirmation of Expression Plasmid pCoVS2_pcDNA3.1(+)

After verification of the full-length insert in pCoVS2_pUC19, the spike EF fragment was transferred back into the pcDNA3.1(+) mammalian expression vector to create an expression vector for the optimized spike ectodomain-foldon fusion construct. To this end, pCoVS2_pUC19 DNA was digested with restriction enzymes NdeI and XbaI, whereas pcDNA3.1(+) vector DNA was digested with NdeI, XbaI, and BamHI. Restriction digestions, electrophoresis and purification of digestion products, ligation, transformation, culturing of transformed bacteria, plasmid DNA extraction, and sequence confirmation of the full-length optimized spike ectodomain construct in pCoVS2_pcDNA3.1(+) clones were performed as previously described.

Multi Site-Directed Mutagenesis Strategy for Introducing Kappa Variant Mutations

The final objective of the project was to generate a Kappa (lineage B.1.167.1) variant spike expression plasmid by introducing seven Kappa mutations into the optimized spike CDS-ectodomain-foldon fusion construct.

Mutagenesis oligos for the Kappa mutations (Appendix, Figure A2) were designed, and produced with the SYNTAX System as previously described. Two rounds of multi-site mutagenesis were performed using the QuikChange Multi Site-Directed Mutagenesis Kit (Agilent Technologies) as previously described. Putative pCoVS2K_pUC19 clones were screened and mutagenesis sites were confirmed by Sanger sequencing as previously described.

Generation and Sequence Confirmation of the Kappa Variant Expression Plasmid pCoVS2K_pcDNA3.1(+)

After sequence verification of the full-length insert in pCoVS2K_pUC19 (performed as previously described), the spike ectodomain-foldon fragment was cloned back into the pcDNA3.1(+) vector backbone to create the optimized Kappa variant spike expression plasmid pCoVS2K_pcDNA3.1(+). Restriction digestions, electrophoresis and purification of digestion products, ligation, transformation, culturing of transformed bacteria, plasmid DNA extraction, and sequence confirmation of the full-length optimized Kappa spike ectodomain construct in pCoVS2K_pcDNA3.1(+) clones were performed as previously described.

REFERENCES

1. World Health Organization. Tracking SARS-CoV-2 variants. https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/.
2. Centers for Disease Control and Prevention. SARS-CoV-2 Variant Classifications and Definitions. https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html#Concern.
3. Stanford University Coronavirus Antiviral & Resistance Database. https://covdb.stanford.edu/page/mutation-viewer/#variants.genome.viewer.
4. Benton DJ, et al. Protein Data Bank 7BNM. doi: 10.2210/pdb7bnm/pdb.
5. DNA Script. Enzymatic DNA Synthesis Technical Note (TN900100). https://www.dnascript.com/wp-content/uploads/2021/04/DNAScript_EDS_3Pillars_TN900100.pdf.
6. DNA Script. SYNTAX System Brochure (PB100296). https://www.dnascript.com/wp-content/uploads/2021/07/DNAScript_SYNTAX-_System_Brochure_PB100296.pdf.
7. Wrapp D et al. Science 2020; 367:1260-1263. doi: 10.1126/science.abb2507.
8. Li J, et al. Viral Immunol. 2013; 26:126-132. doi: 10.1089/vim.2012.0076.
9. https://www.thermofisher.com/order/catalog/product/V79020#/V79020.
10. IDF strain (BetaCoV/France/IDF0372/2020. National Reference Centre for Respiratory Viruses; Institut Pasteur, Paris, France.
11. NCBI Reference Sequence: NC_045512.2. https://www.ncbi.nlm.nih.gov/nuccore/1798174254.
12. https://www.addgene.org/50005/.
13. Plante JA, et al. Nature 2021; 592:116-121. doi: 10.1038/s41586-020-2895-3.
14. Walls AC, et al. Cell 2020; 180:281-292.e6. doi: 10.1016/j.cell.2020.11.032.
15. www.agilent.com/genomics/qcpd.

APPENDIX: MUTAGENESIS AND SEQUENCING PRIMERS

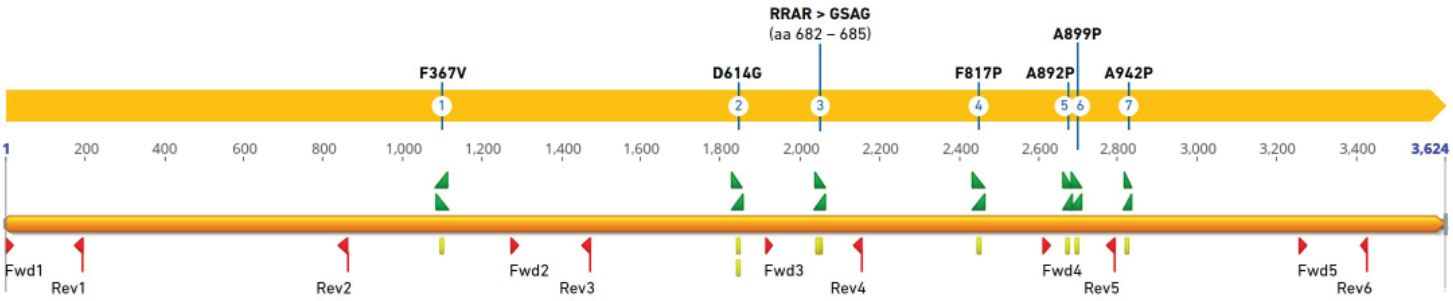

FIGURE A1. MUTATIONS INTRODUCED IN PHASE 1 TO OPTIMIZE EXPRESSION AND STABILITY OF RECOMBINANT SARS-COV-2 SPIKE PROTEIN. Blue numbers (in yellow bar) and yellow rectangles represent the four sets of mutations introduced into the 3,624 bp spike ectodomain (amino acids 1 - 1,208; orange bar) via two rounds of multi site-directed mutagenesis to create pCoVS2_pUC19 (the precursor to the expression plasmid pCoVS2_pcDNA3.1[+]). Green triangles designate the 5'- and 3'-boundaries of mutagenesis oligos produced with the SYNTAX System, whereas red triangles correspond to Sanger sequencing primers used to confirm successful mutagenesis and verify the full-length spike EF sequence.

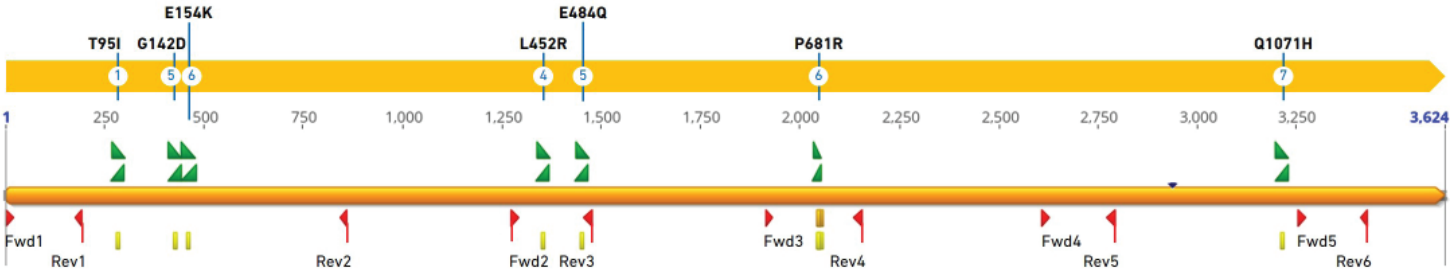

FIGURE A2. KAPPA (LINEAGE B1.167.1) VARIANT MUTATIONS INTRODUCED INTO THE OPTIMIZED SPIKE PROTEIN ECTODOMAIN IN PHASE 2. Blue numbers (in yellow bar) and yellow rectangles represent the seven Kappa mutations introduced into the optimized 3,624 bp spike ectodomain (in pCoVS2_pUC19; orange bar) via two rounds of multi site-directed mutagenesis to create pCoVS2K_pUC19 and the corresponding expression plasmid, pCoV2SK_pcDNA3.1(+). Green triangles designate the 5'- and 3'-boundaries of mutagenesis oligos produced with the SYNTAX System, whereas red triangles correspond to Sanger sequencing primers used to confirm successful mutagenesis and verify the full-length spike EF sequence.

contact@DNAScript.com
www.DNAScript.com

ML1135. Version 1.0.
November 2021.

US Headquarters
DNA Script, Inc.
650 Gateway Blvd.
South San Francisco, CA 94080
+1 (650) 457-0844

Europe Headquarters
DNA Script SAS
67 avenue de Fontainebleau
94270, Le Kremlin-Bicêtre, France
+33 1 56 20 56 00