

3.2.S.1.2 STRUCTURE

The molecular sequence of CX-024414, including the 5' cap, the 5' untranslated region (UTR), the Open Reading Frame (ORF), the 3' UTR, and the 3' polyA tail, is provided in [Figure 1](#). CX-024414 is the mRNA that encodes for the pre-fusion stabilized Spike protein of 2019-novel Coronavirus (SARS-CoV-2).

Figure 1: Molecular Sequence of CX-024414

```
5' 7MeGpppG2'omeGGAAUUAAGAGAGAAAAGAAGAGUAGAAGAAUUAAGACCCCGGCGCCGCCACCAUGUUCGUGUUCUGGUGUGUCGCCCCU  
GGUGAGCAGCCAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAGCUACACCCGGGGGUCUACUACCCGACAAGGU  
GUUCCGGAGCAGCGUCCUGCAGCACCAGGACCCUUCUUCAGCAACGUGACCUGGUUCCACGCAUCCACGUGAGCGGCACCAA  
CGGCACCAGCGGUUCGACCCAAUACAACCCUGGUGCGGGAUCUGCCCCAGGGCUUCUACCCAGCACCAGCAAGAGCAACAUCAUCCGGGCGUGAUCU  
CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUAUCGUGAAUAACGCCACCAACGUGGUGAUAAGGUGUGCGAGUUCAGUUCUGCAACGA  
CCCCUUCUGGGCGUGUACUACCAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCGGGUGUACAGCAGCGCCAACCAUCGACCCUUCGAGUA  
CGUGAGCCAGCCUUCUGAUGGACCCUGGAGGGCAAGCAGGGCAACUUAAGAACCUGCGGGAGUUCGUGUUAAGAACAUCGACGGCUACUUCAA  
GAUCUACAGCAAGCAACCCAAUACAACCCUGGUGCGGGAUCUGCCCCAGGGCUUCUACCCUUGGAGCCUUGGAGCCUUGGAGCCUUCGCGCAUCA  
CAUCACCCGGUUCAGACCCUGCUGGCCCCGACCCGGAGCUACCCAGCCCCAGGGCAGCAGCAGCAGCGGGUGGACAGCAGGCGCGGCGUCUUCUA  
CGUGGGCUACCGCAGCCCCGACCUUCUGCGUAGUACAACGAGAACGGCACCACUACCCAGCAGCGUGGACUGGCCUUGGACCCUUCGAGCGA  
GACCAAGUGCACCUCGAAGAGCUUACCCUGGAGAAGGGCAUCUACCCAGCAGCAGCAACUUCGGGUGCAGCCCCAGGAGCAUCGUGCGGUUCCC  
CAACAUCACCAACCCUUCGCGGCCUUCGCGAGGUGUUAACCCGAGCCUUCAGCGGUGUACGCCUUGGAACCCGGAAGCGGAUCAGCAACUCGCU  
GGCCGACUACAGCGUGCUUACAACAGCGCCAGCUACAGCACCUUCAAGUGCUACGGCGUGAGCCCCACCAAGCUGAAGCAGCCUUGUCUUCACCAA  
CGUGUACGCCGACAGCUUCGUGAUCGUGGGCAGCAGGUGCGGCAUCGACCCCGCCAGCAGGCAAGAUCCGCGACUACAACUACAAGCUGCC  
CGACGACUUCACCGGCGUGGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAGGUGGGGGCAACUACAACUACCCUGUACCGGCGUUCGCGAA  
GAGCAACCCUAGCCUUCGAGCGGGACAUACAGCACCAGAGUACUACCGGCGUCCACCCUUGCAACCCGCGUUGGAGGGGCGUUCAGUACU  
CCCUCUGCAGAGCUACGGCUUCAGCCCCAACGGCGUGGGUACCCAGCCUACCCGGGUGGUGUGCUGAGCUUCGAGCUGCAGCAGCCCCAGC  
CACCUGUGUGGCCCAAGAAGAGCACCACCCUGGUGAAGAACAAGUGCGUAAUUAACUUAACCGCCUUAACCGGCACCGGCGUGCUGACCGA  
GAGCAACAAGAAAUUCUGGCCUUCAGCAGUUCGGCCGGGACUUCGCCGACACCAGCAGCGUGGCGGGAUCCCCAGACCCUGGAGAUCUUGGA  
CAUCACCCUUCGAGCUUCGGCGCGUGAGCGUGAUCACCCAGGCACCAACAGCAAGACAGCGUGGCGGCUUACUACCCAGGACGUGAACUCGAC  
CGAGGUGCCCGUGGCAUCCACGCCGACAGCUGACACCCACCUGCGGGUUCUACAGCACCAGCAGCAACGUGUUCAGACCCGGGCGGUUGCCU  
GAUCGGCGCCGAGCAGUGAACAACAGCUACGAGUGCGACAUCCCAUCGGCGCCGCAUCUGUGCCAGCUACCCAGACCCAGACCAAUUCACCCCG  
GAGGGCAAGGAGCGUGGCCAGCCAGAGCAUACUAGCCUACACCAUAGGACUUGGGCGCCGAGAACAGCGUGGCUACAGCAACAACAGCAUCGCCAU  
CCCCAACCAUCACCAUACCGGUGACCACCGAGAUUUCGCGGCGGAGAGUACAGCAAGACAGCGUGGACCAUGUACAUUCGCGGCGACAG  
CACCGAUGCAGCAACCUGCUGCUGCAGUACGGCAGCUUCGACCCAGCUGAACCGGGCCUGACCCGGCAUCGCCUGGAGCAGGACAAGAACAC  
CCAGGAGGUGUUCGCCAGGUGAAGCAGAUUCAAGACCCUCCAUCAAGGACUUCGGGGCUUACAUCUAGCCAGAUUCUGCCGACCCAG  
CAAGCCAGCAAGCGGAGCUUACUAGGAGCCUGCUGUUAACAAGGUGACCCUAGCCGACCCGCGUUAUCAAGCAGUACGGCGACUGCCUCCG  
CGCAUAGCCCGCCGGGACCUAGCAGGCGCCAGAAUUAACGCGCUGACCGUGCUGCCUCCUUCGUGACCCGAGGAGACGAGACCAUCGAC  
CAGCGCCUUGUAGCCGGAACCAUACAGCGCGUGGACUUCGGCGCUGGAGCCGCUUCGAGAUCCCUUGCCAUAGCAGUAGGCUUACCGGU  
CAACGGCAUCGGCGUGACCCAGAACGUGCUGUACGAGAACAGAGCUGAUCGCCAACCAAGUUAACAGCGCCAUCGGCAAGAUCAGGACAGCCU  
GAGCAGCACCUCUAGCGCCUUGGGCAAGCUGCAGGACGUGGUGAACCAGAACGCCCAGGCCUGAACACCCUUGGUAAGCAGCUGAGCAGCAACU  
CGGGCCCAUCAGCAGCGUGGUAACGACAUCCUGAGCCGGCUGGACCUUCGAGGCGGAGGUGCAGAUACCCGGCGUACUACUGCCGGCGUGCA  
GAGCCUGCAGAACCAACACCGUACGUAUCGCCGCGGAGCAGUUAACCGGCGCAGCAACCGCCUUGGAGGAGCUGGACAAGUACUUAAGAUAACAGC  
GGCCAGAGCAAGCGGUGGACUUCUGCGGCAAGGGCUACCCAGCUGAUGAGCUUCCCCAGAGCGCACCCACGGAGUGGUGUUCUGCAGCUGAC  
CUACGUGCCCGCCAGGAGAAAGAAUUAACACCCAGCCAGCUCUGCCAGCAGCGCAAGGCCCAUUCUCCCGGAGGGGCGUGUUCGUGAGCAA  
CGGCACCCACUGGUCUGACCCAGCGGAACUUCUACGAGCCCCAGAUCAUACCCAGCACAACACCUUCGUGAGCGGCAACUCGACGUGGUGAU  
CGGCAUCGUGAACAACACCGUGUACGUAUCUCCUGCAGCCGAGCUGGAGCAGCUUACAGGAGGAGCUGGACAAGUACUUAAGAUAACAGCAGC  
CGACGUGGACCCUGGGCAGCAUCAGCGGAUCAACGCCAGCGUGGUGAACAUCAGAAAGGAGAUCAUCGGCUGAACGAGGUGGCCAAGAACCUGAA  
CGAGAGCCUGAUCGACCGCAGGAGCUGGGCAAGUACGAGCAGUACAUCAAGUGGCCUUGUACUUCGCGGCGCCUUCGCGGCCUAGCUGC  
CAUCGUGAUGGUGACCAUCAUGCUGUGCUGCAUGACCCAGCUGCAGCUGCCUGAAGGGCUGUUGCAGCUGCGGCGAGCUGCUGCAAGUUCGACGA  
GGACGACAGCCGAGCCGUGUGAAGGGCGUGAAGCUCACUACCCUGAUAAUAGGCGUGGAGCCUUGGUGGCCUAGCUCUUCGCCCCUUGGGCCUC  
CCCCAGCCCCUUCUCCUUCGACCCGUACCCCGGUGUCUUAUAAUAAAGUCUGAGUGGGCGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAUCUAGoh3'
```

Where: A, C, G and U = AMP, CMP, GMP and N1-Me-ΨMP, respectively;
Me = methyl; p = inorganic phosphate.

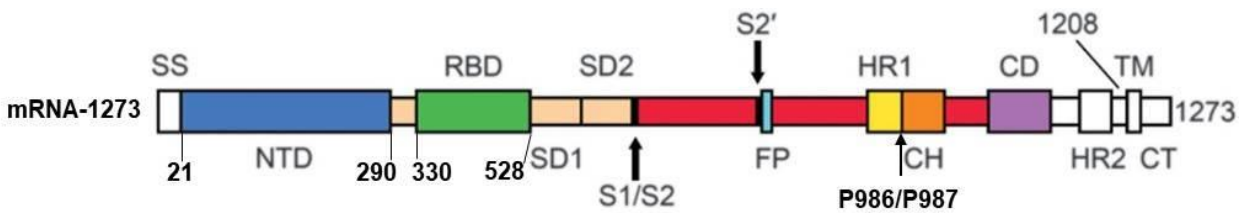
mRNA Sequence Molecular Weight (free acid)	1,329,683 Daltons	
mRNA Sequence Molecule Length	4,101 nucleotides	
mRNA Sequence Elements	ORF:	3819 nucleotides
	5'UTR:	Cap + 57 nucleotides
	3'UTR:	119 nucleotides
	PolyA tail:	105 nucleotides

The protein sequence encoded by CX-024414 was designed by the Applicant and is consistent with a Coronavirus spike protein. The Open Reading Frame (ORF design) is consistent with all current guidelines as provided by WHO on based on the genomic sequence of the SARS-CoV-2 that was made public by China on January 11, 2020. The UTRs are consistent with current recommendations within the Applicant’s proprietary mRNA platform.

mRNA platform and S1S2 protein sequence.

The single mRNA that encodes the full-length SARS-CoV-2 spike (S) protein modified with 2 proline substitutions within the heptad repeat 1 domain (S 2P) to stabilize the S protein into the prefusion conformation. The S protein is stabilized in the so-called pre-fusion conformation by two amino acid mutations, K986P and V987P.

Figure 2: Protein Sequence of CX-024414



3.2.S.1.2.1 Open Reading Frame (ORF) and Incorporation of Pseudouridine

Table 1 presents in more detail each part of CX-024414. The codons used to encode the engineered prefusion S1/S2 protein were generated using Moderna’s proprietary software, which makes the codon usage utilized unique (See Figure 3).

Table 1: Table of Features

Element	Description	Position
Cap	5'-cap1 structure (m7G-5'-ppp-5'-Gm) (additional information in Figure 4)	1 - 2
5' UTR	The 5'UTR sequence in mRNA-1273 is based on the 5' UTR from Warren et al., 2010 that confers robust protein expression. It has been further optimized with a GC-rich sequence near the start codon to increase the fidelity of translation initiation at the designated start codon.	3 - 58
ORF	Codon-optimized sequence encoding full-length SARS-CoV-2 spike (S) glycoprotein containing mutations K986P and V987P (nucleosides 2956-2961) to stabilize pre-fusion conformation; stop codons: 3888-3896 (underlined)	59 - 3880
3' UTR	The 3'UTR is derived from the human HBA1 gene, which has been associated with prolonged mRNA half-life (Waggoner 2003). It has been further optimized by additional of two further stop codons to ensure complete translation termination and removal of an AUG in the 3'UTR to eliminate the risk of any aberrant translation.	3881 - 3996
polyA tail	A 100-nucleotide poly(A)-tail followed by a 5-nucleotide XbaI scar.	3997 - 4101

Figure 3: Schematic Map - Detailed Graphic of the Composition of the Full mRNA Sequence



Cap 5' UTR ORF 3' UTR polyA tail
Nucleotides 2956-2961 in the ORF are corresponding to 2 Proline mutations in prefusion stabilized construct.
All uridines in the mRNA are replaced by 1-methylpseudouridine.

Nucleotides 2956-2961 in the ORF are corresponding to 2 Proline mutations in prefusion stabilized construct. All uridines in the mRNA are replaced by 1-methylpseudouridine. N1-methyl-pseudouridine modified mRNA has been incorporated into mRNA-1273, as it has been in previous Moderna vaccines tested in clinical trials. Recent reports have demonstrated that use of N1-methyl-pseudouridine modified mRNA results in robust protein translation (Corbett et al., 2020a; Corbett et al., 2020b; Jackson et al., 2020; John et al., 2018). mRNA with this nucleoside-modification was found to circumvent TLR7/8 activation, thus decreasing innate immune activation (Espeseth et al., 2020; Kariko et al., 2005), while still eliciting strong immune responses both preclinically and clinically (Corbett et al., 2020a; Corbett et al., 2020b; Espeseth et al., 2020; Jackson et al., 2020; John et al., 2018)

Figure 4: Cap Structure of CX-024414

