

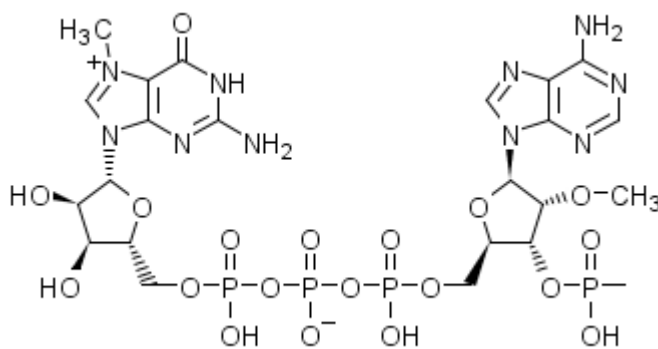


11868

Description

Messenger RNA encoding full-length SARS-Cov-2 spike glycoprotein.

Schematic



5'- capping structure

cap G¹A² = m⁷G⁺-5'-ppp-5'-Am

Table of features

Element	Description	Position
Cap1	A modified 5'-cap1 structure (m ⁷ G ⁺ -5'-ppp-5'-Am)	1-2
5'-UTR	5'-untranslated region (artificial nucleotides)	3-16
S glycoprotein	Coding region for the SARS-CoV-2 spike (S) glycoprotein Position of start codon (17-19) Position of stop codon (3836-3838)	17-3838



3'-UTR	untranslated region (3' UTR) comprising the following element:	3839-3912
	Human alpha-globin 3' UTR sequence element	3861-3904
PolyA	Poly (A) sequence	3913-3976
3' Tail	poly (C)-rich sequence, followed by histone stem Loop sequence and nucleotides from restriction enzyme recognition sites	3977-4040

Sequence / Séquence / Secuencia

gaaggagaaag cuuaccaugu ucguguuccu gguccuccug cccugguga 50
gcucccagug cgugaaccuc accaccgca cgcagcugcc gcccgccuac 100
accaacagcu ucacccgggg cgucuacuac cccgacaagg uguuccgcuc 150
cagcgugcug cacuccacc aggaccucu ccugcccuuc uucagcaacg 200
ucacgugguu ccacgccauc cagcuguccg ggaccaacgg caccaagcgg 250
uucgacaacc cggugcugcc cuucaacgac ggggucuacu ucgcgagcac 300
cgagaagucc acaucaucc gcggcuggau cuucgggacg acccucgaca 350
gcaagaccca gucccugcug aucgugaaca acgccaccaa cgugguauca 400
aaggugugcg aguuccaguu cugcaacgac cccuuccucg gcguguacua 450
ccacaagaac aacaagagcu ggauggaguc cgaguuccgg gucuacagcu 500
ccgccaacaa cugcacguuc gaguacguga gccagcccuu ccugauggac 550
cuggagggca agcaggggaa cuucaagaac cuccgcgagu ucguguuaa 600
gaacaucgac ggcuacuua agaucuacuc caagcacacc ccgaucaacc 650
ugguccggga ccugcccag gguucagcg cccucgagcc ccugguggac 700
cugcccucg gcaucaacau caccgcuuc cagaccuccc uggcccugca 750
ccgguccuac cucacgccg gggacagcag cuccggcugg accgcgggcg 800
ccgcccua cuacguggg uaccugcagc cccgcaccuu ccugcucaag 850
uacaacgaga acggcaccu caggacgcg gucgcugcg cccuggacc 900
ccugagcgag accaagugca ccucaaguc cuuaccgug gagaagggga 950
ucuaccgac gagcaacuuc cgggugcagc ccaccgaguc caucguccgc 1000
uuccgaaca ucaccaaccu gugcccuuc ggcgaggugu ucaacgccac 1050
ccgguucgac agcguacuac ccuggaacg caagcggaucc uccaacugcg 1100
uggcggacua cagcgugcug uacaacuccg ccagcuucuc cacguucaag 1150
ugcuacggcg ucagccccac caagcucaac gaccugugcu ucaccaacgu 1200
guacgccgac uccuucguga uccgcgggga cgagguccgg cagaucgcc 1250
ccggccagac cgggaagauc gccgacuaca acuacaagcu gccggacgac 1300
uucacgggcu gcgugaucgc guggaacagc aacaaccucg acuccaaggu 1350
cgggggcaac uacaacuacc uguaccgcu guuccgcaag agcaaccuca 1400
agcccuucga gcgggacauc uccaccgaga ucuaccaggc cggcagcacc 1450
cccugcaacg ggguggagg cuucaacugc uacuuccccc ugcaguccua 1500



cgguuccag	ccgaccaacg	gcguggggu	ccagcccuac	cgguuggucg	1550
ugcugagcuu	cgagcuccug	cacgccccg	ccacggugug	cgcccccaag	1600
aaguccacca	accuggucaa	gaacaagugc	gugaacuuca	acuucaacgg	1650
ccucaccggg	accggcgugc	ugaccgagag	caacaagaag	uuccugccgu	1700
uccagcaguu	cgggcgcgac	aucgccgaca	cgaccgacgc	gguccgcgac	1750
cccagaccc	uggagaucuu	cgacaucacc	cccugcagcu	ucggcggggu	1800
guccgugauc	acgccgggca	ccaacaccag	caaccagguc	gccgugcugu	1850
accaggacgu	gaacugcacc	gagguccccg	uggccaacca	cgccgaccag	1900
cugacgccca	ccuggcgggu	guacuccacc	ggcagcaacg	ucuuccagac	1950
ccgcccggg	ugccucaucg	gcgcgagca	cgugaacaac	uccuacgagu	2000
gcgacaucce	caucggggcc	ggcaucugcg	ccagcuacca	gacgcagacc	2050
aacuccccgc	ggcgcgcccg	gagcguggcg	ucccagagca	ucaucgccua	2100
caccaugucc	cugggcgccg	agaacagcgu	cgccuacucc	aacaacagca	2150
ucgccaucce	caccaacuuc	acgaucuccg	ugaccaccga	gauccugccc	2200
gugagcauga	ccaagacgag	cgucgacugc	accauguaca	ucugcgggga	2250
cuccaccgag	ugcagcaacc	uccugcugca	guacggcucc	uucugcacc	2300
agcucaaccg	cgcgugacg	gggaucgccc	uggagcagga	caagaacacc	2350
caggaggugu	ucgcccaggu	caagcagauc	uacaagacgc	ccccgaucaa	2400
ggacuucggc	ggguucaacu	ucagccagau	ccugcccgac	ccuccaagc	2450
ccagcaagcg	guccuaucau	gaggaccucc	uguucaacaa	ggugaccucg	2500
gccgacgccg	gcuucaucaa	gcaguacggc	gacugccucg	gggacaucgc	2550
ggcccgcgac	cugaucugcg	cccagaaguu	caacggccug	accgugcucc	2600
cgccccugcu	gaccgacgag	augaucgccc	aguacaccag	cgccccucg	2650
gcggggacca	ucacguccgg	cuggaccuuc	ggggccggcg	ccgcccugca	2700
gaucuccuuc	gcgaugcaga	uggccuaccg	guucaacggc	aucgggguca	2750
cccagaacgu	gcucuacgag	aaccagaagc	ugaucgcaa	ccaguucaac	2800
agcgccaucg	gcaagaacca	ggacucccug	agcuccaccg	ccagcgcgcu	2850
cgggaagcug	caggacgugg	ucaaccagaa	cgcccaggcc	cugaacacgc	2900
ucgugaagca	gcuguccagc	aacuucggcg	ccaucuccag	cgugcugaac	2950
gacaucucuc	cccgccugga	cccgccgag	gccgaggugc	agaucgaccg	3000
gcugaucacc	ggcgccucc	agagccugca	gaccuacgug	accagcagc	3050
ugaucggggc	ggccgagauc	cgcgccagcg	ccaaccucgc	cgcgaccaag	3100
auguccgagu	gcguccuggg	ccagagcaag	cgguuggacu	ucugcggcaa	3150
ggguaccac	cugauguccu	ucccccagag	cgccccgcac	ggcguggucu	3200
uccuccacgu	gacguacguc	cccgccagag	agaagaacuu	caccaccgcc	3250
cccgcgaucu	gccacgacgg	gaaggcccac	uucccccgcg	agggcguguu	3300
cguguccaac	gggaccacu	gguucgucac	gcagcggaac	uucuacgagc	3350
cgcagaucuu	caccaccgac	aacaccuucg	ugagcggcaa	cugcgacgug	3400
gucaucggca	ucgugaacaa	cacgguguac	gacccccugc	agcccagcgu	3450
ggacuuccuuc	aaggaggagc	ucgacaagua	cuucaagaac	cacaccagcc	3500
ccgacgucga	ccugggggac	aucuccggca	ucaacgccag	cguggugaac	3550
auccagaagg	agaucgaccg	ccugaacgag	gucgccaaga	accucaacga	3600
guccugauc	gaccugcagg	agcucgggaa	guacgagcag	uacaucaagu	3650
ggccguggua	caucuggcug	ggcuucaucg	ccggccugau	cgcgaucgug	3700
auggugacca	ucaugcucug	cugcaugacc	agcugcugcu	ccugccugaa	3750



WHO
International Nonproprietary Names Programme

9/2020

ggggugcugc	agcugcggcu	ccugcugcaa	guucgacgag	gacgacagcg	3800
agcccguccu	gaagggggug	aagcuccacu	acacgugagg	acuaguuaua	3850
agacugacua	gcccgauagg	ccucccaacg	ggcccuccuc	cccuccuugc	3900
accgagauua	auaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	3950
aaaaaaaaaa	aaaaaaaaaa	aaaaaaugca	uccccccccc	cccccccccc	4000
cccccccccc	ccaaaggcuc	uuucagagc	caccagaauu		4040