

Supplementary data. Sequence alignments used for homology modelling. Conserved residues are labelled in dark green, same residues in lighter green and residues with similar properties in yellow.

SARS-CoV-2 S protein – close state

Sequence identity: 66%

SARS_CoV_2	14	QCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLFFFSNVTWFAIHVSGTNGTKRFDNPFVLPFDNGVYFASTEKSNIIIRGWIIFGTTLDSKTQSLLIVNNATNV
6VXX_A	27	-----AYTNSFTRGVYYPDKVFRSSVLHSTQDLFLFFFSNVTWFAIH-----DNFVLPFDNGVYFASTEKSNIIIRGWIIFGTTLDSKTQSLLIVNNATNV
5X58_B	18	RCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTDQLFLFFFSNVTGFTIN-----HTFDNPFVLPFDNGVYFASTEKSNIIIRGWIIFGTTLDSKTQSLLIVNNATNV
SARS_CoV_2	127	VIRKCEQFQNDPFLGVYHKNKSWMESEFRVYSSANNCTFEVVSQPFIMDLEGKQGNFKNLREFVFNKIDGYFKIYSKHTPINLVRDLPOGFSALEELVDLEIGINITREFQT
6VXX_A	127	VIRKCEQFQNDPFLGV-----NCTFEVVS-----FKNLREFVFNKIDGYFKIYSKHTPINLVRDLPOGFSALEELVDLEIGINITREFQT
5X58_B	124	VIRACNFELEDNPFPAV----SKPMGTQTHMTIFDNAFNCTFEYISDAFSLDVSEKSGNFKNLREFVFNKIDGYFKIYSKHTPINLVRDLPOGFSALEELVDLEIGINITREFQT
SARS_CoV_2	241	LLALHRSYLTPGDSSSGWTAGAAAYVGYLQRTTELLKYNENGITITDAVDCALDELSETKCTLKSFVEKGIYQTSNFRVQPTESIVRFPNITNLCPPGEVFNATREFASVYANN
6VXX_A	241	LLALH-----AAAYVGYLQRTTELLKYNENGITITDAVDCALDELSETKCTLKSFVEKGIYQTSNFRVQPTESIVRFPNITNLCPPGEVFNATREFASVYANN
5X58_B	234	ILTAFS-----TWGTSAAAYFVGYLKHITTEMLKYDENGITITDAVDCSQNLEALKCSVKSEFIDKGIYQTSNFRVQPSGDDVRFNITNLCPPGEVFNATREFASVYANN
SARS_CoV_2	355	RKRISNCVADYSVLNNSASFSTFKCYGVSPTKLNLCFTNVYADSFVIRGDEVQIAPGQTKGIADYNKLPDDFTCCVIAWNSNNLDSKVGCGNYNYLRLFRKSNLKPFRERDI
6VXX_A	355	RKRISNCVADYSVLNNSASFSTFKCYGVSPTKLNLCFTNVYADSFVIRGDEVQIAPGQTKGIADYNKLPDDFTCCVIAWNSNNLDSK--GNYNYLRL-----KPFERDI
5X58_B	342	RKRISNCVADYSVLNNSASFSTFKCYGVSATKLNLCFTSHVYADSFVVGKDDVRQIAPGQTKGIADYNKLPDDFTCCVIAWNRNIDATSTGNYNYLRLVLRHGKLRKPFERDI
SARS_CoV_2	469	STEIYQAGSTPCNGVEGFNCFYFLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNGLTGTGVLTESNKKELPFQGFGRDIADTTDAVRDFQTL
6VXX_A	489	-----YFELQSYGFQPTN--VGYPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNGLTGTGVLTESNKKELPFQGFGRDIADTTDAVRDFQTL
5X58_B	456	SNVPFSPDGKPCPT--PALNCFWFLNDYGFYTTTGIGYQPYRVVLSFELLHAPATVCGPKKSTDLIKNQCVNFNGLTGTGVLTPSSKRFQPFQGFGRDVSDFTDSVRDEKTS
SARS_CoV_2	583	EILDITPCSGGGVSITPGTNTSNQAVLYQDVNCTEVPVAIHADQLTPTRVYSTGNSVEQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTSNPRRARSVASQSIIAYT
6VXX_A	583	EILDITPCSGGGVSITPGTNTSNQAVLYQDVNCTEV-----NVEQTRAGCLIGAEHVNNSYECDIPIGAGICASYQT-----SQSIIAYT
5X58_B	569	EILDITPCSGGGVSITPGTNTASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTRAGCLIGAEHVDTSYECDIPIGAGICASYHT-----KSTIAYT
SARS_CoV_2	697	MSLGAENSVAYSNNIAIPTNTTISVTTEILFVSMTKTSVDITMYICGDSTECNLLLOYGSGFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDS
6VXX_A	697	MSLGAENSVAYSNNIAIPTNTTISVTTEILFVSMTKTSVDITMYICGDSTECNLLLOYGSGFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDS
5X58_B	679	MSLGADSSIAYSNNTIAIPTNTTISITTEVMEFVSMAKTSVDITNMYICGDSTECANLLLOYGSGFCTQLNRALSGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDEL
SARS_CoV_2	811	KPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLCAQKFNGLTVLPPLLTDemiaQYTSALLAGTITSGWTFGAGAAIQIPFAMQMAVRFNGIGVTQNVLYENQKIIA
6VXX_A	811	KPSKRSFIEDLLFNKVT-----KFNGLTVLPPLLTDemiaQYTSALLAGTITSGWTFGAGAAIQIPFAMQMAVRFNGIGVTQNVLYENQKIIA
5X58_B	793	KPTKRSFIEDLLFNKVT-----DLCAQKFNGLTVLPPLLTDemiaAAYTAALVSGTATAGWTFGAGAAIQIPFAMQMAVRFNGIGVTQNVLYENQKIIA
SARS_CoV_2	925	NQFNSAIGKIQDSLSTASALGRLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSK
6VXX_A	925	NQFNSAIGKIQDSLSTASALGRLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSK
5X58_B	907	NQFNKAISIQIESLITTTSTALGRLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSK
SARS_CoV_2	1039	RVDVFCCKGYHLSFPQSAPHGVVFLHVTYVPAQEKNFETAPAIKCHDGAHFPREGVVFVNGTHWFVTQRNFYEPOIITTDNTFVSGNCDVIGIVNNTVYDPLQPELDS
6VXX_A	1039	RVDVFCCKGYHLSFPQSAPHGVVFLHVTYVPAQEKNFETAPAIKCHDGAHFPREGVVFVNGTHWFVTQRNFYEPOIITTDNTFVSGNCDVIGIVNNTVYDPLQPELDS
5X58_B	1021	RVDVFCCKGYHLSFPQAAPHGVVFLHVTYVPSQERNFETAPAIKCHGKAYFPREGVVFVNGTSWFITQRNFYSPOIITTDNTFV-----

SARS-CoV-2 S protein – open state

Sequence identity: 10%

In this case, the sequence identity is very low because the template structures used to model SARS-CoV-2 S protein in the open conformation with PDB ID 6LZG and 6M0J are the structures of the S protein RBD domain. Therefore, when the sequences of these structures are aligned with the sequences of the template structures with PDB ID 6VSB, 6VYB and 5X5B – which are the structures of an entire S protein monomer – the only region of the alignment in which the sequences of all the template structures align is in the section of the sequences corresponding the RBD.

```

SARS_CoV_2  27  AYTNSFTAGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWTHAIHVSCTGCTGKRFNPVLPFNDGVYFASTSEKSNIRGWIFGTTLDSKTQSLILVNATNVVVKVCFQPCNDPF
6VSB_A      27  AYTNSFTAGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWTH-----FDPVLPFNDGVYFAST--NIIRGWIFGTTLDSKTQSLILVNATNVVVKVCFQPCNDPF
6VYB_B      27  AYTNSFTAGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWTH-----NPVLPFNDGVYFASTSEKSNIRGWIFGTTLDSKTQSLILVNATNVVVKVCFQPCNDPF
6LZG_B      333 -----
6M0J_E      333 -----
5X5B_A      15  QHTSSM-RGVYYPDEIFRSDTLTLTQDLFLPFFSNVTGFTIN-----HTFNPVLPFNDGVYFASTSEKSNVVRGWFVSGTMMNKSQSVIIINNSTNVVIRACNELCDNPF

SARS_CoV_2  141 LGVYYHKNKSNWMESEFRVYSSANNCTFEYVSQPFLLMDLEGKQGNFKNREFVFNIDGYFKIYSKHTPINLVRDLPGQGSFALEPLVDLPIGINITRFQTLALHRSYLTGPDSS
6VSB_A      141 LG-----EFRVYSSANNCTFEYVSQPFLL-----KNLREFVFNIDGYFKIYSKHTPINLVRDLPGQGSFALEPLVDLPIGINITRFQTLALH-----
6VYB_B      164 -----NNCTFEYVS-F-----KNLREFVFNIDGYFKIYSKHTPIN--DLPQGSFALEPLVDLPIGINITRFQTLALH-----
6LZG_B      333 -----
6M0J_E      333 -----
5X5B_A      121 FAV-----SKPMGTQTHMTFDNAFNCTFEYISDAFSLDVSERKSGNFHRLREFVFNIDGYFKIYSKHTPINLVRDLPGQGSFALEPLVDLPIGINITRFQTLALHRSYLTGPDSS

SARS_CoV_2  256 SGWTAGAAAYVGYLQPRTEFLKYENGTITDAVDCALDPLSETKCTLSKSTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGGEVNATREFASVIAWNRKRISNCVADYSVLN
6VSB_A      246 ----RGAAYVGYLQPRTEFLKYENGTITDAVDCALDPLSETKCTLSKSTVEKGIYQTSNFRVQPTESIVR----LCPPFGGEVNATREFASVIAWNRKRISNCVADYSVLN
6VYB_B      263 ----AAYVGYLQPRTEFLKYENGTITDAVDCALDPLSETKCTLSKSTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGGEVNATREFASVIAWNRKRISNCVADYSVLN
6LZG_B      333 -----TNLCPFGGEVNATREFASVIAWNRKRISNCVADYSVLN
6M0J_E      333 -----TNLCPFGGEVNATREFASVIAWNRKRISNCVADYSVLN
5X5B_A      223 -TWGTSAAYFVGYLKPTEFLKYENGTITDAVDCSQNLAEKLCVSKVSEIDKGIYQTSNFRVQPTESIVRFPNITNLCPFGGEVNATREFASVIAWNRKRISNCVADYSVLN

SARS_CoV_2  371 SASFSTFKCYGVSPTKLNDLCFTNVYADSVIRGDEVQRQIAPCGQTGKIADNYKLPDDFTGCVIAWNSNNLDSKVGGNINYLRLFRKSNLKPFPERDISTEIQAGSTPCNGVEG
6VSB_A      371 SASFSTFKCYGVSPTKLNDLCFTNVYADSVIRGDEVQRQIAPCGQTGKIADNYKLPDDFTGCVIAWNSNNLDSKVGGNINYLRLFRKSNLKPFPERDISTEIQAGSTPCNGVEG
6VYB_B      371 SASFSTFKCYGVSPTKLNDLCFTNVYADSVIRGDEVQRQIAPCGQTGKIADNYKLPDDFTGCVIAWNSNNLDSKVGGNINYLRLFRKSNLKPFPERDISTEIQAGSTPCNGVEG
6LZG_B      371 SASFSTFKCYGVSPTKLNDLCFTNVYADSVIRGDEVQRQIAPCGQTGKIADNYKLPDDFTGCVIAWNSNNLDSKVGGNINYLRLFRKSNLKPFPERDISTEIQAGSTPCNGVEG
6M0J_E      371 SASFSTFKCYGVSPTKLNDLCFTNVYADSVIRGDEVQRQIAPCGQTGKIADNYKLPDDFTGCVIAWNSNNLDSKVGGNINYLRLFRKSNLKPFPERDISTEIQAGSTPCNGVEG
5X5B_A      331 STPESTFKCYGVSPTKLNDLCFTNVYADSVIRGDEVQRQIAPCGQTGKIADNYKLPDDFTGCVIAWNSNNLDSKVGGNINYLRLFRKSNLKPFPERDISTEIQAGSTPCNGVEG

SARS_CoV_2  486 FNCYFPLQSYGQPTNGVGYQYPRVVVLSPELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPTQLEILDITPCSFGGVSVITP
6VSB_A      491 ----PLQSYGQPTNGVGYQYPRVVVLSPELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPTQLEILDITPCSFGGVSVITP
6VYB_B      491 ----PLQSYGQPTNGVGYQYPRVVVLSPELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPTQLEILDITPCSFGGVSVITP
6LZG_B      486 FNCYFPLQSYGQPTNGVGYQYPRVVVLSPELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPTQLEILDITPCSFGGVSVITP
6M0J_E      486 FNCYFPLQSYGQPTNGVGYQYPRVVVLSPELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPTQLEILDITPCSFGGVSVITP
5X5B_A      445 LNCYFPLQSYGQPTNGVGYQYPRVVVLSPELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTPSKRFPQFQFGRDVSDFTDSVRDPTKSEILDITPCSFGGVSVITP

SARS_CoV_2  601 GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTRVYSTGSNVFQTRAGCLIGAEHVNNSEYCDIPIGAGICASYQTQTSNPRRARSVAQSISIATYMSLGAENSVAYSNNSTAIPT
6VSB_A      601 GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTRVYSTGSNVFQTRAGCLIGAEHVNNSEYCDIPIGAGICASYQTQTSNPRRARSVAQSISIATYMSLGAENSVAYSNNSTAIPT
6VYB_B      601 GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTRVYSTGSNVFQTRAGCLIGAEHVNNSEYCDIPIGAGICASYQTQTSNPRRARSVAQSISIATYMSLGAENSVAYSNNSTAIPT
6LZG_B      545 GTNASEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNVFPQTAGCLIGAEHVNTSEYCDIPIGAGICASYH-----TKSIVAYTMSLGAENSIAYSNNTIATP
6M0J_E      545 GTNASEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNVFPQTAGCLIGAEHVNTSEYCDIPIGAGICASYH-----TKSIVAYTMSLGAENSIAYSNNTIATP
5X5B_A      545 GTNASEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNVFPQTAGCLIGAEHVNTSEYCDIPIGAGICASYH-----TKSIVAYTMSLGAENSIAYSNNTIATP

SARS_CoV_2  716 TNFTISVTTEILPVSMKTSTSDCTMYICGDSSTECNSLLQYGSFCTQLNRLATGIAVEQDKNTQEVFAQVKQIYKTPPIKDPGGFNFSGILPDPSKRSFYEDLLFNKVTLD
6VSB_A      716 TNFTISVTTEILPVSMKTSTSDCTMYICGDSSTECNSLLQYGSFCTQLNRLATGIAVEQDKNTQEVFAQVKQIYKTPPIKDPGGFNFSGILPDPSKRSFYEDLLFNKVTLD
6VYB_B      716 TNFTISVTTEILPVSMKTSTSDCTMYICGDSSTECNSLLQYGSFCTQLNRLATGIAVEQDKNTQEVFAQVKQIYKTPPIKDPGGFNFSGILPDPSKRSFYEDLLFNKVTLD
6LZG_B      716 TNFTISVTTEILPVSMKTSTSDCTMYICGDSSTECNSLLQYGSFCTQLNRLATGIAVEQDKNTQEVFAQVKQIYKTPPIKDPGGFNFSGILPDPSKRSFYEDLLFNKVTLD
6M0J_E      716 TNFTISVTTEILPVSMKTSTSDCTMYICGDSSTECNSLLQYGSFCTQLNRLATGIAVEQDKNTQEVFAQVKQIYKTPPIKDPGGFNFSGILPDPSKRSFYEDLLFNKVTLD
5X5B_A      647 TNFTISVTTEILPVSMKTSTSDCTMYICGDSSTECNSLLQYGSFCTQLNRLATGIAVEQDKNTQEVFAQVKQIYKTPPIKDPGGFNFSGILPDPSKRSFYEDLLFNKVTLD

SARS_CoV_2  831 AGFIKQYQDCLGDIARDLCAQKFNGLTVLPPLLTDEMAIAQYTSALLAGTITSGWTFGAGAAQIPFAMQAYRFGNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASAL
6VSB_A      853 -----QKFNGLTVLPPLLTDEMAIAQYTSALLAGTITSGWTFGAGAAQIPFAMQAYRFGNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASAL
6VYB_B      854 -----KFNGLTVPPLLTDEMAIAQYTSALLAGTITSGWTFGAGAAQIPFAMQAYRFGNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASAL
6LZG_B      854 -----
6M0J_E      854 -----
5X5B_A      759 -----DLCAQKFNGLTVLPPLLTDEMAIAQYTSALLAGTITSGWTFGAGAAQIPFAMQAYRFGNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASAL

SARS_CoV_2  946 GKLDQDVVNQAQALNTLVKQLSSNFGAISSVNLNLSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSCEVLGQSKRVDPCGKGHYLMSFFQSAAPHGV
6VSB_A      946 GKLDQDVVNQAQALNTLVKQLSSNFGAISSVNLNLSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSCEVLGQSKRVDPCGKGHYLMSFFQSAAPHGV
6VYB_B      946 GKLDQDVVNQAQALNTLVKQLSSNFGAISSVNLNLSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSCEVLGQSKRVDPCGKGHYLMSFFQSAAPHGV
6LZG_B      946 GKLDQDVVNQAQALNTLVKQLSSNFGAISSVNLNLSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSCEVLGQSKRVDPCGKGHYLMSFFQSAAPHGV
6M0J_E      946 GKLDQDVVNQAQALNTLVKQLSSNFGAISSVNLNLSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSCEVLGQSKRVDPCGKGHYLMSFFQSAAPHGV
5X5B_A      857 GKLDQDVVNQAQALNTLVKQLSSNFGAISSVNLNLSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSCEVLGQSKRVDPCGKGHYLMSFFQSAAPHGV

SARS_CoV_2  1061 VFLHVTYVPAQEKNTTAPAICHGDKAHFPREGVFPVNGTHWFVTQRFNIEPQIITDNTFVSGNCDVIGIVNNTVYDPLQPELDS
6VSB_A      1061 VFLHVTYVPAQEKNTTAPAICHGDKAHFPREGVFPVNGTHWFVTQRFNIEPQIITDNTFVSGNCDVIGIVNNTVYDPLQPELDS
6VYB_B      1061 VFLHVTYVPAQEKNTTAPAICHGDKAHFPREGVFPVNGTHWFVTQRFNIEPQIITDNTFVSGNCDVIGIVNNTVYDPLQPELDS
6LZG_B      1061 VFLHVTYVPAQEKNTTAPAICHGDKAHFPREGVFPVNGTHWFVTQRFNIEPQIITDNTFVSGNCDVIGIVNNTVYDPLQPELDS
6M0J_E      1061 VFLHVTYVPAQEKNTTAPAICHGDKAHFPREGVFPVNGTHWFVTQRFNIEPQIITDNTFVSGNCDVIGIVNNTVYDPLQPELDS
5X5B_A      972 VFLHVTYVPAQEKNTTAPAICHGDKAHFPREGVFPVNGTHWFVTQRFNIEPQIITDNTFVSGNCDVIGIVNNTVYDPLQPELDS

```

SARS-CoV S protein – close state

Sequence identity: 100%

100% pP=92.5 SARS_CoV_1 5X5B_A	18 18	RCTTFDDVQAPNYTQHTSSMRGVYYPDRIFRSDTLYLTQDILFLPFYSNVVTGFHTINHTF.NPVTIPKDGITYFAATEKSNVVRGVFGSTMMNKSQSIVIIINNSTNVVI RCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDILFLPFYSNVVTGFHTINHTF.NPVTIPKDGITYFAATEKSNVVRGVFGSTMMNKSQSIVIIINNSTNVVI RCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDILFLPFYSNVVTGFHTINHTF.NPVTIPKDGITYFAATEKSNVVRGVFGSTMMNKSQSIVIIINNSTNVVI
SARS_CoV_1 5X5B_A	126 126	RACNFELCDNPFPAVSKPMGTQTHMTIFDPAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFNKNDGFLVYVKGYQPIDVVRDLPSGFNTLKPFLKPLGINITNFRAI RACNFELCDNPFPAVSKPMGTQTHMTIFDPAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFNKNDGFLVYVKGYQPIDVVRDLPSGFNTLKPFLKPLGINITNFRAI RACNFELCDNPFPAVSKPMGTQTHMTIFDPAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFNKNDGFLVYVKGYQPIDVVRDLPSGFNTLKPFLKPLGINITNFRAI
SARS_CoV_1 5X5B_A	235 235	LTAFS.....WGTSAAAYFVGYLKPTTFMLKYDENGITTDVAVDCSQNPLAELKCSVKSEIDKGIYQTSNFRVVPVSGDVVR.....LCPFGVEFNATKFPVSVAWERK LTAFS.....WGTSAAAYFVGYLKPTTFMLKYDENGITTDVAVDCSQNPLAELKCSVKSEIDKGIYQTSNFRVVPVSGDVVR.....LCPFGVEFNATKFPVSVAWERK LTAFS.....WGTSAAAYFVGYLKPTTFMLKYDENGITTDVAVDCSQNPLAELKCSVKSEIDKGIYQTSNFRVVPVSGDVVR.....LCPFGVEFNATKFPVSVAWERK
SARS_CoV_1 5X5B_A	344 344	KISNCVADYSVLNSTFEFTKCYGVSATKLNLDLCSNVYADSEVVKGDVVRQIAPGQTGVADIYNYKLDDFMCGVLAWNRNIDATSTGNNYKRYRLRHKLRPFE KISNCVADYSVLNSTFEFTKCYGVSATKLNLDLCSNVYADSEVVKGDVVRQIAPGQTGVADIYNYKLDDFMCGVLAWNRNIDATSTGNNYKRYRLRHKLRPFE KISNCVADYSVLNSTFEFTKCYGVSATKLNLDLCSNVYADSEVVKGDVVRQIAPGQTGVADIYNYKLDDFMCGVLAWNRNIDATSTGNNYKRYRLRHKLRPFE
SARS_CoV_1 5X5B_A	453 453	RDISNVFSPDGKPCPTPALNCYWPLNDYGFYTTTGIGYQPYRVVLSFELINAPATVCGPKLSTDLIKNOCVNFNENGLTGTGVLTPSSKRFPQFPQGRDVSDFDTS RDISNVFSPDGKPCPTPALNCYWPLNDYGFYTTTGIGYQPYRVVLSFELINAPATVCGPKLSTDLIKNOCVNFNENGLTGTGVLTPSSKRFPQFPQGRDVSDFDTS RDISNVFSPDGKPCPTPALNCYWPLNDYGFYTTTGIGYQPYRVVLSFELINAPATVCGPKLSTDLIKNOCVNFNENGLTGTGVLTPSSKRFPQFPQGRDVSDFDTS
SARS_CoV_1 5X5B_A	562 562	VRDPTSEILDISPSCFSGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFTQAGCLIGAEHVDTSYECDIPIGAGICASYHT..... VRDPTSEILDISPSCFSGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFTQAGCLIGAEHVDTSYECDIPIGAGICASYHT..... VRDPTSEILDISPSCFSGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFTQAGCLIGAEHVDTSYECDIPIGAGICASYHT.....
SARS_CoV_1 5X5B_A	671 672	.KSIVAYTMSLGADSSIAYSNNITAIPTNFSISITTEVMFVSMKTSVDCNMICYGDSSTECANLLQYGSFCTQLNRALSGIAEQDRNTRREVFAQVKQMYKTPTLKYE KSIVAYTMSLGADSSIAYSNNITAIPTNFSISITTEVMFVSMKTSVDCNMICYGDSSTECANLLQYGSFCTQLNRALSGIAEQDRNTRREVFAQVKQMYKTPTLKYE KSIVAYTMSLGADSSIAYSNNITAIPTNFSISITTEVMFVSMKTSVDCNMICYGDSSTECANLLQYGSFCTQLNRALSGIAEQDRNTRREVFAQVKQMYKTPTLKYE
SARS_CoV_1 5X5B_A	780 780	GGFNFSQILPDPLKPTKRSFTIEDLLFNKVT.....DLICAQKFNGLTVLPPLITDDMIAAYTAALVSGTATAGTFFGAGAAIQIPFAMQMAFYF GGFNFSQILPDPLKPTKRSFTIEDLLFNKVT.....DLICAQKFNGLTVLPPLITDDMIAAYTAALVSGTATAGTFFGAGAAIQIPFAMQMAFYF GGFNFSQILPDPLKPTKRSFTIEDLLFNKVT.....DLICAQKFNGLTVLPPLITDDMIAAYTAALVSGTATAGTFFGAGAAIQIPFAMQMAFYF
SARS_CoV_1 5X5B_A	889 889	NGIGVTQNVLYENQKQIANQFNKAIQSIOESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVNLILSLRDKVEAEVQIDRLITGRLSQLTQVYVQQLIRA NGIGVTQNVLYENQKQIANQFNKAIQSIOESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVNLILSLRDKVEAEVQIDRLITGRLSQLTQVYVQQLIRA NGIGVTQNVLYENQKQIANQFNKAIQSIOESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVNLILSLRDKVEAEVQIDRLITGRLSQLTQVYVQQLIRA
SARS_CoV_1 5X5B_A	998 998	AEIRASANLAATKMSCEVLGQSKRVDFCGKGYHLMSPFOAAPHGVVFLHVTYVPSQERNFTTAPACHGKAYFPREGVVFNGTSSWFTQNRNFFSPQIITTDNTFV AEIRASANLAATKMSCEVLGQSKRVDFCGKGYHLMSPFOAAPHGVVFLHVTYVPSQERNFTTAPACHGKAYFPREGVVFNGTSSWFTQNRNFFSPQIITTDNTFV AEIRASANLAATKMSCEVLGQSKRVDFCGKGYHLMSPFOAAPHGVVFLHVTYVPSQERNFTTAPACHGKAYFPREGVVFNGTSSWFTQNRNFFSPQIITTDNTFV

SARS-CoV S protein – open state

Sequence identity: 100%

100% pP=93.7 SARS_CoV 5X5B_B	18 1	RCTTFDDVQAPNYTQHTSSMRGVYYPDRIFRSDTLYLTQDILFLPFYSNVVTGFHTINHTF.NPVTIPKDGITYFAATEKSNVVRGVFGSTMMNKSQSIVIIINNSTNVVI RCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDILFLPFYSNVVTGFHTINHTF.NPVTIPKDGITYFAATEKSNVVRGVFGSTMMNKSQSIVIIINNSTNVVI RCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDILFLPFYSNVVTGFHTINHTF.NPVTIPKDGITYFAATEKSNVVRGVFGSTMMNKSQSIVIIINNSTNVVI
SARS_CoV 5X5B_B	126 109	RACNFELCDNPFPAVSKPMGTQTHMTIFDPAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFNKNDGFLVYVKGYQPIDVVRDLPSGFNTLKPFLKPLGINITNFRAI RACNFELCDNPFPAVSKPMGTQTHMTIFDPAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFNKNDGFLVYVKGYQPIDVVRDLPSGFNTLKPFLKPLGINITNFRAI RACNFELCDNPFPAVSKPMGTQTHMTIFDPAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFNKNDGFLVYVKGYQPIDVVRDLPSGFNTLKPFLKPLGINITNFRAI
SARS_CoV 5X5B_B	235 218	LTAFS.....WGTSAAAYFVGYLKPTTFMLKYDENGITTDVAVDCSQNPLAELKCSVKSEIDKGIYQTSNFRVVPVSGDVVRFPNITNLCPFGVEFNATKFPVSVAWERK LTAFS.....WGTSAAAYFVGYLKPTTFMLKYDENGITTDVAVDCSQNPLAELKCSVKSEIDKGIYQTSNFRVVPVSGDVVRFPNITNLCPFGVEFNATKFPVSVAWERK LTAFS.....WGTSAAAYFVGYLKPTTFMLKYDENGITTDVAVDCSQNPLAELKCSVKSEIDKGIYQTSNFRVVPVSGDVVRFPNITNLCPFGVEFNATKFPVSVAWERK
SARS_CoV 5X5B_B	344 323	KISNCVADYSVLNSTFEFTKCYGVSATKLNLDLCSNVYADSEVVKGDVVRQIAPGQTGVADIYNYKLDDFMCGVLAWNRNIDATSTGNNYKRYRLRHKLRPFE KISNCVADYSVLNSTFEFTKCYGVSATKLNLDLCSNVYADSEVVKGDVVRQIAPGQTGVADIYNYKLDDFMCGVLAWNRNIDATSTGNNYKRYRLRHKLRPFE KISNCVADYSVLNSTFEFTKCYGVSATKLNLDLCSNVYADSEVVKGDVVRQIAPGQTGVADIYNYKLDDFMCGVLAWNRNIDATSTGNNYKRYRLRHKLRPFE
SARS_CoV 5X5B_B	453 432	RDISNVFSPDGKPCPTPALNCYWPLNDYGFYTTTGIGYQPYRVVLSFELINAPATVCGPKLSTDLIKNOCVNFNENGLTGTGVLTPSSKRFPQFPQGRDVSDFDTS RDISNVFSPDGKPCPTPALNCYWPLNDYGFYTTTGIGYQPYRVVLSFELINAPATVCGPKLSTDLIKNOCVNFNENGLTGTGVLTPSSKRFPQFPQGRDVSDFDTS RDISNVFSPDGKPCPTPALNCYWPLNDYGFYTTTGIGYQPYRVVLSFELINAPATVCGPKLSTDLIKNOCVNFNENGLTGTGVLTPSSKRFPQFPQGRDVSDFDTS
SARS_CoV 5X5B_B	562 541	VRDPTSEILDISPSCFSGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFTQAGCLIGAEHVDTSYECDIPIGAGICASYHT..... VRDPTSEILDISPSCFSGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFTQAGCLIGAEHVDTSYECDIPIGAGICASYHT..... VRDPTSEILDISPSCFSGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFTQAGCLIGAEHVDTSYECDIPIGAGICASYHT.....
SARS_CoV 5X5B_B	671 642	.KSIVAYTMSLGADSSIAYSNNITAIPTNFSISITTEVMFVSMKTSVDCNMICYGDSSTECANLLQYGSFCTQLNRALSGIAEQDRNTRREVFAQVKQMYKTPTLKYE KSIVAYTMSLGADSSIAYSNNITAIPTNFSISITTEVMFVSMKTSVDCNMICYGDSSTECANLLQYGSFCTQLNRALSGIAEQDRNTRREVFAQVKQMYKTPTLKYE KSIVAYTMSLGADSSIAYSNNITAIPTNFSISITTEVMFVSMKTSVDCNMICYGDSSTECANLLQYGSFCTQLNRALSGIAEQDRNTRREVFAQVKQMYKTPTLKYE
SARS_CoV 5X5B_B	780 750	GGFNFSQILPDPLKPTKRSFTIEDLLFNKVT.....DLICAQKFNGLTVLPPLITDDMIAAYTAALVSGTATAGTFFGAGAAIQIPFAMQMAFYF GGFNFSQILPDPLKPTKRSFTIEDLLFNKVT.....DLICAQKFNGLTVLPPLITDDMIAAYTAALVSGTATAGTFFGAGAAIQIPFAMQMAFYF GGFNFSQILPDPLKPTKRSFTIEDLLFNKVT.....DLICAQKFNGLTVLPPLITDDMIAAYTAALVSGTATAGTFFGAGAAIQIPFAMQMAFYF
SARS_CoV 5X5B_B	889 838	NGIGVTQNVLYENQKQIANQFNKAIQSIOESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVNLILSLRDKVEAEVQIDRLITGRLSQLTQVYVQQLIRA NGIGVTQNVLYENQKQIANQFNKAIQSIOESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVNLILSLRDKVEAEVQIDRLITGRLSQLTQVYVQQLIRA NGIGVTQNVLYENQKQIANQFNKAIQSIOESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVNLILSLRDKVEAEVQIDRLITGRLSQLTQVYVQQLIRA
SARS_CoV 5X5B_B	998 947	AEIRASANLAATKMSCEVLGQSKRVDFCGKGYHLMSPFOAAPHGVVFLHVTYVPSQERNFTTAPACHGKAYFPREGVVFNGTSSWFTQNRNFFSPQIITTDNTFV AEIRASANLAATKMSCEVLGQSKRVDFCGKGYHLMSPFOAAPHGVVFLHVTYVPSQERNFTTAPACHGKAYFPREGVVFNGTSSWFTQNRNFFSPQIITTDNTFV AEIRASANLAATKMSCEVLGQSKRVDFCGKGYHLMSPFOAAPHGVVFLHVTYVPSQERNFTTAPACHGKAYFPREGVVFNGTSSWFTQNRNFFSPQIITTDNTFV