Supplemental information

Modeling coronavirus spike protein dynamics: implications for immuno-genicity and immune escape

Genevieve Kunkel, Mohammad Madani, Simon J. White, Paulo H. Verardi, and Anna Tarakanova

Supplementary Information

Modeling Coronavirus Spike Protein Dynamics: Implications for Immunogenicity and Immune Escape

G. Kunkel[1], M. Madani[1], S. J. White[2], P. H. Verardi[3], A. Tarakanova[1,4*]

1. Department of Mechanical Engineering, University of Connecticut, Storrs, CT, USA
2. Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT, USA
3. Department of Pathobiology and Veterinary Science, University of Connecticut, Storrs, CT, USA
4. Department of Biomedical Engineering, University of Connecticut, Storrs, CT, USA

* Corresponding author: anna.tarakanova@uconn.edu

1. SARS-CoV-2 Protein Sequences

*Mutations are in yellow

** Spike protein names correspond to what is listed in the literature through experimental studies or by database identifier. If no name exists, then the spike protein mutant is assigned a name. For proteins that are assigned a name, the naming conventions of its associated experimental study are used. If this information is not listed, then different families of mutations are separated by a period and mutation is assigned a numerical identifier. For example, the 6VXX SARS-CoV-2 mutant sequence contains a signal peptide and trimerization motif, thus its name convention is SC2.S1.TM1. Should another spike protein have the same, for example, trimerization motif mutation, it will also have the ".TM1" identifier. Similar families of mutations include dashes.

Uniport P0DTC2 (Wild Type Sequence)
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSG
TNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVY
YHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD
LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITD
AVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISN
CVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG
CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNG
VGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTT
DAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT
RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFT
ISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKT
PPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLL
TDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQD
SLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYV
TQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTA
PAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKE
ELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIA
GLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

Name: BiPro
PDB: 6VSB (2)
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSG
TNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVY

YHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD
LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITD
AVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISN
CVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG
CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNG
VGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTT
DAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT
RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSP**GSAS**SVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFT
ISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKT
PPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLL
TDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQD
SLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLD**PP**EAEVQIDRLITGRLQSLQTYVT
QQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAP
AICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEE
LDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQ**GSGYIPEAPRDGQAY
VRKDGEWVLLSTFLGRSLEVLFQGPGHHHHHHHHSAWSHPQFEKGGGSGGGGSGGSAWSHPQFEK**

Unresolved: 1-26, 67-78, 96-98, 143-155,177-186, 247-260, 330-334, 444-490, 501-502, 621-640, 673-686, 812-814,
829-850, 1147-1288

Name: HexaPro
PDB: 6XKL (3)
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSG
TNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVY
YHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD
LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITD
AVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISN
CVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG
CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNG
VGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTT
DAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT
RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSP**GSAS**SVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFT
ISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKT
PPIKDFGGFNFSQILPDPSKPSKRS**P**IEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLL
TDEMIAQYTSALLAGTITSGWTFGAG**P**ALQIPF**P**MQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDS
LSST**P**SALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQ
QLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAI
CHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELD
KYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQ**GSGYIPEAPRDGQAYVR
KDGEWVLLSTFLGRSLEVLFQGPGHHHHHHHHSAWSHPQFEKGGGSGGGGSGGSAWSHPQFEK**

Unresolved: 1-26, 67-78, 96-98, 143-155,177-186, 247-260, 330-334, 444-490, 501-502, 621-640, 673-686, 812-814,
829-850, 1147-1288

Name: BiPro-1
PDB:6Z97 (4)
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSG
TNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVY
YHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD
LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITD
AVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISN
CVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG
CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNG
VGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTT
DAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT
RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSP**GSAS**SVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFT

ISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKT
PPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLL
TDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQD
SLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLD<mark>PP</mark>EAEVQIDRLITGRLQSLQTYVT
QQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAP
AICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEE
LDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQ<mark>GSGYIPEAPRDGQAY</mark>
<mark>VRKDGEWVLLSTFLGRSLEVLFQGPGHHHHHHHHGSAWSHPQFEKGGGSGGGSGGSAWSHPQFEK</mark>

Unresolved: 1-26, 70-81,114-115, 144-187, 243-262, 621-640, 677-689, 828-850, 1148-1288

Name: SC2.S1.TM1
PDB: 6VYB & 6VXX (5)
M<mark>GILPSPGMPALLSLVSLLSVLLMGCVAETGT</mark>QCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQ
DLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATN
VVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFK
NIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVG
YLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVF
NATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPG
QTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVE
GFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTE
SNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNEVAVLYQDVNCTEVPVAIHAD
QLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPS<mark>GAGS</mark>VASQSIIAYTMSL
GAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVE
QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAA
RDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLY
ENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLD<mark>PP</mark>EAE
VQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVV
FLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIV
NNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGK
YEQYIK<mark>GSGRENLYFQGGGGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHH</mark>

Unresolved: 1-26, 70-81, 114-115, 144-185, 243-262, 443-489, 502, 621-640, 677-689, 812, 828-854, 1148-1281

Name: SC2.S2. TM1-1
PDB: 6ZGG (6)
M<mark>GILPSPGMPALLSLVSLLSVLLMGCVAETGM</mark>FVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPD
KVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSK
TQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQ
GNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSG
WTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVR
FPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF
VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEI
YQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFN
FNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQD
VNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARS
VASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFC
TQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFI
KQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYR
FNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVL
NDILSRLD<mark>PP</mark>EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHL
MSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT
FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKN
LNESLIDLQELGKYEQ<mark>SGRENLYFQGGGGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHH</mark>

Unresolved: 1-13, 71-75, 618-640, 677-688, 828-848, 941-943, 1147-1287

Name: SC2.N1.C1.2P.TM2
PDB: 6XF6 (7)
GPQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNP
VLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWM
ESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEP
LVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLS
ETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLY
NSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNN
LDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTL
EILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAE
HVNNSYECDIPIGAGICASYQTQTNSP<mark>GSAS</mark>SVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPV
SMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGF
NFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQY
TSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASAL
GKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLD<mark>PP</mark>EAEVQIDRLITGRLQSLQTYVTQQLIRAAEI
RASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKA
HFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNH
TSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQ<mark>GSGYIPEAPRDGQAYVRKDGEW</mark>
<mark>VLLSTFLGRSGGGLVPQQSGGLNDIFEAQKIEWHEG</mark>

Unresolved: 12-26, 70-81, 114-115, 144-165, 173-185, 243-262, 621-640, 677-690, 828-854, 1148-1266

Name: SC2.C2. 1P. TM3
PDB: 7AD1 (8)
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSG
TNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVY
YHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD
LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITD
AVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISN
CVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG
CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNG
VGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTT
DAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQNVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT
RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSP<mark>SRAG</mark>SVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFT
ISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKT
PPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLL
TDEMIAQYTSALLAGTITSGWTFGAGPALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQD
SLSST<mark>P</mark>SALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKPEAEVQIDRLITGRLQSLQTYVT
QQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAP
AICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEE
LDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQ<mark>GSGYIPEAPRDGQAY</mark>
<mark>VRKDGEWVLLSTFLGRSLEVLFQGPGSLPETGGGSDYKDDDDKGGGGSGGGGSGGGGSGGGGSGGGGSH</mark>
<mark>HHHHH</mark>

Unresolved: 1-26, 70-87, 114-115, 132-165, 173-185, 243-262,443-448, 477-489, 502-503, 621-640, 677-689, 812,
828-854, 1148-1297

Name: u1S2q
PDB: 6X2B (9)
VNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFN
DGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFR
VYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPI
GINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTL
KSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFST

FKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELL
HAPATVCGPKKSTN<mark>L</mark>VKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDILDITDAVRDPQTLEILDITPC
SFGGVSVITPGTNTSNEVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYE
CDIPIGAGICASYQTQTNSP<mark>GSAS</mark>SVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSV
DCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPD
PSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKYIGLTVLPPLLTDEMIAQYTSALLAGT
<mark>I</mark>TSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVV
NQNAQALNTLVKQLSSNFGAISSVLNDILSRLD<mark>PP</mark>EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLA
ATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVF
VSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG
DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQ<mark>GSGYIPEAPRDGQAYVRKDGEWVLLSTFLGR</mark>
<mark>SLEVLFQGPGHHHHHHHSAWSHPQFEKGGGSGGGGSGGSAWSHPQFEK</mark>

Unresolved: 16-26, 68-81, 114-115, 144-185, 243-262, 443-489, 502, 621-640, 677-689, 812, 828-854, 1148-1288

Name: SC2.C1. 2P
PDB: 7CN9 (10)
QCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLP
FNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEF
RVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDL
PIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC
TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASF
STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKV
GGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFEL
LHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITP
CSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY
ECDIPIGAGICASYQTQTNSP<mark>GSAG</mark>SVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTS
VDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILP
DPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLA
GTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQD
VVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLD<mark>PP</mark>EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASAN
LAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREG
VFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDP

Unresolved: 14-26, 180-182, 444-489, 622-640, 673-685, 812-852

Name: SC2.C1.2P.TM4
PDB: 6XM0 (11)
QCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLP
FNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEF
RVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDL
PIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC
TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASF
STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKV
GGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFEL
LHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITP
CSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY
ECDIPIGAGICASYQTQTNSP<mark>GSAS</mark>SVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTS
VDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILP
DPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLA
GTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQD
VVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLD<mark>PP</mark>EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASAN
LAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREG
VFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVD

LGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQGSGYIPEAPRDGQAYVRKDGEWVLLSTFL
GRSLEVLFQGPGHHHHHHHHSAWSHPQFEKGGGSGGGGSGGSAWSHPQFEK

Unresolved: 14-26, 70-79, 144-158, 174-185, 251-263, 445-446, 677-688, 829-848, 1148-1288

Name: SC2. TM4-1
PDB: 6XR8 (12)
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSG
TNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVY
YHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD
LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITD
AVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISN
CVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG
CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNG
VGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTT
DAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT
RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFT
ISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKT
PPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLL
TDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQD
SLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYV
TQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTA
PAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKE
ELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIA
GLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYTLESGGGSAWSHPQFEKGGGS
GGGSGGSSAWSHPQFEK

Name: SC2.C1. 2P. TM4-2
PDB: 6XM4 (11)
QCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLP
FNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEF
RVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDL
PIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC
TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASF
STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKV
GGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFEL
LHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITP
CSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY
ECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTS
VDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILP
DPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLA
GTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQD
VVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASAN
LAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREG
VFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVD
LGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQGSGYIPEARDGQAYVRKDGEWVLLSTFLG
RSLEVLFQGPGHHHHHHHHSAWSHPQFEKGGGSGGGGSGGSAWSHPQFEK

Unresolved: 14-26, 70-79, 144-185, 251-263, 445-446, 677-688, 829-848, 1148-1288

Name: SC2.C1.TM4-2
PDB: 7KDH (13)
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSG
TNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVY
YHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD

LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITD
AVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISN
CVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG
CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNG
VGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTT
DAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT
RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSP<mark>GSAS</mark>SVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFT
ISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKT
PPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLL
TDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQD
SLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYV
TQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTA
PAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKE
ELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQ<mark>GSGYIPEAPRDGQA</mark>
<mark>YVRKDGEWVLLSTFLGRSLEVLFQGPGHHHHHHHHSAWSHPQFEKGGGSGGGGSGGSAWSHPQFEK</mark>

Unresolved: 1-26, 70-81, 144-185, 243-262, 443-447, 471-489, 502, 621-640, 677-689, 812-852, 1148-1288

Name: BiPro-0
PDB: 6ZP7 (14)
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSG
TNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVY
YHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD
LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITD
AVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISN
CVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG
CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNG
VGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTT
DAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT
RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFT
ISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKT
PPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLL
TDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQD
SLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLD<mark>PP</mark>EAEVQIDRLITGRLQSLQTYVT
QQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAP
AICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEE
LDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAG
LIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

Unresolved: 40-53, 105-134, 147-160, 216-234, 419-420, 594-612, 650-661, 802-824, 1122-1273

Name: SC2.TM5
PDB ID: 7C2L (35)
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSG
TNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVY
YHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD
LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITD
AVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISN
CVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG
CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNG
VGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTT
DAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT
RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRGSASSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNF
TISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYK
TPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPL
LTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQ

DSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTY
VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTT
APAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFK
EELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFI
AGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT<mark>LEDYKDDDDK</mark>

2.   SARS-CoV Protein Sequences

SARS-CoV WT (Uniprot ID P59594) (15)
MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFLPFYSNVTGFHTIN
HTFGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIIINNSTNVVIRACNFELCDNPFFAVSKPMG
TQTHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFKNKDGFLYVYKGYQPIDVVRDLPSGFNTLK
PIFKLPLGINITNFRAILTAFSPAQDIWGTSAAAYFVGYLKPTTFMLKYDENGTITDAVDCSQNPLAELKCSV
KSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFST
FKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATS
TGNYNYKYRYLRHGKLRPFERDISNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFEL
LNAPATVCGPKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDISPC
SFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQAGCLIGAEHVDTSYE
CDIPIGAGICASYHTVSLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTEVMPVSMAKTSVDCN
MYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPL
KPTKRSFIEDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGT
ATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGKLQDV
VNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANL
AATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGV
FVFNGTSWFITQRNFFSPQIITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG
DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTS
CCSCLKGACSCGSCCKFDEDDSEPVLKGVKLHYT

Name: SC1. TM1
PDB: 6ACD (35)
MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFLPFYSNVTGFHTIN
HTFGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIIINNSTNVVIRACNFELCDNPFFAVSKPMG
TQTHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFKNKDGFLYVYKGYQPIDVVRDLPSGFNTLK
PIFKLPLGINITNFRAILTAFSPAQDIWGTSAAAYFVGYLKPTTFMLKYDENGTITDAVDCSQNPLAELKCSV
KSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFST
FKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATS
TGNYNYKYRYLRHGKLRPFERDISNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFEL
LNAPATVCGPKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDISPC
SFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQAGCLIGAEHVDTSYE
CDIPIGAGICASYHTVSLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTEVMPVSMAKTSVDCN
MYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPL
KPTKRSFIEDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGT
ATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGKLQDV
VNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANL
AATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGV
FVFNGTSWFITQRNFFSPQIITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG
DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWSHPQFEK

Name: SC1.S1.TM2
PDB: 6NB6 (35)
MGILPSPGMPALLSLVSLLSVLLMGCVAETGTSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTL
YLTQDLFLPFYSNVTGFHTINHTFDNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIIINNSTNVVI
RACNFELCDNPFFAVSKPMGTQTHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFKNKDGFLYV
YKGYQPIDVVRDLPSGFNTLKPIFKLPLGINITNFRAILTAFSPAQDTWGTSAAAYFVGYLKPTTFMLKYDE
NGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKFPSVYAW

ERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVIADYNYK
LPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERDISNVPFSPDGKPCTPPALNCYWPLNDY
GFYTTTGIGYQPYRVVVLSFELLNAPATVCGPKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGR
DVSDFTDSVRDPKTSEILDISPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGN
NVFQTQAGCLIGAEHVDTSYECDIPIGAGICASYHTVSLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNF
SISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQMY
KTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGLTVL
PPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAI
SQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSL
QTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTYVPSQERN
FTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSF
KEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKGSGRENLYF
QGGGGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHH

3. **MERS-CoV Protein Sequences**

MERS-CoV WT (Uniprot ID K9N5Q8) (16)
MIHSVFLLMFLLTPTESYVDVGPDSVKSACIEVDIQQTFFDKTWPRPIDVSKADGIIYPQGRTYSNITITYQGL
FPYQGDHGDMYVYSAGHATGTTPQKLFVANYSQDVKQFANGFVVRIGAAANSTGTVIISPSTSATIRKIYP
AFMLGSSVGNFSDGKMGRFFNHTLVLLPDGCGTLLRAFYCILEPRSGNHCPAGNSYTSFATYHTPATDCSD
GNYNRNASLNSFKEYFNLRNCTFMYTYNITEDEILEWGITQTAQGVHLFSSRYVDLYGGNMFQFATLPVY
DTIKYYSIIPHSIRSIQSDRKAWAAFYVYKLQPLTFLLDFSVDGYIRRAIDCGFNDLSQLHCSYESFDVESGV
YSVSSFEAKPSGSVVEQAEGVECDFSPLLSGTPPQVYNFKRLVFTNCNYNLTKLLSLFSVNDFTCSQISPAAI
ASNCYSSLILDYFSYPLSMKSDLSVSSAGPISQFNYKQSFSNPTCLILATVPHNLTTITKPLKYSYINKCSRFLS
DDRTEVPQLVNANQYSPCVSIVPSTVWEDGDYYRKQLSPLEGGGWLVASGSTVAMTEQLQMGFGITVQY
GTDTNSVCPKLEFANDTKIASQLGNCVEYSLYGVSGRGVFQNCTAVGVRQQRFVYDAYQNLVGYYSDDG
NYYCLRACVSVPVSVIYDKETKTHATLFGSVACEHISSTMSQYSRSTRSMLKRRDSTYGPLQTPVGCVLGL
VNSSLFVEDCKLPLGQSLCALPDTPSTLTPRSVRSVPGEMRLASIAFNHPIQVDQLNSSYFKLSIPTNSFGVTQ
EYIQTTIQKVTVDCKQYVCNGFQKCEQLLREYGQFCSKINQALHGANLRQDDSVRNLFASVKSSQSSPIIPG
FGGDFNLTLLEPVSISTGSRSARSAIEDLLFDKVTIADPGYMQGYDDCMQQGPASARDLICAQYVAGYKVL
PPLMDVNMEAAYTSSLLGSIAGVGWTAGLSSFAAIPFAQSIFYRLNGVGITQQVLSENQKLIANKFNQALGA
MQTGFTTTNEAFHKVQDAVNNNAQALSKLASELSNTFGAISASIGDIIQRLDVLEQDAQIDRLINGRLTTLN
AFVAQQLVRSESAALSAQLAKDKVNECVKAQSKRSGFCGQGTHIVSFVVNAPNGLYFMHVGYYPSNHIEV
VSAYGLCDAANPTNCIAPVNGYFIKTNNTRIVDEWSYTGSSFYAPEPITSLNTKYVAPQVTYQNISTNLPPPL
LGNSTGIDFQDELDEFFKNVSTSIPNFGSLTQINTTLLDLTYEMLSLQQVVKALNESYIDLKELGNYTYYNK
WPWYIWLGFIAGLVALALCVFFILCCTGCGTNCMGK
LKCNRCCDRYEEYDLEPHKVHVH

Name: MC. SD. TM1
PDB: 5X5C & 5X5F (36)
YVDVGPDSVKSACIEVDIQQTFFDKTWPRPIDVSKADGIIYPQGRTYSNITITYQGLFPYQGDHGDMYVYSA
GHATGTTPQKLFVANYSQDVKQFANGFVVRIGAAANSTGTVIISPSTSATIRKIYPAFMLGSSVGNFSDGKM
GRFFNHTLVLLPDGCGTLLRAFYCILEPRSGNHCPAGNSYTSFATYHTPATDCSDGNYNRNASLNSFKEYF
NLRNCTFMYTYNITEDEILEWGITQTAQGVHLFSSRYVDLYGGNMFQFATLPVYDTIKYYSIIPHSIRSIQS
DRKAWAAFYVYKLQPLTFLLDFSVDGYIRRAIDCGFNDLSQLHCSYESFDVESGVYSVSSFEAKPSGSVVE
QAEGVECDFSPLLSGTPPQVYNFKRLVFTNCNYNLTKLLSLFSVNDFTCSQISPAAIASNCYSSLILDYFSYPL
SMKSDLSVSSAGPISQFNYKQSFSNPTCLILATVPHNLTTITKPLKYSYINKCSR<mark>L</mark>LSDDRTEVPQLVNANQY
SPCVSIVPSTVWEDGDYYRKQLSPLEGGGWLVASGSTVAMTEQLQMGFGITVQYGTDTNSVCPKLEFAND
TKIASQLGNCVEYSLYGVSGRGVFQNCTAVGVRQQRFVYDAYQNLVGYYSDDGNYYCLRACVSVPVSVI
YDKETKTHATLFGSVACEHISSTMSQYSRSTRSMLKRRDSTYGPLQTPVGCVLGLVNSSLFVEDCKLPLGQ
SLCALPDTPSTLTPRSV<mark>S</mark>SVPGEMRLASIAFNHPIQVDQLNSSYFKLSIPTNFSFGVTQEYIQTTIQKVTVDCK
QYVCNGFQKCEQLLREYGQFCSKINQALHGANLRQDDSVRNLFASVKSSQSSPIIPGFGGDFNLTLLEPVSIS
TGSRSARSAIEDLLFDKVTIADPGYMQGYDDCMQQGPASARDLICAQYVAGYKVLPPLMDVNMEAAYTS
SLLGSIAGVGWTAGLSSFAAIPFAQSIFYRLNGVGITQQVLSENQKLIANKFNQALGAMQTGFTTTNEAFQK
VQDAVNNNAQALSKLASELSNTFGAISASIGDIIQRLDVLEQDAQIDRLINGRLTTLNAFVAQQLVRSESAA

LSAQLAKDKVNECVKAQSKRSGFCGQGTHIVSFVVNAPNGLYFMHVGYYPSNHIEVVSAYGLCDAANPT
NCIAPVNGYFIKTNNTRIVDEWSYTGSSFYAPEPITSLNTKYVAPQVTYQNISTNLPPPLLGNSTGIDFQDELD
EFFKNVSTSIPNFGSLTQINTTLLDLTYEMLSLQQVVKALNESYIDLKELGNYTYYNKEFRLVPRGSPGSGYI
PEAPRDGQAYVRKDGEWVLLSTFLGHHHHHH

Name: MC.TM2
PDB: 5W9K (37)
MIHSVFLLMFLLTPTESYVDVGPDSVKSACIEVDIQQTFFDKTWPRPIDVSKADGIIYPQGRTYSNITITYQGL
FPYQGDHGDMYVYSAGHATGTTPQKLFVANYSQDVKQFANGFVVRIGAAANSTGTVIISPSTSATIRKIYP
AFMLGSSVGNFSDGKMGRFFNHTLVLLPDGCGTLLRAFYCILEPRSGNHCPAGNSYTSFATYHTPATDCSD
GNYNRNASLNSFKEYFNLRNCTFMYTYNITEDEILEWFGITQTAQGVHLFSSRYVDLYGGNMFQFATLPVY
DTIKYYSIIPHSIRSIQSDRKAWAAFYVYKLQPLTFLLDFSVDGYIRRAIDCGFNDLSQLHCSYESFDVESGV
YSVSSFEAKPSGSVVEQAEGVECDFSPLLSGTPPQVYNFKRLVFTNCNYNLTKLLSLFSVNDFTCSQISPAAI
ASNCYSSLILDYFSYPLSMKSDLSVSSAGPISQFNYKQSFSNPTCLILATVPHNLTTITKPLKYSYINKCSRFLS
DDRTEVPQLVNANQYSPCVSIVPSTVWEDGDYYRKQLSPLEGGGWLVASGSTVAMTEQLQMGFGITVQY
GTDTNSVCPKLEFANDTKIASQLGNCVEYSLYGVSGRGVFQNCTAVGVRQQRFVYDAYQNLVGYYSDDG
NYYCLRACVSVPVSVIYDKETKTHATLFGSVACEHISSTMSQYSRSTRSMLKRRDSTYGPLQTPVGCVLGL
VNSSLFVEDCKLPLGQSLCALPDTPSTLTPASVGSVPGEMRLASIAFNHPIQVDQLNSSYFKLSIPTNFSFGVT
QEYIQTTIQKVTVDCKQYVCNGFQKCEQLLREYGQFCSKINQALHGANLRQDDSVRNLFASVKSSQSSPIIP
GFGGDFNLTLLEPVSISTGSRSARSAIEDLLFDKVTIADPGYMQGYDDCMQQGPASARDLICAQYVAGYKV
LPPLMDVNMEAAYTSSLLGSIAGVGWTAGLSSFAAIPFAQSIFYRLNGVGITQQVLSENQKLIANKFNQALG
AMQTGFTTTNEAFHKVQDAVNNNAQALSKLASELSNTFGAISASIGDIIQRLDPPEQDAQIDRLINGRLTTL
NAFVAQQLVRSESAALSAQLAKDKVNECVKAQSKRSGFCGQGTHIVSFVVNAPNGLYFMHVGYYPSNHIE
VVSAYGLCDAANPTNCIAPVNGYFIKTNNTRIVDEWSYTGSSFYAPEPITSLNTKYVAPQVTYQNISTNLPPP
LLGNSTGIDFQDELDEFFKNVSTSIPNFGSLTQINTTLLDLTYEMLSLQQVVKALNESYIDLKELGNYTYGSG
YIPEAPRDGQAYVRKDGEWVLLSTFLGRSLEVLFQ

4.   Supplementary Table 1

See attached file titled <u>Supplementary_Table_1.xlsx</u> for extensive review of individual binding properties
and epitopes of over 40 antibodies and their related pdb identifiers. References for the table are located in
this document.

5.   Protein sequence alignment and comparison

Multiple   sequence   alignment   of   spike   protein   sequences   are   attached   in   file
Spike_Protein_Alignment_Tarakanova.clustal_num

6.   Dynamic domain & NMA videos

Like the Dynamic Domain index files, videos showing the NMA trajectory and highlighted dynamic
domains are labeled [PDB ID].mp4

7.   Results of artificial controls

The graphical results of the artificial controls used in **Results section 3.4** are listed below:

i.   SC2.S2.TM1-1' & WT'-A

Combined Unresolved regions (WT'A)-: 1-26, 70-81, 114-115, 144-185, 243-262, 443-489, 502, 621-640, 677-689, 812, 828-854, 1148-end

Combined unresolved regions (SC2.S2.TM1-1'): 1-26, 70-81,114-115,144-185,243-262, 443-489,502,618-640, 677-698, 812, 828-859, 1148-end

ii.   **SC2.C1.2P' & SC2.C1.2P.TM4',**
Combined unresolved regions:12-26, 70-81, 114-115, 144-165, 173-185, 243-263, 445-446, 621-640, 613-640, 673-690, 812-854, 1148-end
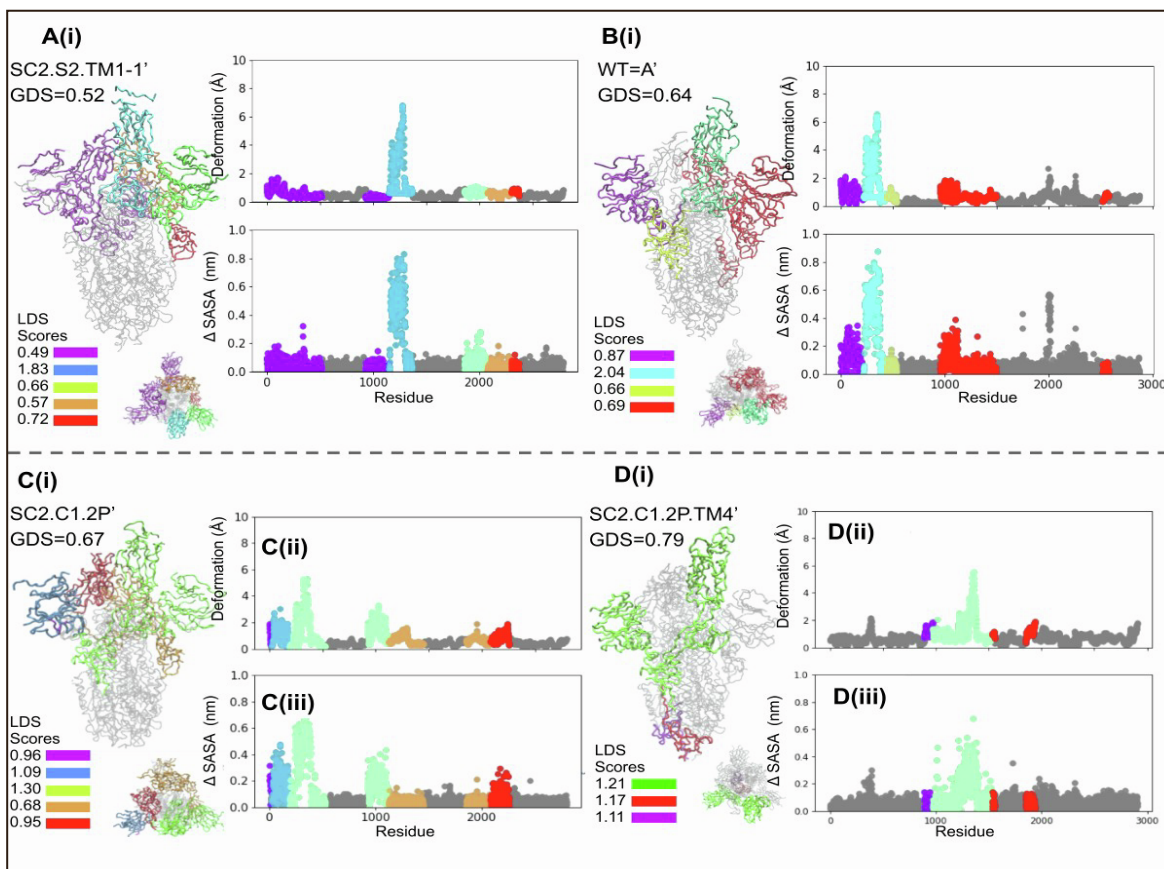
SI Figure 1: The domain dynamics associated with **(A)** WT'-A, **(B)** SC2.S2.TM1-1', **(C)** SC2.C1.2P' and **(D)** SC2.C1.2P.TM4' ANMs. The PDB ID, global dynamics score (GDS), local dynamics scores (LDSs), deformation profile **(ii)**, and Δ SASA profile **(iii)** is listed for each structure. On each 3-D structure and profile, identified dynamic domains are labeled in different colors and their LDSs are listed in each legend. The corresponding dynamic domain breakdowns are located in **Figure 4** and **Figure 1.**

8.   Derived features used for thermal stability model training

The features used to train the thermal stability predictor are those contained in the original combined data set and also those computed by the following resources. The Amber software (39) was used to find bond length, bond angle, dihedral angles, van der walls contributions, electrostatic contributions, polar solvation, total gas free energy, total solvation free energy, and total system energy. The FoldX software (40) was used to find solvation energy for polar groups, solvation energy for apolar groups, water bridge hydrogen bonding, intra-molecule hydrogen bonding, electrostatic interactions between charged

groups, and atomic clash overlaps. The remaining features are outlined in **Table S3**. Amino acid-related biological features are found using AAindex (42). Disorder and Relative Surface Accessibility are found using the SCRATCH webserver (43). All remaining features in **Table S3** are found using the Expasy ProtScale tool (41).

Table S3: Summary of the biochemical, structural, and biological features used for thermal stability predictor training.

Biochemical Features	Structural features	Biological Features
Molecular Weight	Solvent Accessibility	Mobility of Amino Acids
Hydrophobicity Index	3-State Secondary Structure	Codon Amount per Amino Acid
Side Chain pKa	Bulkiness	
Frequencies of buried and exposed sequence	Buried Area from Standard to Folded State	
Electrostatic Charge	Disorder	
	Flexibility Index	
	Relative Surface Accessibility	
	Absolute Surface Accessibility	

9.   Analysis of different PDB structures corresponding to a single sequence

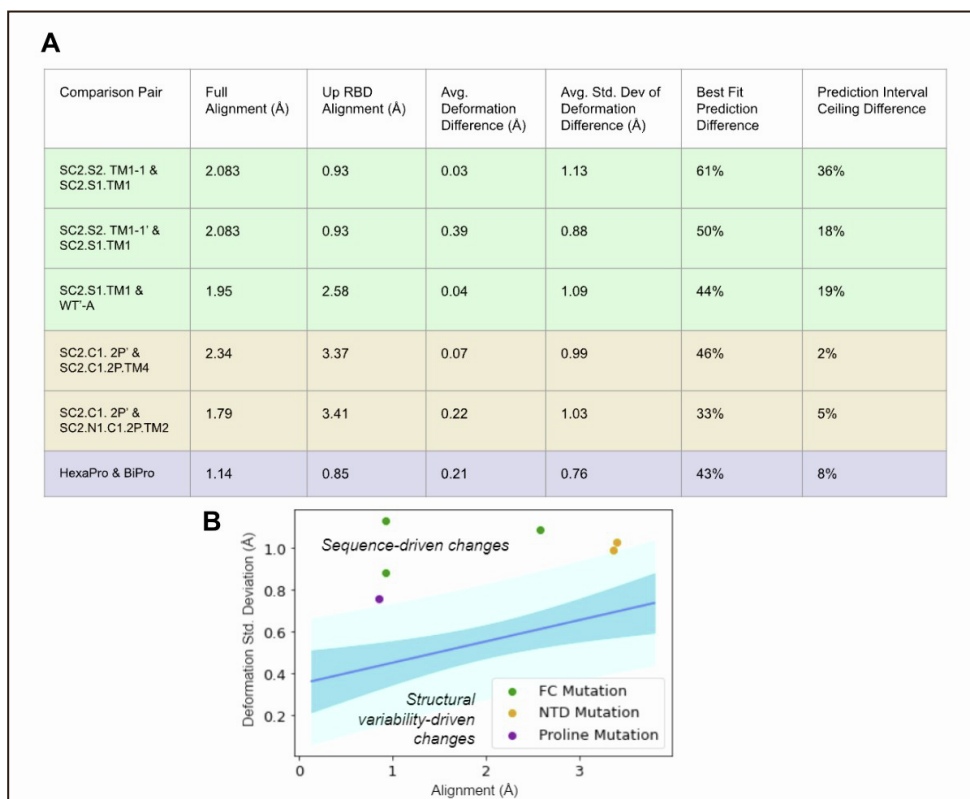### i.   Statistical threshold model from PDB structures

To resolve whether ANM models can accurately capture differences in dynamic behavior between proteins of different sequence, we selected multiple structures from the PDB that correspond to the same sequence to analyze the correlation between protein structure alignment and the difference in their dynamical signature. We measure the difference between dynamical signatures by the average difference and standard deviation between per-residue protein deformations computed from the ANM-derived trajectory. We explore different levels of protein structural alignment to establish a threshold for measured structural alignment and dynamical similarity. This allows us to evaluate if compared protein structures (see **Results Section 3.4**) experience changes that can in fact be attributed to sequence, or are artifacts of the ANM model. The PDBs used as case studies are listed in **Supplementary Table 2** (see attached file Supplementary_Table_2.xlsx) and this information is summarized in **Figure 3.** Taken together, the results from the PDB case studies establish a model that describes how NMA dynamics should differ between structures of the same sequence as the structural alignment between compared PDBs differs. Linear regression is used to fit the correlated data to a best fit line, then to calculate the corresponding 95% confidence interval and prediction interval (**SI Figure 2-B**). The prediction interval indicates an estimate of an interval in which a future observation will fall and the confidence interval indicates a range of values that is likely to contain the true mean of the population. From this model, we can understand if the dynamical differences measured between two compared proteins with different sequences are likely to be driven by sequence differences or structural differences (i.e. model artifacts). Data points that fall above the best fit line may attribute their dynamical differences to sequence. Data points below the best fit line indicate that dynamical difference may be an artifact of the model.
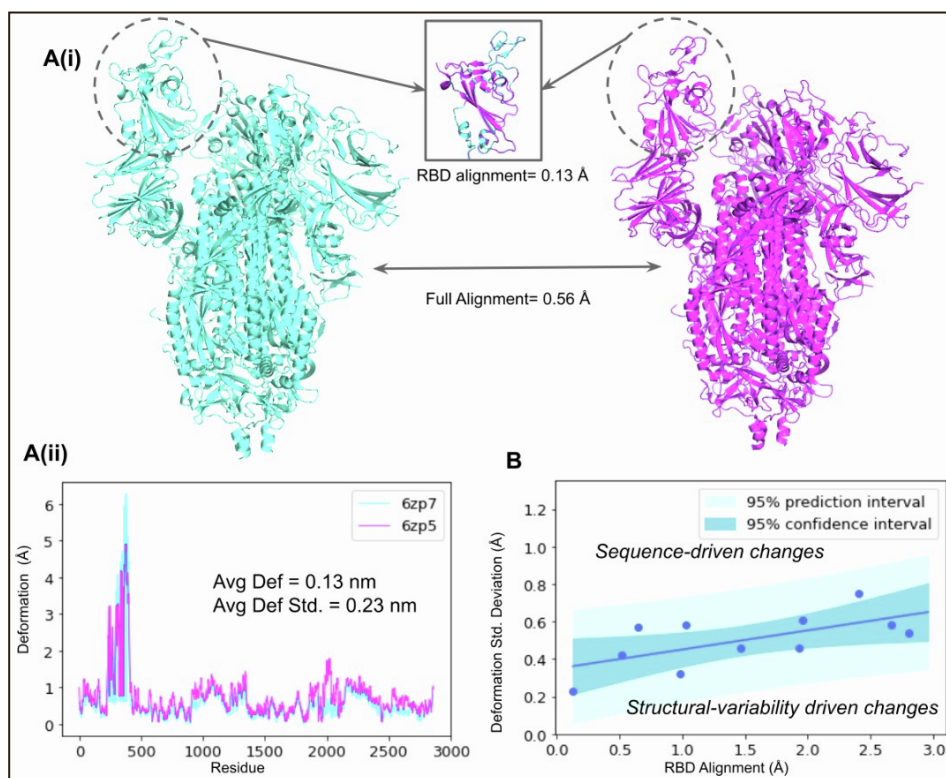
SI Figure 2: Proteins selected for a single case study **(A)**, visualization of their structural alignment **(i)** and alignment of their NMA deformations **(ii)**. This pair presents high structural alignment and level of dynamics correspondence and creates a single data point that is plotted with the remaining results of the case studies **(B)**. The plot shows the protein deformation trend variability as a function of the static RBD alignment **(B)**. Note that low alignment values indicate better structure correspondence and high alignment values indicate poorer structure correspondence.

ii. Contribution of sequence and structure to mutant dynamic variability

After establishing a model that characterizes the relationship between PDB structural alignment between pairs of structures and their resulting dynamical differences, we can understand the contribution of sequence-driven dynamical changes in mutants versus changes that arise from e.g. experimental processing variability (i.e. model artifacts). To do this, we perform a similar analysis outlined in the previous section with mutant structures and their controls. If these fall below the best fit line then it is likely that the resulting NMA dynamical differences are driven by experimental processing differences and not sequence. However, if data points fall above the best fit line then it is likely that the differences in NMA deformation profiles, as measured by the standard deviation of deformation difference, is due to the structural differences that arise from changes in protein sequence. Using the linear regression model, we calculate the percent difference between the true mutant standard deviation values with the predicted points on the best fit line and the 95% prediction interval ceiling (**SI Figure 3**). We see that all mutant standard deviation values fall at least 30% above the best fit values for the same RBD alignment and fall at least 2% above the prediction interval ceiling. This suggests that all of the results that we see from comparing the dynamics of mutant proteins and their controls (see **Section 3.4**) are due to the structural consequences of protein sequence.

A

Comparison Pair	Full Alignment (Å)	Up RBD Alignment (Å)	Avg. Deformation Difference (Å)	Avg. Std. Dev of Deformation Difference (Å)	Best Fit Prediction Difference	Prediction Interval Ceiling Difference
SC2.S2. TM1-1 & SC2.S1.TM1	2.083	0.93	0.03	1.13	61%	36%
SC2.S2. TM1-1' & SC2.S1.TM1	2.083	0.93	0.39	0.88	50%	18%
SC2.S1.TM1 & WT'-A	1.95	2.58	0.04	1.09	44%	19%
SC2.C1. 2P' & SC2.C1.2P.TM4	2.34	3.37	0.07	0.99	46%	2%
SC2.C1. 2P' & SC2.N1.C1.2P.TM2	1.79	3.41	0.22	1.03	33%	5%
HexaPro & BiPro	1.14	0.85	0.21	0.76	43%	8%

B

SI Figure 3: Mutant comparisons made to understand contribution of sequence to dynamic variability. **(A)** Full results are listed and **(B)** the standard deviation of the difference between the deformation curves is plotted against RBD structural alignment in reference to the linear regression model from **SI Figure 2-B**. Proteins that are used to study each group of mutations are highlighted in different colors. Green represents furin cleavage mutations, orange represents NTD mutations, and purple represents proline mutations. All proteins comparisons made for the mutant proteins fall well above the best fit line and 95% intervals, see best fit and prediction interval differences **(A)**, indicating that dynamical differences between mutant and control proteins are driven by sequence.

iii.   Molecular dynamics simulation of WT structure

a.   Simulation Methods

A 100 ns classical molecular dynamics (MD) simulation was performed on the atomistic WT S protein structure resolved by Amaro et al *(44)*. The Amaro structure is missing the first 13 residues, and these were included in the starting structure by way of homology modeling using the Robetta web server (*2*). The full sequence was modeled using each chain of the WT MD model as a template. Next, the resulting homology chains were aligned with the WT model, and the 13 residues of the homology model were attached to each chain of the WT model. The resulting RBD-up model, termed WT'-B, is used as the input structure for simulation. The MD simulation was carried out using the GROMACS 2019 version MD software (*49*) and performed on the TACC Stampede2 HPC platform.

We used the CHARMM36m forcefield, which is uniquely suited for flexible proteins and other biomolecular structure *(55)*. For all runs, the protein was simulated in a neutral cubic water box of size 45^3 nm^3 using periodic boundary conditions. Prior to the final production run, an energy minimization, NVT equilibration, and NPT equilibration runs were performed. The energy minimization run was conducted using the steepest descent algorithm. The NVT and NPT equilibration runs were both simulated for 100 ps each. During the NVT run, the protein and environment were equilibrated at 300 K using the V-Rescale

thermostat—Berendsen thermostat with velocity rescaling—for temperature coupling (*50*). Next, during the NPT run, the system was equilibrated at 1 atm using Parrinello-Rahman barostat for pressure coupling (*51*). The final MD production was run for a time length of 100 ns at 300 K using a 2 femtosecond time step and the same thermostat and barostat that were employed for the equilibration runs. The Particle Mesh Edwald (PME) algorithm *(52)* is used for long range electrostatics (grid spacing at 1.6) and Verlet *(53)* algorithm for neighbor searching (short range cutoffs at 1.0). The LINCS algorithm (*54*) was used to constrain hydrogen and covalent bonds that are explicitly stated in the protein topology.

b. WT Simulation and ANM Results

We explore different configurations of the WT structure in simulation to confirm if different snapshots from an MD trajectory may impact predicted ANM dynamics. We chose 2 frames from the WT'-B simulation, one in which the up RBD is rotated outward compared to the starting structure and one where the RBD is rotated inward—frames 2400 and 152 respectively. These specific structures were chosen due to the RBD orientation and because they exhibit high RMSD measurements (indicating lower structural alignment). We conduct the same analysis as above and the results are compared to the case study model in **SI Figure 2B**. The alignment analysis was conducted three separate times on (1) the full length protein, (2) protein with the commonly unresolved C-terminus tail removed, and (3) protein with additional commonly unresolved regions removed, to make parallels to the results presented in **Figure 2.** Each data point is the result of one of the more rotated RBD structures taken from the WT'-B simulation trajectory compared to the baseline (Amaro et al.) structure, resulting in 6 total data points. The results for the WT'-B analysis are summarized in **SI Figure 4.** In **SI Figure 4-B**, we see that the majority of points fall below the best fit line, consistent with the fact that structural differences drive dynamical changes rather than a change in sequence. There are 2 points resulting from the proteins with regions removed that fall above the best fit line. However, these fall within the expected range of data variability (95% confidence interval) meaning that the reflected NMA dynamics of the MD snapshots correspond well enough to that of the baseline structure. Overall, these results indicate that reliable results are obtained when using the Amaro WT structure, used in all primary WT analysis, given that it is able to capture the dynamics presented in structurally heterogeneous MD snapshots.
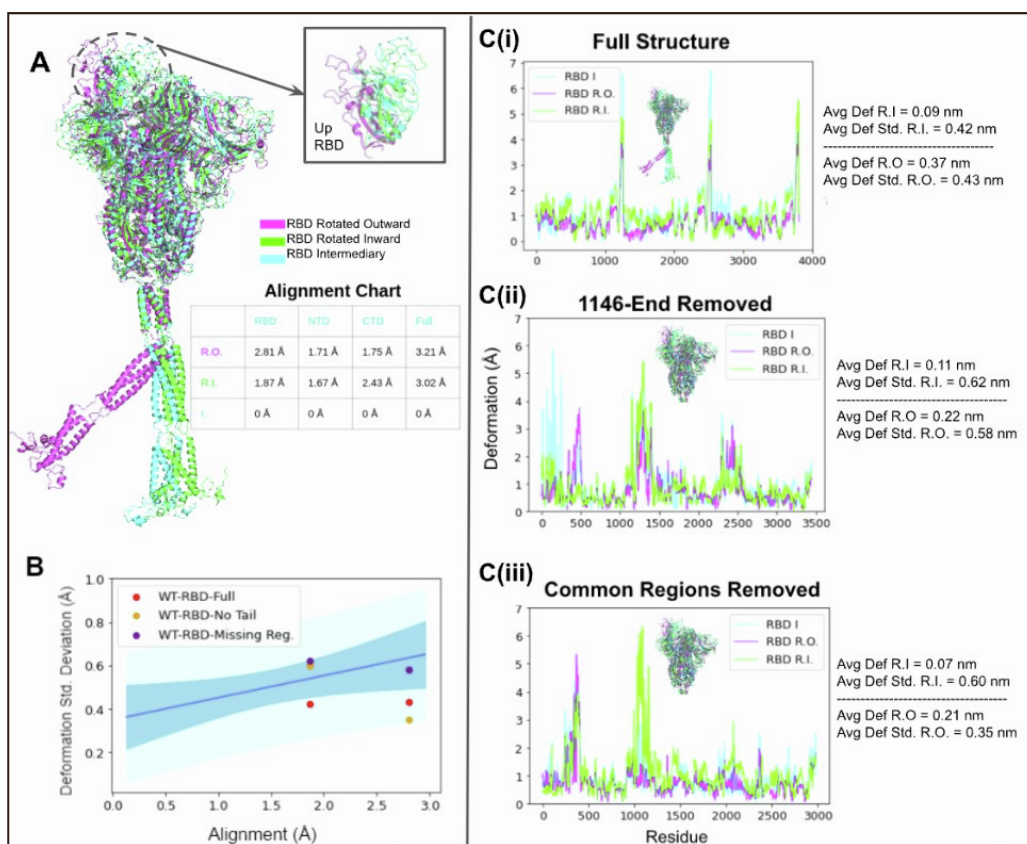
SI Figure 4: Visualization of WT'-B MD snapshots as compared to the baseline WT structure **(A)**. The results of the WT alignment analysis are plotted in reference to the linear regression model that is shown in **SI Figure 2-B (B)**. The deformation curves and their resulting alignment measurements are listed **(C)** for the full structure **(i)** structure with tail removed **(ii)** and structure with additional regions removed **(iii)**. These results show that the WT model used in our main analysis is able to capture the dynamic variability that may come from using an alternative MD model.

10. Supplementary Videos

Videos S1-S20 show the NMA trajectory and dynamic domain results for each analyzed spike protein. **Table S4** shows the PDB ID of the protein corresponding to each video.

Video ID	Protein PDB ID
S1	5W9K
S2	6NB6
S3	6VSB
S4	6VYB
S5	6X2B
S6	6X2C
S7	6XF6
S8	6XKL
S9	6XM0
S10	6XR8
S11	6ZGG
S12	6ZP7
S13	7AD1

S14	7BYR
S15	7C2L
S16	7CN9
S17	7JW0
S18	7K4N
S19	7K43
S20	7KDH

References

1. Wu, F., S. Zhao, B. Yu, Y.-M. Chen, W. Wang, Z.-G. Song, Y. Hu, Z.-W. Tao, J.-H. Tian, Y.-Y. Pei, M.-L. Yuan, Y.-L. Zhang, F.-H. Dai, Y. Liu, Q.-M. Wang, J.-J. Zheng, L. Xu, E.C. Holmes, and Y.-Z. Zhang. 2020. A new coronavirus associated with human respiratory disease in China. *Nature*. 579:265–269.
2. Wrapp, D., N. Wang, K.S. Corbett, J.A. Goldsmith, C.-L. Hsieh, O. Abiona, B.S. Graham, and J.S. McLellan. 2020. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*. 367:1260–1263.
3. Hsieh, C.-L., J.A. Goldsmith, J.M. Schaub, A.M. DiVenere, H.-C. Kuo, K. Javanmardi, K.C. Le, D. Wrapp, A.G.-W. Lee, Y. Liu, C.-W. Chou, P.O. Byrne, C.K. Hjorth, N.V. Johnson, J. Ludes-Meyers, A.W. Nguyen, J. Park, N. Wang, D. Amengor, J.A. Maynard, I.J. Finkelstein, and J.S. McLellan. 2020. Structure-based Design of Prefusion-stabilized SARS-CoV-2 Spikes. *bioRxiv*. 2020.05.30.125484.
4. Huo, J., Y. Zhao, J. Ren, D. Zhou, H.M.E. Duyvesteyn, H.M. Ginn, L. Carrique, T. Malinauskas, R.R. Ruza, P.N.M. Shah, T.K. Tan, P. Rijal, N. Coombes, K.R. Bewley, J.A. Tree, J. Radecke, N.G. Paterson, P.P. Supasa, J. Mongkolsapaya, G.R. Screaton, M. Carroll, A. Townsend, E.E. Fry, R.J. Owens, and D.I. Stuart. 2020. Neutralization of SARS-CoV-2 by Destruction of the Prefusion Spike. *Cell Host & Microbe*.
5. Walls, A.C., Y.-J. Park, M.A. Tortorici, A. Wall, A.T. McGuire, and D. Veesler. 2020. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell*. 181:281–292.e6.
6. Wrobel, A.G., D.J. Benton, P. Xu, C. Roustan, S.R. Martin, P.B. Rosenthal, J.J. Skehel, and S.J. Gamblin. 2020. SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects. *Nature Structural & Molecular Biology*. 27:763–767.
7. Bank, R.P.D. RCSB PDB - 6XF6: Cryo-EM structure of a biotinylated SARS-CoV-2 spike probe in the prefusion state (1 RBD up). .
8. Juraszek, J., L. Rutten, S. Blokland, P. Bouchier, R. Voorzaat, T. Ritschel, M.J.G. Bakkers, L.L.R. Renault, and J.P.M. Langedijk. 2020. Stabilizing the Closed SARS-CoV-2 Spike Trimer. *bioRxiv*. 2020.07.10.197814.
9. Henderson, R., R.J. Edwards, K. Mansouri, K. Janowska, V. Stalls, S.M.C. Gobeil, M. Kopp, D. Li, R. Parks, A.L. Hsu, M.J. Borgnia, B.F. Haynes, and P. Acharya. 2020. Controlling the SARS-CoV-2 spike glycoprotein conformation. *Nature Structural & Molecular Biology*. 27:925–933.
10. A Carbohydrate-Binding Protein from the Edible Lablab Beans Effectively Blocks the Infections of Influenza Viruses and SARS-CoV-2: Cell Reports. .
11. Zhou, D., H.M.E. Duyvesteyn, C.-P. Chen, C.-G. Huang, T.-H. Chen, S.-R. Shih, Y.-C. Lin, C.-Y. Cheng, S.-H. Cheng, Y.-C. Huang, T.-Y. Lin, C. Ma, J. Huo, L. Carrique, T. Malinauskas, R.R. Ruza, P.N.M. Shah, T.K. Tan, P. Rijal, R.F. Donat, K. Godwin, K.R. Buttigieg, J.A. Tree, J. Radecke, N.G. Paterson, P. Supasa, J. Mongkolsapaya, G.R. Screaton, M.W. Carroll, J. Gilbert-Jaramillo, M.L. Knight, W. James, R.J. Owens, J.H. Naismith, A.R. Townsend, E.E. Fry, Y. Zhao, J. Ren, D.I. Stuart, and K.-Y.A. Huang. 2020. Structural basis for the neutralization of SARS-CoV-2 by an antibody from a convalescent patient. *Nature Structural & Molecular Biology*. 27:950–958.
12. Cai, Y., J. Zhang, T. Xiao, H. Peng, S.M. Sterling, R.M. Walsh, S. Rawson, S. Rits-Volloch, and B. Chen. 2020. Distinct conformational states of SARS-CoV-2 spike protein. *Science*. 369:1586–1592.
13. Gobeil, S.M.-C., K. Janowska, S. McDowell, K. Mansouri, R. Parks, K. Manne, V. Stalls, M. Kopp, R. Henderson, R.J. Edwards, B.F. Haynes, and P. Acharya. 2020. D614G mutation alters SARS-CoV-2 spike conformational dynamics and protease cleavage susceptibility at the S1/S2 junction. *bioRxiv*. 2020.10.11.335299.
14. Melero, R., C.O.S. Sorzano, B. Foster, J.-L. Vilas, M. Martínez, R. Marabini, E. Ramírez-Aportela, R. Sanchez-Garcia, D. Herreros, L. del Caño, P. Losana, Y.C. Fonseca-Reyna, P. Conesa, D. Wrapp, P. Chacon, J.S. McLellan, H.D. Tagare, and J.-M. Carazo. 2020. Continuous flexibility analysis of SARS-CoV-2 Spike prefusion structures. *bioRxiv*. 2020.07.08.191072.

15. Rota, P.A., M.S. Oberste, S.S. Monroe, W.A. Nix, R. Campagnoli, J.P. Icenogle, S. Peñaranda, B. Bankamp, K. Maher, M. Chen, S. Tong, A. Tamin, L. Lowe, M. Frace, J.L. DeRisi, Q. Chen, D.Wang, D.D. Erdman, T.C.T. Peret, C. Burns, T.G. Ksiazek, P.E. Rollin, A. Sanchez, S. Liffick, B. Holloway, J. Limor, K. McCaustland, M. Olsen-Rasmussen, R. Fouchier, S. Günther, A.D.M.E. Osterhaus, C. Drosten, M.A. Pallansch, L.J. Anderson, and W.J. Bellini. 2003. Characterization of a Novel Coronavirus Associated with Severe Acute Respiratory Syndrome. *Science*. 300:1394–1399.

16. Bermingham, A., M.A. Chand, C.S. Brown, E. Aarons, C. Tong, C. Langrish, K. Hoschler, K. Brown, M. Galiano, R. Myers, R.G. Pebody, H.K. Green, N.L. Boddington, R. Gopal, N. Price, W. Newsholme, C. Drosten, R.A. Fouchier, and M. Zambon. 2012. Severe respiratory illness caused by a novel coronavirus, in a patient transferred to the United Kingdom from the Middle East, September 2012. *Euro Surveill*. 17:20290.

17. Wu, N.C., M. Yuan, H. Liu, C.-C.D. Lee, X. Zhu, S. Bangaru, J.L. Torres, T.G. Caniels, P.J.M.Brouwer, M.J. van Gils, R.W. Sanders, A.B. Ward, and I.A. Wilson. 2020. An alternative binding mode of IGHV3-53 antibodies to the SARS-CoV-2 receptor binding domain. *bioRxiv*. 2020.07.26.222232.

18. Liu, L., P. Wang, M.S. Nair, J. Yu, M. Rapp, Q. Wang, Y. Luo, J.F.-W. Chan, V. Sahi, A. Figueroa, X.V. Guo, G. Cerutti, J. Bimela, J. Gorman, T. Zhou, Z. Chen, K.-Y. Yuen, P.D. Kwong, J.G. Sodroski, M.T. Yin, Z. Sheng, Y. Huang, L. Shapiro, and D.D. Ho. 2020. Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature*. 584:450–456.

19. Yuan, M., H. Liu, N.C. Wu, C.-C.D. Lee, X. Zhu, F. Zhao, D. Huang, W. Yu, Y. Hua, H. Tien, T.F. Rogers, E. Landais, D. Sok, J.G. Jardine, D.R. Burton, and I.A. Wilson. 2020. Structural basis of a shared antibody response to SARS-CoV-2. *Science*. 369:1119–1123.

20. Cao, Y., B. Su, X. Guo, W. Sun, Y. Deng, L. Bao, Q. Zhu, X. Zhang, Y. Zheng, C. Geng, X. Chai, R. He, X. Li, Q. Lv, H. Zhu, W. Deng, Y. Xu, Y. Wang, L. Qiao, Y. Tan, L. Song, G. Wang, X. Du, N. Gao, J. Liu, J. Xiao, X. Su, Z. Du, Y. Feng, C. Qin, C. Qin, R. Jin, and X.S. Xie. 2020. Potent Neutralizing Antibodies against SARS-CoV-2 Identified by High-Throughput Single-Cell Sequencing of Convalescent Patients' B Cells. *Cell*. 182:73-84.e16.

21. Barnes, C.O., C.A. Jette, M.E. Abernathy, K.-M.A. Dam, S.R. Esswein, H.B. Gristick, A.G.Malyutin, N.G. Sharaf, K.E. Huey-Tubman, Y.E. Lee, D.F. Robbiani, M.C. Nussenzweig, A.P. West, and P.J. Bjorkman. 2020. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature*. 1–9.

22.. Du, S., Y. Cao, Q. Zhu, P. Yu, F. Qi, G. Wang, X. Du, L. Bao, W. Deng, H. Zhu, J. Liu, J. Nie, Y. Zheng, H. Liang, R. Liu, S. Gong, H. Xu, A. Yisimayi, Q. Lv, B. Wang, R. He, Y. Han, W. Zhao, Y. Bai, Y. Qu, X. Gao, C. Ji, Q. Wang, N. Gao, W. Huang, Y. Wang, X.S. Xie, X. Su, J. Xiao, and C. Qin. 2020. Structurally Resolved SARS-CoV-2 Antibody Shows High Efficacy in Severely Infected Hamsters and Provides a Potent Cocktail Pairing Strategy. *Cell*. 183:1013-1023.e13.

23. Hansen, J., A. Baum, K.E. Pascal, V. Russo, S. Giordano, E. Wloga, B.O. Fulton, Y. Yan, K. Koon, K. Patel, K.M. Chung, A. Hermann, E. Ullman, J. Cruz, A. Rafique, T. Huang, J. Fairhurst, C. Libertiny, M. Malbec, W. Lee, R. Welsh, G. Farr, S. Pennington, D. Deshpande, J. Cheng, A. Watty, P. Bouffard, R. Babb, N. Levenkova, C. Chen, B. Zhang, A.R. Hernandez, K. Saotome, Y. Zhou, M. Franklin, S. Sivapalasingam, D.C. Lye, S. Weston, J. Logue, R. Haupt, M. Frieman, G. Chen, W. Olson, A.J. Murphy, N. Stahl, G.D. Yancopoulos, and C.A. Kyratsous. 2020. Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science*. 369:1010–1014.

24. Lv, Z., Y.-Q. Deng, Q. Ye, L. Cao, C.-Y. Sun, C. Fan, W. Huang, S. Sun, Y. Sun, L. Zhu, Q. Chen, N. Wang, J. Nie, Z. Cui, D. Zhu, N. Shaw, X.-F. Li, Q. Li, L. Xie, Y. Wang, Z. Rao, C.-F. Qin, and X. Wang. 2020. Structural basis for neutralization of SARS-CoV-2 and SARS-CoV by a potent therapeutic antibody. *Science*. 369:1505–1509.

25. Bank, R.P.D. RCSB PDB - 6ZHD: H11-H4 bound to Spike. .

26. Tortorici, M.A., M. Beltramello, F.A. Lempp, D. Pinto, H.V. Dang, L.E. Rosen, M. McCallum, J. Bowen, A. Minola, S. Jaconi, F. Zatta, A. De Marco, B. Guarino, S. Bianchi, E.J. Lauron, H. Tucker, J. Zhou, A. Peter, C. Havenar-Daughton, J.A. Wojcechowskyj, J.B. Case, R.E. Chen, H. Kaiser, M. Montiel-Ruiz, M. Meury, N. Czudnochowski, R. Spreafico, J. Dillen, C. Ng, N. Sprugasci, K. Culap, F. Benigni, R.

Abdelnabi, S.-Y.C. Foo, M.A. Schmid, E. Cameroni, A. Riva, A. Gabrieli, M. Galli, M.S. Pizzuto, J. Neyts, M.S. Diamond, H.W. Virgin, G. Snell, D. Corti, K. Fink, and D. Veesler. 2020. Ultrapotent human antibodies protect against SARS-CoV-2 challenge via multiple mechanisms. *Science*. 370:950–957.

27. Chi, X., R. Yan, J. Zhang, G. Zhang, Y. Zhang, M. Hao, Z. Zhang, P. Fan, Y. Dong, Y. Yang, Z. Chen, Y. Guo, J. Zhang, Y. Li, X. Song, Y. Chen, L. Xia, L. Fu, L. Hou, J. Xu, C. Yu, J. Li, Q. Zhou, and W. Chen. 2020. A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science*. 369:650–655.

28. Hurlburt, N.K., Y.-H. Wan, A.B. Stuart, J. Feng, A.T. McGuire, L. Stamatatos, and M. Pancera. 2020. Structural basis for potent neutralization of SARS-CoV-2 and role of antibody affinity maturation. *bioRxiv*. 2020.06.12.148692.

29. A Therapeutic Non-self-reactive SARS-CoV-2 Antibody Protects from Lung Pathology in a COVID-19 Hamster Model: Cell. .

30. Bank, R.P.D. RCSB PDB - 7C8W: Structure of sybody MR17 in complex with the SARS-CoV-2 S receptor-binding domain (RBD). .

31. Ju, B., Q. Zhang, J. Ge, R. Wang, J. Sun, X. Ge, J. Yu, B. Zhou, S. Song, X. Tang, J. Yu, J. Lan, J. Yuan, H. Wang, J. Zhao, S. Zhang, Y. Wang, X. Shi, and L. Zhang. 2020. Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature*. 584:1–8.

32. Pinto, D., Y.-J. Park, M. Beltramello, A.C. Walls, M.A. Tortorici, S. Bianchi, S. Jaconi, K. Culap, F. Zatta, A.D. Marco, A. Peter, B. Guarino, R. Spreafico, E. Cameroni, J.B. Case, R.E. Chen, C. Havenar-Daughton, G. Snell, A. Telenti, H.W. Virgin, A. Lanzavecchia, M.S. Diamond, K. Fink, D. Veesler, and D. Corti. 2020. Structural and functional analysis of a potent sarbecovirus neutralizing antibody. *bioRxiv*. 2020.04.07.023903.

33. Piccoli, L., Y.-J. Park, M.A. Tortorici, N. Czudnochowski, A.C. Walls, M. Beltramello, C. Silacci-Fregni, D. Pinto, L.E. Rosen, J.E. Bowen, O.J. Acton, S. Jaconi, B. Guarino, A. Minola, F. Zatta, N. Sprugasci, J. Bassi, A. Peter, A.D. Marco, J.C. Nix, F. Mele, S. Jovic, B.F. Rodriguez, S.V. Gupta, F. Jin, G. Piumatti, G.L. Presti, A.F. Pellanda, M. Biggiogero, M. Tarkowski, M.S. Pizzuto, E. Cameroni, C. Havenar-Daughton, M. Smithey, D. Hong, V. Lepori, E. Albanese, A. Ceschi, E. Bernasconi, L. Elzi, P. Ferrari, C. Garzoni, A. Riva, G. Snell, F. Sallusto, K. Fink, H.W. Virgin, A. Lanzavecchia, D. Corti, and D. Veesler. 2020. Mapping Neutralizing and Immunodominant Sites on the SARS-CoV-2 Spike Receptor-Binding Domain by Structure-Guided High-Resolution Serology. *Cell*. 0.

34. Custódio, T.F., H. Das, D.J. Sheward, L. Hanke, S. Pazicky, J. Pieprzyk, M. Sorgenfrei, M.A. Schroer, A.Y. Gruzinov, C.M. Jeffries, M.A. Graewert, D.I. Svergun, N. Dobrev, K. Remans, M.A. Seeger, G.M. McInerney, B. Murrell, B.M. Hällberg, and C. Löw. 2020. Selection, biophysical and structural analysis of synthetic nanobodies that effectively neutralize SARS-CoV-2. *Nature Communications*. 11:5588.

35. Song, W., M. Gui, X. Wang, and Y. Xiang. 2018. Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. *PLOS Pathogens*. 14:e1007236.

36. Yuan, Y., D. Cao, Y. Zhang, J. Ma, J. Qi, Q. Wang, G. Lu, Y. Wu, J. Yan, Y. Shi, X. Zhang, and G.F. Gao. 2017. Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nature Communications*. 8:15092.

37. McLellan, J.S., M. Chen, M.G. Joyce, M. Sastry, G.B.E. Stewart-Jones, Y. Yang, B. Zhang, L. Chen, S. Srivatsan, A. Zheng, T. Zhou, K.W. Graepel, A. Kumar, S. Moin, J.C. Boyington, G.Y. Chuang, C. Soto, U. Baxa, A.Q. Bakker, H. Spits, T. Beaumont, Z. Zheng, N. Xia, S.Y. Ko, J.P. Todd, S. Rao, B.S. Graham, and P.D. Kwong. 2013. Structure-based design of a fusion glycoprotein vaccine for respiratory syncytial virus. *Science*. 342:592–598.

38. Chi, X., R. Yan, J. Zhang, G. Zhang, Y. Zhang, M. Hao, Z. Zhang, P. Fan, Y. Dong, Y. Yang, Z. Chen, Y. Guo, J. Zhang, Y. Li, X. Song, Y. Chen, L. Xia, L. Fu, L. Hou, J. Xu, C. Yu, J. Li, Q. Zhou, and W. Chen. 2020. A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science*. 369:650–655.

39. Walke, R.C., J.K. Longworth, R.H. Little, A. Venter, and B.W. Miller. 2004. The practical application of the AMBER software tool to support environmental decision making. *Multimedia Modeling*. 10.

40. Delgado, J., L.G. Radusky, D. Cianferoni, and L. Serrano. 2019. FoldX 5.0: working with RNA, small molecules and a new graphical interface. *Bioinformatics*. 35:4168–4169.

41. Gasteiger, E., C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, and A. Bairoch. 2005. Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM, editor. The Proteomics Protocols Handbook. Totowa, NJ: Humana Press. pp. 571–607.

42. Kawashima, S., P. Pokarowski, M. Pokarowska, A. Kolinski, T. Katayama, and M. Kanehisa. 2008. AAindex: amino acid index database, progress report 2008. *Nucleic Acids Res*. 36:D202-205.

43. Cheng, J., A.Z. Randall, M.J. Sweredoski, and P. Baldi. 2005. SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Res*. 33:W72-76.

44. L. Casalino *et al.*, Beyond Shielding: The Roles of Glycans in SARS-CoV-2 Spike Protein. *bioRxiv*, (2020).

45. *Protein structure prediction and analysis using the Robetta server | Nucleic Acids Research | Oxford Academic*.

46. T. Zhou *et al.*, Cryo-EM Structures Delineate a pH-Dependent Switch that Mediates Endosomal Positioning of SARS-CoV-2 Spike Receptor-Binding Domains. *bioRxiv*, 2020.2007.2004.187989 (2020).

47. S. M. C. Gobeil *et al.*, D614G mutation alters SARS-CoV-2 spike conformational dynamics and protease cleavage susceptibility at the S1/S2 junction. *bioRxiv*, 2020.2010.2011.335299 (2020).

48. M. C. Gobeil Sophie *et al.*, Effect of natural mutations of SARS-CoV-2 on spike structure, conformation, and antigenicity. *Science* **373**, eabi6226.

49. M. J. Abraham *et al.*, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19-25 (2015).

50. G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity-rescaling. *The Journal of Chemical Physics* **126**, 014101 (2007).

51. R. Martonák, A. Laio, M. Parrinello, Predicting Crystal Structures: The Parrinello-Rahman Method Revisited. *Physical review letters* **90**, 075503 (2003).

52. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: AnN·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*. 98(12):10089-10092, doi: 10.1063/1.464397, https://dx.doi.org/10.1063/1.464397.

53. Toxvaerd, S., O. J. Heilmann, and J. C. Dyre. 2012. Energy conservation in molecular dynamics simulations of classical systems. *The Journal of Chemical Physics*. 136(22):224106, doi: 10.1063/1.4726728, https://dx.doi.org/10.1063/1.4726728.

54. B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **18**, 1463-1472 (1997).

55. Huang, J., S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller, and A. D. MacKerell. 2017. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature Methods*. 14(1):71-73, doi: 10.1038/nmeth.4067, https://www.nature.com/articles/nmeth.4067