Supplementary Information for

Free energy landscapes from SARS-CoV-2 spike glycoprotein simulations suggest that RBD opening can be modulated via interactions in an allosteric pocket

Lucy Fallon[1,2#], Kellon A.A. Belfon[1,2#], Lauren Raguette[1,2], Yuzhang Wang[1,2], Darya Stepanenko[1,4], Abbigayle Cuomo[2], Jose Guerra[5], Stephanie Budhan[2], Sarah Varghese[6], Christopher Corbo[3], Robert C. Rizzo[1,4], and Carlos Simmerling[1,2]*

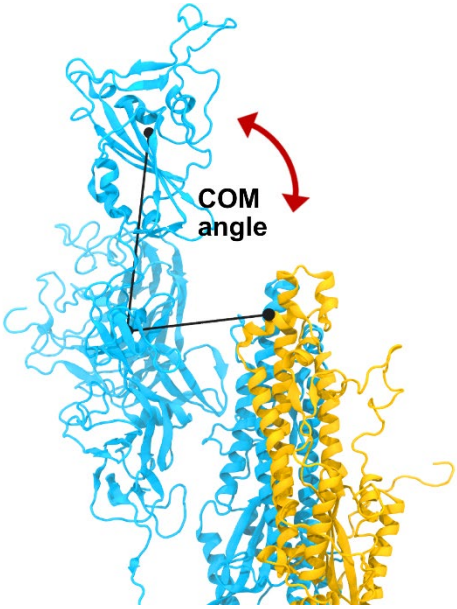
[1]Laufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook, New York 11794, United States
[2]Department of Chemistry, Stony Brook University, Stony Brook, New York 11794, United States
[3]Graduate Program in Molecular and Cellular Pharmacology, Stony Brook University, Stony Brook, New York 11794, United States
[4]Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, New York 11794, United States
[5]Department of Biochemistry and Cell Biology, Stony Brook University, Stony Brook, New York 11794, United States
[6]Undergraduate Program in Biology, Stony Brook University, Stony Brook, New York 11794, United States

# These authors contributed equally to this work.

Organization

---

1. Figures



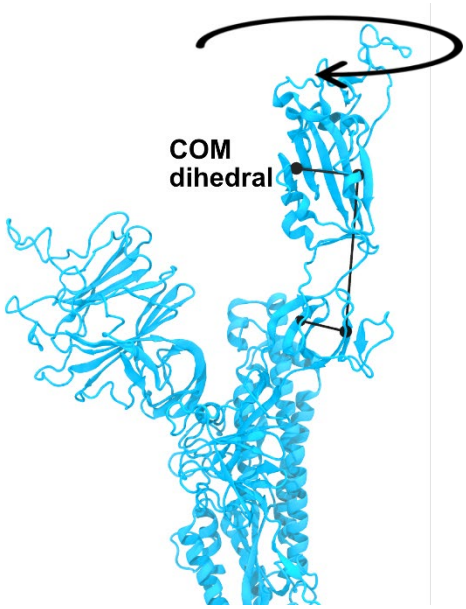


Figure S1– Center of mass (COM) based collective variables (CVs) used to quantify RBD motion relative to the remainder of the spike. (Left) the COM angle measures movement of the RBD up and away from the core central helices, and (right) the COM dihedral measures rotation of the RBD relative to the CTD1 domain. Portions of 2 protomers are shown, with coloring matching **Figure 4**. Amino acid ranges used to define the COM groups (black circles) are shown in **Table S4**. Additional details are provided in Methods.

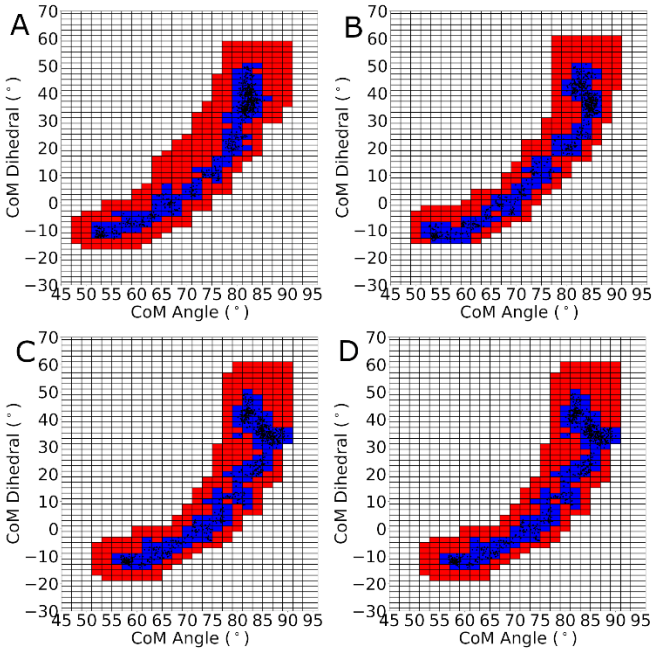


Figure S2: 2D grid of CoM angle and CoM dihedral collective variables. The colored grids were used as windows for US simulations to calculate the free energy for RBD opening in A) WT, B) K528A, C) A522L, and D) A522V S-protein. The initial grids used as US windows before grid expansion are shown in blue, additional grids used as new windows for expansion are shown in red. The SMD trajectory snapshots are mapped onto the 2D grid as black dots.

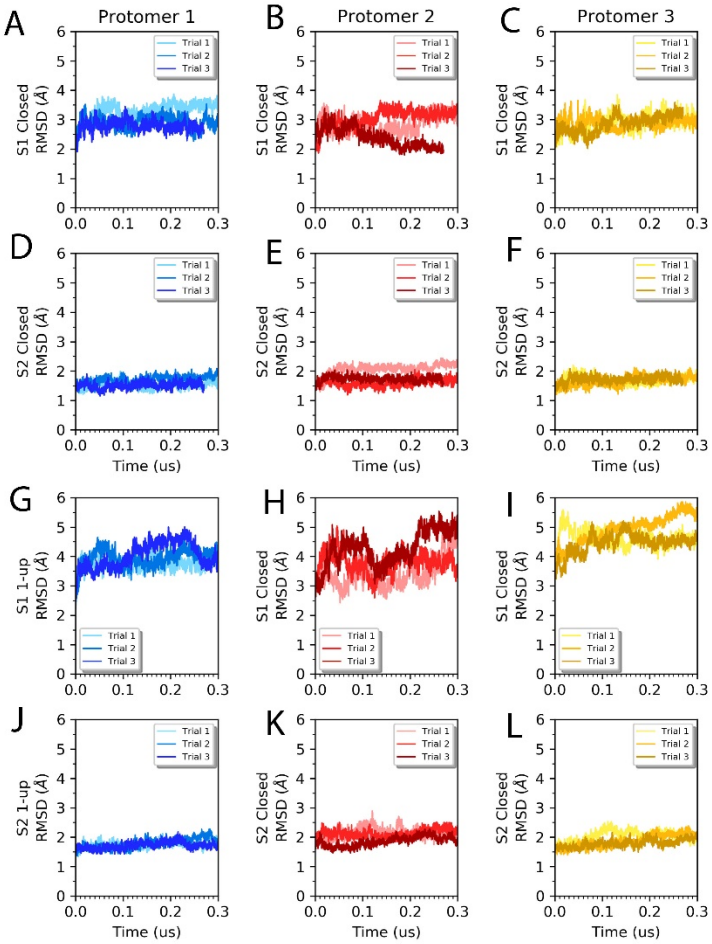


Figure S3. RMSD of backbone atoms of each protomer (blue, red, yellow) over three MD runs of the spike over 270, 300 and 300 ns. (A,B,C) S1, closed spike system; (D,E,F) S2, closed; (G,H,I) S1, 1-up; (J,K,L) S2, 1-up. Each shade in a plot represents one of three independent trials. Columns represent protomers 1, 2 and 3 with colors matching **Figure 4**. RMSD is relative to the respective reference structure.

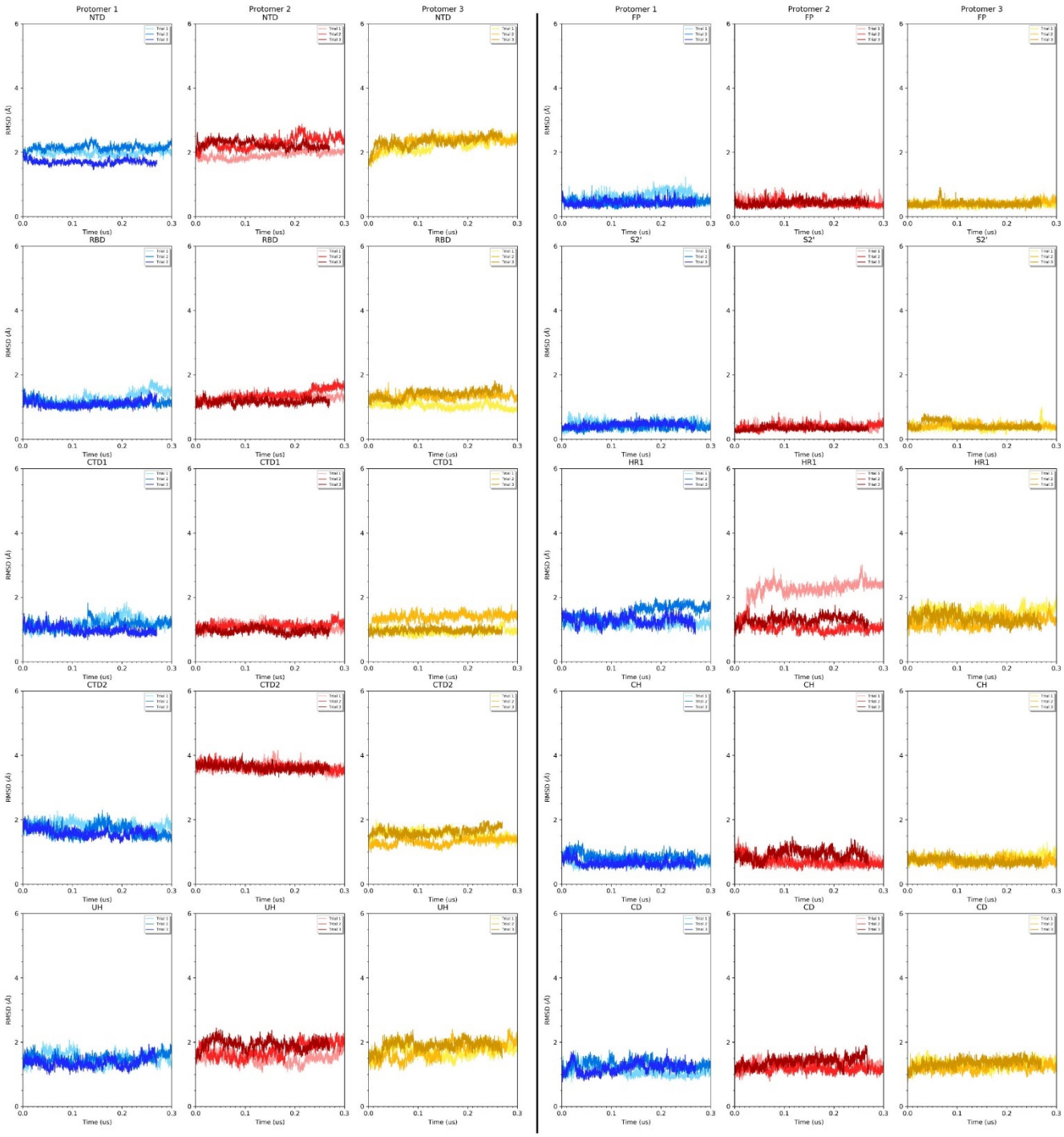


Figure S4. Domain RMSDS during MD simulations of the closed spike. Residue ranges for each domain are provided in Table S3. Rows correspond to specific domains. Columns show data for each protomer, with 3 independent runs in different shades. RMSD is relative to the respective reference structure.

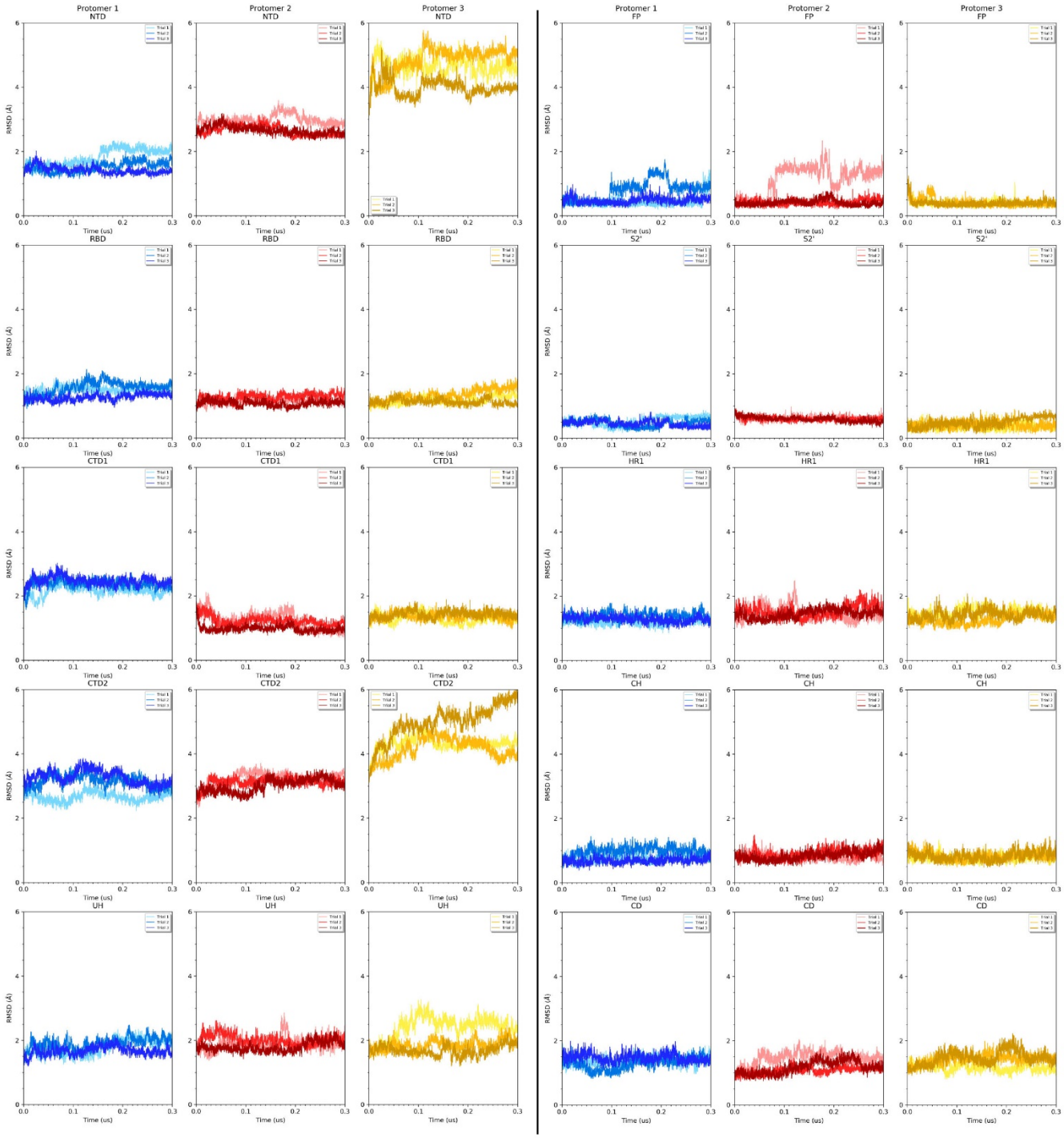


Figure S5. Domain RMSDS during MD simulations of the 1-up spike. Residue ranges for each domain are provided in Table S3. Columns show data for each protomer, with 3 independent runs in different shades. Protomer 1 has the RBD in the open position. RMSD is relative to the respective reference structure.

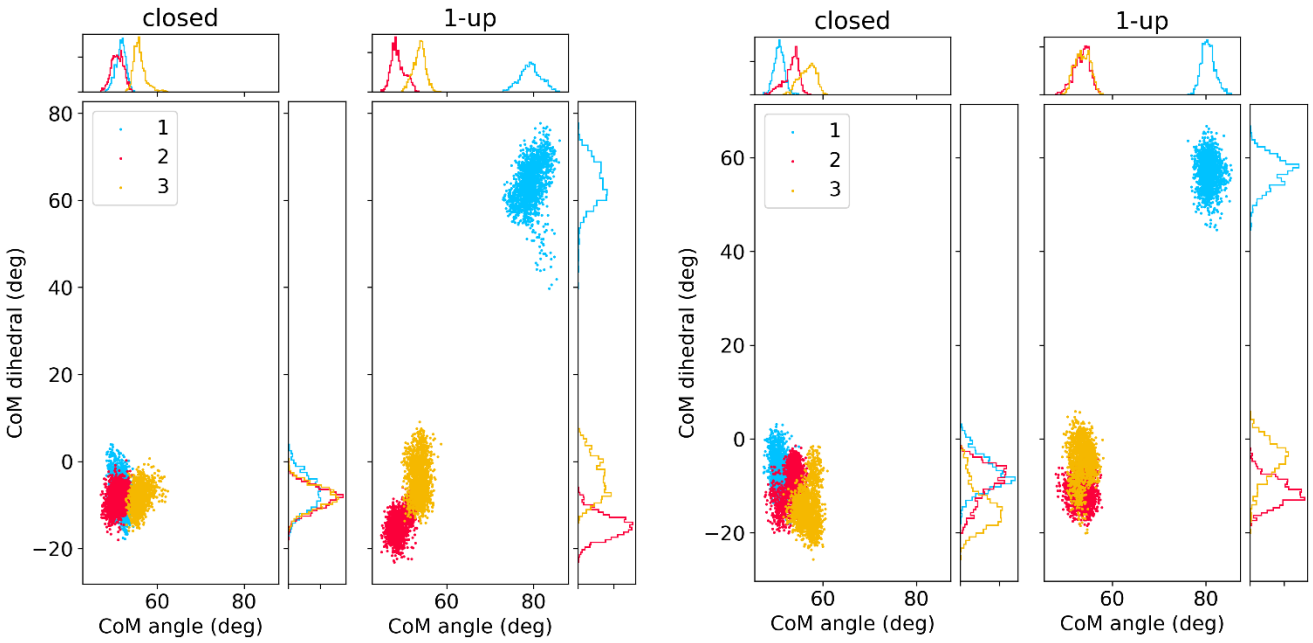


Figure S6. Collective variables for 2 additional independent runs of closed and 1-up spike. Simulation times: 300 ns (left), 270 ns (right).

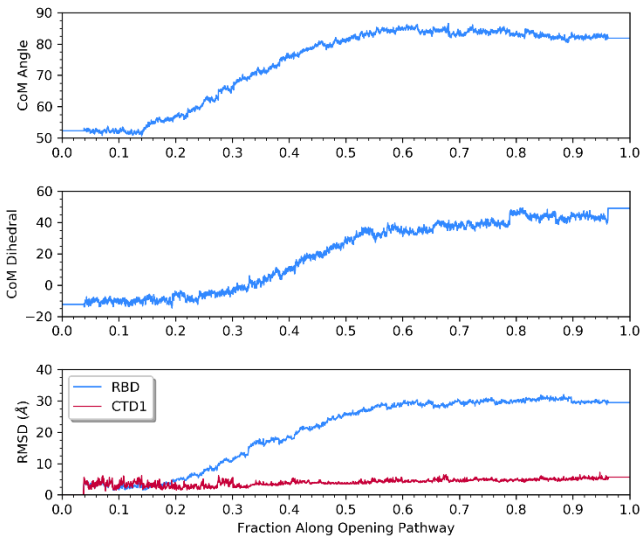


Figure S7. Spike structure properties during the RBD opening pathway, with the closed state on the left and the 1-up spike on the right. (Top) RBD opening angle, (middle) RBD rotation dihedral, (bottom) RMSD values for RBD and CTD1 domains, after best-fit to central helices (CH).

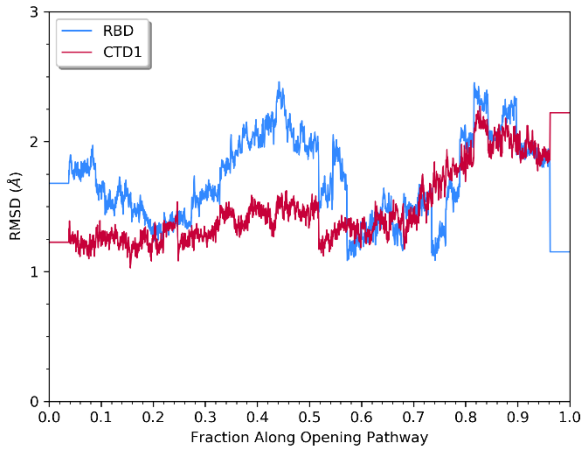


Figure S8. RMSD of RBD and CTD1 domains as a function of opening pathway, with each domain best-fit to the same atoms in the open reference model, using the residues listed in **Table S3**.

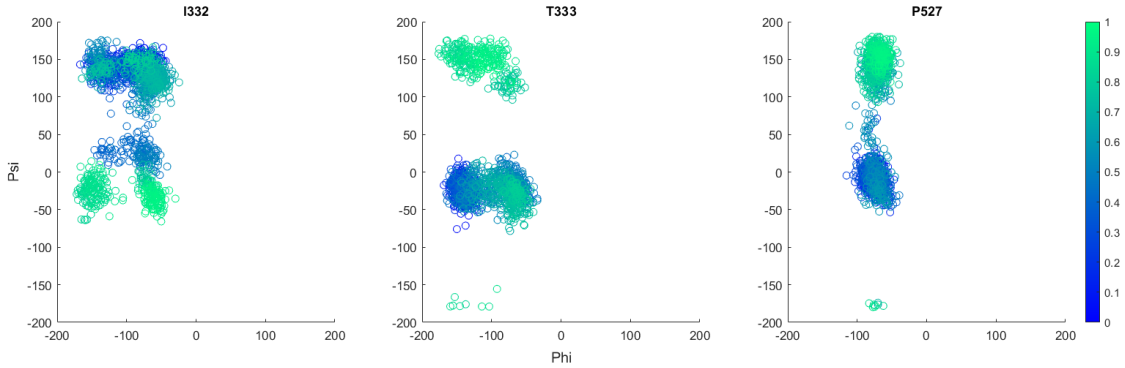


Figure S9. Selected backbone dihedrals in the RBD-CTD1 connector strands as a function of RBD opening pathway. Snapshots are shown as circles, with the opening progress indicated by the color range.

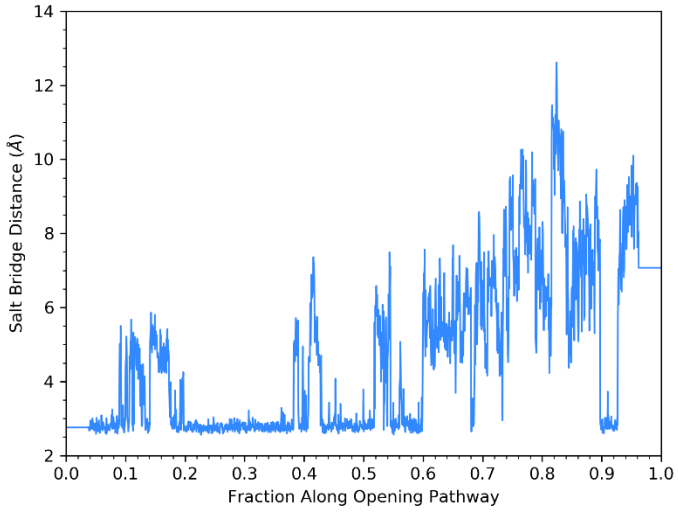


Figure S10. Distance between sidechains of K528 (Nζ) and D389 (Cδ) as a function of RBD opening pathway.

Figure S11. Sequence conservation for amino acids in the hinge pocket region (including positions 324-591), across 21 different coronaviruses.

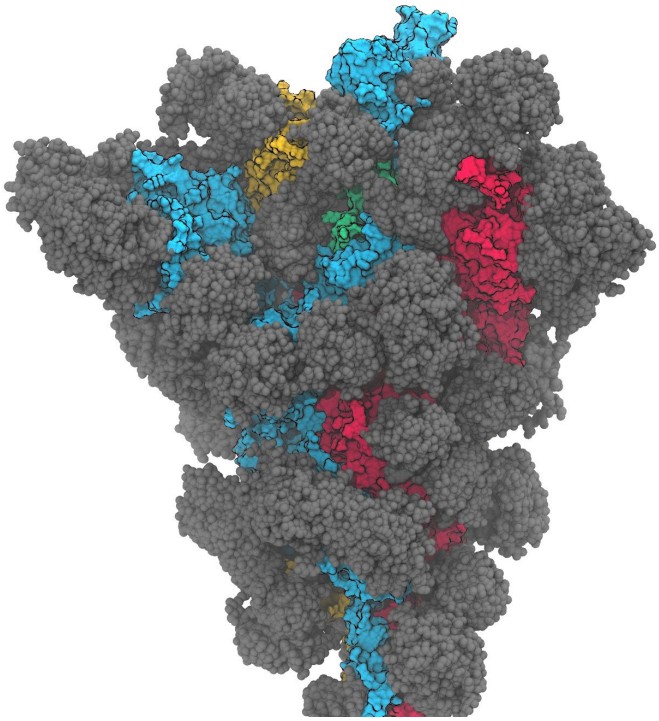


Figure S12. Spike structure in MD simulations, comparable to **Figure 3**. The gray glycans effectively shield the exterior of the conserved hinge region (highlighted in green).

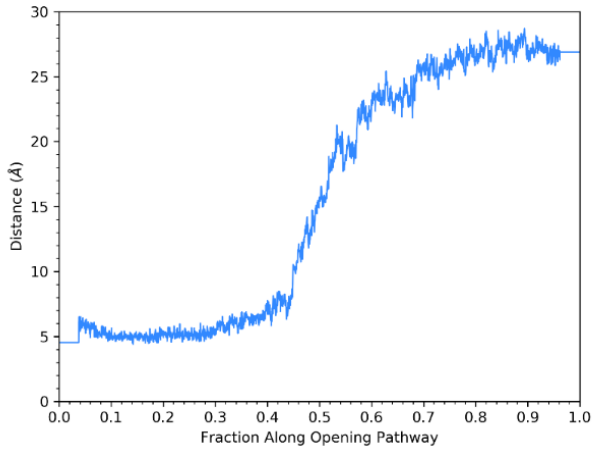


Figure S13. Distance between Cα atoms of A520 in RBD and Q564 in CTD1 as a function of RBD opening pathway for full spike;

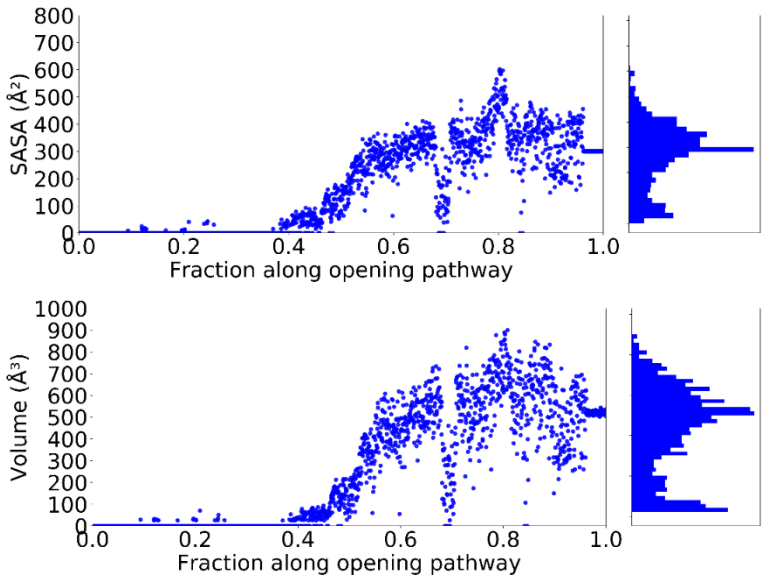


Figure S14: The change in hinge pocket volume and SASA as a function of RBD opening pathway. Data were calculated from the NEB snapshots using fpocket.[1]

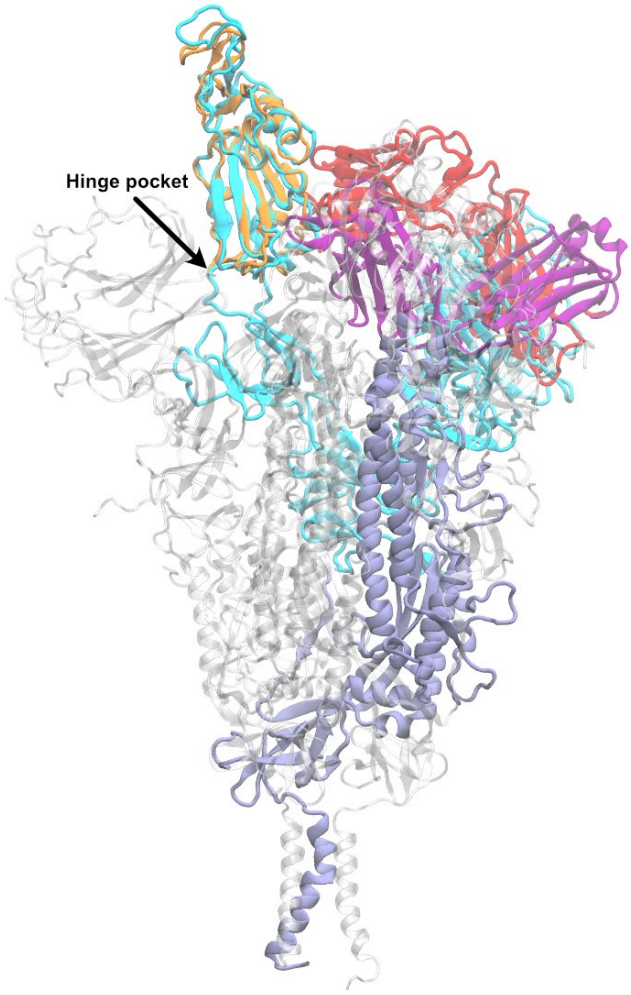


Figure S15. Spike protein from 1-up MD (1 protomer shown with S1 in cyan and S2 in gray, other protomers transparent), with hinge pocket indicated in relation to location of binding of cryptic antibody CR3022 from crystal structure 6W41[2] (RBD in orange, antibody in red/purple). The crystal structure contained only the RBD, not the entire spike. The RBD domains of the 2 systems were best-fit.

2. Tables

Table S1. Structures and types of glycans on each chain.[3]

Type	Chain A	Chain B	Chain C	Sequence
M5	N61, N122, N657	N61, N165, N603, N709, N717, N1074	N61, N122, N603	aDMan(1→6)[aDMan(1→3)]aDMan(1→6)[aDMan(1→3)]bDMan(1→4)bDGlcNAc(1→4)bDGlcNAc(1→)
M6	N709, N801		N717	aDMan(1→6)[aDMan(1→3)]aDMan(1→6)[aDMan(1→2)aDMan(1→3)]bDMan(1→4)bDGlcNAc(1→4)bDGlcNAc(1→)
M7		N801		aDMan(1→2)aDMan(1→2)aDMan(1→3)[aDMan(1→6)[aDMan(1→3)]aDMan(1→6)]bDMan(1→4)bDGlcNAc(1→4)bDGlcNAc(1→)
M8	N234			aDMan(1→2)aDMan(1→6)[aDMan(1→3)]aDMan(1→6)[aDMan(1→2)aDMan(1→2)aDMan(1→3)]bDMan(1→4)bDGlcNAc(1→4)bDGlcNAc(1→)
M9		N234	N234	aDMan(1→2)aDMan(1→6)[aDMan(1→2)aDMan(1→3)]aDMan(1→6)[aDMan(1→2)aDMan(1→2)aDMan(1→3)]bDMan(1→4)bDGlcNAc(1→4)bDGlcNAc(1→)
A2	N616, N1158	N1098	N74, N282, N1158	bDGlcNAc(1→2)aDMan(1→6)[bDGlcNAc(1→2)aDMan(1→3)]bDMan(1→4)bDGlcNAc(1→4)bDGlcNAc(1→)
A3	N74			bDGlcNAc(1→6)[bDGlcNAc(1→2)]aDMan(1→6)[bDGlcNAc(1→2)aDMan(1→3)]bDMan(1→4)bDGlcNAc(1→4)bDGlcNAc(1→)
FA1	N1134	N343		bDGlcNAc(1→2)aDMan(1→3)[aDMan(1→6)]bDMan(1→4)bDGlcNAc(1→4) [aLFuc(1→6)]bDGlcNAc(1→)
FA2	N17, N331, N343, N603, N1098	N122, N331, N616	N149, N343, N616, N709, N801, N1074, N1134	bDGlcNAc(1→2)aDMan(1→6)[bDGlcNAc(1→2)aDMan(1→3)]bDMan(1→4)bDGlcNAc(1→4)[aLFuc(1→6)]bDGlcNAc(1→)
FA2G2S1	N149, N1074	N1158	N165	aDNeu5Ac(2→6)bDGal(1→4)bDGlcNAc(1→2)aDMan(1→6)[bDGal(1→4)bDGlcNAc(1→2)aDMan(1→3)]bDMan(1→4)bDGlcNAc(1→4) [aLFuc(1→6)]bDGlcNAc(1→)
FA2G2S2	N165			xaDNeu5Ac(2→6)bDGal(1→4)bDGlcNAc(1→2)aDMan(1→6)[aDNeu5Ac(2→6)bDGal(1→4)bDGlcNAc(1→2)aDMan(1→3)]bDMan(1→4)bDGlcNAc(1→4) [aLFuc(1→6)]bDGlcNAc(1→)
FA3	N282	N17, N149, N1134	N17	bDGlcNAc(1→6)[bDGlcNAc(1→2)]aDMan(1→6)[bDGlcNAc(1→2)aDMan(1→3)]bDMan(1→4)bDGlcNAc(1→4)[aLFuc(1→6)]bDGlcNAc(1→)

FA3G3S1		N282	N331	aDNeu5Ac(2→6)bDGal(1→4)bDGlcNAc(1→2)aDMan(1→3)[bDGal(1→4)bDGlcNAc(1→6) [bDGal(1→4)bDGlcNAc(1→2)]aDMan(1→6)]bDMan(1→4)bDGlcNAc(1→4) [aLFuc(1→6)]bDGlcNAc(1→)
FA3G3S2		N74		aDNeu5Ac(2→6)bDGal(1→4)bDGlcNAc(1→6)[bDGal(1→4)bDGlcNAc(1→2)]aDMan(1→6) [aDNeu5Ac(2→6)bDGal(1→4)bDGlcNAc(1→2)aDMan(1→3)]bDMan(1→4)bDGlcNAc(1→4) [aLFuc(1→6)]bDGlcNAc(1→)
FA4	N1173	N1173	N1173	bDGlcNAc(1→6)[bDGlcNAc(1→2)]aDMan(1→6)[bDGlcNAc(1→4) [bDGlcNAc(1→2)]aDMan(1→3)]bDMan(1→4)bDGlcNAc(1→4)[aLFuc(1→6)]bDGlcNAc(1→)
FA4G4S1	N1194	N1194	N1194	aDNeu5Ac(2→6)bDGal(1→4)bDGlcNAc(1→6)[bDGal(1→4)bDGlcNAc(1→2)]aDMan(1→6) [bDGal(1→4)bDGlcNAc(1→4) [bDGal(1→4)bDGlcNAc(1→2)]aDMan(1→3)]bDMan(1→4)bDGlcNAc(1→4) [aLFuc(1→6)]bDGlcNAc(1→)
Hybrid G1	N717	N657	N657	bDGal(1→4)bDGlcNAc(1→2)aDMan(1→3)[aDMan(1→6) [aDMan(1→3)]aDMan(1→6)]bDMan(1→4)bDGlcNAc(1→4)bDGlcNAc(1→)
Hybrid G1S1			N1098	aDNeu5Ac(2→6)bDGal(1→4)bDGlcNAc(1→2)aDMan(1→3)[aDMan(1→6) [aDMan(1→3)]aDMan(1→6)]bDMan(1→4)bDGlcNAc(1→4)bDGlcNAc(1→)
O1			T323	bDGal(1→3)aDGalNAc(1→)
O2	T323, S325			aDNeu5Ac(2→3)bDGal(1→3)aDGalNAc(1→)
O3		T323		aDNeu5Ac(2→3)bDGal(1→3)[aDNeu5Ac(2→6)]aDGalNAc(1→)

Table S2. Disulfide bonds in the "amended models". Rows in bold indicate the five disulfide bonds in the RBD and CTD1 domains, also included in the RBD+CTD1 construct.

SARS-CoV-2
C131 - C166
C291 - C301
C336 - C361
C379 - C432
C391 - C525
C480 - C488
C538 - C590
C617 - C649
C662 - C671
C730 - C768
C743 - C749
C1032 - C1043
C1082 - C1126

Table S3. Reside ranges used for RMSD calculations

Subunit	Domain/ Region of interest	Residues in domain	Residues for RMSD
S1	NTD	16-285	27-66, 81-140, 164-172, 186-196, 200-211, 215-242, 263-285
	RBD	335-530	335-454, 462-466, 491-515, 522-530
	CTD1	315-334, 531-591	315-334, 531-591
	CTD2	592-685	592-620, 641-676
S2	UH	686-813	691-811, 813
	S2' (KRSF)	814-817	814-817
	FP	818-825	818-825
	HR1	867-984	867-984
	CH	985-1036	985-1036
	CD	1037-1137	1037-1137

Table S4. Amino acids used to define center-of-mass collective variables for the RBD closed-open transition. The group is defined by the center of mass of all Cα atoms of all amino acids in the range.

CoM groups	1	2	3	4
Angle	F338-L517	E324-V327, C538-L585	T747*-Q755*	
Dihedral	E324-V327, C538-F543, L546-T549	A575-D578, L582-D586	N354-I358, T396-S399	S375-C379, C432-W436

Asterisks indicates the neighboring protomer on the clockwise side when viewing the spike from the S1 side (top side)

Table S5. List of sequences used to calculate conservation scores. We gratefully acknowledge the following Authors from the Originating laboratories responsible for obtaining the specimens and the Submitting laboratories where genetic sequence data were generated and shared via the GISAID Initiative, on which this research is based. All submitters of data may be contacted directly via www.gisaid.org[4]

Accession ID	**Virus Name**	**Location**	**Date Collected**	**Host**	**Originating Lab**	**Submitting Lab**	**Authors**
EPI_ISL_402124	hCoV-19/Wuhan/WIV04/2019 **(SARS-CoV-2)**	Asia / China / Hubei / Wuhan	2019-12-30	*Homo sapiens*	Wuhan Jinyintan Hospital	Wuhan Institute of Virology, Chinese Academy of Sciences	Peng Zhou, Xing-Lou Yang, Ding-Yu Zhang, Lei Zhang, Yan Zhu, Hao-Rui Si, Zhengli Shi

EPI_ISL_410539	hCoV-19/pangolin/Guangxi/P1E/2017	Asia / China / Guangxi	2017	Manis javanica	Beijing Institute of Microbiology and Epidemiology	Beijing Institute of Microbiology and Epidemiology	Wu-Chun Cao; Tommy Tsan-Yuk Lam; Na Jia; Ya-Wei Zhang; Jia-Fu Jiang; Bao-Gui Jiang
EPI_ISL_410541	hCoV-19/pangolin/Guangxi/P5E/2017	Asia / China / Guangxi	2017	Manis javanica	Beijing Institute of Microbiology and Epidemiology	Beijing Institute of Microbiology and Epidemiology	Wu-Chun Cao; Tommy Tsan-Yuk Lam; Na Jia; Ya-Wei Zhang; Jia-Fu Jiang; Bao-Gui Jiang
EPI_ISL_410540	hCoV-19/pangolin/Guangxi/P5L/2017	Asia / China / Guangxi	2017	Manis javanica	Beijing Institute of Microbiology and Epidemiology	Beijing Institute of Microbiology and Epidemiology	Wu-Chun Cao; Tommy Tsan-Yuk Lam; Na Jia; Ya-Wei Zhang; Jia-Fu Jiang; Bao-Gui Jiang
EPI_ISL_410538	hCoV-19/pangolin/Guangxi/P4L/2017	Asia / China / Guangxi	2017	Manis javanica	Beijing Institute of Microbiology and Epidemiology	Beijing Institute of Microbiology and Epidemiology	Wu-Chun Cao; Tommy Tsan-Yuk Lam; Na Jia; Ya-Wei Zhang; Jia-Fu Jiang; Bao-Gui Jiang
EPI_ISL_410542	hCoV-19/pangolin/Guangxi/P2V/2017	Asia / China / Guangxi	2017	Manis javanica	Beijing Institute of Microbiology and Epidemiology	Beijing Institute of Microbiology and Epidemiology	Wu-Chun Cao; Tommy Tsan-Yuk Lam; Na Jia; Ya-Wei Zhang; Jia-Fu Jiang; Bao-Gui Jiang
EPI_ISL_410721	hCoV-19/pangolin/Guangdong/1/2019	Asia / China / Guangdong	2019	Manis javanica	South China Agricultural University	South China Agricultural University	Yongyi Shen, Lihua Xiao, Wu Chen
NC_019843.3	Middle East respiratory syndrome-related coronavirus			*Homo sapiens*			van Boheemen,S., de Graaf,M., Lauber,C., Bestebroer,T.M., Raj,V.S.,Zaki,A.M., Osterhaus,A.D., Haagmans,B.L., Gorbalenya,A.E.,Snijder,E.J. and Fouchier,R.A.
NP_073551.1	Human coronavirus 229E			*Homo sapiens*			Thiel,V., Herold,J. and Siddell,S.G.

YP_009555241.1	Human coronavirus OC43			*Homo sapiens*			St-Jean,J.R., Jacomy,H., Desforges,M., Vabret,A., Freymuth,F. And Talbot,P.J.
AVP78042.1	Bat SARS-like coronavirus			*Rhinolophus pusillus*			Hu,D., Zhu,C., Ai,L., He,T., Wang,Y., Ye,F., Yang,L., Ding,C.,Zhu,X., Lv,R., Zhu,J., Hassan,B., Feng,Y., Tan,W. and Wang,C.
APO40579.1	Severe acute respiratory syndrome-related coronavirus			*Rhinolophus sp.*			Tao,Y. and Tong,S.
YP_003858584.1	Bat coronavirus BM48-31/BGR/2008			*Rhinolophus blasii*			Drexler,J.F., Corman,V.M. and Drosten,C.
AHX37558.1	*Rhinolophus affinis Coronavirus*			*Rhinolophus affinis*			He,B., Zhang,Y., Xu,L., Yang,W., Yang,F., Feng,Y., Xia,L., Zhou,J., Zhen,W., Feng,Y., Guo,H., Zhang,H. and Tu,C.
AAZ67052.1	Bat Coronavirus RaTG13			*Bat*			Li,W., Shi,Z., Yu,M., Ren,W., Smith,C., Epstein,J.H., Wang,H.,Crameri,G., Hu,Z., Zhang,H., Zhang,J., McEachern,J., Field,H.,Daszak,P., Eaton,B.T., Zhang,S. and Wang,L.F.
ABD75332.1	Bat SARS CoV Rm1/2004			*Rhinolophus macrotis*			Li,W., Shi,Z., Yu,M., Ren,W., Smith,C., Epstein,J.H., Wang,H., Crameri,G., Hu,Z., Zhang,H., Zhang,J., McEachern,J., Field,H., Daszak,P., Eaton,B.T., Zhang,S. and Wang,L.F.
YP_009825051.1	SARS Coronavirus			*Homo sapiens*			He,R., Dobie,F., Ballantine,M.,

							Leeson,A., Li,Y., Bastien,N., Cutts,T., Andonov,A., Cao,J., Booth,T.F., Plummer,F.A., Tyler,S., Baker,L. and Li,X.
ATO98108.1	Bat SARS-like Coronavirus			*Aselliscus stoliczkanu*			Hu,B., Zeng,L.P., Yang,X.L., Ge,X.Y., Zhang,W., Li,B., Xie,J.Z., Shen,X.R., Zhang,Y.Z., Wang,N., Luo,D.S., Zheng,X.S., Wang,M.N., Daszak,P., Wang,L.F., Cui,J. and Shi,Z.L.
ATO98145.1	Bat SARS-like Coronavirus			*Rhinolophus ferrumequinum*			Hu,B., Zeng,L.P., Yang,X.L., Ge,X.Y., Zhang,W., Li,B., Xie,J.Z., Shen,X.R., Zhang,Y.Z., Wang,N., Luo,D.S., Zheng,X.S., Wang,M.N., Daszak,P., Wang,L.F., Cui,J. and Shi,Z.L.
ATO98157.1	Bat SARS-like Coronavirus			*Rhinolophus sinicus*			Hu,B., Zeng,L.P., Yang,X.L., Ge,X.Y., Zhang,W., Li,B., Xie,J.Z., Shen,X.R., Zhang,Y.Z., Wang,N., Luo,D.S., Zheng,X.S., Wang,M.N., Daszak,P., Wang,L.F., Cui,J. and Shi,Z.L.
AAY88866.1	Bat SARS Coronavirus HKU3-1			*Bat*			Lau,S.K., Woo,P.C., Li,K.S., Huang,Y., Tsoi,H.W., Wong,B.H., Wong,S.S., Leung,S.Y., Chan,K.H. and Yuen,K.Y.
AGC74165.1	Bat Coronavirus Rp/Shaanxi2011			*Rhinolophus pusillus*			Yang,L., Wu,Z., Ren,X., Yang,F., He,G., Zhang,J., Dong,J., Sun,L., Zhu,Y., Du,J., Zhang,S. and Jin,Q.

AGC74176.1	Bat Coronavirus Cp/Yunnan2011			*Chaerephon plicata*			Yang,L., Wu,Z., Ren,X., Yang,F., He,G., Zhang,J., Dong,J., Sun,L., Zhu,Y., Du,J., Zhang,S. and Jin,Q.
AID16716.1	Bat SARS-like coronavirus			*Rhinolophus monoceros*			Lin,X.D., Wang,W., Hao,Z.Y., Wang,Z.X., Guo,W.P., Guan,X.Q., Wang,M.R., Wang,H.W., Zhou,R.H., Li,M.H., Tang,G.P., Wu,J., Holmes,E.C. and Zhang,Y.Z.
AIA62277.1	BtRf-BetaCoV/JL2012			*Rhinolophus ferrumequinum*			Wu,Z., Yang,L., Ren,X., He,G., Zhang,J., Yang,J., Qian,Z., Dong,J., Sun,L., Zhu,Y., Du,J., Yang,F., Zhang,S. and Jin,Q.
AIA62310.1	BtRs-BetaCoV/HuB2013			*Rhinolophus sinicus*			Wu,Z., Yang,L., Ren,X., He,G., Zhang,J., Yang,J., Qian,Z., Dong,J., Sun,L., Zhu,Y., Du,J., Yang,F., Zhang,S. and Jin,Q.
QHR63300.2	Bat Coronavirus RaTG13			*Rhinolophus affinis*			Zhou,P., Yang,X.-L., Wang,X.-G., Hu,B., Zhang,L., Zhang,W.,Si,H.-R., Zhu,Y., Li,B., Huang,C.-L., Chen,H.-D., Chen,J., Luo,Y., Guo,H., Jiang,R.-D., Liu,M.-Q., Chen,Y., Shen,X.-R., Wang,X., Zheng,X.-S., Zhao,K., Chen,Q.-J., Deng,F., Liu,L.-L., Yan,B., Zhan,F.-X., Wang,Y.-Y., Xiao,G.-F. and Shi,Z.-L.
AVP78031.1	Bat SARS-like Coronavirus			*Rhinolophus pusillus*			Hu,D., Zhu,C., Ai,L., He,T., Wang,Y., Ye,F., Yang,L., Ding,C., Zhu,X., Lv,R., Zhu,J.,

							Hassan,B., Feng,Y., Tan,W. and Wang,C.

Table S6. Calculated sequence conservation percentages, across 28 coronavirus sequences, for the RBD and CTD1.

position	% conserved	sequence (SARS-CoV-2 listed first)
324	7.	EIIIIIEGVAQTTKQQGQVKQQTQVQDQ
325	43.	SSSSSSSSSDSEEEEEDEDEEEEEDESS
326	36.	IIIIIIIVIVIVVVVVVVVVVVVVVVIV
327	79.	VVVVVVVVVVYVVVVIVVVVVIVIIAVVV
328	93.	RRRRRRRESRRRRRRRRRRRRRRRRRR
329	89.	FFFFFFFQLRFFFFFFFFFFFFFFFFF
330	93.	PPPPPPPAPKPPPPPPPPPPPPPPPP
331	89.	NNNNNNNE-PNNNNNNNNNNNNNNNNN
332	89.	IIIIIIIG-LIIIIIIIIIIIIIIIII
333	89.	TTTTTTTV-PTTTTTTTTTTTTTTTTT
334	89.	NNNNNNNE-NNNQNNNNNNNNNNNNNN
335	50.	LLLLLLL---VLLLRRLRLLRRRRVRLV
336	96.	CCCCCCCC-CCCCCCCCCCCCCCCCCC
337	89.	PPPPPPPD-NPPPPPPPPPPPPPPPPP
338	93.	FFFFFFFF-IFFFFFFFFFFFFFFFFF
339	43.	GGGGGGGS-EHGNGDDGDDGDDDDDGH
340	43.	EEEEEEEP-AKQEEKKEKAEKKRKKREK
341	93.	VVVVVVVLVWVVVVVVVVVVVVVVVVV
342	89.	FFFFFFFLYLFFFFFFFFFFFFFFFF
343	93.	NNNNNNNSHNNNNNNNNNNNNNNNNN
344	82.	AAAAAAAGKDAAIAAAAAAAAAAVAAAA
345	61.	TSSSSSTTHKTSTTTSTTTTTTSTTSTT
346	46.	RKKKKKT-TSRNSTRRKRRTRRRRRTR
347	93.	FFFFFFFPFVFFFFFFFFFFFFFFFF
348	29.	AAAAAAAPIPPPPPPPPPPPPPPPAP
349	79.	SSSSSSSQVSSSSSNNSSSSNSSNSSSS
350	93.	VVVVVVVVLPVVVVVVVVVVVVVVVVV
351	96.	YYYYYYYYYLYYYYYYYYYYYYYYYY
352	89.	AAAAAAAN-NAAAAAAAAAAAAAAAAA
353	93.	WWWWWWWF-WWWWWWWWWWWWWWWWW
354	29.	NNNNNNNK-EEEEEEEEEEEEEEEENE
355	96.	RRRRRRRR-RRRRRRRRRRRRRRRRRR
356	43.	KKKKKKKL-KTLMKTTKTVKTTTTTKT
357	43.	RRRRRRRV-TKRRRKKKKRKKKKKRK
358	89.	IIIIIIIF-FIIIIIIIIIIIIIIII

359	89.	SSSSSSST-SSSTSSSSSSSSSSSSSS
360	54.	NNNNNNNN-NDDNNDDNDNNDDDDDDND
361	96.	CCCCCCCCC-CCCCCCCCCCCCCCCCCCC
362	82.	VVVVVVVN-NIVVVVVVVVVVVVVVVVI
363	89.	AAAAAAAYVFAAAAAAAAAAAAAAAAA
364	93.	DDDDDDDNDNDDDDDDDDDDDDDDDD
365	89.	YYYYYYYLFMYYYYYYYYYYYYYYYYY
366	46.	SSSSSSSTKSTASSTTSTTSTTTTTST
367	86.	VVVVVVVKPSVVVVVVVVAVVVVVVVV
368	82.	LLLLLLLLQLFLLLLLLLFLLLLFLLF
369	89.	YYYYYYYLSMYYYYYYYYYYYYYYYYY
370	89.	NNNNNNNSGSNNNNNNNNNNNNNNNN
371	89.	SSSSSSSLGFSSSSSSSSSSSSSSSS
372	7.	ATTTTTTFGITSATTTTTTTTTTTTTT
373	89.	SSSSSSSS-QSSSSSSFSSSSSSSSSS
374	89.	FFFFFFFV-AFFFFFFFFFFFFFFFFF
375	89.	SSSSSSSN-DSSSSSSSSSSSSSSSSS
376	89.	TTTTTTTD-STTTTTTTTTTTTTTTTT
377	96.	FFFFFFFF-FFFFFFFFFFFFFFFFFF
378	86.	KKKKKKKTKTKKQKKKKKKKKKKNKKK
379	100.	CCCCCCCCCCCCCCCCCCCCCCCCCCCC
380	89.	YYYYYYYSFNYYYYYYYYYYYYYYYYY
381	89.	GGGGGGGQNNGGGGGGGGGGGGGGGG
382	89.	VVVVVVVICIVVVVVVVVVVVVVVVVV
383	93.	SSSSSSSSYDSSSSSSSSSSSSSSSS
384	86.	PPPPPPPPPAPPPAPPAPPAPPPPPPP
385	43.	TTTTTTTAAASTTISSTSSTSSSSSTS
386	93.	KKKKKKKAGKKKKKKKKKKKKKKKKK
387	89.	LLLLLLLI-ILLLLLLLLLLLLLLLLL
388	46.	NNNNNNNAVYINNNIINIINIIIIIINI
389	89.	DDDDDDDSNGDDDDDDDDDDDDDDDDD
390	89.	LLLLLLLNIMLLLLLLLLLLLLLLLL
391	96.	CCCCCCCCTCCCCCCCCCCCCCCCCCCC
392	93.	FFFFFFFYLFFFFFFFFFFFFFFFFF
393	71.	TTTTTTTSASTSSSTTSTTSTTTTTTT
394	43.	NNNNNNNSNSSSSSNSSNSSNSSSSSSNS
395	89.	VVVVVVVL-IVVVVVVVVVVVVVVVVV
396	89.	YYYYYYYIFTYYYYYYYYYYYYYYYYY
397	89.	AAAAAAALNIAAAAAAAAAAAAAAAAA
398	96.	DDDDDDDDEDDDDDDDDDDDDDDDDD
399	39.	SSSSSSSYTKTYYSTTSTTSTTTTTST
400	96.	FFFFFFFF-FFFFFFFFFFFFFFFFFF

401	46.	VVVVVVVS-ALVVVLLVLLVLLLLLLVL
402	54.	IVVVVVVY-IIVVVIIVIIVIIIIIIII
403	50.	RKKKKKRPKPRKKKRRKRRKRRRRRRTR
404	50.	GGGGGGGLGNFGGGSSGSFGSSFSFSGF
405	46.	DDDDDDDSPGSDDDSSDSSDSSSSSSDS
406	71.	EEEEEEEMLREDDDEEDEEDEEEEEEEE
407	89.	VVVVVVVKCKVVVVVVVVVVVVVVVVVV
408	89.	RRRRRRRSVVRRRRRRRRRRRRRRRRR
409	89.	QQQQQQQDDDQQQQQQQQQQQQQQQQQ
410	50.	IIIIIIILTLVIIIVVIVVIVVIVVVIV
411	89.	AAAAAAASSQAAAAAAAAAAAAAAAAA
412	89.	PPPPPPPVHLPPPPPPPPPPPPPPPP
413	86.	GGGGGGGSFGGAAGGGGGGGGGGGGG
414	57.	QQQQQQQSTNQQQQEEQEEQEEEEQEQQ
415	93.	TTTTTTTATLTTTTTTTTTTTTTTTT
416	96.	GGGGGGGGKGGGGGGGGGGGGGGGGG
417	7.	KVVVVVRPYYVVVVVVVVVVVVVVVKV
418	93.	IIIIIIIIVLIIIIIIIIIIIIIIII
419	93.	AAAAAAASAQAAAAAAAAAAAAAAAA
420	89.	DDDDDDDQVSDDDDDDDDDDDDDDDD
421	93.	YYYYYYYFYFYYYYYYYYYYYYYYYY
422	96.	NNNNNNNN-NNNNNNNNNNNNNNNNN
423	96.	YYYYYYYY-YYYYYYYYYYYYYYYYY
424	93.	KKKKKKKK-RKKKKKKKKKKKKKKK
425	89.	LLLLLLLQ-ILLLLLLLLLLLLLLLL
426	89.	PPPPPPPS-DPPPPPPPPPPPPPPPP
427	89.	DDDDDDDF-TDDDDDDDDDDDDDDDD
428	86.	DDDDDDDS-TDDDDDDDDDDDEDDDD
429	89.	FFFFFFFN-AFFFFFFFFFFFFFFFF
430	79.	TTTTTTTP-TTTTMTTMTTLTTTTITT
431	89.	GGGGGGGTASGGGGGGGGGGGGGGGG
432	96.	CCCCCCCCNCCCCCCCCCCCCCCCCC
433	93.	VVVVVVVLVQVVVVVVVVVVVVVVVV
434	79.	IIIIIIIIGLILILIILIILIIIIIII
435	89.	AAAAAAALRYAAAAAAAAAAAAAAAA
436	93.	WWWWWWWAWYWWWWWWWWWWWWWWW
437	93.	NNNNNNN--NNNNNNNNNNNNNNNNN
438	29.	SSSSSSSV--TTTTTTTTTTTTTTST
439	18.	NVVVVVNPAPANNRAARAANAAAAAAKA
440	21.	NKKKKKNHSAKSSNKQNQKSKNNKKKHK
441	11.	LQQQQQLNIAQVLIQQIQQKHQQQQQIQ
442	89.	DDDDDDDLNNDDDDDDDDDDDDDDDD

443	18.	SAAAAASTTVTSSAQQAKVSTQRIVTAV
444	14.	KLLLLLKTGSGK-TGGTGGSGGGGGGKG
445	11.	VTTTTTVI-V-S-S--S--T------E-
446	29.	GGGGGGGT-S---S--T--S------G-
447	36.	GGDGGGG--S---G--G--G------G-
448	43.	NNNNNNNKNT-GSN--N--N------N-
449	32.	YYYYYYYPCW-NNF--Y--Y------F-
450	39.	N-GGGGNKPNHNENQQNQSNNQQNSYNN
451	82.	Y-YYYYYYFKYFFYYYYYYYYYYYYYYY
452	32.	LLLLLLLSSRFYFKYYKYFLYYYYFYLF
453	93.	YYYYYYYYFFYYYYYYYYYYYYYYYYY
454	89.	RRRRRRRI-GRRRRRRRRRRRRRRRRR
455	32.	LLLLLLLN-FSLRSSSYSSWSSSSSSLS
456	36.	FFFFFFFK-IHFFLHYLSHVHSSHHHFH
457	89.	RRRRRRRC-ERRRRRRRRRRRRRRRRR
458	57.	KKKKKKKS-DSHHHKKHKSRKKKKSKKS
459	39.	SSSSSSSRGSTGGGTEGTSSTETTSTAT
460	11.	NKKKKKNLKVKKKKKKKKKKKKKKKNK
461	86.	LLLLLLLLVFLIILLLLLLLLLLLLLL
462	82.	KKKKKKK--KKKKRKKRKKNKKKKKKK
463	89.	PPPPPPPDNNPPPPPPPPPPPPPPPP
464	82.	FFFFFFFDF-FYYFFFFFFYFFFFFFF
465	86.	EEEEEEERVHEEGEEEEEEEEEEEEE
466	89.	RRRRRRRTKDRRRRRRRRRRRRRRRR
467	89.	DDDDDDD--VDDDDDDDDDDDDDDDD
468	39.	IIIIIIIVFYLILILLILLLLLLLLIL
469	89.	SSSSSSSPGASSSSSSSSSSSSSSSS
470	29.	TTTTTTTQSQSNNNSSNSSNSSSSSTS
471	32.	EEEEEEELVHDVVVDDVDDDDDDEDED
472	32.	IIIIIIIVCCELLPEEPEEIDEEDEDIE
473	36.	YYYYYYYNFF-YFF--F--YG--G-GY-
474	29.	QQQQQQQASK-NNS--S--S------Q-
475	32.	AAAAAAANLA-SPP--P--P------A-
476	32.	GGGGGGGQKP-ASD--D--G------G-
477	29.	SSSSSSSYIK-GGG--G--G------S-
478	25.	TTTTTTTS-N-GGK--K--Q------K-
479	39.	PPPPPPPP-F-TTP--P--S------P-
480	54.	CCCCCCCC-CNCCC--C--C------CN
481	29.	NNNNNNNV-P-SS----------N-
482	36.	GGGGGGGS-CGS-T--T--S------GG
483	14.	VQQQQQV---VIAP--P--A------QV
484	11.	EVVVVVEIPKRSEP--P--I------TR

485	39.	GGGGGGGVGL-QGA--A--G------G-
486	11.	FLLLLLFPGN-LLF--L--P------L-
487	79.	NNNNNNNSCG-GNNNNNNNNNNNNNNN-
488	46.	CCCCCCCTA--CCCGGCGGCGGGGGGC-
489	46.	YYYYYYYVM--YYYVVYVVYVVVVVVY-
490	7.	FYYYYYFWP--EKWRYWRRNYYRYLYY-
491	46.	PPPPPPPEISTPPPTTPTTPTTTTTTPT
492	89.	LLLLLLLDVCLLLLLLLLLLLLLLLLL
493	7.	QEEEEEQGAVSKANSSNSSRSSSSSSYS
494	14.	SRRRRRSDNGTSSDTTDTTPTTTTTTRT
495	93.	YYYYYYYYWSYYYYYYYYYYYYYYYYY
496	50.	GGGGGGGQAGDGGGDDGDDGDDDDDDGD
497	89.	FFFFFFFWYPFFFFFFFFFFFFFFFFF
498	4.	QHHHHHHLSGNTTYYYYYNFNYYNNNYN
499	71.	PPPPPPPV-KPPQTPPTPPTPPPPQPPP
500	46.	TTTTTTTA-NNTSTSSTTNTNSSNNNTN
501	14.	NTTTTTNS-NVVSNVITVVAVVVVVVDV
502	54.	GGGGGGGG-GPGGGPPGPPGPPPPPPGP
503	54.	VVVVVVVS-ILVIIVVIILVVLLVLVVL
504	32.	GNNNNNGTKGEGGGAEGEDGADEAEAGE
505	82.	YYYYYYYVYTYYFYYYYYYHYYYYYYHY
506	89.	QQQQQQQAYCQQQQQQQQQQQQQQQQQQ
507	50.	PPPPPPPMTPAPPPAAPAAPAAAAAAPA
508	29.	YFFFFFYTIATYYYTTYTTYTTTTTTYT
509	89.	RRRRRRREGGRRRRRRRRRRRRRRRRR
510	89.	VVVVVVVQSTVVVVVVVVVVVVVVVVV
511	89.	VVVVVVVLLNVVVVVVVVVVVVVVVVV
512	89.	VVVVVVVQYYVVVVVVVVVVVVVVVVV
513	93.	LLLLLLLMVLLLLLLLLLLLLLLLLLL
514	93.	SSSSSSSGSTSSSSSSSSSSSSSSSS
515	86.	FFFFF-F--PFFFFFFFFFFFFFFFFF
516	89.	EEEEEEE--QEEEEEEEEEEEEEEEE
517	89.	LLLLLLL--ALLLLLLLLLLLLLLLL
518	89.	LLLLLLL--FLLLLLLLLLLLLLLLL
519	4.	HNNNNNN--LNNNNNNNNNNNNNNNN
520	71.	AGGGGGA--GAAAAAAAAAAAAAAAA
521	89.	PPPPPPP--WPPPPPPPPPPPPPPPP
522	89.	AAAAAAA--SAAAAAAAAAAAAAAAA
523	89.	TTTTTTT--DTTTTTTTTTTTTTTTT
524	89.	VVVVVVV--SVVVVVVVVVVVVVVVV
525	93.	CCCCCCC--CCCCCCCCCCCCCCCCC
526	93.	GGGGGGGDGLGGGGGGGGGGGGGGGG

527	89.	PPPPPPPTVQPPPPPPPPPPPPPPPP
528	96.	KKKKKKKKSKKKKKKKKKKKKKKKKK
529	11.	KLLLLLQISALKQLLLLLLLLLLLLKL
530	89.	SSSSSSSAFNSSSSSSSSSSSSSSSS
531	93.	TTTTTTTSMTTTTTTTTTTTTTTTTT
532	14.	NTTTTTNQNDQEEDQQDGQDETSQPENQ
533	89.	LLLLLLL-VILLLLLLLLLLLLLLLLL
534	75.	VVVVVVV-TIVVVIVVIVVIVVIVVVVV
535	86.	KKKKKKKLLLKKKTKKKKKKKKKKKKK
536	71.	NDDDDDNGDGNNNNNNNNNNNNNNNN
537	39.	KKKKKKKNKVQKKQQQQQQQQQQQQQKQ
538	100.	CCCCCCCCCCCCCCCCCCCCCCCCCCCC
539	96.	VVVVVVVVTVVVVVVVVVVVVVVVVVV
540	93.	NNNNNNNEKNNNNNNNNNNNNNNNNN
541	89.	FFFFFFFYYYFFFFFFFFFFFFFFFF
542	93.	NNNNNNNSNDNNNNNNNNNNNNNNNN
543	89.	FFFFFFFLILFFFFFFFFFFFFFFFF
544	89.	NNNNNNNYYYNNNNNNNNNNNNNNNN
545	96.	GGGGGGGGDGGGGGGGGGGGGGGGGG
546	89.	LLLLLLLVVILLLLLLLLLLLLLLLL
547	46.	TTTTTTTSSLKTTTKRTKKTKKKKKTK
548	100.	GGGGGGGGGGGGGGGGGGGGGGGGGG
549	89.	TTTTTTTRVQTTTTTTTTTTTTTTTT
550	100.	GGGGGGGGGGGGGGGGGGGGGGGGGG
551	96.	VVVVVVVVVIVVVVVVVVVVVVVVVV
552	89.	LLLLLLLFIFLLLLLLLLLLLLLLLL
553	89.	TTTTTTTQRVTTTTTTTTTTTTTTTT
554	18.	ETTTTTENVEDSNPETPDDSSADSDKED
555	93.	SSSSSSSCSVSSSSSSSSSSSSSSSS
556	14.	NKKKKKSTNNSTTLSSSSSSSSSSSNS
557	89.	KKKKKKKADAKKKKKKKKKKKKKKKK
558	25.	KQQQQQKVTTRKKRRRRRRRRKKRTRKR
559	93.	FFFFFFFGFYFFFFFFFFFFFFFFFF
560	32.	LLLLLLLVLYQQQQQQQQQQQQQQQLQ
561	46.	PPPPPPPRNNSPPPSSPSSPSSSSSPS
562	89.	FFFFFFF--SFFFFFFFFFFFFFFFF
563	96.	QQQQQQQQIQQQQQQQQQQQQQQQQQQ
564	89.	QQQQQQQRTNQQQQQQQQQQQQQQQQQ
565	93.	FFFFFFFFYLFFFFFFFFFFFFFFFF
566	89.	GGGGGGGVTLGGGGGGGGGGGGGGGG
567	82.	RRRRRRRYSYKRRRRRRRRRRRRRRK
568	96.	DDDDDDDDTDDDDDDDDDDDDDDDDD

569	29.	IIIIIIIASSAVVFTTVTTVTAATATIA
570	11.	ASSSSSAYGNSSSSSSSSSSSSSSSAS
571	89.	DDDDDDDQ--DDDDDDDDDDDDDDDDD
572	29.	TTTTTTTLLNFFFFFFFFFFFFFFFTF
573	82.	TTTTTTTVLLITTTTTTTTTTTTTTTI
574	89.	DDDDDDD--YDDDDDDDDDDDDDDDDD
575	29.	AAAAAAAGGGSSSSSSSSSSSSSSSAS
576	89.	VVVVVVVYFFVVVVVVVVVVVVVVVVV
577	93.	RRRRRRRYKRRRRRRRRRRRRRRRRR
578	96.	DDDDDDDSDDDDDDDDDDDDDDDDDD
579	89.	PPPPPPPDVYPPPPPPPPPPPPPPPP
580	71.	QQQQQQQDTIQKKKQQKQQKQQQQQQQ
581	93.	TTTTTTTNKTTTTTTTTTTTTTTTTT
582	79.	LLLLLLL-GNLFLLLLSLLSLLLLLLL
583	75.	EEEEEEE-TREEEEEEEQDEEEQEQEEE
584	86.	IIIIIIIYITIIIVIIIVIIIIIIIII
585	89.	LLLLLLLYYFLLLLLLLLLLLLLLLL
586	89.	DDDDDDDCSMDDDDDDDDDDDDDDDD
587	96.	IIIIIIILIIIIIIIIIIIIIIIIII
588	46.	TTTTTTTRTRTSASSSSTTSSSSSSTT
589	93.	PPPPPPPAPSPPPPPPPPPPPPPPPP
590	100.	CCCCCCCCCCCCCCCCCCCCCCCCCC
591	86.	SSSSSSSVNYSSSSSSASSSSSSSSSS

REFERENCES

1. Le Guilloux, V.; Schmidtke, P.; Tuffery, P., Fpocket: An open source platform for ligand pocket detection. *BMC Bioinformatics* **2009,** *10* (1), 168.
2. Yuan, M.; Wu, N. C.; Zhu, X.; Lee, C.-C. D.; So, R. T. Y.; Lv, H.; Mok, C. K. P.; Wilson, I. A., A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science* **2020,** *368* (6491), 630-633.
3. Casalino, L.; Gaieb, Z.; Goldsmith, J. A.; Hjorth, C. K.; Dommer, A. C.; Harbison, A. M.; Fogarty, C. A.; Barros, E. P.; Taylor, B. C.; McLellan, J. S.; Fadda, E.; Amaro, R. E., Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein. *ACS Central Science* **2020,** *6* (10), 1722-1734.
4. Elbe, S.; Buckland-Merrett, G., Data, disease and diplomacy: GISAID's innovative contribution to global health. *Global Challenges* **2017,** *1* (1), 33-46.