Significance of the RBD mutations in the SARS-CoV-2 Omicron: from spike opening to antibody escape and cell attachment

Md Lokman Hossen,[a] Prabin Baral,[a] Tej Sharma,[a] Bernard Gerstman,[a,b] and Prem Chapagain*[a,b]

[a]Department of Physics, Florida International University, Miami, FL 33199, USA
[b]Biomolecular Sciences Institute, Florida International University, Miami, FL 33199, USA

* Corresponding Author, Email: chapagap@fiu.edu

Supporting Information

Methods

System preparation

For the RBD interactions with the surrounding domains in the closed-form trimer, the initial frame of the simulation trajectory of the full spike protein trimer from the COVID-19 Data Sets of Amaro lab[1] (https://amarolab.ucsd.edu/covid19.php). This fully glycosylated spike trimer (RBD-down) was prepared from the cryo-EM structure (pdb ID 6VXX[2]). Because a full trimer structure of the Omicron spike protein was unavailable, we introduced the Omicron-specific mutations to the WT using CHARMM-GUI webserver[3, 4]. These Omicron mutations are: G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H in the RBD region (residues 330-530). We used these systems for investigating the RBD opening mechanism. For the WT RBD-ACE2 system, Cryo-EM structure (PDB ID 7A92) was used. For the Omicron RBD-ACE2 system, a recently deposited cryo-EM structure (PDB ID 7T9L)[5] was obtained from the Protein Data Bank. Similarly, the cryo-EM structure (PDB ID 7W9I)[6] was used for the Delta RBD-ACE2 simulations. The RBD-only system for Omicron was prepared with the RBD from the RBD-ACE2 complex (PDB ID 7T9L)[5]. The simulation system was set up with the same procedure as in our earlier studies[7, 8]. All the systems were prepared using the solution builder interface of CHARMM-GUI website[3, 4].

Molecular Dynamics (MD) Simulations

We performed molecular dynamics (MD) simulation following the procedures used in our previous work[8]. Briefly, constant pressure MD simulations were performed with NAMD 2.14[9] using CHARMM36 force-field[10, 11]. Following 10,000-step minimization and 2 ns equilibrations, the production runs were performed at 303.15K temperature and 1 atm pressure using 2fs timestep, and using Particle Mesh Ewald method to treat electrostatics[12, 13] and SHAKE algorithm to constrain covalent bonds involving hydrogen atoms[14]. Visual Molecular Dynamics (VMD)[15] 1.93 was used for the structure and trajectory visualization, as well as for the hydrogen-bond analysis. The information about the simulation systems and simulation lengths are given in Table S1. Hydrogen bond analysis was performed for the RBD-ACE2 complexes for the 100 ns of the trajectories using a cut-off of 3.5 Å and 30°.

Epitope predictions

The Spike protein sequence was retrieved from the NCBI (GenBank: QHD43416.1, accession MN908947.3), which is based upon the 2019-nCoV sequence[16]. Various independent methods were used for each of the MHC-I, MHC-II and B-cell epitopes : ProPred-I, CTLPred, and NetCTL1.2 for the prediction of MHC-I T-cell epitopes[17-19]; NetMHCII2.3 and EpiTOP3.0 for MHC-II T-cell epitopes[20, 21]; and BepiPred and BcePred[22, 23] for the B-cell epitopes. The T-cell epitope predictions are based on quantitative matrix and Quantitative Structure-Activity Relationship models (QSAR)[20], whereas the B-cell prediction method BepiPred uses a random forest algorithm trained on epitopes annotated from antibody-antigen structures and BcePred locates the epitopes using physiological properties such as hydrophilicity, polarity, and exposed surface. In addition to the sequence-based epitope predictions, we also used the spike trimer structure for the structure-based prediction using Ellipro and DiscoTope. Ellipro incorporates the properties such as antigenicity, solvent accessibility, and flexibility of the protein structures to predict linear and conformational epitopes. DiscoTope uses the spatial information, surface accessibility, and statistics of amino acids on the protein 3D structure to predict residue by residue conformational epitopes[24, 25]. The epitopes predicted by various methods are given in Table S3, S4, and S5. The epitopes that involve mutations in the variants of concern are summarized in Table S2 and their antigenicity was predicted using VaxiJen[26].

Table S1. Molecular systems simulated in this work. The WT and Delta simulations of the RBD-only systems in our earlier work (Baral et. al, 2021)[27] were extended from 600ns to 1000ns. All systems include appropriate glycosylation. Glycosylated spike residues: N165, N234, N343, and glycosylated ACE2 residues: N53, N90, N103, N322, N546.

System	Variant/ (PDB ID)	# Residues	# Water	# Ions	Total # Atoms in the System	Box Size (Å)^3	Simulation Time (ns)
RBD-only	WT (6VSB*)	201	24,764	144	77,736	95x95x95	1000
	Delta (Mutated on 6VSB*)	201	26,654	156	84,000	95x95x95	1000
	Omicron (7T9L)	201	26,723	157	83,518	95x95x95	1000
RBD-hACE2	WT (7A92)	796	79,743	273	253,548	140x140x140	100
	Delta (7W9I)	791	75,956	454	242,060	138x138x138	100
	Omicron (7T9L)	796	70,801	420	226,700	134x134x134	100
RBD-down**	WT (6VXX from Amaro Lab)	969	87,924	509	280,520	145x145x145	100
	Omicron (Mutated on 6VXX from Amaro Lab)	969	87,814	518	280,367	145x145x145	100

* From the CHARM-GUI Archive - COVID-19 Proteins Library[28]; RBD structure based on PDB ID 6VSB[29], with WT sequence[16].

**Residues in the truncated systems: Chain A (330-530); Chain B (30-530, 968-1000); Chain C (330-530, 968-1000).

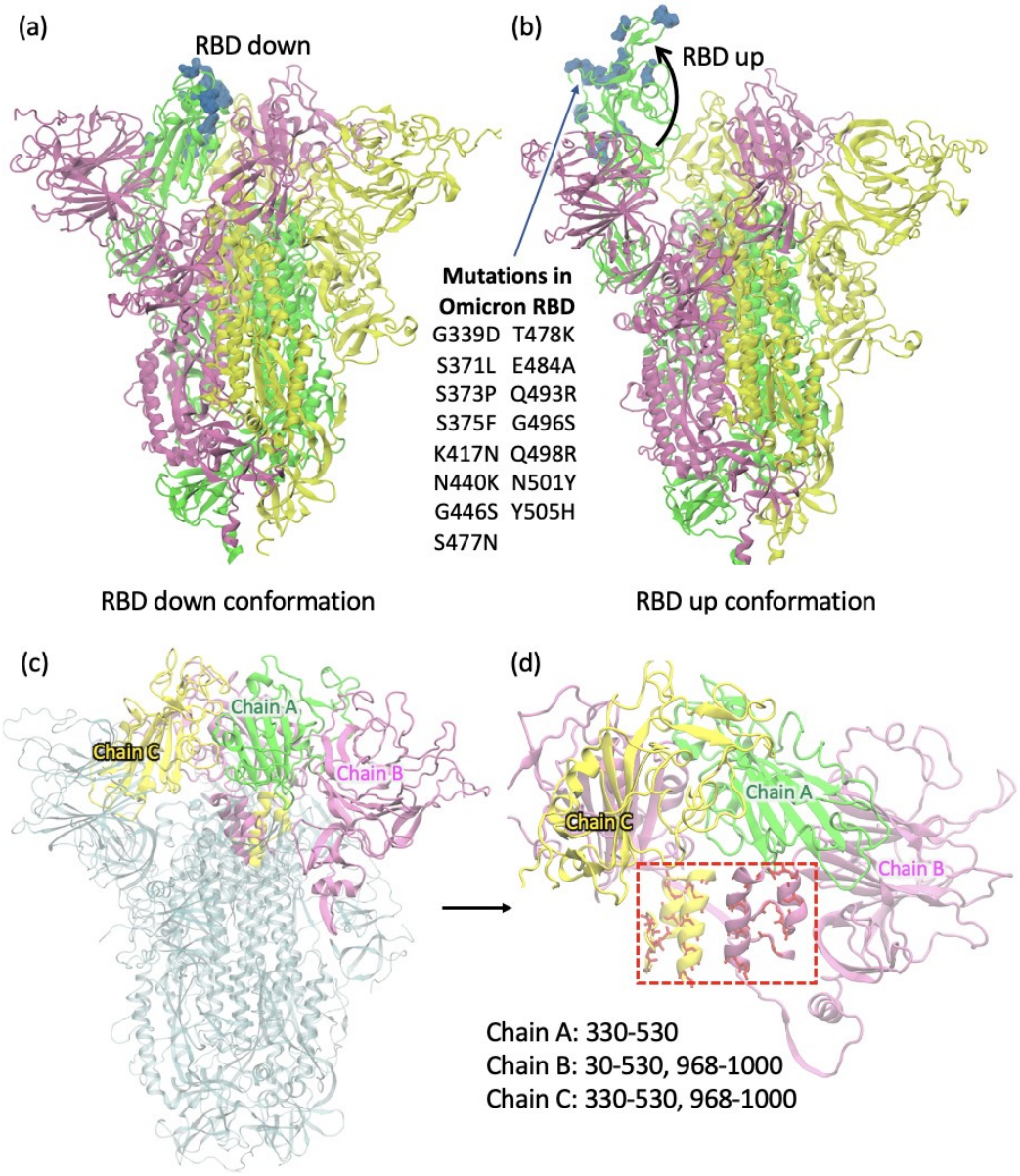


Figure S1. The spike protein trimer of SARS-CoV-2 in its prefusion state in the a) closed (RBD down) and b) open (RBD up) conformations. c) The closed form conformation of the spike protein highlighting the protein domains considered for the MD simulation. The trimer structure used here was the first frame of an MD trajectory obtained from the Amaro Lab[30] (trajectory of spike opening: https://amarolab.ucsd.edu/covid19.php). The RBD of chain A (green) and the surrounding domains (chain B – magenta and chain C – yellow) are highlighted. b) The colored part shows the truncated system consisting of the Chain A RBD and the surrounding domains considered for MD simulations. All C_α atoms >12 Å from the RBD of chain A are harmonically restrained for MD simulations. The helical segments highlighted in red dashed box are the non-contiguous segments of chain B and C that interact with the chain A RBD. For the Omicron system, all mutations within 12 Å of RBD of chain A and the surrounding were considered.

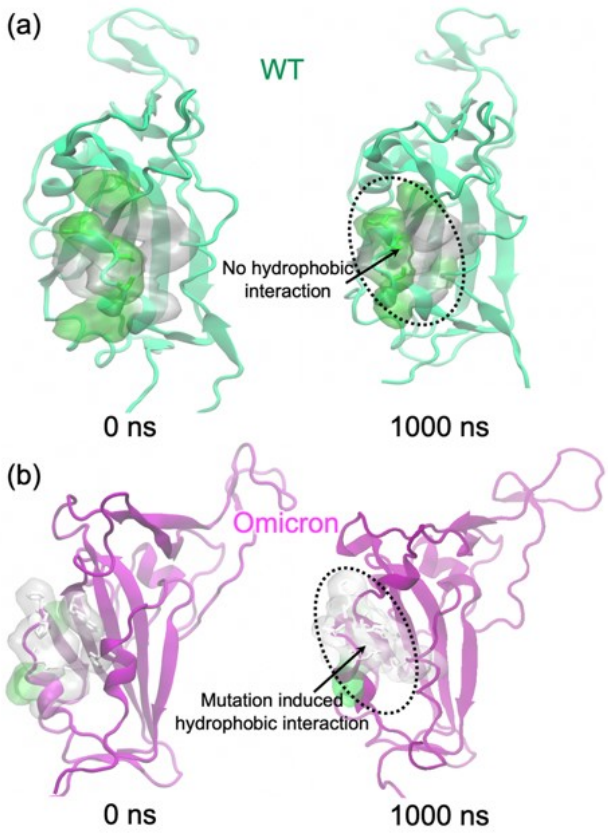


Figure S2. Comparison of the motif structure (residues 364-375) at 0 ns and 1000ns for the a) WT and b) Omicron RBDs. The polar to hydrophobic mutations S371L, S373P, S375F in Omicron allow interactions with the nearby hydrophobic residues. The polar residues are highlighted in green surface and the hydrophobic residues are highlighted in white/gray surface.

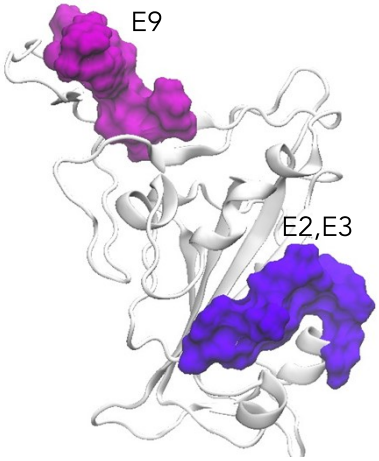


Figure S3. Location of the epitopes E2, E3, and E9 with increased antigenicity in the Omicron RBD. All epitopes have good surface accessibility for the antibody binding.

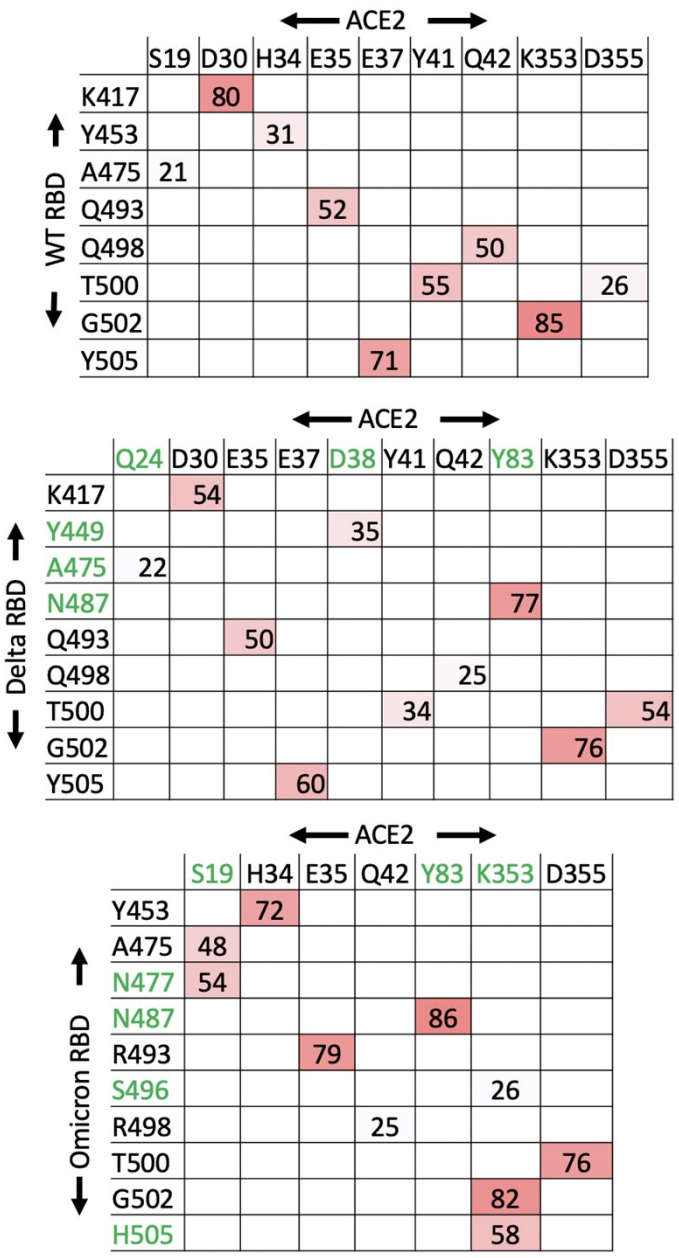
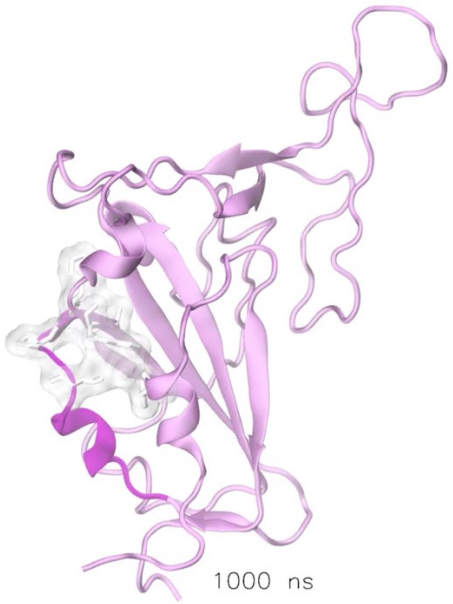


Figure S4. Hydrogen bond pairs for the interprotein interactions between the RBD and ACE2 for the WT, Delta, and Omicron. The matrices for the Delta and Omicron are the same as in Fig. 4a,b but given here for comparison with WT. The matrix for residues participating in the unique hydrogen-bond pairs in Omicron are highlighted in green.

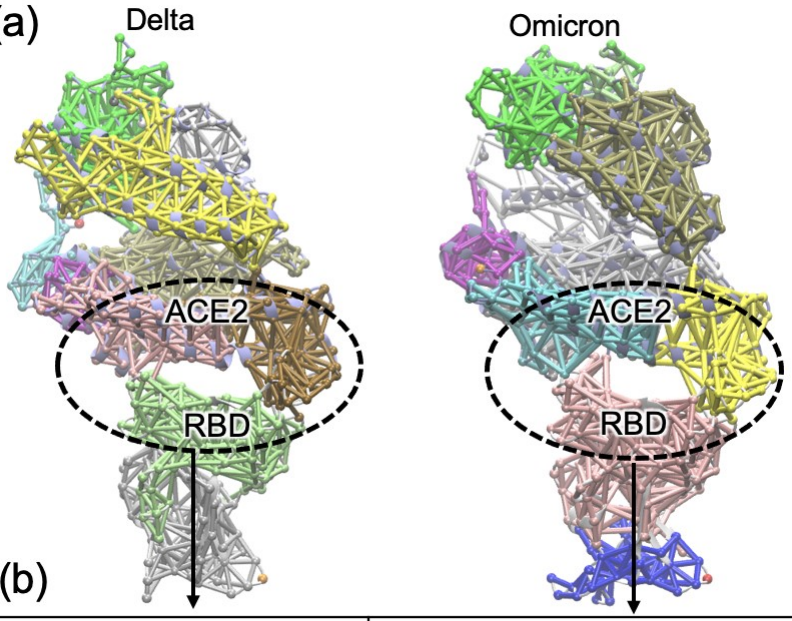


Movie S1. Movie showing the dynamics of the isolated spike RBD of Omicron. The motif comprised of residues 364-375 is shaded in bright purple and the mutated hydrophobic residues L371, P373, and F375 in the motif as well as nearby hydrophobic residues A435, W436, and F342 are shown in white surface representation.

Table S2. Predicted RBD epitopes that involve mutations in different variants of concern - WT, Alpha (B.1.1.7), Beta (B.1.351), Delta (B.1.617.2), Mu (B.1.621), and Omicron (B.1.1.529) and their corresponding antigenicity values. The mutations in the epitope are underlined. Significantly increased antigenicity in Omicron epitopes E2, E3, and E9 are highlighted in green and a moderately decreased antigenicity in E5 is highlighted in red.

Epitope	Residues	Variant	Epitope Sequence	Antigenicity
E1	338-356	WT	FGEVFNATRFASVYAWNRK	0.24
		B.1.1.529	F<u>D</u>EVFNATRFASVYAWNRK	0.28
E2	370-378	WT	NSASFSTFK	0.12
		B.1.1.529	N<u>L</u>A<u>P</u>F<u>F</u>TFK	1.34
E3	372-380	WT	ASFSTFKCY	0.28
		B.1.1.529	A<u>P</u>F<u>F</u>TFKCY	1.20
E4	417-425	WT	KIADYNYKL	1.66
		B.1.351	<u>N</u>IADYNYKL	1.55
		B.1.621	<u>N</u>IADYNYKL	1.55
		B.1.1.529	<u>N</u>IADYNYKL	1.55
E5	437-448	WT	NSNNLDSKVGGN	0.70
		B.1.1.529	NSN<u>K</u>LDSKV<u>S</u>GN	0.24
E6	439-451	WT	NNLDSKVGGNYNY	0.94
		B.1.1.529	N<u>K</u>LDSKV<u>S</u>GNYNY	0.69
E7	447-471	WT	GNYNYLYRLFRKSNLKPFERDISTE	0.16
		B.1.617.2	GNYNY<u>R</u>YRLFRKSNLKPFERDISTE	0.42
E8	455-478	WT	LFRKSNLKPFERDISTEIYQAGST	0.13
		B.1.617.2	LFRKSNLKPFERDISTEIYQAGS<u>K</u>	0.11
		B.1.1.529	LFRKSNLKPFERDISTEIYQAG<u>NK</u>	0.10
E9	483-493	WT	VEGFNCYFPL<u>Q</u>	0.56
		B.1.351	V<u>K</u>GFNCYFPL<u>Q</u>	0.60
		B.1.621	V<u>K</u>GFNCYFPL<u>Q</u>	0.60
		B.1.1.529	V<u>A</u>GFNCYFPL<u>R</u>	1.23
E10	491-505	WT	PL<u>Q</u>SYGFQPTNGVGY	0.34
		B.1.1.7	PL<u>Q</u>SYGFQPT<u>Y</u>GVGY	0.41
		B.1.351	PL<u>Q</u>SYGFQPT<u>Y</u>GVGY	0.41
		B.1.1.529	PL<u>R</u>SY<u>S</u>F<u>R</u>PT<u>Y</u>GVG<u>H</u>	0.41
E11	502-510	WT	GVGY<u>Q</u>PYRV	1.36
		B.1.1.529	GVG<u>H</u><u>Q</u>PYRV	1.01

(a)



(b)

Delta		Omicron	
RBD	ACE2	RBD	ACE2
F456	T27, D30, K31	F456, Y489	K31
Y473	T27	Y489	F28, Y83
A475	Q24, T27	F486	L79, M82, Y83
G476, N487	Q24	N487	Q24, Y83
F486	L79, M82	A475, G476, N477	S19
F486, N487, Y489	Y83	A475, G476, N477	Q24
Y489	T27, K31, F28	F456, Y473, A475	T27
Q498, T500	Y41	N330, D355, R357, T500	Q42
T500	D355, R357	S496, Y501, G502, H505	K353
		Y501, G502	G354
Total Connection = 20		Total Connection = 28	

Figure S5. Community analysis performed for the first 50 ns of the trajectories for the Delta and Omicron variants. a) Each identified community is represented by a different color and the region that span both the RBD and ACE2 are circled. b) The amino acid residues involved in the dynamic network of the communities spanning the two proteins (circled region).

Table S3. MHC-I and MHC-II T-cell epitopes for WT predicted by different methods. The sequences involving mutations in different variants of concern are listed in Table S2.

MHC-I			MHC-II	
Propred-I	CTLPred	NetCTL	NetMHCII2.3	EpiTOP3.0
TLDSKTQSL	NLNESLIDL	LTDEMIAQY	FELLHAPAT	VVIKVCEFQ
TQSLLIVNN	TRFQTLLAL	WTAGAAAYY	FGAGAALQI	MESEFRVYS
FEYVSQPFL	RLDKVEAEV	TSNQVAVLY	VKQLSSNFG	YSKHTPINL
GKQGNFKNL	NSPRRARSV	CVADYSVLY	FQTLLALHR	LQPRTFLLK
TPINLVRDL	ALDPLSETK	KTSVDCTMY	YYPDKVFRS	YNYKLPDDF
VRDLPQGFSALEPL	DLLFNKVTL	STECSNLLL	FLPFFSNVT	VKNKCVNFN
QGFSALEPLVDL	AISSVLNDI	GAEHVNNSY	LALHRSYLT	IAARDLICA
INITRFQTL	NCDVVIGIV	NIDGYFKIY	LLALHRSYL	MAYRFNGIG
FQTLLALHR	KSNIIRGWI	YSSANNCTF	YRLFRKSNL	YRFNGIGVT
GTITDAVDCAL	ARDLICAQK	WMESEFRVY	FIAGLIAIV	ILSRLDKVE
CYGVSPTKL	TEVPVAIHA	SANNCTFEY	YVGYLQPRT	ITGRLQSLQ
IRGDEVRQI	AQKFNGLTV	VASQSIIAY	LIVNNATNV	YVTQQLIRA
KIADYNYKL	QSLLIVNNA	NSFTRGVYY	LIANQFNSA	
VVVLSFELL	ATNVVIKVC	CNDPFLGVY		
NGLTGTGVLTESNKKFL	YQDVNCTEV	FTNVYADSF		
DAVRDPQTL	KNKCVNFNF	GAAAYYVGY		
QTQTNSPRR	LFLPFFSNV	RISNCVADY		
SIIAYTMSL	FERDISTEI	ERDISTEIY		
TNFTISVTT	SVDCTMYIC	ITDAVDCAL		
FCTQLNRAL	AYSNNSIAI	RVDFCGKGY		
GGFNFSQIL	KRSFIEDLL	NATRFASVY		
SKRSFIEDLLFNKVTL	YTNSFTRGV	MTSCCSCLK		
FIKQYGDCL	VYAWNRKRI	STQDLFLPF		
GDCLGDIAA	VSQPFLMDL	VSNGTHWFV		
QKFNGLTVLPPLL	INITRFQTL	GTITSGWTF		
IAQYTSALL	IVNNATNVV	RSFIEDLLF		
GTITSGWTFGAGAAL	VVFLHVTYV	VLKGVKLHY		
AIGKIQDSLSSTASALGKL	TNLCPFGEV	ECSNLLLQY		
VVNQNAQALNTLVKQL	ANNCTFEYV	NSASFSTFK		
NFGAISSVLNDILSRL	CPFGEVFNA	ASFSTFKCY		
QLIRAAEIRASANLA	RVYSTGSNV	QLTPTWRVY		
SAPHGVVFL	EILPVSMTK	LSETKCTLK		
VYDPLQPELDSFKEEL	SKPSKRSFI	FVFKNIDGY		
NHTSPDVDL	GVGYQPYRV	VGGNYNYLY		
RLNEVAKNLNESL	FIKQYGDCL	VLPFNDGVY		
PWYIWLGFI	GTITSGWTF	FVSNGTHWF		
DEDDSEPVL	VIAWNSNNL	YQDVNCTEV		
	QYGDCLGDI			
	FSNVTWFHA			
	EVFAQVKQI			
	NQFNSAIGK			
	QTNSPRRAR			
	NNATNVVIK			
	NATNVVIKV			
	QNAQALNTL			
	QFNSAIGKI			
	WTAGAAAYY			
	FVSNGTHWF			
	RSVASQSII			
	RVVVLSFEL			

Table S4. Sequence based B-cell epitopes predicted by different methods. The sequences involving mutations in different variants of concern are listed in Table S2.

Bepipred	Bcepred
QCVNLTTRTQLPPAYTNSFTRGV	MFVFLVLLPLVSSQCVNLTT
FSNVTWFHAIHVSGTNGTKRFDN	NLTTRTQLPPAYTNSFTRGVYYPDKVFRSS
KSNI	VYYPDKVFRSSVLHSTQDLFLPFFSNV
LGVYYHKNNKSWMESEFRVYSSA	SGTNGTKRFDNPV
DLEGKQGNFKNLRE	YFASTEKSNIIR
HTPINLVRDLPQGFSA	TTLDSKTQSL
YLTPGDSSSGWTA	GVYYHKNNKSWMESEFRVY
FTVE	MDLEGKQGNFKNLREF
YQTSNFRVQP	YFKIYSKHTPIN
NITNLC	DPLSETKCTLKSFTVEKGIYQTSNFRVQPTES
FGEVFNATRFASVYAWNRK	VYAWNRKRISNC
NSASFSTFKCYGVSPTKLNDLCFTNV	RGDEVRQ
GDEVRQIAPGQTGKIADYNYK	KIADYNYKLPDDFT
NNLDSKVGGNYNY	NSNNLDSKVGGN
LFRKSNLKPFERDISTEIYQAGST	GNYNYLYRLFRKSNLKPFERDISTE
VEGFNCYFPLQ	CGPKKSTNLVKNK
FQPTNG	LTESNKKFLP
ELLHAPATVCGPKKSTNLVK	TPGTNTSNQV
KKFLP	EHVNNSYECD
TNTS	ASYQTQTNSPRRARSVA
VNCTEVP	AVEQDKNTQEVF
ADQLTPTWRVYSTGSNVFQT	QVKQIYKTPPIKD
VNNSYECDIP	LPDPSKPSKRSFIED
SYQTQTNSPRRARSVASQS	TLVKQLSS
AYTMSLGAENSVAYSN	SRLDKVEAEVQ
DKNTQ	LQSLQTYVTQQLI
KQIYKTPPIKDFGGF	GQSKRVDFC
LPDPSKPSKR	CHDGKAHFPREGV
LADAGFIKQYGDCLGD	NNTVYDPLQPELDSFKEELDKYFKNHTSPD
VEAEVQI	NIQKEIDRLNEVAKNLNESL
GQSKRVDFC	QELGKYEQYIKWP
FYEPQIITTD	CKFDEDDSEPV
VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGI	EPVLKGVKLHYT
ELGKYE	
SCCKFDEDDSEPVLKGVKL	

Table S5. B-cell epitopes for WT predicted by Ellipro. The sequences involving mutations in different variants of concern are listed in Table S2.

Sequence Based

Residues	Sequence	Score
63-85	TWFHFDNPVLP	0.687
92-105	FASTNIIRGWI	0.745
108-192	TTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGEFRVYSSANNCTFEYVSQPFLKNLREF	0.757
205-218	SKHTPINLVRDLPQ	0.678
236-266	TRFQTLLALHRGAAAYY	0.772
328-381	RLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG	0.765
390-528	LCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSYNYLYRPLQSYGFQPTVGYQPYRVVVLSFELLHAPATVCGPK	0.809
553-564	TESNKKFLPFQQ	0.713
577-585	RDPQTLEIL	0.661
701-717	AENSVAYSNNSIAIPTN	0.664
783-800	AQVKQIYKTPPIKDFGGF	0.631
806-815	LPDPSKR	0.552
883-904	TSGWTFGAGAALQIPFAMQMAY	0.623
1071-1146	QEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELD	0.871

Structure Based

	Residues	Score
1	A:D1139, A:P1140, A:L1141, A:Q1142, A:P1143, A:E1144, A:L1145, A:D1146	0.975
2	A:Y707, A:S708, A:N709, A:N710, A:S711, A:I712, A:A713, A:I714, A:P715, A:T716, A:N717, A:Q1071, A:K1073, A:N1074, A:F1075, A:T1076, A:T1077, A:A1078, A:P1079, A:A1080, A:I1081, A:C1082, A:H1083, A:D1084, A:G1085, A:K1086, A:A1087, A:H1088, A:F1089, A:P1090, A:R1091, A:E1092, A:G1093, A:V1094, A:F1095, A:V1096, A:S1097, A:N1098, A:G1099, A:T1100, A:H1101, A:W1102, A:F1103, A:V1104, A:T1105, A:Q1106, A:R1107, A:F1109, A:Y1110, A:E1111, A:P1112, A:Q1113, A:I1114, A:I1115, A:T1116, A:T1117, A:D1118, A:N1119, A:T1120, A:F1121, A:V1122, A:S1123, A:G1124, A:N1125, A:C1126, A:D1127, A:V1128, A:V1129, A:I1130, A:G1131, A:I1132, A:V1133, A:N1134, A:N1135, A:T1136, A:V1137, A:Y1138	0.845
3	A:L335, A:C336, A:P337, A:F338, A:G339, A:E340, A:V341, A:F342, A:N343, A:A344, A:T345, A:R346, A:F347, A:A348, A:S349, A:V350, A:Y351, A:A352, A:W353, A:N354, A:R355, A:K356, A:R357, A:I358, A:S359, A:N360, A:C361, A:V362, A:A363, A:D364, A:Y365, A:S366, A:V367, A:L368, A:Y369, A:N370, A:S371, A:A372, A:S373, A:F374, A:S375, A:T376, A:F377, A:K378, A:C379, A:Y380, A:L390, A:C391, A:F392, A:T393, A:N394, A:V395, A:Y396, A:A397, A:D398, A:S399, A:F400, A:V401, A:I402, A:R403, A:G404, A:D405, A:E406, A:V407, A:R408, A:Q409, A:I410, A:A411, A:P412, A:G413, A:Q414, A:T415,	0.799

	A:G416, A:K417, A:I418, A:A419, A:D420, A:Y421, A:N422, A:Y423, A:K424, A:L425, A:P426, A:D427, A:D428, A:F429, A:T430, A:G431, A:C432, A:V433, A:I434, A:A435, A:W436, A:N437, A:S438, A:N439, A:N440, A:L441, A:D442, A:S443, A:Y449, A:N450, A:Y451, A:L452, A:Y453, A:R454, A:P491, A:L492, A:Q493, A:S494, A:Y495, A:G496, A:F497, A:Q498, A:P499, A:T500, A:V503, A:G504, A:Y505, A:Q506, A:P507, A:Y508, A:R509, A:V510, A:V511, A:V512, A:L513, A:S514, A:F515, A:E516, A:L517, A:L518, A:H519, A:A520, A:P521, A:A522, A:T523, A:V524, A:C525, A:G526, A:P527, A:K528	
4	A:F559, A:L560, A:P561, A:F562, A:Q563	0.789
5	A:F79, A:D80, A:N81, A:P82, A:V83, A:L84, A:P85, A:I100, A:I101, A:R102, A:G103, A:W104, A:I105, A:T108, A:T109, A:L110, A:D111, A:S112, A:K113, A:T114, A:Q115, A:S116, A:L117, A:L118, A:I119, A:V120, A:N121, A:N122, A:A123, A:T124, A:N125, A:V126, A:V127, A:I128, A:K129, A:V130, A:C131, A:E132, A:F133, A:Q134, A:F135, A:C136, A:N137, A:D138, A:P139, A:F140, A:L141, A:G142, A:E156, A:F157, A:R158, A:V159, A:Y160, A:S161, A:S162, A:A163, A:N164, A:N165, A:C166, A:T167, A:F168, A:E169, A:Y170, A:V171, A:S172, A:Q173, A:P174, A:F175, A:L176, A:T236, A:R237, A:F238, A:Q239, A:T240, A:L241, A:L242, A:A243, A:L244, A:H245, A:R246	0.756
6	A:A27, A:Y28, A:T29, A:R34, A:T63, A:W64, A:F65, A:F92, A:A93, A:S94, A:T95, A:K187, A:N188, A:L189, A:R190, A:E191, A:F192, A:S205, A:K206, A:H207, A:T208, A:P209, A:I210, A:N211, A:L212, A:V213, A:R214, A:D215, A:L216, A:P217, A:Q218, A:S221, A:A222, A:L223, A:G261, A:A262, A:A263, A:A264, A:Y265, A:Y266	0.661
7	A:G744, A:D745, A:S746, A:T747, A:E748, A:C749, A:S750, A:N751, A:L752, A:L753, A:L754, A:Q755, A:G757, A:S758, A:T761, A:R765	0.627
8	A:A783, A:Q784, A:V785, A:K786, A:Q787, A:I788, A:Y789, A:K790, A:T791, A:P792, A:P793, A:I794, A:K795, A:D796, A:F797, A:G798, A:G799, A:F800, A:F802, A:Q872, A:A876, A:A879, A:G880, A:T883, A:S884, A:G885, A:W886, A:T887, A:F888, A:G889, A:A890, A:G891, A:A892, A:A893, A:L894, A:Q895, A:I896, A:P897, A:F898, A:A899, A:M900, A:Q901, A:A903, A:Y904, A:N907, A:G910, A:V911, A:T912, A:Q913, A:N914, A:V915, A:L916, A:Y917, A:E918, A:N919, A:Q920, A:L922, A:N925, A:L1034, A:N1108	0.61
9	A:L984, A:D985, A:P986, A:P987, A:E988, A:E990	0.602
10	A:R328, A:T531, A:N532, A:L533, A:V534, A:K535, A:N536, A:T553, A:E554, A:S555, A:N556, A:K557, A:Q564, A:R577, A:D578, A:P579, A:Q580, A:T581, A:L582, A:E583, A:I584, A:L585	0.597
11	A:L806, A:P807, A:D808, A:P809, A:S810, A:K811	0.578
12	A:A701, A:E702, A:N703, A:S704, A:V705, A:A706	0.576
13	A:L226, A:V227, A:D228, A:L229	0.513
14	A:P863, A:L864, A:L865, A:E868, A:M869	0.511

Table S6. Structure based B-cell epitopes for WT, predicted by Discotope. The RBD residues that are mutated in Omicron are highlighted in boldface and included in Table S2.

Residue#	Residue	Residue#	Residue	Residue#	Residue
281	GLU	503	VAL	917	TYR
282	ASN	505	TYR	918	GLU
415	THR	556	ASN	1071	GLN
420	ASP	558	LYS	1099	GLY
449	TYR	560	LEU	1100	THR
450	ASN	561	PRO	1101	HIS
454	ARG	562	PHE	1111	GLU
491	PRO	703	ASN	1118	ASP
492	LEU	704	SER	1140	PRO
493	GLN	705	VAL	1141	LEU
494	SER	793	PRO	1142	GLN
496	GLY	794	ILE	1143	PRO
498	GLN	809	PRO	1144	GLU
499	PRO	810	SER	1145	LEU
500	THR	914	ASN	1146	ASP

References

1. Sztain, T.; Ahn, S. H.; Bogetti, A. T.; Casalino, L.; Goldsmith, J. A.; Seitz, E.; McCool, R. S.; Kearns, F. L.; Acosta-Reyes, F.; Maji, S.; Mashayekhi, G.; McCammon, J. A.; Ourmazd, A.; Frank, J.; McLellan, J. S.; Chong, L. T.; Amaro, R. E., A glycan gate controls opening of the SARS-CoV-2 spike protein. *Nat Chem* **2021,** *13* (10), 963-968.
2. Walls, A. C.; Park, Y.-J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D., Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **2020,** *181* (2), 281-292. e6.
3. Jo, S.; Kim, T.; Iyer, V. G.; Im, W., CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of computational chemistry* **2008,** *29* (11), 1859-1865.
4. J. Lee, X. C., J.M. Swails, M.S. Yeom, P.K. Eastman, J.A. Lemkul, S. Wei, J. Buckner, J.C. Jeong, Y. Qi, S. Jo, V.S. Pande, D.A. Case, C.L. Brooks III, A.D. MacKerell Jr, J.B. Klauda, and W. Im, CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput* **2016,** *12*, 405-413.
5. Mannar, D.; Saville, J. W.; Zhu, X.; Srivastava, S. S.; Berezuk, A. M.; Tuttle, K. S.; Marquez, C.; Sekirov, I.; Subramaniam, S., SARS-CoV-2 Omicron Variant: ACE2 Binding, Cryo-EM Structure of Spike Protein-ACE2 Complex and Antibody Evasion. *bioRxiv* **2021**, 2021.12.19.473380.
6. Wang, Y.; Liu, C.; Zhang, C.; Wang, Y.; Hong, Q.; Xu, S.; Li, Z.; Yang, Y.; Huang, Z.; Cong, Y., Structural basis for SARS-CoV-2 Delta variant recognition of ACE2 receptor and broadly neutralizing antibodies. *Nature Communications* **2022,** *13* (1), 1-12.

7. Baral, P.; Bhattarai, N.; Hossen, M. L.; Stebliankin, V.; Gerstman, B. S.; Narasimhan, G.; Chapagain, P. P., Mutation-induced changes in the receptor-binding interface of the SARS-CoV-2 Delta variant B.1.617.2 and implications for immune evasion. *Biochem Biophys Res Commun* **2021,** *574*, 14-19.
8. Bhattarai, N.; Baral, P.; Gerstman, B. S.; Chapagain, P. P., Structural and Dynamical Differences in the Spike Protein RBD in the SARS-CoV-2 Variants B. 1.1. 7 and B. 1.351. *The Journal of Physical Chemistry B* **2021**.
9. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K., Scalable molecular dynamics with NAMD. *Journal of computational chemistry* **2005,** *26* (16), 1781-1802.
10. Huang, J.; MacKerell Jr, A. D., CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of computational chemistry* **2013,** *34* (25), 2135-2145.
11. Huang, J.; MacKerell Jr, A. D., Force field development and simulations of intrinsically disordered proteins. *Current opinion in structural biology* **2018,** *48*, 40-48.
12. Darden, T.; York, D.; Pedersen, L., Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. *The Journal of chemical physics* **1993,** *98* (12), 10089-10092.
13. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A smooth particle mesh Ewald method. *The Journal of chemical physics* **1995,** *103* (19), 8577-8593.
14. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J., Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of computational physics* **1977,** *23* (3), 327-341.
15. Humphrey, W.; Dalke, A.; Schulten, K., VMD: visual molecular dynamics. *Journal of molecular graphics* **1996,** *14* (1), 33-38.
16. Wu, F.; Zhao, S.; Yu, B.; Chen, Y.-M.; Wang, W.; Song, Z.-G.; Hu, Y.; Tao, Z.-W.; Tian, J.-H.; Pei, Y.-Y., A new coronavirus associated with human respiratory disease in China. *Nature* **2020,** *579* (7798), 265-269.
17. Singh, H.; Raghava, G., ProPred1: prediction of promiscuous MHC Class-I binding sites. *Bioinformatics* **2003,** *19* (8), 1009-1014.
18. Bhasin, M.; Raghava, G. P., Prediction of CTL epitopes using QM, SVM and ANN techniques. *Vaccine* **2004,** *22* (23-24), 3195-3204.
19. Larsen, M. V.; Lundegaard, C.; Lamberth, K.; Buus, S.; Lund, O.; Nielsen, M., Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC bioinformatics* **2007,** *8* (1), 1-12.
20. Dimitrov, I.; Garnev, P.; Flower, D. R.; Doytchinova, I., EpiTOP—a proteochemometric tool for MHC class II binding prediction. *Bioinformatics* **2010,** *26* (16), 2066-2068.
21. Jensen, K. K.; Andreatta, M.; Marcatili, P.; Buus, S.; Greenbaum, J. A.; Yan, Z.; Sette, A.; Peters, B.; Nielsen, M., Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology* **2018,** *154* (3), 394-406.
22. Jespersen, M. C.; Peters, B.; Nielsen, M.; Marcatili, P., BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic acids research* **2017,** *45* (W1), W24-W29.

23. Saha, S.; Raghava, G. P. S. In *BcePred: prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties*, International Conference on Artificial Immune Systems, Springer: 2004; pp 197-204.
24. Kringelum, J. V.; Lundegaard, C.; Lund, O.; Nielsen, M., Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS computational biology* **2012,** *8* (12), e1002829.
25. Ponomarenko, J.; Bui, H.-H.; Li, W.; Fusseder, N.; Bourne, P. E.; Sette, A.; Peters, B., ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC bioinformatics* **2008,** *9* (1), 1-8.
26. Irini A Doytchinova, D. R. F., VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* **2007**.
27. Baral, P.; Bhattarai, N.; Hossen, M. L.; Stebliankin, V.; Gerstman, B. S.; Narasimhan, G.; Chapagain, P. P., Mutation-induced changes in the receptor-binding interface of the SARS-CoV-2 Delta variant B. 1.617. 2 and implications for immune evasion. *Biochemical and biophysical research communications* **2021,** *574*, 14-19.
28. Woo, H.; Park, S.-J.; Choi, Y. K.; Park, T.; Tanveer, M.; Cao, Y.; Kern, N. R.; Lee, J.; Yeom, M. S.; Croll, T. I., Developing a fully glycosylated full-length SARS-CoV-2 spike protein model in a viral membrane. *The journal of physical chemistry B* **2020,** *124* (33), 7128-7137.
29. Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C.-L.; Abiona, O.; Graham, B. S.; McLellan, J. S., Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020,** *367* (6483), 1260-1263.
30. Sztain, T.; Ahn, S.-H.; Bogetti, A. T.; Casalino, L.; Goldsmith, J. A.; Seitz, E.; McCool, R. S.; Kearns, F. L.; Acosta-Reyes, F.; Maji, S., A glycan gate controls opening of the SARS-CoV-2 spike protein. *Nature Chemistry* **2021,** *13* (10), 963-968.