3-8-2022

Functional evolutionary adaptation of binding dynamics in the SARS-CoV-2/ACE2 interface from bats to humans

Madhusudan Rajendran
mr8236@rit.edu

RIT

Functional evolutionary adaptation of binding dynamics in the SARS-CoV-2/ACE2 interface from bats to humans

By

Madhusudan Rajendran

Thesis Advisor: Dr. Gregory A. Babbitt

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Bioinformatics

Department of Bioinformatics

College of Science

Thomas H. Gosnell School of Life Sciences

Rochester Institute of Technology

Rochester, NY

March 08, 2022

Rochester Institute of Technology
Thomas H. Gosnell School of Life Sciences
Bioinformatics Program

To: Head, Thomas H. Gosnell School of Life Sciences

The undersigned state that Madhusudan Rajendran, a candidate for the Master of Science degree in Bioinformatics, has submitted his thesis and has satisfactorily defended it.

This completes the requirements for the Master of Science degree in Bioinformatics at Rochester Institute of Technology.

Thesis committee members:

Name	**Date**
Gregory A. Babbitt, Ph.D. Thesis Advisor	
Maureen C. Ferran, Ph.D.	
Feng Cui, Ph.D.	

Feng Cui, Ph.D.
Director of Bioinformatics MS Program

475-4115 (voice)
fxcsbi@rit.edu

TABLE OF CONTENTS

Abstract

The COVID-19 pandemic highlights the substantial public health, economic, and societal consequences of virus spillover from a wildlife reservoir. Widespread human transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) also presents a new set of challenges when considering viral spillover from people to naïve wildlife and other animal populations. Here, we used molecular dynamic (MD) simulations to understand the transmission of the SARS-CoV-2 virus from bats to humans and investigate the evolution of the various human variants. More specifically, we used MD simulations to understand the atomic fluctuation dampening at the receptor-binding domain (RBD)/ angiotensin-converting enzyme 2 (ACE2) interface when the RBD of the different SARS-CoV-2 variants is simulated with ACE2 of bat origin (bACE2) or human origin (hACE2). Towards that end, we found that the RaTG13 RBD is more stabilized with strong atomic fluctuation dampening when bound to bACE2 than hACE2.

Interestingly, in terms of the variants being monitored (VBM - beta, kappa, and epsilon variants), we saw very similar binding dynamics and MD profiles between bACE2 and hACE2. Of note, when either bACE2 or hACE2 was simulated with variants of concern (VOC – alpha, delta, and omicron variants) RBD, we saw slightly stabilized atomic fluctuation dampening when the RBD was simulated with hACE2. Our results indicated that the RBD of the newer human SARS-CoV-2 variants does not differ significantly in the atomic fluctuation dampening when interacting with ACE2 of either bat or human origin. As a result, reverse zoonosis events that cause the virus to jump back to bats are highly possible, including the emergence of new and different variants. Therefore, in addition to the ongoing genomic surveillance, we also advise the inclusion of MD simulation surveillance that investigates the binding dynamics of the new variants with the receptors of the different host species.

Introduction

Coronavirus disease (COVID-19) is caused by novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The virus first emerged from the Wuhan province in China (1, 2). Since its emergence, the virus has had a devastating effect on the world's population, resulting in more than 5.5 million deaths worldwide (3). Since being declared a global pandemic by the World Health Organization (WHO) on March 11, 2020, the virus continues to cause devastation, with several countries enduring multiple waves of outbreaks of this viral illness. Infections may be mild or fatal. The most common symptoms include fever, coughing, shortness of breath, difficulty breathing, vomiting, chills, body aches, headache, sore throat, congestion/runny nose, loss of smell or taste, nausea, and diarrhea. The disease can also lead to pneumonia, respiratory failure, heart attacks, liver problems, septic shock, and death (4).

Most of the efforts to develop therapies and vaccines against COVID-19 aim to either decrease the stability or exploit some of the features of the spike (S) protein (5-7). The S protein triggers an immune response and plays an essential role in infecting the host. The S protein is a homotrimer composed of two functions subunits, S1 and S2. The S1 subunit is responsible for receptor binding, while the S2 subunit is responsible for fusing viral and cellular membranes. Most importantly, the S1 subunits contain the receptor-binding domain (RBD). The RBD domain alternates between a closed (down) and open (up) state. The open state exposes the receptor-binding motif in the RBD, thus enabling the interaction and binding with angiotensin-converting enzyme 2 (ACE2) (8). The primary physiological function of the ACE2 is to lower blood pressure by hydrolyzing angiotensin II (a vasoconstrictor) into angiotensin-(1-7) (a vasodilator).

Furthermore, ACE2 is a type I transmembrane protein expressed in various organs, including lungs, hearts, kidneys, and intestines (9). Recent structural studies show that ACE2 is a

homodimer with each monomer consisting of N-terminal peptidase domain, C terminal collectrin-like domain, a single pass transmembrane region, and a short cytoplasmic region (10). The RBD binding region on ACE2 is located in the N-terminal peptidase domain with major contact regions located in the α1 and α2 helices as well as the linker between β3 and β4 strands (11).

The outbreak of COVID-19 was initially linked to a local seafood market in Wuhan, China, where the sale of wild animals has been implicated as the primary source of SARS-CoV-2 infections (12). The Current SARS-CoV-2 virus is known to have 96.2% similarity to the bat coronavirus RaTG13 at the whole genome level (13). Based on the viral genome sequence and further evolutionary analysis, Chinese horseshoe bats of genus Rhinolophus have been considered a natural reservoir host for the SARS-CoV-2 virus (14). However, as of now, no definitive intermediate host has been identified. Based on the isolation of closely related genomes from Malayan pangolins (Manis javanica), they are thought to be possible intermediate hosts of the SARS-CoV-2 (15).

Regardless of the bat origin, the genome of SARS-Cov-2 is prone to mutations. However, unlike other RNA viruses, SARS-CoV-2 encodes a proofreading domain (ExoN) that reduces its mutation rate (16). Therefore, the mutation rate of the SARS-CoV-2 genome has thus been estimated at 1.87 x 10-6 nucleotide substitutions per site per day. This is roughly 5-fold lower than influenza A/H3N2 viruses, which have 10.9 x 10-6 nucleotide substitutions per site per day (17). Thus, across the ~30,000 base pair genome, approximately 20 genetic changes occur per year. Since the emergence of the virus in December 2019, thousands of variants of SARS-CoV-2 have emerged (18). The WHO defined the SARS-CoV-2 variants of concern (VOC) as a variant with increased transmissibility, virulence, and decreased response to available diagnostics,

vaccines, and therapeutics (19). Based on the recent epidemiological updated by WHO, as of January 10, 2022, five SARS-CoV-2 VOCs have been identified since the pandemic's beginning (20). Alpha (B.1.1.7) was the first VOC described in the United Kingdom in late December 2020. Then came the beta (B.1.351) and Gamma (P.1) variants which were first reported in South Africa in December 2020 and Brazil in January 2021, respectively. Until most recently, Delta (B.1.617.2) has been the most dominant variant. It was first reported in India in December 2020. Lastly, the Omicron (B.1.1.529) variant has become the most dominant variant, with its origin in South Africa in late November 2021 (19). All five reported VOCs have mutations in the RBD domain, of which N501Y mutation is common to all variants except the Delta variant.

The virus's origin from the spillover of a zoonotic pathogen, and the broad host range of the virus is partly because ACE2 is found in all major vertebrate groups (21). The ubiquity of ACE2 coupled with the high prevalence of SARS-CoV-2 in the global human population explains the multiple spillback infections since the emergence of the virus in 2019. In spillback infection, the human hosts transmit the SARS-CoV-2 virus to cause infection in non-human animals. In addition to threating wildlife and domestic animals, the repeated spillback infection may lead to the establishment of new animal hosts from which SARS-CoV-2 can then pose a risk of secondary spillover infection to humans through bridge hosts or new established enzootic reservoirs. There is a small number of reports of human to animal transmission of SARS-CoV-2 in pet cats and dogs and gorillas, tigers, lions, and other felines in zoos in the USA, Europe, and South Africa (22, 23). Initial human to animal transmission has resulted in sustained outbreaks in farmed mink in Europe and North America, with likely mink to human transmission reported in the Netherlands (24, 25). Experimental infection of Egyptian fruit bats (Rousettus aegyptiacus) resulted in transient subclinical infection with oral and fecal shedding (26). Given that the

probable sources of SARS-CoV-2 or its progenitor are bat species, the potential risk of reverse zoonotic transmission from humans to bats and the subsequent negative impacts have to be recognized and studied. Furthermore, a significant concern in such secondary spillover events is the appearance of a mutant strain affecting the host range or leading to increased transmissibility in humans, reduced sensitivity to neutralizing antibodies, and reduced vaccine efficacy.

Here, we utilize a simple systematic approach of comparative statistical analyses to study the dynamic motions of proteins generated from accelerated all-atom molecular dynamics (MD) simulations. We utilized an MD simulation software developed by our lab (DROIDS - Detecting Relative Outlier Impacts in Dynamic Simulation or DROIDS 4.0) to look at the binding signature between ACE2 and the RBD domain of the spike protein. Our software employs a site-wise Kullback-Leibler (KL) divergence metric and a multiple test corrected two-sample Kolmogorov-Smirnov (KS) test to compare atomic fluctuations between the unbound RBD and RBD domain bound to ACE2. This site-wise comparison of the virus in its bound and unbound state represents a direct quantitative measure of its functional fitness that can be compared across different variants of concern and different ACE2 orthologs. This enables us to observe the functional effects of mutations in different SARS-CoV-2 strains as they interact with various ACE2 orthologs. As we are primarily interested in how rapidly the virus may have adapted from bats to humans and conversely may no longer be able to effectively interact with bat populations, we focused our analyses on the human ACE2 and its ortholog in the greater horseshoe bat Rhinolophus affinis. We specifically looked at comparative dynamics of structural models of different SARS-CoV-2 variants bound to human ACE2 (hACE2) and bat ACE2(bACE2). In the RaTG13 variant, we see comparatively better binding atomic fluctuations to bACE2 rather than hACE2. However, in the case of several human variants, we do not see a difference in binding

atomic fluctuation between the hACE2 and bACE2. Due to the similarity in binding of the human variants, including VOC, between bACE2 and hACE2, our results indicate that the reverse zoonosis of the human variants back to bats is very high probable, which could lead to the emergence of new variants in both humans and Rhinolophus bats.

Materials and Methods

PDB structure and model preparation

Structures of the RBD of spike protein, hACE2, and Greater Horseshoe Bat Rhinolophus macrotis bACE2 were obtained from the Protein Data Bank (PDB). The summary of the structures used for MD simulations is listed in Table 1. Upon downloading the structures from PDB, any crystallographic reflections and other small molecules used in crystallization were removed. Furthermore, any glycans present in the structure were also removed. During the cleanup of the PDB, any glycans were removed and then later rebuilt using glycam.org's glycoprotein builder so that the PDB file structure regarding atom types was compatible with the Amber 20 preprocessing software tLeAP. See details below. When preparing the structures, we needed each of the variants (RaTG13, Wuhan WT, Alpha, Beta, Delta, Kappa, Epsilon, and Omicron) bound to hACE2 and bACE2. We were able to find structures in the PDB where the RaTG13 variant was bound to bACE2 and the human variants bound to hACE2. Therefore, to model the RaTG13 variant bound to hACE2 and the human variants bound to bACE2, we used UCSF Chimera's MatchMaker superposition tool to properly place the receptor belonging to the opposite species (27). Using pdb4amber (AmberTools20), hydrogen atoms were added, and crystallographic waters were removed (28). Any mossing loop structures in the files were inferred via homology modeling using the "refine loop" command to modeler within UCSF chimera (29, 30).

Table 1: Table summarizing the primary models (RBD variants bound to hACE2/bACE2) used for MD simulations. bACE2 denotes ACE2 of bat origin, and hACE2 denotes ACE2 of human origin.

RBD Variant	PDB ID	Receptor	PDB ID
RaTG13 (BatCoV)	7CN4	bACE2	7C8J
RaTG13 (BatCoV)	7CN4	hACE2	6VW1
Wuhan WT (nCoV)	7C8D	bACE2	7C8J
Wuhan WT (nCoV)	7C8D	hACE2	6VW1
Alpha (B.1.1.7)	7LWV	bACE2	7C8J
Alpha (B.1.1.7)	7LWV	hACE2	6VW1
Beta (B.1.351)	7LYO	bACE2	7C8J
Beta (B.1.351)	7LYO	hACE2	6VW1
Delta (B.1.617.2)	7V7Q	bACE2	7C8J
Delta (B.1.617.2)	7V7Q	hACE2	6VW1
Kappa (B.1.617.1)	7V7E	bACE2	7C8J
Kappa (B.1.617.1)	7V7E	hACE2	6VW1
Epsilon (B.1.429)	7N8H	bACE2	7C8J
Epsilon (B.1.429)	7N8H	hACE2	6VW1
Omicron (B.1.1.529)	7T9J	bACE2	7C8J
Omicron (B.1.1.529)	7T9J	hACE2	6VW1

Model glycosylation

As mentioned previously, glycans present in the original PDB structure were deleted. Predicted glycosylation was rebuilt for the Amber forcefield using the glycoprotein builder on the glycam.org web server (31). The glycans were rebuilt using the GLYCAM-06j-1 force field (32). 2-acetamido-2-deoxy-beta-D-glucopyranose was attached to ASN10 in the RBD of all variants. Similarly, 2-acetamido-2-deoxy-beta-D-glucopyranose was attached to ASN227, ASN264, and ASN503, and 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose was attached to ASN720 of bACE2. Lastly, 2-acetamido-2-deoxy-beta-D-glucopyranose was attached to ASN271, and 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose was attached to ASN221, ASN258, ASN490, and ASN714 of hACE2.

Molecular dynamic simulation protocols

MD simulation protocol was followed as previously described, with slight modifications (33-35). Briefly, for each MD comparison, large replicate sets of accelerated MD simulations were prepared and then conducted using the particle mesh Ewald method implemented on the graphical processor unit (GPU) hardware by pmemd.cuda (Amber20) (36-38). The MD simulations were either performed on a Linux mint 19 operating system (two RTX 2080 Ti or two RTX 3080 Ti) or on high performance computing cluster (A100). All comparative MD analysis via our DROIDS pipeline was based upon 100 replicate sets of 1 nanosecond accelerated MD run (i.e., 100x1ns MD run in each comparative state RBD, e.g., RBD bound to the receptor, unbound RBD). Explicitly solvated protein systems were first prepared using teLeap (AmberTools 20) using the ff14SB protein forcefield, in conjunction with the GLYCAM_06j-1 forcefield (32, 39). Solvation was generated using the Tip3p water model in a

12nm octahedral water box (40). Automated charge neutralization was also done with teLeap software with Na+ and Cl- ions. Each replicate set was preceded by energy minimization, 300 picoseconds of heating to 300K, a ten nanosecond of equilibration, followed by random equilibration intervals for each replicate ranging from 0-0.5 nanoseconds. All simulations were regulated using the Anderson thermostat at 300K and one atmospheric pressure (41). Root mean square atom fluctuations and atom correlations were conducted in CPPTRAJ using the atomicfluct and atomicorr commands (42).

Comparative protein dynamic analyses with DROIDS 4.0 and statistical analyses

Comparative signatures of dampened atom fluctuation during RBD binding to ACE2 were presented as protein site-wise divergence in atom fluctuation in the ACE2 bound versus unbound states for each RBD. Divergences were calculated using the signed symmetric Kullback-Leibler (KL) divergence calculation in DROIDS 4.0. Significance tests and p-values for these site-wise differences were calculated in DROIDS 4.0 using two-sample Kolmogorov-Smirnov tests with the Benjamini-Hochberg multiple test correction in DROIDS 4.0. The mathematical details of DROIDS 4.0 site-wise comparative protein dynamics analysis were published previously by our group and can be found here (33, 35). This code is available at our GitHub web landing: https://gbabbitt.github.io/DROIDS-4.0-comparativeprotein-dynamics/ , which is also available at our GitHub repository https://github.com/gbabbitt/DROIDS-4.0-comparative-protein-dynamics.

Results

Stronger binding of RatG13 RBD to bACE2 than to hACE2

We performed MD simulations of RaTG13 RBD bound and unbound to bACE2 and RaTG13 RBD bound and unbound to hACE2. To compare atomic fluctuations between RatG13 RBD bound and unbound structures, we used site-wise KL divergence along with multiple test corrected two-sample KS tests. The more negative the KL divergence value of a specific amino acid residue, the stronger the dampening of atomic fluctuations due to the RBD interactions with bACE2/hACE2. As one would expect, the bat RaTG13 RBD has a better binding with stronger amino acid-specific interactions with the bACE2 (Figure 1A, 1C). However, in the case of hACE2, dampening of atomic fluctuations is lesser at those specific sites due to ACE2 being of human origin (Figure 1A, 1D). Interestingly, the amino acid residues involved with the interactions of both bACE2 and hACE2 are very similar. This was observed in the normalized KL divergence graph, where the normalized KL divergence values per amino acid are similar between bACE2 and hACE2 (Figure 2A). Other studies have found that 26 residues of the bACE2 and nine residues of the RaTG13 are present at the interface. These residues create 12 H-bonds, two salt bridges, and 157 non-bonded contacts (43). The residues of RatG13 RBD that are strongly dampened by bACE2 include K417, L455, F456, S477, N487, and D501 (Figure 1A). The weaker dampening of atomic fluctuations of RaTG13 RBD and hACE2 is primarily due to the lesser number of interactions between the RBD and hACE2. Compared to bACE2, hACE2 only makes 113 non-bonded contacts and 9 H-bonds (44). Lastly, MD comparison of RaTG13 bound to bACE2 and RaTG13 bound to hACE2 show that almost half of the amino acids in the RBD of RaTG13 behave statistically different (Figure 1B). When comparing RaTG13 RBD bound to hACE2 with RaTG13 RBD bound to bACE2, the significance tests are conducted site-

wise. Therefore, a separate test was conducted at each given amino acid site to compare the significant difference in fluctuations of the backbone atoms. Thus, a separate D value from a two-sample KS test is obtained for each amino acid site, and a multiple test correction (Benjamini-Hochberg) was applied to adjust the p-value to account for multiple significance tests. Interestingly, there is no statistical difference in the atomic fluctuation dampening of some of the residues that interact with ACE2. In contrast, a majority of the interacting residues have statistically different atomic fluctuations between bACE2 and hACE2 ($p < 0.001$) (Figure 1B).

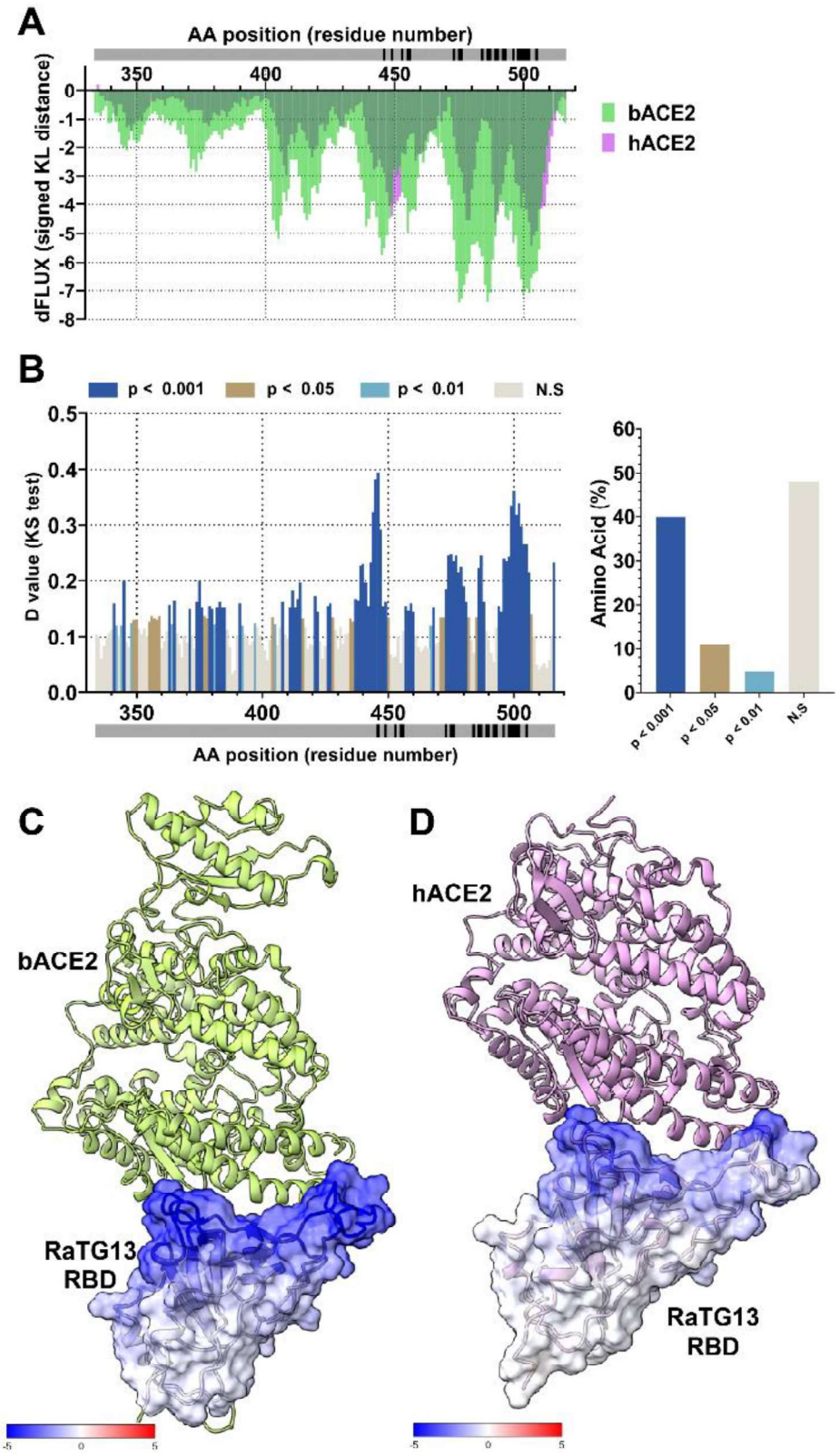

A
AA position (residue number)
350
400
450
500
dFLUX (signed KL distance)
0
-1
-2
-3
-4
-5
-6
-7
-8
bACE2
hACE2
B
p < 0.001
p < 0.05
p < 0.01
N.S
D value (KS test)
0.5
0.4
0.3
0.2
0.1
0.0
350
400
450
500
AA position (residue number)
Amino Acid (%)
60
50
40
30
20
10
0
p < 0.001
p < 0.05
p < 0.01
N.S
C
bACE2
RaTG13
RBD
-5
0
5
D
hACE2
RaTG13
RBD
-5
0
5

Figure 1: Analysis of atomic fluctuation differences of RaTG13 RBD bound to bACE2 and hACE2. (A) Sequence positional plotting of dampening of atom motion on RaTG13 RBD by bat ACE2 (bACE2, green) and human ACE2 (hACE2, pink). (B, left panel) Multiple test corrected two-sample KS tests of significance for the difference in atomic fluctuations of RaTG13 RBD bound to bACE2 and RaTG13 RBD bound to hACE2. The grey bar denotes the RBD domain amino acid backbone with the location of RBD residues interacting with ACE2 shown in black. (B, right panel) Percent of amino acid of the RaTG13 RBD with different levels of significance. N.S. denotes no significance. The change in atom fluctuation is due to the (C) bACE2 and (D) hACE2 interactions with RaTG13 RBD (PDB 7CN4). Dark blue denotes a KL divergence value of -5, with red denoting a KL divergence value of +5. bACE2 (PDB 7C8J) is shown in green, and hACE2 PDB 6VW1) shown in pink.

Compared with VBM, VOCs have better binding to hACE2 than to bACE2

In addition to RaTG13 RBD bound and unbound to bACE2 and hACE2, we performed MD simulations of other variants. These variants include the original Wuhan-Hu-1 strain, alpha (B.1.1.7), beta (B.1.351), delta (B.1.617.2), epsilon (B.1.427), kappa (B.1.1617.1), and omicron (B.1.629). Similar to RaTG13, the MD simulation of the different variants included the RBD bound and unbound to bACE2 and hACE2. Upon calculating the site-wise KL divergence for the RBD bound/unbound of the different variants, we normalized the KL divergence value. The residue with very little atomic fluctuation dampening was set to 0, and the residue with the strongest atomic fluctuation dampening was set to 100 (Figure 2). Both RaTG13 and the original Wuhan-Hu-1 RBDs have a similar binding dynamic between the bACE2 and hACE2. The peaks of normalized KL divergence values between the bACE2 and hACE2 of the two variants being identical corresponds to the similar amino acid residues interacting with the bACE2 and hACE2

(Figure 2A, 2B). We performed a similar analysis with the variants being monitored (VBM). The VBM includes beta, epsilon, and kappa variants. Like the original bat progenitor and the first human variant, the VBM show identical normalized KL divergence plots between the hACE2 and bACE2. Even at the sites corresponding to the VBM mutation, the KL divergence values appear to be very similar (Figure 2D, 2F, 2G). Lastly, we also looked at the VOC's interactions with bACE2 and hACE2. The VOC included alpha, delta, and omicron. Surprisingly, with the alpha variant, we see very similar normalized KL divergence plots between the bACE2 and hACE2 (Figure 2C). However, we see quite a difference in the delta and omicron variant plots. The normalized KL divergence values show additional peaks in the simulations with the hACE2. These additional peaks correspond to delta and omicron mutations (Figure 2E, 2H).

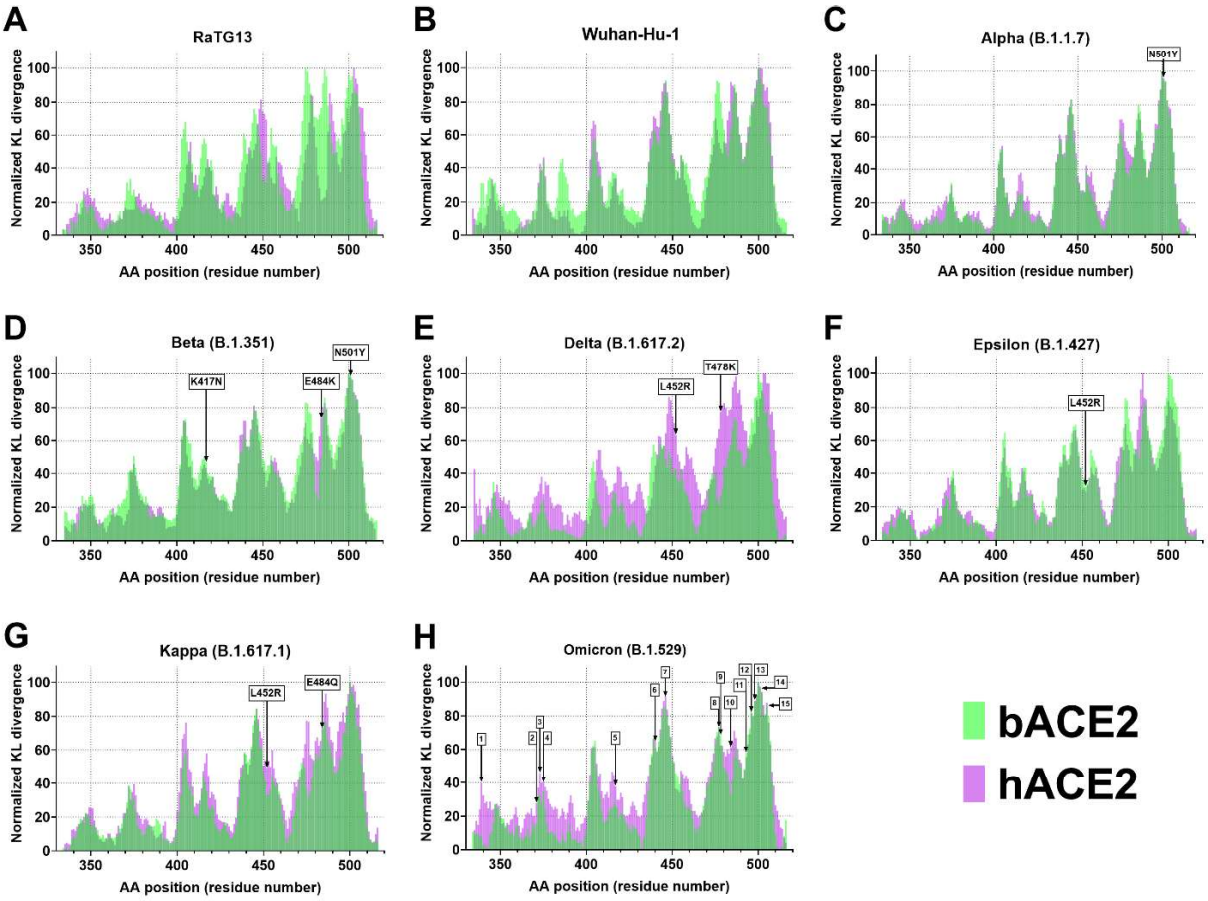


Figure 2: Binding interaction of the different SARS-CoV-2 variants with bACE2 and hACE2. Sequence positional plotting of normalized dampening of atom motion on (A) RaTG13, (B) Wuhan-Hu-1, (C) Alpha (B.1.1.7), (D) Beta (B.1.351), (E) Delta (B.1.617.2), (F) Epsilon (B.1.427), (G) Kappa (B.1.617.1) and (H) Omicron (B.1.529) RBD by bat ACE2 (bACE2, green) and human ACE2 (hACE2, pink). The amino mutations denoted by the arrows correspond to the variant mutation. The list of omicron mutations in (H) include (1) G339D, (2) S371L, (3) S373P, (4) S375F, (5) K417N, (6) N440K, (7) G446S, (8) S477N, (9) T478K, (10) E484A, (11) Q493K, (12) G496S, (13) Q498R, (14) N501Y, (15) Y505H.

Lastly, to quantify the interaction of the different variants with bACE2 and hACE2, we also calculated the area under the curve (AUC) of the non-normalized KL divergence values, true atomic fluctuation dampening due to the RBD interaction with ACE2. As expected, the bat progenitor strain, RaTG13, has a higher AUC with bACE2 than hACE2. Since the Wuhuna-Hu-1 RBD is the first variant to make the transmission from bats to humans, the AUC profile is similar for both bACE2 and hACE2 (Figure 3A). Interestingly, all VOC (alpha, delta, omicron) have a higher AUC for the hACE2 than bACE2. When a paired t-test was performed to compare the AUC values of the VOC bound to bACE2 and hACE2, we saw a significantly higher AUC profile of the VOC with hACE2 (Figure 3A, 3B). And lastly, there is no clear trend for the AUC profile of the VBM. This is also clearly seen with no significant difference in AUC values of VBM between bACE2 and hACE2 (Figure 3A, 3C).

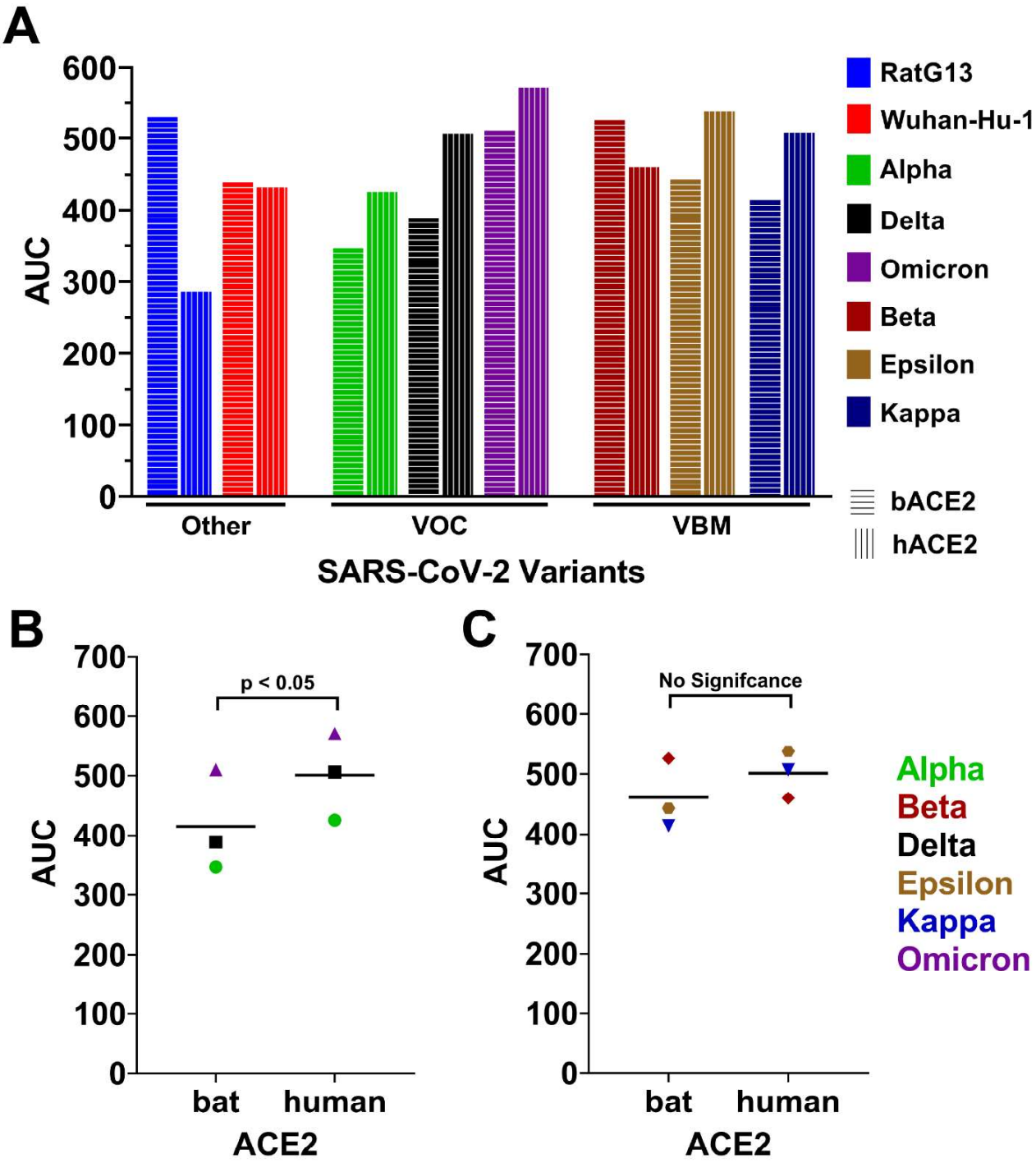

A
600
500
400
300
200
100
0
AUC
Other
VOC
VBM
SARS-CoV-2 Variants
RatG13
Wuhan-Hu-1
Alpha
Delta
Omicron
Beta
Epsilon
Kappa
bACE2
hACE2
B
700
600
500
400
300
200
100
0
AUC
p < 0.05
bat
human
ACE2
C
700
600
500
400
300
200
100
0
AUC
No Signifcance
bat
human
ACE2
Alpha
Beta
Delta
Epsilon
Kappa
Omicron

Figure 3: Binding dynamic differences between VOC and VBM with ACE2. (A) Area under the curve (AUC) values of the non-normalized KL divergence values of the different SARS-CoV-2 RBD bound and unbound to bACE2 and hACE2. Alpha, delta, and omicron variants are classified as variants of concern (VOC); beta, epsilon, and kappa variants are classified as variants being monitored (VBM). RaTG13 and Wuhan-Hu-1 are classified as others. Data points in (B)VOC and (C)VBM correspond to the AUC value of the different variants bound bACE2 and hACE2. The horizontal line corresponds to the mean value. A paired t-test was conducted to calculate the statistical significance.

Stronger binding of Delta RBD to hACE2 than to bACE2

The delta variant (B.1.617.2) was first identified in India in October 2020. The two mutations that make up the RBD of the delta variant include L452R and T478K. The site-wise KL divergence interaction of the delta RBD between bACE2 and hACE2, the hACE2 certainly has better binding dynamics than the bACE2 (Figure 4A). KS test between the delta RBD bound to bACE2 and delta RBD bound to hACE2 shows very few amino acids with statistically different atomic fluctuations (35%). About 65% of the amino acid of delta RBD behave very similarly when either bound to bACE2 or hACE2. At the sites corresponding to the delta variant mutation (L452R and T478K), we see a high level of significance in the difference in atomic fluctuation between the bACE2 and hACE2 (Figure 4B). As mentioned previously, when comparing delta RBD bound to hACE2 with delta RBD bound to bACE2, significance tests were conducted in a site-wise manner, with a D value from a two-sample KS test calculated for each amino acid, and Benjamini-Hochberg multiple test correction was applied to adjust the p-values. The test correction was done to account for the multiple significance tests. Lastly, compared to bACE2,

the binding of the hACE2 also has a dampening effect on the residues that are farther away from the RBD/ACE2 interface (Figure 4C and 4D).

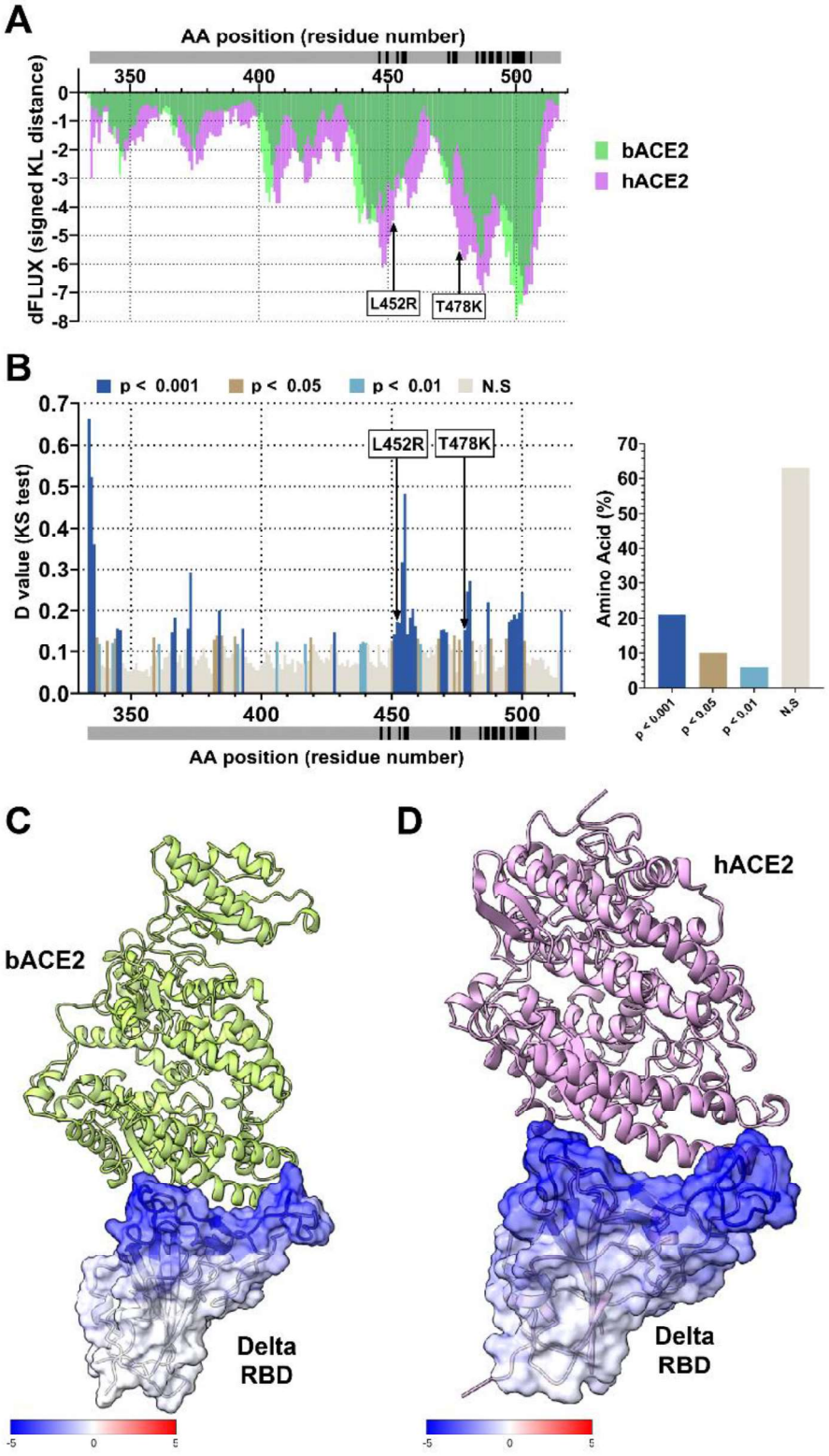


Figure 4: Analysis of atomic fluctuation differences of Delta (B.1.617.2) RBD bound to bACE2 and hACE2. (A) Sequence positional plotting of dampening of atom motion on delta RBD by bat ACE2 (bACE2, green) and human ACE2 (hACE2, pink). (B, left panel) Multiple test corrected two-sample KS tests of significance for the difference in atomic fluctuations of delta RBD bound to bACE2 and RaTG13 RBD bound to hACE2. The grey bar denotes the RBD domain amino acid backbone with RBD residues of interaction with ACE2 shown in black. (B, right panel) Percent of amino acid of the delta RBD with different levels of significance. N.S. denotes no significance. Arrows in (A) and (B) correspond to the delta variant mutations (L452R and T478K). The change in atom fluctuation is due to the (C) bACE2 and (D) hACE2 interactions with delta RBD (PDB 7V7Q). Dark blue denotes a KL divergence value of -5, with red denoting a KL divergence value of +5. bACE2 (PDB 7C8J) is shown in green, and hACE2 (PDB 6VW1) is pink.

Omicron RBD mutations influence binding to hACE2

Unlike the delta variant, the omicron variant (B.1.529) was most recently identified in South Africa in November 2021. Compared to all other variants, omicron variants include 15 mutations in the RBD. Comparison of the atomic fluctuation of the omicron RBD bound to bACE2 and hACE2 revealed certain amino acids with stronger dampening of atomic fluctuations when bound to hACE2 (Figure 5A). Furthermore, statistically significant differences in the atomic fluctuation of the omicron RBD bound to bACE2 and hACE2 are only observed in 30% of the amino acids of the omicron. Of those different amino acid positions, 9 of them correspond to the omicron RBD mutations (G339D, S371L, S373P, S345F, K417N, E484A, Q493K, G496S, and Q498R) (Figure 5B). Similar to the delta variant, the RBD of the omicron variant

was furthermore stabilized when bound to hACE2. This can be seen by the dampening of residues by hACE2 that are farther away from the RBD/hACE2 interface (Figure 5D). However, in the case of the bACE2, there is only dampening happening at the RBD/bACE2 interface (Figure 5C).

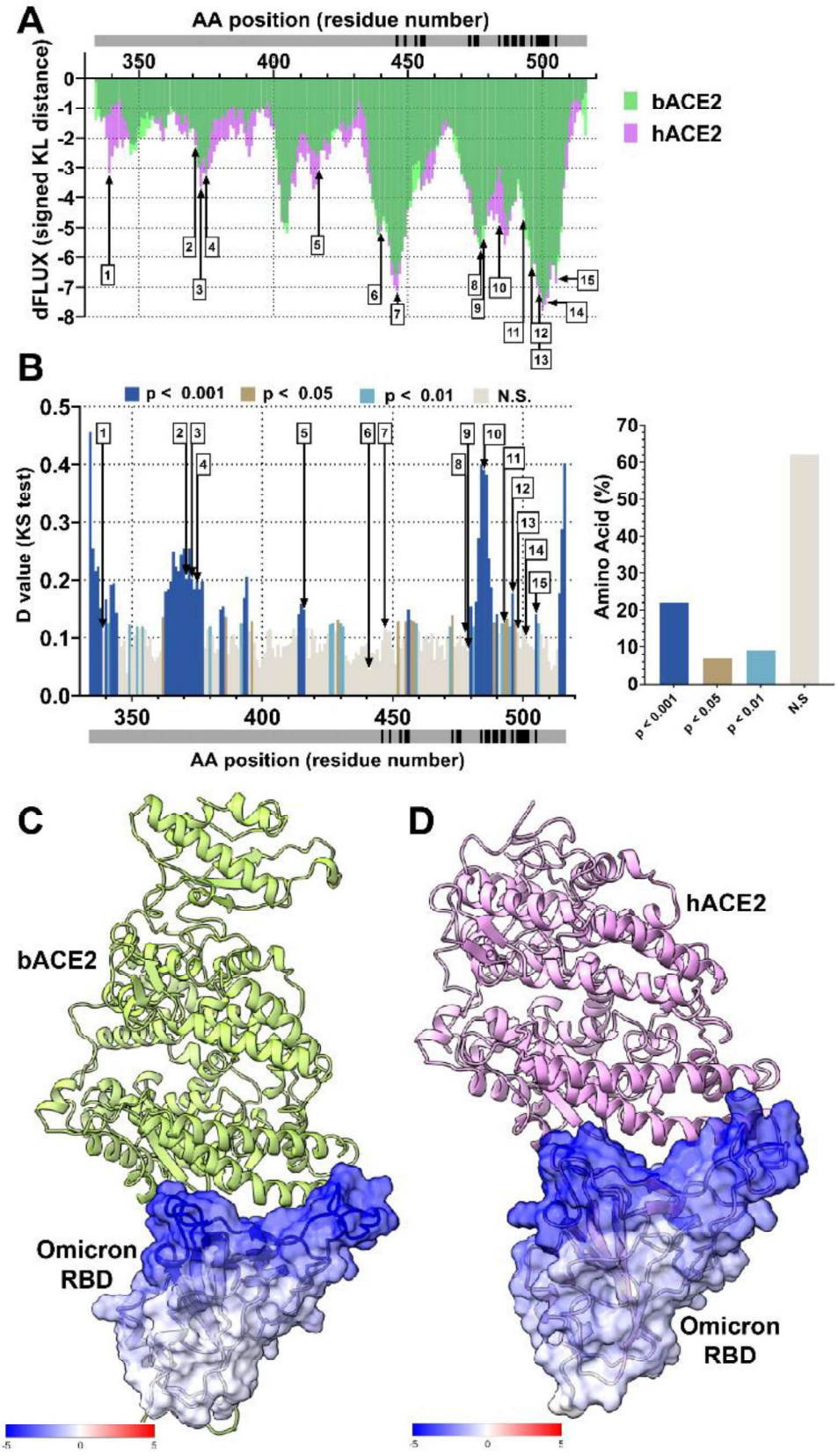

A
AA position (residue number)
350
400
450
500
dFLUX (signed KL distance)
bACE2
hACE2
B
p < 0.001
p < 0.05
p < 0.01
N.S.
D value (KS test)
AA position (residue number)
Amino Acid (%)
p < 0.001
p < 0.05
p < 0.01
N.S
C
bACE2
Omicron RBD
D
hACE2
Omicron RBD

Figure 5: Analysis of atomic fluctuation differences of Omicron (B.1.529) RBD bound to bACE2 and hACE2. (A) Sequence positional plotting of dampening of atom motion on omicron RBD by bat ACE2 (bACE2, green) and human ACE2 (hACE2, pink). (B, left panel) Multiple test corrected two-sample KS tests of significance for the difference in atomic fluctuations of omicron RBD bound to bACE2 and omicron RBD bound to hACE2. The grey bar denotes the RBD domain amino acid backbone with RBD residues of interaction with ACE2 shown in black. (B, right panel) Percent of amino acid of the omicron RBD with different levels of significance. N.S. denotes no significance. Arrows in (A) and (B) correspond to the omicron variant mutations. The list of omicron mutations includes (1) G339D, (2) S371L, (3) S373P, (4) S375F, (5) K417N, (6) N440K, (7) G446S, (8) S477N, (9) T478K, (10) E484A, (11) Q493K, (12) G496S, (13) Q498R, (14) N501Y, (15) Y505H. The change in atom fluctuation is due to the (C) bACE2 and (D) hACE2 interactions with omicron RBD (PDB 7T9K). Dark blue denotes a KL divergence value of -5, with red denoting a KL divergence value of +5. bACE2 (PDB 7C8J) is shown in green, and hACE2 (PDB 6VW1) shown in pink.

Discussion / Conclusion

Using our comparative MD simulation pipeline, we compared the binding profile of bACE2 and hACE2 with the RBD domains of different SARS-CoV-2 variants (RaTG13, Wuhan-Hu-1, alpha, beta, delta, kappa, epsilon, and omicron). MD simulations of RatG13 spike protein revealed a better binding profile and stronger dampening of amino acids on the interface of RBD and bACE2 compared to hACE2. In the case of the RBD from Wuhan-Hu-1, the original 2019 SARS-CoV-2 virus, we see no difference in the binding profile between bACE2 and hACE2. Lastly, we also observed two different profiles with the VBM and VOC. In the case of VBM (beta, kappa, and epsilon variants), we see a similar binding profile and atomic

fluctuation dampening in the RBD/bACE2 and RBD/hACE2 interfaces. On the other hand, VOCs (alpha, delta, and omicron variants) show preferable binding to hACE2 than to bACE2, indicating higher AUC values when RBD is bound to hACE2 than RBD bound to bACE2.

The SARS-CoV-2 virus was first reported from pneumonia patients of the Wuhan city in China's Hubei province. The spillover of SARS-CoV-2 from animals to humans occurred at the beginning of December 2019, when some of the pneumonia patients were involved in the wet animal market in the Hunnan district (45). Genomic sequences, homology of ACE2 receptor, and single intact ORF on gene 8 of the virus indicate bats as the natural reservoir of these viruses. However, an unknown animal is yet to be identified as an intermediate host (46, 47). It should be noted that even though the initial spread of the disease was due to a spillover even, the rapid spread of the disease was primarily due to human-to-human transmission (45).

The receptor usage by the corona virus has been well known to be a significant determinant of host range, tissue tropism, and pathogenesis. Therefore, it is reasonable to assume that SARS-CoV-2 can infect humans, bats, and other species. As a matter of fact, several in vivo infection and seroconversion studies have confirmed that SARS-CoV-2 can infect rhesus monkeys, feline, ferret, and canines (48, 49). Our MD simulation has shown that RaTG13 RBD can bind to both bACE2 and hACE2, however to bACE2 with stronger binding, shown by severe dampening of atomic fluctuations in the bACE2/RBD interface (Figure 1A). Even though hACE2 doesn’t severely dampen the atomic fluctuations, the KL divergence profile is very similar to that of bACE2, indicating the flexibility of the virus to jump hosts. Additionally, it has also been shown that the RaG13RBD Y493 is speculated to confer a potential steric clash to hACE2 (50).

As mentioned previously, for zoonotic spillover events to occur, humans must be exposed to the viruses. This can occur through direct contact with viruses excreted from infected bats or bridge hosts or through other contacts with infected animals such as slaughtering or butchering. The nature and intensity of the bat–human interface are critical to determining spillover risk. Human behavior is a primary determinant of exposure, which may increase contact with bats or with other animals (bridge hosts) that may expose susceptible humans. Little is known about the specific conditions of coronavirus spillovers. Still, human behaviors that may increase viral exposure include activities such as bat hunting and consumption, guano farming, and wildlife trading (51-53). Due to recent spillover events, we see an identical binding profile with both bACE2 and hACE2 in the case of Wuhan-Hu-1 RBD (figure 2B). In the case of Wuhan-Hu-1 RBD, our MD simulation shows identical dampening of the atomic fluctuations and that the amino acid backbone in the RBD interacts with both bACE2 and hACE2 identically.

Regardless, in the case of human SARS-CoV-2 strains, we see a slightly different trend. In the case of VOC, including alpha, delta, and omicron variants, we see a slightly higher binding profile with hACE2 than bACE2 (Figure 3A, 3B, 4, and 5). However, unlike RaTG13 RBD, we do not see very strong differences in the atomic fluctuations in the bACE2/hACE2 interface with the VOC RBDs. We don't see any significant differences in the binding profile with the VBM, including beta, epsilon and kappa variants. At some areas of the RBD, the atomic fluctuations of the VBM are very similar between bACE2 and hACE2. The ability of the recent VOC and VBM to bind to both hACE2 and bACE2 with only slight differences supports the reverse zoonosis theory. The transmission of SARS-CoV-2 from humans to numerous animals and conducted in vitro infection experiments make it clear that the virus can infect and be transmitted between a wide range of distantly related mammal species. For example, case reports

on cats (Felis catus) living in the same household with COVID-19 patients in Europe, Asia, North America, and South America revealed that these animals could be infected with SARS-CoV-2, showing clinical manifestations ranging from asymptomatic to severe respiratory illness (54, 55). The reports show that 14% of tested cats in Hong Kong were SARS-CoV-2 positive by RT-PCR (56).

Furthermore, the seroprevalence screening performed among pets living in SARS-CoV-2-positive households in Italy demonstrated that 3.3% of dogs and 5.8% of cats were seropositive (57). The high seroprevalence and SARS-CoV-2 detection rates in cats and, to some extent, in dogs indicate that these animals can be infected with SARS-CoV-2 (58). Several other zoo animals, like tigers, lions, cougars, and gorillas, were found to test positive for the virus. Farmed minks are highly susceptible to SARS-CoV-2 infection, and, in some cases, they have transmitted the virus back to humans. SARS-CoV-2-positive minks were detected in 290 fur farms in Denmark, 69 mink fur farms in the Netherlands, 13 of 40 mink farms in Sweden, 23 out of 91 mink farms in Greece, 17 fur farms in the USA, four farms in Lithuania, two farms in Canada, and one fur farm in Italy, Latvia, Poland, France, and Spain (59, 60). As a result of the virus being able to infect multiple species and is also able to jump hosts, there are concerns that the introduction and circulation of new virus strains in humans could result in modifications of transmissibility or virulence and decreased treatment and vaccine efficacy.

In conclusion, our MD simulations identified that the original bat progenitor RaTG13 RBD shows preferential binding to its host bACE2 receptor than hACE2. However, some of the recent human variants show differential binding between bACE2 and hACE2. Lastly, the VOC RBD show slightly higher binding to hACE2 than the bACE2. These findings provide evidence that recent SARS-CoV-2 variants may infect bats and that the extensive species diversity of bats

may have profound effects on SARS-CoV-2 evolution. Overall, our results indicate the possible inclusion of MD-based surveillance of the virus and the viral receptor, in addition to genomic surveillance.

Identifying protein sites contributing to vaccine escape via statistical comparisons of short-term molecular dynamics simulations

Abstract

The identification of viral mutations that confer escape from antibodies is crucial for understanding the interplay between immunity and viral evolution. We describe a molecular dynamics (MD) based approach that scales well to a desktop computer with a high-end modern graphics processor and enables the user to identify protein sites that are prone to vaccine escape in a viral antigen. We first implement our MD pipeline to employ site-wise calculation of Kullback-Leibler divergence in atom fluctuation over replicate sets of short-term MD production runs thus enabling a statistical comparison of the rapid motion of influenza hemagglutinin (HA) in both the presence and absence of three well-known neutralizing antibodies. Using this simple comparative method applied to motions of viral proteins, we successfully identified in silico all previously empirically confirmed sites of escape in influenza HA, predetermined via selection experiments and neutralization assays. Upon the validation of our computational approach, we then surveyed potential hot spot residues in the receptor binding domain of the SARS-CoV-2 virus in the presence of COVOX-222 and S2H97 antibodies. We identified many single sites in the antigen-antibody interface that are similarly prone to potential antibody escape and that match many of the known sites of mutations arising in the SARS-CoV-2 variants of concern. In the omicron variant, we find only minimal adaptive evolutionary shifts in the functional binding profiles of both antibodies. In summary, we provide a fast and accurate computational method to monitor hot spots of functional evolution in antibody binding footprints.

Introduction

In current attempts to prevent the spread of the Coronavirus disease 2019 (COVID-19) pandemic, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), clinicians and scientists have focused their efforts on the development of vaccines that are hoped to induce broad, long-lasting, neutralizing antibodies. However, the selective pressures imposed by the presence of the neutralizing antibodies in the host population can also drive the evolution of viruses towards adaptations that allow them to escape neutralization. For example, in the case of influenza viruses, immunity is provided by antibodies that target the hemagglutinin (HA), responsible for viral attachment and viral fusion to the host cell. However, these antibodies drive selection for amino acid substitutions in the HA, causing the influenza virus to rapidly evolve every year (61). Similarly, the antibody selection by the host immune system can also drive the emergence of new SARS-CoV-2 variants. Therefore, when developing vaccines that elicit antibodies against a broad range of strains, research efforts should also be aimed at identifying potential mutations that can facilitate viral escape from the neutralization effects of specific antibodies (62, 63).

The traditional approach to identifying these mutations is by passaging the virus in the presence of antibodies in a directed selection experiment, followed by validation of the variants that arise with neutralization assays. For example, in influenza viruses, escape mutant selection using a panel of monoclonal antibodies (mAbs) was used to identify the five major antigenic regions, Sa, Sb, Ca1, Ca2, and Cb (64-66). However, a significant drawback of this approach is that the directed selection typically only favors one of the many potential mutations that can escape a given antibody. Another approach is to test antibody binding to a panel of viral variants. In a typical 500 residue viral protein, there are about 104 potential single amino acid mutants

(67). Creating all individual mutants and then testing the mutants against the antibodies is an impossible task, thus causing researchers to limit themselves to exploring only a small portion of protein space (e.g., examining only mutations to alanine) (68, 69). Such studies cannot give a complete picture of the mutational spectra that can allow a virus to escape neutralization by a given antibody (70).

Functional evolutionary studies of viral vaccine escape are often supplemented with protein structural determination via x-ray crystallography or cryogenic electron microscopy. While structural biology can provide a static image of how an antibody physically contacts a viral protein, it cannot provide complete information regarding which amino acid sites are more prone to single replacement mutations (i.e. hot-spot sites or residues). Prediction of hot-spot residues is crucial as these sites on the antibody-antigen complex have a strong propensity to disrupt binding interactions within the antibody-antigen interface (71, 72). Recently, it has been demonstrated, via site-directed selection experiments and neutralization assays, that these hot-spot regions share a common biophysical feature. They all tend to harbor single amino acid sites that have significant large affects upon binding interactions in the antibody-antigen interface (73). Given this common feature of hot-spot residues, we hypothesize that some simple analyses of computer-based dynamics simulations of the antibody-antigen interface might help predict potential vaccine escape mutations before they happen, allowing for important functional context to real-time sequence-based surveillance of current and future pandemics.

Here we utilize a relatively simple method of comparative statistical analysis of molecular dynamic (MD) simulations developed by our lab (DROIDS - Detecting Relative Outlier Impacts in Dynamic Simulation or DROIDS 4.0) that employs a site-wise Kullback-Leibler (KL) divergence metric and a multiple test corrected two sample Kolmogorov-Smirnov

(KS) test to successfully validate previously known sites of antibody escape in the influenza HA (33-35, 73, 74). We then utilized our site-wise comparative MD approach to identify potential sites prone to antibody escape in the spike protein of SARS-CoV-2. Specifically, in the omicron variant we were able to identify sites in the receptor binding domain (RBD) that support the binding efficiency to two general neutralizing antibodies and its competitive binding to the natural receptor, human angiotensin-converting enzyme 2 (hACE2). In summary, we present a method to accurately identify hot-spot residues that are prone to single point mutations with large functional effects upon the antibody-antigen binding interface and thus are likely pre-adapted to allow for vaccine escape. Identifying such residues in silico will be essential for pre-screening the antigenic consequences of viral genetic variations and designing better vaccines that induce long-term and broadly neutralizing antibodies against viral pathogens.

Methods

PDB structure, glycosylation, and model preparation

The protein structures used for the primary models for analyzing the molecular dynamics of antibody interactions with HA and the SARS-CoV-2-RBD are listed in Figure 6A. Any crystallographic reflections were removed along with any other small molecules used in crystallization. Hydrogens were added, and crystallographic waters were removed using pdb4amber (AmberTools18) (28). Glycosylation was deleted using the swapaa function in UCSF Chimera 1.14 (27). Predicted glycosylation were rebuilt for the Amber force field using the glycoprotein builder on the glycam.org webserver and the GLYCAM06j-1 force field (31, 32). Any missing loop structure in the files was inferred via homology modeling using the 'refine loop' command to Modeler within UCSF Chimera (29, 30). The globular head domain of the influenza HA is stable in its monomeric form. However, the stalk domain needs to be in a trimer

to be stable (75). Therefore, MD simulations of PDB 4GMS was performed using trimmed (residues 57-270) monomeric form of the head domain. For MD simulations of the 3ZTJ and 4HLZ, the trimeric form of the stalk domain was used. The antibodies were trimmed, and only the heavy and light chains of the fragment antigen-binding (Fab) were used. Lastly, for the Omicron (B.1.1.529) variant simulations, we used the swapaa function to model the 15 RBD mutations onto 7OR9.

A

PDB ID	Antigen	Antibody	AA and position	-KL divergence	D value	p value
4GMS	A/Victoria/3/1975 H3 Head domain	S139/1	K156	-4.65	0.810	p < 0.001
			G158	-5.81	0.895	p < 0.001
			N193	-4.81	0.807	p < 0.001
3ZTJ	A/California/4/2009 H1 Stalk domain	FI6v3	E47	-2.41	0.647	p < 0.001
4HLZ	A/Japan/305/1957 H2 Stalk domain	C179	-	-	-	-
7OR9	SARS-CoV-2 RBD	COVOX-222	D405	-2.57	0.572	p < 0.001
			K417	-3.27	0.678	p < 0.001
			K458	-3.49	0.703	p < 0.001
			G476	-5.50	0.888	p < 0.001
			F486	-4.52	0.800	p < 0.001
			N501	-4.55	0.708	p < 0.001
7M7W	SARS-CoV-2 RBD	S2H97	N394	-4.25	0.760	p < 0.001
			D428	-3.81	0.692	p < 0.001
			K462	-3.37	0.723	p < 0.001
			E484	-5.20	0.837	p < 0.001
			L518	-6.56	0.945	p < 0.001
PDB ID	**Protein 1**	**Protein 2**	**AA and position**	**-KL divergence**	**D value**	**p value**
6VW1	SARS-CoV-2 RBD mutated to include VOC mutations	ACE2	K417N	-0.053	0.140	p < 0.001
			L452R	-0.206	0.155	p < 0.001
			T478K	-0.129	0.190	p < 0.001
			E484K	-0.204	0.165	p < 0.001
			N501Y	-0.208	0.130	p < 0.05

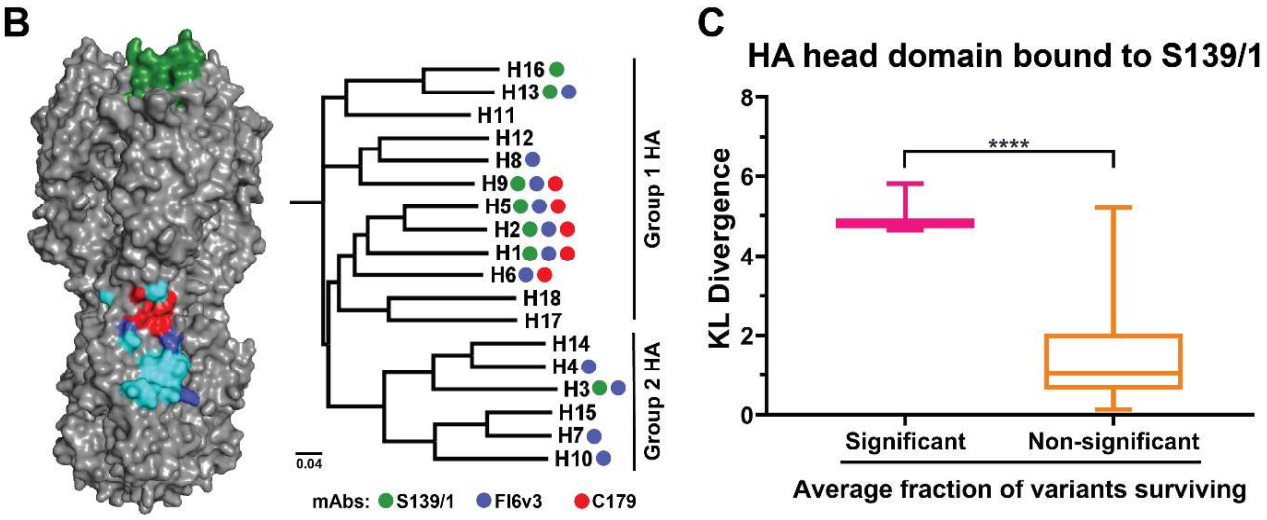


Figure 6: Description of PDB files and antibodies used in this study, and the comparison between KL divergence and average fraction of variants surviving. (A) Table summarizing the protein structure used for primary models for analyzing the molecular dynamics of antibody interactions with the influenza HA and SARS-CoV-2 RBD, and ACE2 interactions with wildtype and mutated SARS-CoV-2 RBD (includes VOC mutations). The table includes amino acid positions and the corresponding -KL divergence value denoting atomic fluctuations dampening for antibody-antigen/protein 1-protein 2 MD simulations. D value and the level of significance for the corresponding amino acid position are also given. (B) S139/1 (green), FI6v3 (blue), and C179 (red) epitopes are mapped onto the HA trimer, shown in grey (PDB 1RVX). Overlapping epitopes between FI6v3 and C179 are shown in cyan. Next to the HA trimer is a phylogenetic tree of HA subtypes. Circles denote reported antibody binding or neutralization against that subtype (C) Box and whisker plots showing the KL divergence values between amino acid sites that had significant average fraction of variants surviving and non-significant average fraction of variants in the presence of monoclonal antibody S139/1 (73). Only sites K156, G158, and N193 of the HA head domain had significant average fraction surviving viral particles.

Molecular dynamics simulation protocols

All comparative molecular dynamics analysis via our DROIDS pipeline was based upon 100 replicate sets of 1 nanosecond accelerated MD runs (i.e., 100x1ns MD runs in each comparative state, e.g., unbound vs. bound). MD simulations were conducted using the particle mesh Ewald method via accelerated MD (pmemd.cuda) in Amber18 and the ff14SB protein and GLYCAM06j-1 force fields and implemented on two RTX 2080 Ti graphics processor units

controlled via Linux Mint 19 operating system (32, 36-39, 76, 77). These replicate sets were preceded by energy minimization, 300 picoseconds of heating to 300K, and ten nanoseconds of equilibration, followed by random equilibration intervals for each replicate ranging from 0-0.5 nanoseconds. All protein systems were prepared via tLeAP (Ambertools18) and explicitly solvated, and charge neutralized with Na+ and Cl- ions in a Tip3P octahedral water box set to 12 nm beyond the surface of each protein with periodic boundaries (28, 78). All simulations were regulated using the Anderson thermostat at 300K and one atmospheric pressure (41). Root mean square atom fluctuations and atom correlations were conducted in CPPTRAJ using the atomicfluct and atomicorr commands (42).

Comparative protein dynamics analyses with DROIDS 4.0 and statistical analyses

Comparative signatures of dampened atom fluctuation during antibody binding were presented as protein site-wise divergence in atom fluctuation in the antibody-bound versus unbound states for each viral target protein. Divergences were calculated using the signed symmetric Kullback-Leibler (KL) divergence calculation in DROIDS 4.0. Significance tests and p-values for these site-wise differences were calculated in DROIDS 4.0 using two-sample Kolmogorov-Smirnov tests with the Benjamini-Hochberg multiple test correction in DROIDS 4.0. The mathematical details of DROIDS 4.0 site-wise comparative protein dynamics analysis were published previously by our group and can be found here (33, 35). This code is available at our GitHub web landing: https://gbabbitt.github.io/DROIDS-4.0-comparative-protein-dynamics/, is also available at our GitHub repository https://github.com/gbabbitt/DROIDS-4.0-comparative-protein-dynamics.

Results and Discussion

Using our comparative MD analysis pipeline (DROIDS 4.0), we successfully identified and validated previously known antibody escape sites in the head and stalk domains of the influenza HA. We then expanded the scope of our study to identify potential mutational sites prone to antibody escape in the spike protein of SARS-CoV-2 RBD and its recent omicron genetic variant.

Initially, we applied a Mann-Whitney U-test to compare KL divergence values, generated using MD simulations, to amino acid residues with significant average fraction of variants surviving and non-significant average fraction of variants surviving in the presence of S139/1, (Figure 6C) (73). In our comparative analyses of MD simulations, we found significantly higher KL divergence in the average fraction of variants surviving during directed selection ($p < 0.001$), suggesting that our KL divergence metric might prove as a useful quantitative measure in discriminating sites with mutations leading to vaccine escape. We then applied our approach to anti-HA antibodies with a wide range of breadth against different influenza strains. S139/1 is a one such monoclonal antibody known to neutralize both group 1(H1, H2, H5, H9, and H13) and group 2 (H3) influenza viruses (Figure 6B) (79). Crystallization studies have revealed that the antibody targets highly conserved residues in the RBD of the head domain (80) (Figure 6B). Comparative MD simulations of HA bound to S139/1 and unbound have revealed strong dampening of atom fluctuations occurring at sites K156, G158, and N193 of the HA, with -KL divergence values -4.65 ($D = 0.810$, $p < 0.001$), -5.81 ($D = 0.895$, $p < 0.001$), and -4.81 ($D = 0.807$, $p < 0.001$), respectively (Figure 6A and 7A). Directed selection using mAb S139/1 revealed that these three residues are sites of strong escape (Figure 7B) (73). As expected, the three amino acid residues with the highest negative KL divergence in our comparison of atom

fluctuations from MD simulations also fall directly in the empirically determined physical binding footprint of the antibody. And they are the same three sites where previous works have selected escape mutants in the H1, H2, and H3 HAs (79, 80). Residue 156 is a part of the 150 loop of the influenza HA and forms electrostatic interactions by inserting itself into the acidic pocket in the Fab formed by residues GluH35 and GluH50. Another 150 loop binding determinant is residue 158, which has the largest dampening of atomic fluctuations. Gly158 is closely stacked on the light chain complementarity determining region 3 (LCDR3) of S139/1 and is further stabilized by a main chain hydrogen bond between S159 and AsnL92. Any mutations at these positions will cause clashes at the antibody interface. S139/1 also recognizes members of the 190 helix. In the residues of 190 helix, Asn193 plays a significant role in antibody recognition and is buried by heavy chain complementarity determining regions (HCDR) 1-3 of S139/1 (79). Furthermore, it should be noted that site 193 was known to interact with the host receptor molecule (sialic acid moiety), suggesting the contribution of this residue to receptor binding of HA (81, 82). Therefore, isolates with mutated residues at site 193 have reduced viral fitness due to their inability to bind to the host receptors (83). Selective pressure created by S139/1 on residue 193 has other secondary effects on the virus, such as the reduced binding activity of the influenza virus.

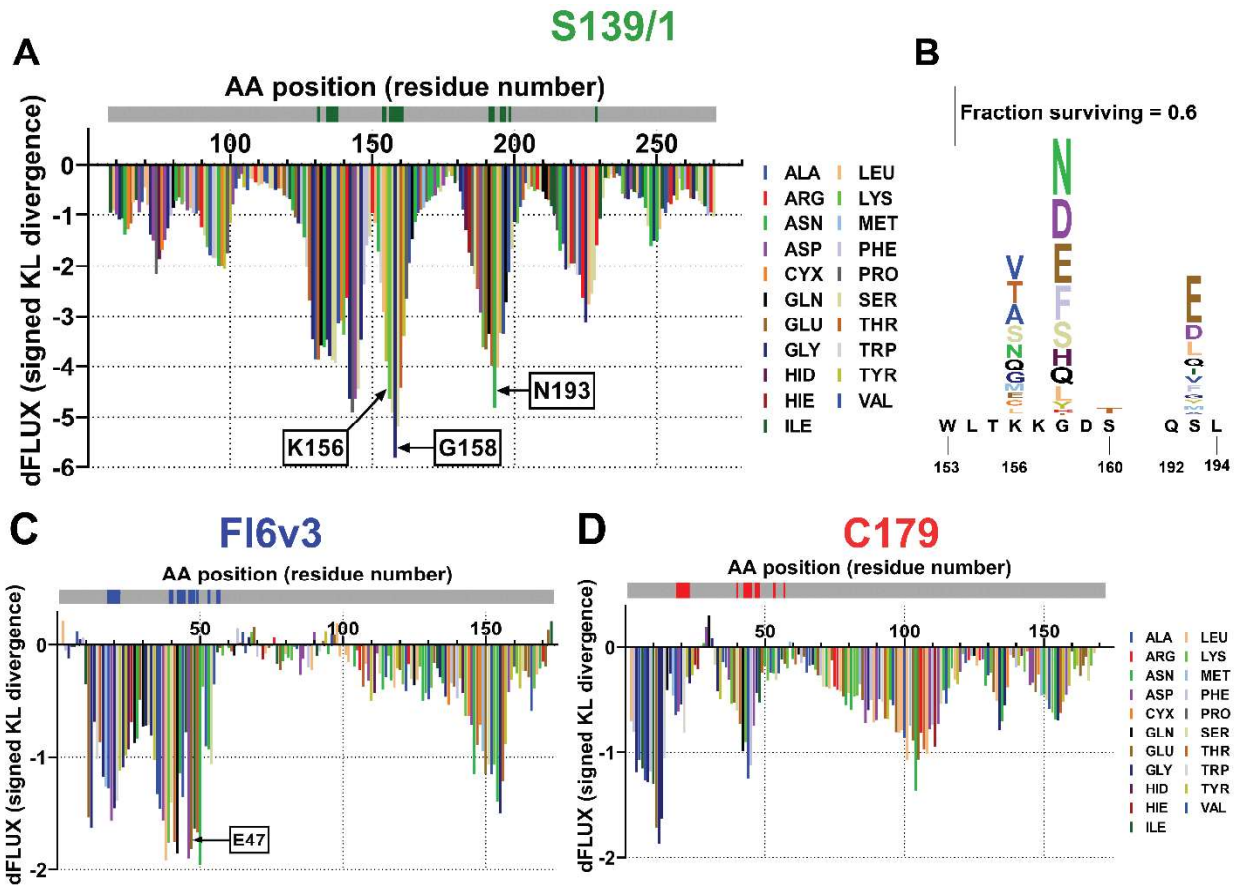


Figure 7: Epitope prediction and validation of hot-spot residues of the influenza hemagglutinin (HA) in the presence of neutralizing antibodies. Sequence positional plotting of dampening of atom motion on the influenza HA head domain by (A) S139/1 and on the HA stalk domain by (C) FI6v3 and (D) C179. The sequence profile of the -KL divergence between S139/1 and the head domain produces strong negative peaks in (A) at K156, G158, and N193. A modest negative peak is observed in the stalk domain in (C) at E47 in the presence of FI6v3. HA1 numbering is used to identify the amino acid positions in (A). HA2 numbering is used to identify the amino acid position in (C) and (D). The grey bar on top of the -KL divergence plots denotes the HA amino acid backbone with the location of (A) S139/1 epitopes shown in green, (C) FI6v3 epitopes in blue, and (D) C179 epitopes in red. (B) Logo plots of S139/1 show the amino acid position that has the largest effect. Letter heights are proportional to the excess fraction of virions

with that mutation that survive the antibody, as indicated by scale bars. The logo plot was prepared using deep mutational scanning data from Doud, M.B. *et al.* (73).

Lastly, with mAb S139/1, atomic fluctuation dampening is also seen in residues 133-134 of the influenza HA (Figure 7A). These residues are part of the 130 loop of the HA, and make a significant portion of the RBD of the influenza virus. Lastly, in addition to atomic fluctuation dampening at residues 133-134 and at residues 156/158, we also see dampening at 145-147 (Figure 7A). Upon further examination, there lie no epitopes of S139/1 between residues 145-147. Therefore, we believe that this dampening is an artifact caused purely by the residue's location between two epitope sites (130 loop and 150 loop) of S139/1. In summary, the three sites determined to have the largest physical effect upon antibody binding in our simulations (i.e. showing the largest negative KL divergence) are exactly the same three sites most likely to evolve antibody escape under directed selection experiments.

We also chose two broad antibodies that target the stalk domain of the HA (FI6v3 and C179) to further validate our computational approach to identify potential sites of escape. FI6v3 was first isolated by high-throughput screening of immortalized antibody-secreting cells and was found to bind to both group 1 (H1, H2, H5, H6, H8, H9, and H13) and group 2 (H3, H4, H7, and H10) viruses (Figure 6B) (84). Antibody C179 was first isolated from a mouse that had been immunized with the H2N2 virus and was later found to cross-neutralize H1, H2, H5, H6, and H9 subtypes (Figure 6B) (85). Both FI6v3 and C179 have epitopes that lie in the stalk domain and are known to interfere with membrane fusion (Figure 6B). Our MD simulations do not reveal dampening of atomic fluctuations in the stalk domain in the presence of FI6v3 or C179. When

we look at the difference in atomic fluctuations of the stalk antibodies on the same scale as S139/1, we only see a few sites with strong dampening (Figures 7C and 7D).

Antibody selection revealed very similar results as our MD simulation, with only few sites in the stalk domain with slightly increased fraction of variants surviving mAbs C179 and (73). As a result, the authors concluded that the stalk domain is less capable of escaping antibodies by single mutations (73). In the presence of FI6v3, both directed selection and MD simulation show a small bump at site 47 ($KL = -2.41$, $D = 0.647$, $p < 0.001$) that is of importance (Figures 6A and 7C) (73). Mutational studies have demonstrated that the introduction of E47R in the stalk domain has increased the resistance to FI6v3 (86). Like other studies, we also conclude that the HA stalk is intolerant of mutations, and confirmed by us here due to the lack of sites with dampened fluctuations. The absence of sites with large divergences in atomic fluctuations in our comparative MD simulations would seem to confirm that the anti-stalk antibodies studied here readily target sites with high mutational tolerance as suggested by other site directed experiments (87). Another possible explanation could be that the binding energetics at protein-protein interfaces can be asymmetrically distributed across all sites, thus preventing us from identifying the mutational tolerant sites (71, 72).

We implemented our computational approach to identify potential hot spot residues in the RBD of the spike protein of SARS-CoV-2 in the presence of two recently published antibodies against the virus: COVOX-222 and S2H97 (Figure 8A). COVOX-222 is known to bind to different residues than S2H97 and is known to neutralize strains P.1 (Gamma) from Brazil, B.1.351 (Beta) from South Africa, and B.1.1.7 (Alpha) from the United Kingdom (88). Termed "super-antibody," S2H97 is known to bind with high affinity across all sarbecovirus clades and prophylactically protects hamsters from viral challenge (89, 90). In the presence of COVOX-

222, we see the most dampening of atomic fluctuations at residues G476 (KL = -5.50, D = 0.888, $p < 0.001$), F486 (KL = -4.52, D = 0.800, $p < 0.001$), and N501 (KL = -4.55, D = 0.708, $p < 0.001$) of the RBD (Figure 6A and 8B). We see modest amount of dampening at residues D405 (KL = -2.57, D = 0.572, $p < 0.001$), K417 (KL = -3.27, D = 0.678, $p < 0.001$), and K458 (KL = -3.49, D = 0.703, $p < 0.001$) of the RBD (Figure 6A and 8B).

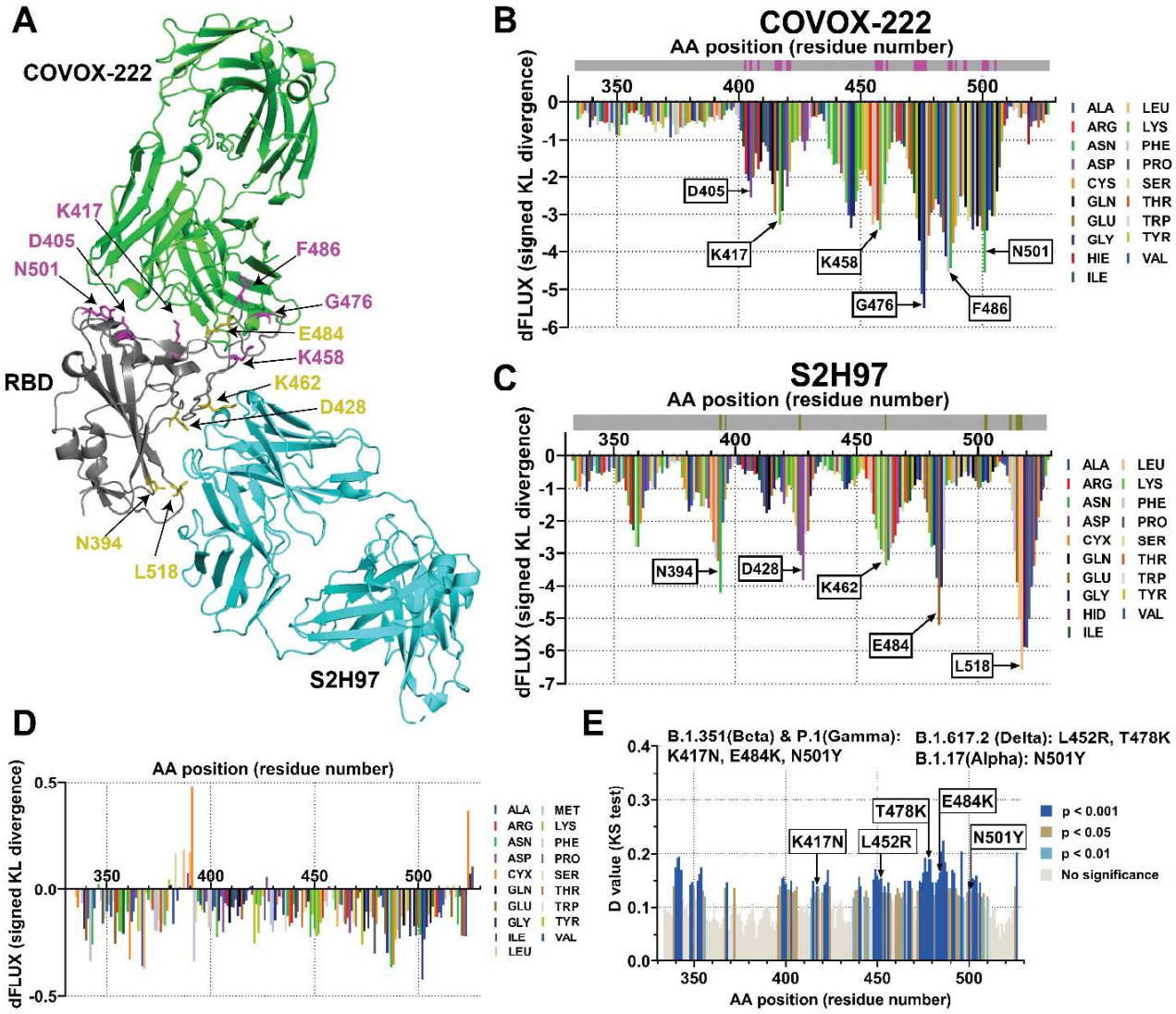


Figure 8: Identification of hot-spot residues that escape neutralizing antibodies and the importance of the variants of concern mutations. (A) Crystal structure of COVOX-222 (green, PDB 7OR9) and S2H97(cyan, PDB 7M7W) superimposed onto the structure of RBD of SARS-CoV-2 (grey, PDB 7M7W) (88, 89). COVOX-222 epitopes with the greatest KL divergence dampening are highlighted in pink, and S2H97 epitopes with the greatest KL divergence dampening are highlighted in olive. Sequence positional plotting of dampening of atom motion on the RBD of SARS-CoV-2 by (B) COVOX-222 and (C) S2H97. Amino acid positions with moderate to modest dampening of atomic fluctuation are identified in (B) and (C). The grey bar on the top of the KL divergence plots denotes the RBD domain amino acid backbone with the location of COVOX-222 epitopes shown in pink and the S2H97 epitopes shown in olive. (D) The site-wise KL divergence profiles showing the dampening of atom motion between the wildtype and mutated RBD of the spike protein of SARS-CoV-2 in the presence of human angiotensin-converting enzyme 2 (ACE2). The wildtype RBD was computationally mutated to include mutations from the variants of concern (VOC). (E) Multiple test-corrected two sample KS tests of significance for the impact of the mutations are also shown. Amino acid positions that correspond to the VOC mutations are highlighted and show higher level significance than other positions. Mutations that correspond to the four VOC in the RBD of spike protein are noted on top of (E).

It should be noted that the residues picked up by our computational method are part of the epitopes of COVOX-222. Of all the residues identified by the comparison of MD simulation, the residues with the greatest dampening of atomic fluctuation have a higher chance of being classified as hot-spot residues. They are more prone to mutate, thus allowing the virus to escape

COVOX-222. Residue 417, one of the residues with moderate dampening of atomic fluctuation, makes a weak salt-bridge interaction with the HCDR3 residue E99 of COVOX-222 (88). And residue N501, one with residues with the most dampening of atomic fluctuations, is known to interact with LCDR1 residue P30 of the antibody via a stacking interactions (88, 91). In the case of the mAb S2H97, we see a moderate dampening of atomic fluctuations at sites D428 (KL = -3.81, D = 0.692, $p < 0.001$) and K462 (KL = -3.37, D = 0.723, $p < 0.001$) . Sites N394 (KL = -4.25, D = 0.760, $p < 0.001$), E484 (KL = -5.20, D = 0.837, $p < 0.001$), L518 (KL = -6.56, D = 0.945, $p < 0.001$) have the most dampening of atomic fluctuations (Figure 6A and 8C). Four of the five sites, except E484, fall in the epitope footprint of S2H97 (89). Like mAb COVOX-222, we predict that the sites with the most dampening of atomic fluctuations (i.e. negative KL divergence) are more prone to functionally evolve under the selection pressure of the vaccine, thus allowing the virus to potentially escape the binding of S2H97. Even though, E484 does not fall in the S2H97 footprint, it has been shown that the mutation of this site is known to enhance immune escape form neutralizing antibodies, and also increase affinity to hACE2 (92).

Many of the residues identified as potential sites of escape, either in the presence of COVOX-222 or S2H97, overlap with mutations seen in the variants of concern (VOC). At present, there are mainly five kinds of VOC: Alpha, Beta, Gamma, Delta, and Omicron. We employed comparative MD simulation of the SARS-CoV-2 RBD bound to hACE2 (PDB 6VW1 trimmed to only include the viral RBD) modeled with and without the recently arising VOC mutations in the alpha to delta variant strains. Three of the VOC mutations (417, 484, and 501) were identified as potential escape sites using our approach (Figure 6A, 8B, and 8C). In addition to the three sites, we confirmed (i.e., 417, 484, and 501), we also saw significant differences in the binding atomic fluctuations at the positions 452 (KL = -0.206, D = 0.155, $p < 0.001$) and 478

(KL = -0.129, D = 0.190, $p < 0.001$) of RBD (Figures 6A, 8D and 8E). Investigation with pseudoviruses possessing RBD mutations harbored by VOC demonstrated that plasma neutralizing activity of vaccinated individuals showed a one-fold to three-fold significant decreases against E484K, N501Y, or the K417N + E484K + N501Y triple mutant (93). These results confirm what we have observed using our comparative MD analysis (Figures 6A and 8E). Furthermore, evidence from clinical trials on the impact of VOC on vaccine efficacy confirms what we have observed as well. For example, the ChAdOx1 nCoV-19 vaccine and the single-dose JNJ-78436735 (Johnson & Johnson/Janssen) vaccine have reduced vaccine efficacy by 10.4% and 57%, respectively, against the B.1.351 variant, which contains K417N, E484K, and N501Y mutations (94-96). Therefore, when evaluating vaccine efficacy and the effect of neutralizing antibodies against the SARS-CoV-2 virus, the focus should be given to amino acid positions that are prone to mutate and escape the antibody, as these positions might cause the emergence of new VOCs.

Lastly, we also modeled the most recent VOC, Omicron. The omicron RBD is known to harbor 15 different mutations, with several mutations linked to greater transmissibility, lower vaccine efficiency, and increased risk of reinfection (Figure 9B) (97). To further understand the evolution of the variant, we ran comparative MD simulations with the Omicron variant bound and unbound to mAb COVOX-222. In the wildtype RBD we see a dampening of atomic fluctuation at residues that correspond to the antibody footprint (Figure 9A, 9B). Interestingly, in the omicron RBD, we see some increase in the atomic fluctuation dampening at several of the residues that correspond to the COVOX-222 epitopes (Figure 9). Furthermore, some of the mutations that correspond to the omicron RBD overlap with the COVOX-222 antibody footprint. These sites include K417N (KL = -2.60, D = 0.631, $p < 0.001$), S477N (KL = -3.83, D = 0.762, p

< 0.001), Q493K (KL = -2.84, D = 0.762, p < 0.001), N501Y (KL = -3.25, D = 0.737, p < 0.001), and Y505H (KL = -3.16, D = 0.77, p < 0.001) (Figure 9, 10). Upon closer examination, the residues that overlap between the antibody footprint and the omicron mutations, are sites where the hACE2 is known to interact with the RBD. As a result, the increase in atomic fluctuation dampening might not be because of the direct evolution of stronger binding to the antibody, but also due to the evolution of increased binding affinity to hACE2. Other studies have shown that the mutations in the omicron variant increase the number of salt bridges and hydrophobic interactions between RBD and hACE2, resulting in a higher binding efficiency to hACE2 (98). Additionally, it has also been shown that the structural changes in the RBD domain, caused by the mutations, reduce the antibody interactions (99, 100). We present a similar analysis of the functional binding of the omicron variant with S2H97 (Supplemental Figure 2) which also indicates minimal overall effects of the genetic changes on potential antibody neutralization, with the exception of E484A, T478K and S477N sites, in which omicron appears to have lost interaction with S2H97. In summary, we observe that introduction of mutations sites in the omicron RBD have only slightly altered the COVOX-222 and S2H97 interactions with the viral RBD. Most importantly, our computational method appears to be an effective way to quantify these changes.

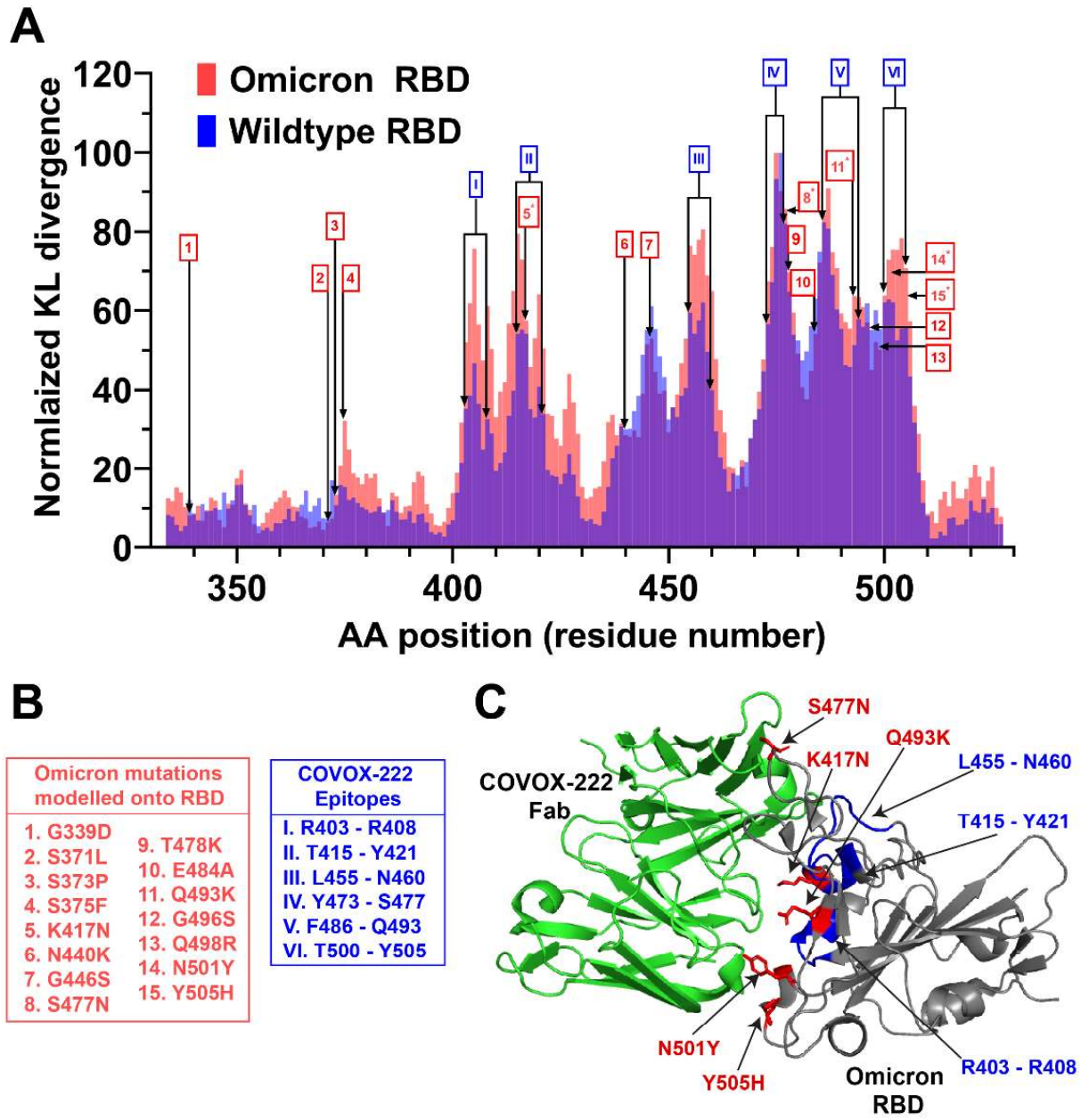
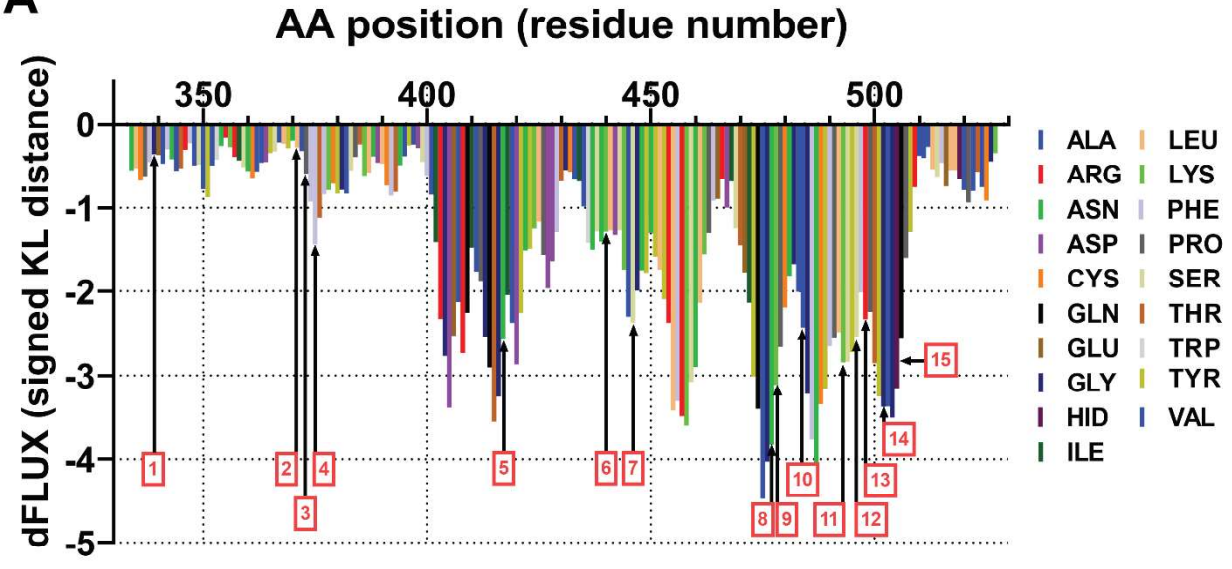


Figure 9: MD simulations with the Omicron variant reveal sites that promote similar binding affinity to hACE2. (A) Sequence positional plotting of the normalized dampening of atom motions on the Omicron RBD (red) and the wildtype RBD (blue) by monoclonal antibody COVOX-222. KL divergence values for normalization of the wild type RBD were obtained from Figure 3B, and for the Omicron RBD were obtained from supplemental figure 1A. The omicron RBD mutations are labeled in red (1-15). The sites corresponding to the COVOX-222 epitope on

the wildtype RBD on labeled in blue (I – VII). The amino acid residues that correspond to the numbers and roman numerals are listed in (B). In the omicron RBD, we see an increase in atomic fluctuation peaks at several of the epitope sites. Sites that overlap between the omicron mutations and COVOX-222 epitope (shown by *) are sites of ACE2 interaction. (B) List of omicron mutations modeled onto RBD, and the list of COVOX-222 epitopes. (C) Crystal structure of COVOX-222 Fab (green) superimposed onto the structure of the Omicron RBD (grey) (PDB 7NXA). Some of the amino acid residues that correspond to the COVOX-222 epitopes, which show an increase in atomic fluctuation dampening, are shown in blue. Several of the amino acid residue that correspond to omicron mutations with increased atomic fluctuation dampening, which are sites of ACE2 interactions, are shown in red.

A



B

PDB ID	Antigen	Antibody	AA and position	-KL divergence	D value	p value
7NXA	Omicron RBD	COVOX-222	K417N (5)	-2.60	0.631	p < 0.001
			S477N (8)	-3.83	0.762	p < 0.001
			Q493K (11)	-2.84	0.762	p < 0.001
			N501Y (14)	-3.25	0.737	p < 0.001
			Y505H (15)	-3.16	0.771	p < 0.001

Figure 10: Molecular dynamic simulation of the omicron variant. (A) Sequence positional plotting of dampening of atom motion on omicron RBD domain by COVOX-222. Mutation sites correspond to the Omicron RBD are labeled in red (1-15) and are listed in Figure 4B. Table summarizing the protein structure used for primary models for analyzing the molecular dynamics of COVOX-222 interaction with Omicron RBD. The table also includes amino acid positions, the corresponding -KL divergence value denoting atomic fluctuations dampening, D value and the corresponding level of significance for the KL divergence values.

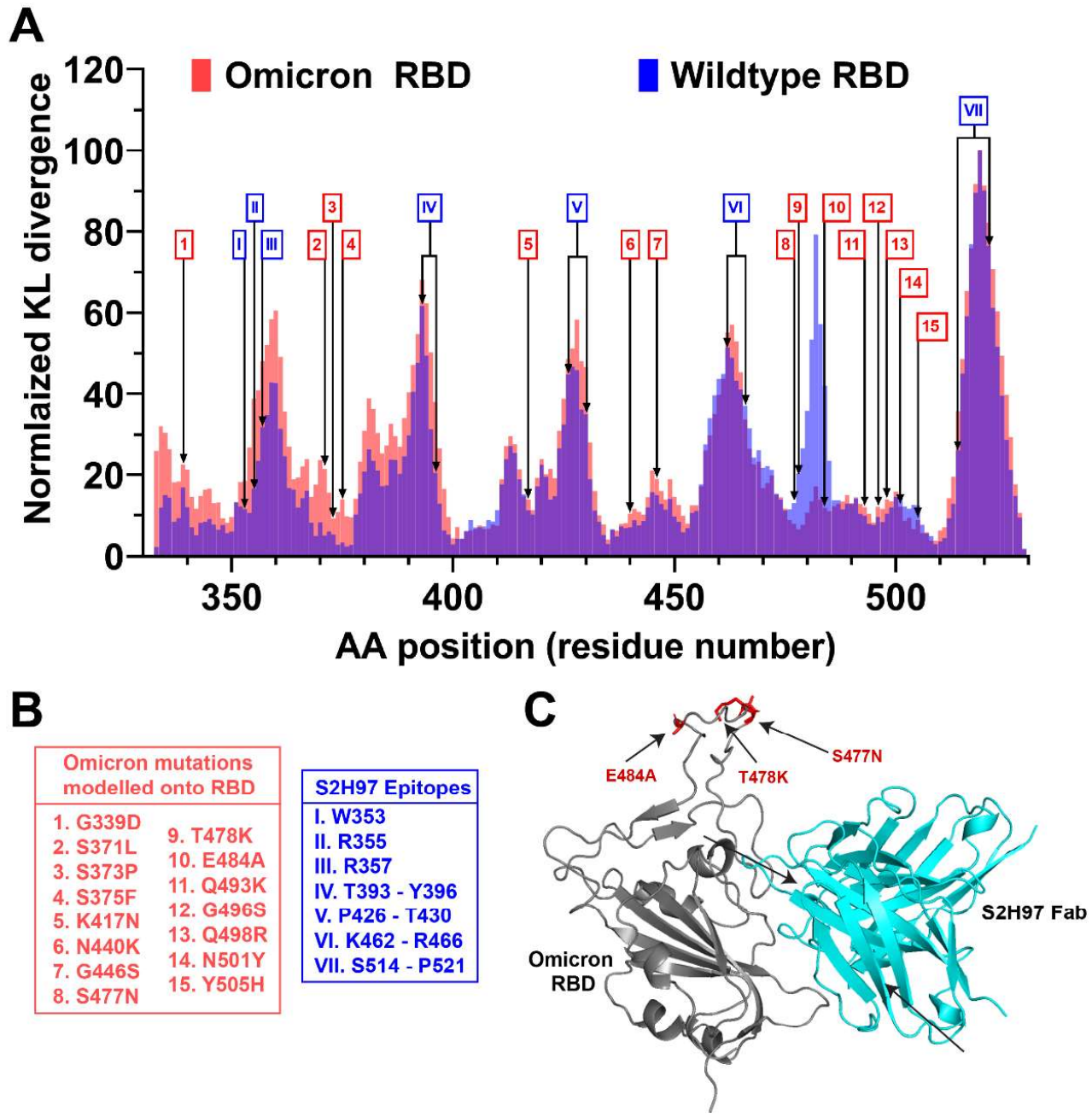


Figure 11: MD simulations with the Omicron variant reveal sites that promote similiar binding affinity to hACE2. (A) Sequence positional plotting of the normalized dampening of atom motions on the Omicron RBD (red) and the wildtype RBD (blue) by monoclonal antibody S2H97.. The omicron RBD mutations are labeled in red (1-15). The sites corresponding to the S2H97 epitope on the wildtype RBD on labeled in blue (I – VII). The amino acid residues that

correspond to the numbers and roman numerals are listed in (B). In the omicron RBD, we see alteration in atomic fluctuation peaks at several of the epitope sites.. (B) List of omicron mutations modeled onto RBD, and the list of S2H97 epitopes. (C) Crystal structure of S2H97 Fab (cyan) superimposed onto the structure of the Omicron RBD (grey) (PDB 7M7W). Some of the amino acid residues that correspond to the S2H97 epitopes, show an increase in atomic fluctuation dampening, are shown in blue. Several of the amino acid residue that correspond to omicron mutations with increased atomic fluctuation dampening, which are sites of ACE2 interactions, are shown in red. NOTE: Major changes occurring E4848A, T478K, S477N.

Directed selection experiments followed by confirmation through neutralization assays are the classical approach for identifying the location of the hot-spot residues. Not only are the selection experiments time consuming, they only identify one of many mutations that escape an antibody (101). Compared to the classical method, our simple computational approach can identify the locations of sites prone to vaccine escape in a matter of days. Additionally it identifies positions of all amino acids that could mutate to escape the antibody in a single in silico experiment. And compared to other computational tools, our MD-based approach does not require large and potentially biased machine learning training data sets. Furthermore, the de novo prediction is based on a given experimental structure, which enables an unprecedented close synergy between our computational approach and existing laboratory methods for identifying potential routes of viral evolution leading to enhanced transmissibility and vaccine escape. Additionally, some of the other tools used in epitope prediction involve sequence-based methods, machine-learning methods, and structure-based methods. The epitope surface accessibility to antibody binding is generally used by the sequence-based methods (102). The availability of

antigen sequence is crucial for the sequence-based method; however, the predicted epitope residues are not grouped into the corresponding epitopes (103). The machine learning (ML) based epitope prediction methods include several steps: a collection of datasets with clean and comprehensive data, extraction of antigen features of the sequence (e.g., physicochemical properties, evolutionary information, amino acid composition), and training the model using ML algorithms (104). Some of the commonly used ML tools for epitope prediction are ABCPred (uses artificial neural network), COBEpro (uses support vector machine), EPSVR (uses support vector regression method), and BepiPred (based on random forest algorithm) (102, 105-107). And lastly, structure-based methods identify epitopes by antigen structure and epitope-related propensity scales, including specific physicochemical properties and geometric attributes (108, 109). As the evolution of antibody-antigen interactions seem largely driven by single-site mutations with large functional effects and because antibody-antigen binding has quite obvious biophysical impacts on molecular dynamics, as would the formation of any protein complex, we offer a simple and effective alternative computational method to studying vaccine escape in silico using readily available MD simulation software combined with a classic and easily interpretable approach to the statistical comparison of distributions of local atom fluctuations on virus proteins in the presence/absence of antibodies.

In summary, we have computationally identified hot-spot residues known to mutate in a viral antigen in the presence of a neutralizing antibody. We first validated our approach using the influenza spike protein and three well-characterized antibodies against the influenza HA (73). We then implemented our approach to identify sites in the RBD of SARS-CoV-2 that are known to mutate in the presence of neutralizing mAbs S2H97 and COVOX-222 (88, 89). We further identified that residues known to mutate in the presence of the antibody overlap with the

mutations seen in the VOC. And lastly, we also identified sites in the omicron variant mutations that enhance binding efficiency to hACE2. While the method described here is not a complete substitute for laboratory-based methods we believe it can greatly complement these methods by allowing time and cost-saving through the computational prescreening of the underlying biophysics that may drive outcomes of many potential lab experiments. Determining which viral mutations escape from antibodies will be crucial for designing future therapeutics and vaccines and assessing future antigenic implications of ongoing viral evolution.

References

1. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T, Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B. 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 395:497-506.
2. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KSM, Lau EHY, Wong JY, Xing X, Xiang N, Wu Y, Li C, Chen Q, Li D, Liu T, Zhao J, Liu M, Tu W, Chen C, Jin L, Yang R, Wang Q, Zhou S, Wang R, Liu H, Luo Y, Liu Y, Shao G, Li H, Tao Z, Yang Y, Deng Z, Liu B, Ma Z, Zhang Y, Shi G, Lam TTY, Wu JT, Gao GF, Cowling BJ, Yang B, Leung GM, Feng Z. 2020. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. N Engl J Med 382:1199-1207.
3. WHO. 2021. WHO Coronavirus (COVID-19) Dashboard. https://covid19.who.int/. Accessed
4. Esakandari H, Nabi-Afjadi M, Fakkari-Afjadi J, Farahmandian N, Miresmaeili SM, Bahreini E. 2020. A comprehensive review of COVID-19 characteristics. Biol Proced Online 22:19.
5. Yang J, Wang W, Chen Z, Lu S, Yang F, Bi Z, Bao L, Mo F, Li X, Huang Y, Hong W, Yang Y, Zhao Y, Ye F, Lin S, Deng W, Chen H, Lei H, Zhang Z, Luo M, Gao H, Zheng Y, Gong Y, Jiang X, Xu Y, Lv Q, Li D, Wang M, Li F, Wang S, Wang G, Yu P, Qu Y, Yang L, Deng H, Tong A, Li J, Wang Z, Shen G, Zhao Z, Li Y, Luo J, Liu H, Yu W, Yang M, Xu J, Wang J, Li H, Wang H, Kuang D, et al. 2020. A vaccine targeting the RBD of the S protein of SARS-CoV-2 induces protective immunity. Nature 586:572-577.
6. Salvatori G, Luberto L, Maffei M, Aurisicchio L, Roscilli G, Palombo F, Marra E. 2020. SARS-CoV-2 SPIKE PROTEIN: an optimal immunological target for vaccines. J Transl Med 18:222.
7. Saghazadeh A, Rezaei N. 2020. Towards treatment planning of COVID-19: Rationale and hypothesis for the use of multiple immunosuppressive agents: Anti-antibodies, immunoglobulins, and corticosteroids. Int Immunopharmacol 84:106560.

8. Gur M, Taka E, Yilmaz SZ, Kilinc C, Aktas U, Golcuk M. 2020. Conformational transition of SARS-CoV-2 spike glycoprotein between its closed and open states. J Chem Phys 153:075101.
9. Jia HP, Look DC, Shi L, Hickey M, Pewe L, Netland J, Farzan M, Wohlford-Lenane C, Perlman S, McCray PB. 2005. ACE2 receptor expression and severe acute respiratory syndrome coronavirus infection depend on differentiation of human airway epithelia. J Virol 79:14614-21.
10. Donoghue M, Hsieh F, Baronas E, Godbout K, Gosselin M, Stagliano N, Donovan M, Woolf B, Robison K, Jeyaseelan R, Breitbart RE, Acton S. 2000. A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. Circ Res 87:E1-9.
11. Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. 2020. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science 367:1444-1448.
12. Lake MA. 2020. What we know so far: COVID-19 current clinical knowledge and research. Clin Med (Lond) 20:124-127.
13. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL, Chen HD, Chen J, Luo Y, Guo H, Jiang RD, Liu MQ, Chen Y, Shen XR, Wang X, Zheng XS, Zhao K, Chen QJ, Deng F, Liu LL, Yan B, Zhan FX, Wang YY, Xiao GF, Shi ZL. 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579:270-273.
14. Zhang T, Wu Q, Zhang Z. 2020. Probable Pangolin Origin of SARS-CoV-2 Associated with the COVID-19 Outbreak. Curr Biol 30:1346-1351.e2.
15. Xiao K, Zhai J, Feng Y, Zhou N, Zhang X, Zou JJ, Li N, Guo Y, Li X, Shen X, Zhang Z, Shu F, Huang W, Li Y, Chen RA, Wu YJ, Peng SM, Huang M, Xie WJ, Cai QH, Hou FH, Chen W, Xiao L, Shen Y. 2020. Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. Nature 583:286-289.
16. Tahir M. 2021. Coronavirus genomic nsp14-ExoN, structure, role, mechanism, and potential application as a drug target. J Med Virol 93:4258-4264.
17. Vasilarou M, Alachiotis N, Garefalaki J, Beloukas A, Pavlidis P. 2021. Population Genomics Insights into the First Wave of COVID-19. Life (Basel) 11.
18. Biswas SK, Mudi SR. 2020. Genetic variation in SARS-CoV-2 may explain variable severity of COVID-19. Med Hypotheses 143:109877.
19. Choi JY, Smith DM. 2021. SARS-CoV-2 Variants of Concern. Yonsei Med J 62:961-968.
20. WHO. 2022. Tracking SARS-CoV-2 variants. https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/. Accessed January 13.
21. Chou CF, Loh CB, Foo YK, Shen S, Fielding BC, Tan TH, Khan S, Wang Y, Lim SG, Hong W, Tan YJ, Fu J. 2006. ACE2 orthologues in non-mammalian vertebrates (Danio, Gallus, Fugu, Tetraodon and Xenopus). Gene 377:46-55.
22. Leroy EM, Ar Gouilh M, Brugère-Picoux J. 2020. The risk of SARS-CoV-2 transmission to pets and other wild and domestic animals strongly mandates a one-health strategy to control the COVID-19 pandemic. One Health 10:100133.
23. McAloose D, Laverack M, Wang L, Killian ML, Caserta LC, Yuan F, Mitchell PK, Queen K, Mauldin MR, Cronk BD, Bartlett SL, Sykes JM, Zec S, Stokol T, Ingerman K, Delaney MA, Fredrickson R, Ivančić M, Jenkins-Moore M, Mozingo K, Franzen K, Bergeson NH, Goodman L, Wang H, Fang Y, Olmstead C, McCann C, Thomas P,

Goodrich E, Elvinger F, Smith DC, Tong S, Slavinski S, Calle PP, Terio K, Torchetti MK, Diel DG. 2020. From People to. mBio 11.
24. Oude Munnink BB, Sikkema RS, Nieuwenhuijse DF, Molenaar RJ, Munger E, Molenkamp R, van der Spek A, Tolsma P, Rietveld A, Brouwer M, Bouwmeester-Vincken N, Harders F, Hakze-van der Honing R, Wegdam-Blans MCA, Bouwstra RJ, GeurtsvanKessel C, van der Eijk AA, Velkers FC, Smit LAM, Stegeman A, van der Poel WHM, Koopmans MPG. 2021. Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. Science 371:172-177.
25. Oreshkova N, Molenaar RJ, Vreman S, Harders F, Oude Munnink BB, Hakze-van der Honing RW, Gerhards N, Tolsma P, Bouwstra R, Sikkema RS, Tacken MG, de Rooij MM, Weesendorp E, Engelsma MY, Bruschke CJ, Smit LA, Koopmans M, van der Poel WH, Stegeman A. 2020. SARS-CoV-2 infection in farmed minks, the Netherlands, April and May 2020. Euro Surveill 25.
26. Schlottau K, Rissmann M, Graaf A, Schön J, Sehl J, Wylezich C, Höper D, Mettenleiter TC, Balkema-Buschmann A, Harder T, Grund C, Hoffmann D, Breithaupt A, Beer M. 2020. SARS-CoV-2 in fruit bats, ferrets, pigs, and chickens: an experimental transmission study. Lancet Microbe 1:e218-e225.
27. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25:1605-12.
28. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ. 2005. The Amber biomolecular simulation programs. J Comput Chem 26:1668-88.
29. Fiser A, Do RK, Sali A. 2000. Modeling of loops in protein structures. Protein Sci 9:1753-73.
30. Sali A, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779-815.
31. Woods-Group. 2005-2021. GLYCAM Web, Complex Carbohydrate Research Center, University of Georgia, Athens, GA. http://legacy.glycam.org.
32. Kirschner KN, Yongye AB, Tschampel SM, González-Outeiriño J, Daniels CR, Foley BL, Woods RJ. 2008. GLYCAM06: a generalizable biomolecular force field. Carbohydrates. J Comput Chem 29:622-55.
33. Babbitt GA, Fokoue EP, Evans JR, Diller KI, Adams LE. 2020. DROIDS 3.0-Detecting Genetic and Drug Class Variant Impact on Conserved Protein Binding Dynamics. Biophys J 118:541-551.
34. Babbitt GA, Lynch ML, McCoy M, Fokoue EP, Hudson AO. 2020. Function and evolution of B-Raf loop dynamics relevant to cancer recurrence under drug inhibition. J Biomol Struct Dyn:1-16.
35. Babbitt GA, Mortensen JS, Coppola EE, Adams LE, Liao JK. 2018. DROIDS 1.20: A GUI-Based Pipeline for GPU-Accelerated Comparative Protein Dynamics. Biophys J 114:1009-1017.
36. Ewald PP. 1921. Die Berechnung optischer und elektrostatischer Gitterpotentiale. Annalen der Physik 369:253-287.
37. Pierce LC, Salomon-Ferrer R, Augusto F de Oliveira C, McCammon JA, Walker RC. 2012. Routine Access to Millisecond Time Scale Events with Accelerated Molecular Dynamics. J Chem Theory Comput 8:2997-3002.

38. Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, Walker RC. 2013. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. J Chem Theory Comput 9:3878-88.
39. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. 2015. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J Chem Theory Comput 11:3696-713.
40. Jorgensen WL, Chandrasekar J, Madura JD. 1983. Comparison of simple potential functions for simulating liquid water. Journal of Chemical Physics 79.
41. Andersen HC. 1980. Molecular dynamics simulations at constant pressure and/or temperature. The Journal of Chemical Physics. 72(4):2384-2393.
42. Roe DR, Cheatham TE. 2013. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. J Chem Theory Comput 9:3084-95.
43. Rehman Z, Fahim A, Bhatti MF. 2021. Scouting the receptor-binding domain of SARS coronavirus 2: a comprehensive immunoinformatics inquisition. Future virology 16.
44. Liu K, Pan X, Li L, Yu F, Zheng A, Du P, Han P, Meng Y, Zhang Y, Wu L, Chen Q, Song C, Jia Y, Niu S, Lu D, Qiao C, Chen Z, Ma D, Ma X, Tan S, Zhao X, Qi J, Gao GF, Wang Q. 2021. Binding and molecular basis of the bat coronavirus RaTG13 virus to ACE2 in humans and other species. Cell 184:3438-3451.e10.
45. Heymann DL, Shindo N, Hazards WSaTAGfI. 2020. COVID-19: what is next for public health? Lancet 395:542-545.
46. Wan Y, Shang J, Graham R, Baric RS, Li F. 2020. Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. J Virol 94.
47. Rothan HA, Byrareddy SN. 2020. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. J Autoimmun 109:102433.
48. Munster VJ, Feldmann F, Williamson BN, van Doremalen N, Pérez-Pérez L, Schulz J, Meade-White K, Okumura A, Callison J, Brumbaugh B, Avanzato VA, Rosenke R, Hanley PW, Saturday G, Scott D, Fischer ER, de Wit E. 2020. Respiratory disease in rhesus macaques inoculated with SARS-CoV-2. Nature 585:268-272.
49. Shi J, Wen Z, Zhong G, Yang H, Wang C, Huang B, Liu R, He X, Shuai L, Sun Z, Zhao Y, Liu P, Liang L, Cui P, Wang J, Zhang X, Guan Y, Tan W, Wu G, Chen H, Bu Z. 2020. Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS-coronavirus 2. Science 368:1016-1020.
50. Wrobel AG, Benton DJ, Xu P, Roustan C, Martin SR, Rosenthal PB, Skehel JJ, Gamblin SJ. 2020. SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects. Nat Struct Mol Biol 27:763-767.
51. Neri A, Aygen M, Zukerman Z, Bahary C. 1980. Subjective assessment of sexual dysfunction of patients on long-term administration of digoxin. Arch Sex Behav 9:343-7.
52. Seligson U, Söderlund C. 1986. ERCP and serum alkaline phosphatase in pancreatic carcinoma. Acta Chir Scand 152:309-12.
53. Shivaprakash KN, Sen S, Paul S, Kiesecker JM, Bawa KS. 2021. Mammals, wildlife trade, and the next global pandemic. Curr Biol 31:3671-3677.e3.
54. de Morais HA, Dos Santos AP, do Nascimento NC, Kmetiuk LB, Barbosa DS, Brandão PE, Guimarães AMS, Pettan-Brewer C, Biondo AW. 2020. Natural Infection by SARS-CoV-2 in Companion Animals: A Review of Case Reports and Current Evidence of Their Role in the Epidemiology of COVID-19. Front Vet Sci 7:591216.

55. Michelitsch A, Hoffmann D, Wernike K, Beer M. 2020. Occurrence of Antibodies against SARS-CoV-2 in the Domestic Cat Population of Germany. Vaccines (Basel) 8.
56. Barrs VR, Peiris M, Tam KWS, Law PYT, Brackman CJ, To EMW, Yu VYT, Chu DKW, Perera RAPM, Sit THC. 2020. SARS-CoV-2 in Quarantined Domestic Cats from COVID-19 Households or Close Contacts, Hong Kong, China. Emerg Infect Dis 26:3071-3074.
57. Patterson EI, Elia G, Grassi A, Giordano A, Desario C, Medardo M, Smith SL, Anderson ER, Prince T, Patterson GT, Lorusso E, Lucente MS, Lanave G, Lauzi S, Bonfanti U, Stranieri A, Martella V, Solari Basano F, Barrs VR, Radford AD, Agrimi U, Hughes GL, Paltrinieri S, Decaro N. 2020. Evidence of exposure to SARS-CoV-2 in cats and dogs from households in Italy. Nat Commun 11:6231.
58. Fritz M, Rosolen B, Krafft E, Becquart P, Elguero E, Vratskikh O, Denolly S, Boson B, Vanhomwegen J, Gouilh MA, Kodjo A, Chirouze C, Rosolen SG, Legros V, Leroy EM. 2021. High prevalence of SARS-CoV-2 antibodies in pets from COVID-19+ households. One Health 11:100192.
59. Domańska-Blicharz K, Orłowska A, Smreczak M, Niemczuk K, Iwan E, Bomba A, Lisowska A, Opolska J, Trębas P, Potyrało P, Kawiak-Sadurska M, Rola J. 2021. Mink SARS-CoV-2 Infection in Poland - Short Communication. J Vet Res 65:1-5.
60. Fenollar F, Mediannikov O, Maurin M, Devaux C, Colson P, Levasseur A, Fournier PE, Raoult D. 2021. Mink, SARS-CoV-2, and the Human-Animal Interface. Front Microbiol 12:663815.
61. Smith DJ, Lapedes AS, de Jong JC, Bestebroer TM, Rimmelzwaan GF, Osterhaus AD, Fouchier RA. 2004. Mapping the antigenic and genetic evolution of influenza virus. Science 305:371-6.
62. Luksza M, Lässig M. 2014. A predictive fitness model for influenza. Nature 507:57-61.
63. Krammer F, Palese P. 2015. Advances in the development of influenza virus vaccines. Nat Rev Drug Discov 14:167-82.
64. Gerhard W, Yewdell J, Frankel ME, Webster R. 1981. Antigenic structure of influenza virus haemagglutinin defined by hybridoma antibodies. Nature 290:713-7.
65. Caton AJ, Brownlee GG, Yewdell JW, Gerhard W. 1982. The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). Cell 31:417-27.
66. Webster RG, Laver WG. 1980. Determination of the number of nonoverlapping antigenic areas on Hong Kong (H3N2) influenza virus hemagglutinin with monoclonal antibodies and the selection of variants with potential epidemiological significance. Virology 104:139-48.
67. Haddox HK, Dingens AS, Bloom JD. 2016. Experimental Estimation of the Effects of All Amino-Acid Mutations to HIV's Envelope Protein on Viral Replication in Cell Culture. PLoS Pathog 12:e1006114.
68. Pierce BG, Keck ZY, Lau P, Fauvelle C, Gowthaman R, Baumert TF, Fuerst TR, Mariuzza RA, Foung SKH. 2016. Global mapping of antibody recognition of the hepatitis C virus E2 glycoprotein: Implications for vaccine design. Proc Natl Acad Sci U S A 113:E6946-E6954.
69. Walker LM, Phogat SK, Chan-Hui PY, Wagner D, Phung P, Goss JL, Wrin T, Simek MD, Fling S, Mitcham JL, Lehrman JK, Priddy FH, Olsen OA, Frey SM, Hammond PW, Kaminsky S, Zamb T, Moyle M, Koff WC, Poignard P, Burton DR, Investigators PGP.

2009. Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. Science 326:285-9.
70. Polonis VR, Brown BK, Rosa Borges A, Zolla-Pazner S, Dimitrov DS, Zhang MY, Barnett SW, Ruprecht RM, Scarlatti G, Fenyö EM, Montefiori DC, McCutchan FE, Michael NL. 2008. Recent advances in the characterization of HIV-1 neutralization assays for standardized evaluation of the antibody response to infection and vaccination. Virology 375:315-20.
71. Jin L, Fendly BM, Wells JA. 1992. High resolution functional analysis of antibody-antigen interactions. J Mol Biol 226:851-65.
72. Dall'Acqua W, Goldman ER, Lin W, Teng C, Tsuchiya D, Li H, Ysern X, Braden BC, Li Y, Smith-Gill SJ, Mariuzza RA. 1998. A mutational analysis of binding interactions in an antigen-antibody protein-protein complex. Biochemistry 37:7981-91.
73. Doud MB, Lee JM, Bloom JD. 2018. How single mutations affect viral escape from broad and narrow antibodies to H1 influenza hemagglutinin. Nat Commun 9:1386.
74. Rynkiewicz P, Lynch ML, Cui F, Hudson AO, Babbitt GA. 2021. Functional binding dynamics relevant to the evolution of zoonotic spillovers in endemic and emergent. J Biomol Struct Dyn:1-19.
75. Aartse A, Eggink D, Claireaux M, van Leeuwen S, Mooij P, Bogers WM, Sanders RW, Koopman G, van Gils MJ. 2021. Influenza A Virus Hemagglutinin Trimer, Head and Stem Proteins Identify and Quantify Different Hemagglutinin-Specific B Cell Subsets in Humans. Vaccines (Basel) 9.
76. Darden T, York D, Pedersen L. 1993. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. 98(10089-10092.
77. Götz AW, Williamson MJ, Xu D, Poole D, Le Grand S, Walker RC. 2012. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. J Chem Theory Comput 8:1542-1555.
78. Wang J, Wang W, Kollman PA, Case DA. 2006. Automatic atom type and bond type perception in molecular mechanical calculations. J Mol Graph Model 25:247-60.
79. Yoshida R, Igarashi M, Ozaki H, Kishida N, Tomabechi D, Kida H, Ito K, Takada A. 2009. Cross-protective potential of a novel monoclonal antibody directed against antigenic site B of the hemagglutinin of influenza A viruses. PLoS Pathog 5:e1000350.
80. Lee PS, Yoshida R, Ekiert DC, Sakai N, Suzuki Y, Takada A, Wilson IA. 2012. Heterosubtypic antibody recognition of the influenza virus hemagglutinin receptor binding site enhanced by avidity. Proc Natl Acad Sci U S A 109:17040-5.
81. Eisen MB, Sabesan S, Skehel JJ, Wiley DC. 1997. Binding of the influenza A virus to cell-surface receptors: structures of five hemagglutinin-sialyloligosaccharide complexes determined by X-ray crystallography. Virology 232:19-31.
82. Fleury D, Barrère B, Bizebard T, Daniels RS, Skehel JJ, Knossow M. 1999. A complex of influenza hemagglutinin with a neutralizing antibody that binds outside the virus receptor binding site. Nat Struct Biol 6:530-4.
83. Matrosovich M, Tuzikov A, Bovin N, Gambaryan A, Klimov A, Castrucci MR, Donatelli I, Kawaoka Y. 2000. Early alterations of the receptor-binding properties of H1, H2, and H3 avian influenza virus hemagglutinins after their introduction into mammals. J Virol 74:8502-12.
84. Corti D, Voss J, Gamblin SJ, Codoni G, Macagno A, Jarrossay D, Vachieri SG, Pinna D, Minola A, Vanzetta F, Silacci C, Fernandez-Rodriguez BM, Agatic G, Bianchi S,

Giacchetto-Sasselli I, Calder L, Sallusto F, Collins P, Haire LF, Temperton N, Langedijk JP, Skehel JJ, Lanzavecchia A. 2011. A neutralizing antibody selected from plasma cells that binds to group 1 and group 2 influenza A hemagglutinins. Science 333:850-6.

85. Smirnov YA, Lipatov AS, Gitelman AK, Okuno Y, Van Beek R, Osterhaus AD, Claas EC. 1999. An epitope shared by the hemagglutinins of H1, H2, H5, and H6 subtypes of influenza A virus. Acta Virol 43:237-44.
86. Roubidoux EK, Carreño JM, McMahon M, Jiang K, van Bakel H, Wilson P, Krammer F. 2021. Mutations in the Hemagglutinin Stalk Domain Do Not Permit Escape from a Protective, Stalk-Based Vaccine-Induced Immune Response in the Mouse Model. mBio 12.
87. Heaton NS, Sachs D, Chen CJ, Hai R, Palese P. 2013. Genome-wide mutagenesis of influenza virus reveals unique plasticity of the hemagglutinin and NS1 proteins. Proc Natl Acad Sci U S A 110:20248-53.
88. Dejnirattisai W, Zhou D, Supasa P, Liu C, Mentzer AJ, Ginn HM, Zhao Y, Duyvesteyn HME, Tuekprakhon A, Nutalai R, Wang B, López-Camacho C, Slon-Campos J, Walter TS, Skelly D, Costa Clemens SA, Naveca FG, Nascimento V, Nascimento F, Fernandes da Costa C, Resende PC, Pauvolid-Correa A, Siqueira MM, Dold C, Levin R, Dong T, Pollard AJ, Knight JC, Crook D, Lambe T, Clutterbuck E, Bibi S, Flaxman A, Bittaye M, Belij-Rammerstorfer S, Gilbert SC, Carroll MW, Klenerman P, Barnes E, Dunachie SJ, Paterson NG, Williams MA, Hall DR, Hulswit RJG, Bowden TA, Fry EE, Mongkolsapaya J, Ren J, Stuart DI, Screaton GR. 2021. Antibody evasion by the P.1 strain of SARS-CoV-2. Cell 184:2939-2954.e9.
89. Starr TN, Czudnochowski N, Liu Z, Zatta F, Park YJ, Addetia A, Pinto D, Beltramello M, Hernandez P, Greaney AJ, Marzi R, Glass WG, Zhang I, Dingens AS, Bowen JE, Tortorici MA, Walls AC, Wojcechowskyj JA, De Marco A, Rosen LE, Zhou J, Montiel-Ruiz M, Kaiser H, Dillen JR, Tucker H, Bassi J, Silacci-Fregni C, Housley MP, di Iulio J, Lombardo G, Agostini M, Sprugasci N, Culap K, Jaconi S, Meury M, Dellota E, Abdelnabi R, Foo SC, Cameroni E, Stumpf S, Croll TI, Nix JC, Havenar-Daughton C, Piccoli L, Benigni F, Neyts J, Telenti A, Lempp FA, Pizzuto MS, Chodera JD, et al. 2021. SARS-CoV-2 RBD antibodies that maximize breadth and resistance to escape. Nature 597:97-102.
90. Kwon D. 2021. This 'super antibody' for COVID fights off multiple coronaviruses, *on* Nature. **https://www.nature.com/articles/d41586-021-01917-9#correction-0**. Accessed November 19.
91. Shrestha LB, Tedla N, Bull RA. 2021. Broadly-Neutralizing Antibodies Against Emerging SARS-CoV-2 Variants. Front Immunol 12:752003.
92. Barton MI, MacGowan SA, Kutuzov MA, Dushek O, Barton GJ, van der Merwe PA. 2021. Effects of common mutations in the SARS-CoV-2 Spike RBD and its ligand, the human ACE2 receptor on binding affinity and kinetics. Elife 10.
93. Wang Z, Schmidt F, Weisblum Y, Muecksch F, Barnes CO, Finkin S, Schaefer-Babajew D, Cipolla M, Gaebler C, Lieberman JA, Oliveira TY, Yang Z, Abernathy ME, Huey-Tubman KE, Hurley A, Turroja M, West KA, Gordon K, Millard KG, Ramos V, Da Silva J, Xu J, Colbert RA, Patel R, Dizon J, Unson-O'Brien C, Shimeliovich I, Gazumyan A, Caskey M, Bjorkman PJ, Casellas R, Hatziioannou T, Bieniasz PD, Nussenzweig MC. 2021. mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. Nature 592:616-622.

94. Mahase E. 2021. Covid-19: Where are we on vaccines and variants? BMJ 372:n597.
95. Emary KRW, Golubchik T, Aley PK, Ariani CV, Angus B, Bibi S, Blane B, Bonsall D, Cicconi P, Charlton S, Clutterbuck EA, Collins AM, Cox T, Darton TC, Dold C, Douglas AD, Duncan CJA, Ewer KJ, Flaxman AL, Faust SN, Ferreira DM, Feng S, Finn A, Folegatti PM, Fuskova M, Galiza E, Goodman AL, Green CM, Green CA, Greenland M, Hallis B, Heath PT, Hay J, Hill HC, Jenkin D, Kerridge S, Lazarus R, Libri V, Lillie PJ, Ludden C, Marchevsky NG, Minassian AM, McGregor AC, Mujadidi YF, Phillips DJ, Plested E, Pollock KM, Robinson H, Smith A, Song R, et al. 2021. Efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine against SARS-CoV-2 variant of concern 202012/01 (B.1.1.7): an exploratory analysis of a randomised controlled trial. Lancet 397:1351-1362.
96. Madhi SA, Baillie V, Cutland CL, Voysey M, Koen AL, Fairlie L, Padayachee SD, Dheda K, Barnabas SL, Bhorat QE, Briner C, Kwatra G, Ahmed K, Aley P, Bhikha S, Bhiman JN, Bhorat AE, du Plessis J, Esmail A, Groenewald M, Horne E, Hwa SH, Jose A, Lambe T, Laubscher M, Malahleha M, Masenya M, Masilela M, McKenzie S, Molapo K, Moultrie A, Oelofse S, Patel F, Pillay S, Rhead S, Rodel H, Rossouw L, Taoushanis C, Tegally H, Thombrayil A, van Eck S, Wibmer CK, Durham NM, Kelly EJ, Villafana TL, Gilbert S, Pollard AJ, de Oliveira T, Moore PL, Sigal A, et al. 2021. Efficacy of the ChAdOx1 nCoV-19 Covid-19 Vaccine against the B.1.351 Variant. N Engl J Med 384:1885-1898.
97. Torjesen I. 2021. Covid-19: Omicron may be more transmissible than other variants and partly resistant to existing vaccines, scientists fear. BMJ 375:n2943.
98. Ford CT, Machado DJ, Janies DA. 2021. Predictions of the SARS-CoV-2 Omicron Variant (B.1.1.529) Spike Protein Receptor-Binding Domain Structure and Neutralizing Antibody Interactions. bioRxiv.
99. Kumar S, Thambiraja TS, Karuppanan K, Subramaniam G. 2021. Omicron and Delta Variant of SARS-CoV-2: A Comparative Computational Study of Spike protein. doi:https://doi.org/10.1101/2021.12.02.470946.
100. Golcuk M, Yildiz A, Gur M. 2021. The Omicron Variant Increases the Interactions of SARS-CoV-2 Spike Glycoprotein with ACE2. bioRxiv.
101. Leon PE, Wohlbold TJ, He W, Bailey MJ, Henry CJ, Wilson PC, Krammer F, Tan GS. 2017. Generation of Escape Variants of Neutralizing Influenza Virus Monoclonal Antibodies. J Vis Exp.
102. Saha S, Raghava GP. 2006. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins 65:40-8.
103. Yao B, Zheng D, Liang S, Zhang C. 2013. Conformational B-cell epitope prediction on antigen protein structures: a review of current algorithms and comparison with common binding site prediction methods. PLoS One 8:e62249.
104. Sun P, Ju H, Liu Z, Ning Q, Zhang J, Zhao X, Huang Y, Ma Z, Li Y. 2013. Bioinformatics resources and tools for conformational B-cell epitope prediction. Comput Math Methods Med 2013:943636.
105. Sweredoski MJ, Baldi P. 2009. COBEpro: a novel system for predicting continuous B-cell epitopes. Protein Eng Des Sel 22:113-20.
106. Liang S, Zheng D, Standley DM, Yao B, Zacharias M, Zhang C. 2010. EPSVR and EPMeta: prediction of antigenic epitopes using support vector regression and multiple server results. BMC Bioinformatics 11:381.

107. Jespersen MC, Peters B, Nielsen M, Marcatili P. 2017. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res 45:W24-W29.
108. Saha S, Raghava GPS. 2004. BcePred: Prediction of Continuous B-Cell Epitopes in Antigenic Sequences Using Physico-chemical Properties, p 197-204. *In* Nicosia G, Cutello V, Bentley P, Timmis J (ed), Artificial Immune Systems. Springer, Italy.
109. Negi SS, Braun W. 2009. Automated detection of conformational epitopes using phage display Peptide sequences. Bioinform Biol Insights 3:71-81.