Computational and Structural Biotechnology Journal 23 (2024) 759–770

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

Mini-review

Binding affinity between coronavirus spike protein and human ACE2 receptor

Marcus Ho-Hin Shum [a,b,c], Yang Lee [b,d], Leighton Tam [b,c], Hui Xia [e,f], Oscar Lung-Wa Chung [e], Zhihong Guo [f], Tommy Tsan-Yuk Lam [a,b,c,d,*]

[a] State Key Laboratory of Emerging Infectious Diseases, The University of Hong Kong, Hong Kong, China
[b] School of Public Health, The University of Hong Kong, Hong Kong, China
[c] Laboratory of Data Discovery for Health (D24H), Hong Kong Science Park, Hong Kong, China
[d] Centre for Immunology and Infection (C2i), Hong Kong Science Park, Hong Kong, China
[e] Department of Chemistry, South University of Science and Technology of China, China
[f] Department of Chemistry, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

ARTICLE INFO

Keywords:
Coronavirus
Spike protein
Binding affinity
Receptor use
Human ACE2

ABSTRACT

Coronaviruses (CoVs) pose a major risk to global public health due to their ability to infect diverse animal species and potential for emergence in humans. The CoV spike protein mediates viral entry into the cell and plays a crucial role in determining the binding affinity to host cell receptors. With particular emphasis on α- and β-coronaviruses that infect humans and domestic animals, current research on CoV receptor use suggests that the exploitation of the angiotensin-converting enzyme 2 (ACE2) receptor poses a significant threat for viral emergence with pandemic potential. This review summarizes the approaches used to study binding interactions between CoV spike proteins and the human ACE2 (hACE2) receptor. Solid-phase enzyme immunoassays and cell binding assays allow qualitative assessment of binding but lack quantitative evaluation of affinity. Surface plasmon resonance, Bio-layer interferometry, and Microscale Thermophoresis on the other hand, provide accurate affinity measurement through equilibrium dissociation constants (K_D). *In silico* modeling predicts affinity through binding structure modeling, protein-protein docking simulations, and binding energy calculations but reveals inconsistent results due to the lack of a standardized approach. Machine learning and deep learning models utilize simulated and experimental protein-protein interaction data to elucidate the critical residues associated with CoV binding affinity to hACE2. Further optimization and standardization of existing approaches for studying binding affinity could aid pandemic preparedness. Specifically, prioritizing surveillance of CoVs that can bind to human receptors stands to mitigate the risk of zoonotic spillover.

1. Background

Coronaviruses (CoVs) are enveloped, positive-sense, single-stranded RNA viruses belonging to the family *Coronaviridae*, subfamily *Orthocoronavirinae* [1,2]. The CoV genome is linear and monopartite, ranging in size from 27,000 bp to 32,000 bp, consisting of four main structural proteins: spike (S), envelope (E), membrane (M), and nucleocapsid (N) [2]. There are four genera within *Orthocoronavirinae* namely *alphacoronavirus* (α-CoV), *betacoronavirus* (β-CoV), *gammacoronavirus* (γ-CoV) and *deltacoronavirus* (δ-CoV) (Fig. 1) [3].

CoVs infecting humans and other mammals such as camels, felines and canines mainly belong to two genera, the α-CoVs and β-CoVs (Fig. 1), which are believed to have originated in bats and rodents [4]. While the α-CoVs are detected in different genera of bats [4,5], they are also known to cause respiratory and enteric infection across a wide range of mammals (Fig. 1). α-CoVs such as the transmissible gastroenteritis virus (TGEV), porcine respiratory coronavirus (PRCV), porcine epidemic diarrhea virus (PEDV) and Swine Acute Diarrhea Syndrome Coronavirus (SADS-CoV) present as respiratory and enteric infections in pigs [6–10]. Canids and felids, including domestic dogs and cats, and wild raccoon dogs, may harbor the canine coronavirus (CCoV) and feline coronavirus (FCoV) [11–13]. α-CoVs have also been identified in mustelids, including ferrets and mink, as well as viverrids, such as civets [13–16]. Recently, novel α-CoVs have been identified in rats and shrews

* Corresponding author at: State Key Laboratory of Emerging Infectious Diseases, The University of Hong Kong, Hong Kong, China.
E-mail address: ttylam@hku.hk (T.T.-Y. Lam).

https://doi.org/10.1016/j.csbj.2024.01.009
Received 15 September 2023; Received in revised form 14 January 2024; Accepted 15 January 2024
Available online 17 January 2024

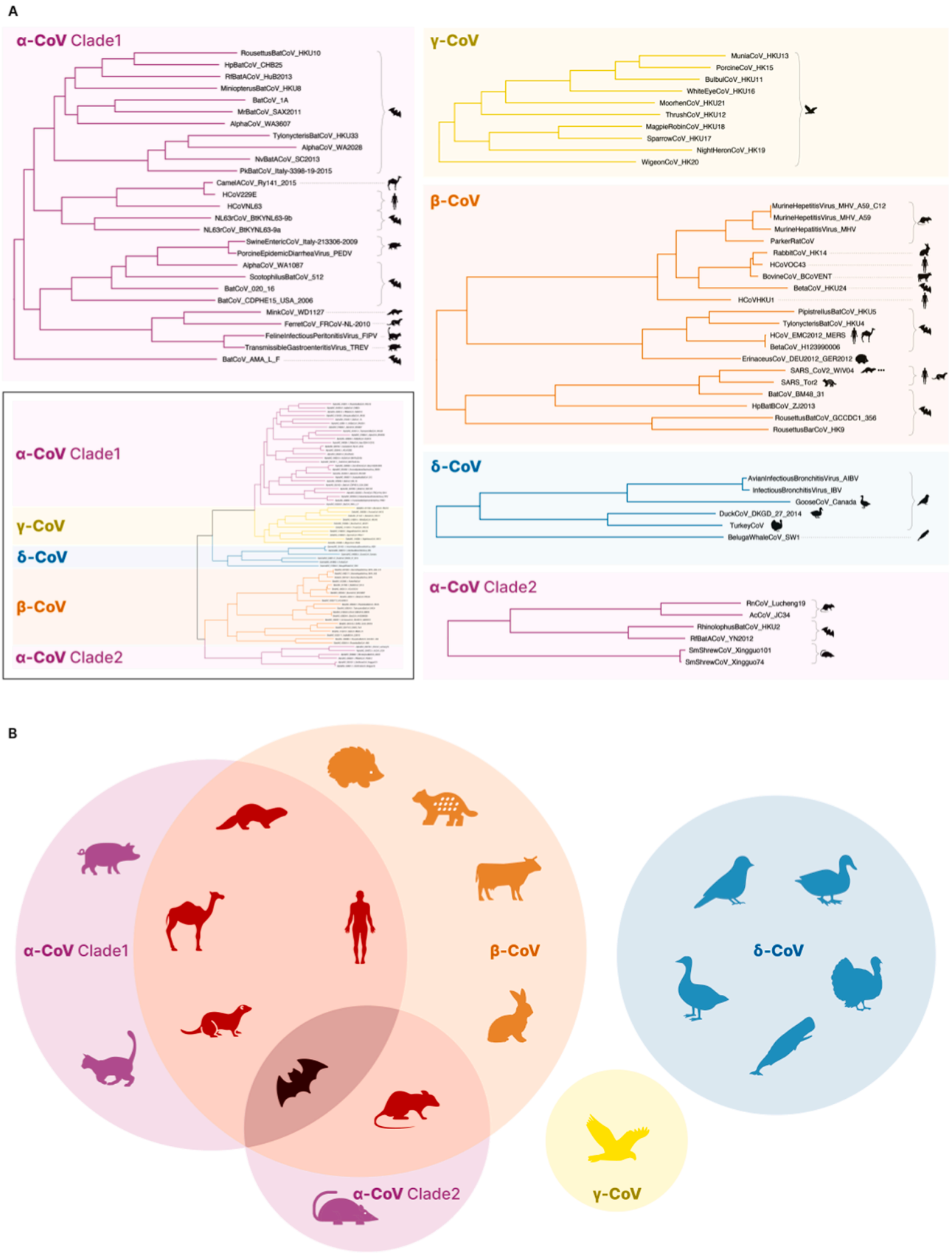

Fig. 1. A.) Phylogenetic trees constructed using the spike protein of all *Orthocoronavirinae* with complete genome (genome size from 27,000 to 32,000 bp). The phylogenies were inferred by IQTREE2 using GTR+I+R4 substitution model. The host of each CoVs is indicated with the label at the tip of each taxon. B.) The hosts where the four genera of *Orthocoronavirinae* are identified.

in China indicating their transmission in rodents and eulipotyphlans [17,18]. Two seasonal human coronaviruses (HCoVs) which cause mild respiratory infection namely HCoV-NL63 and HCoV-229E also belong to the α-CoV family [19–21].

β-CoVs have been associated with three of the most serious and fatal epidemics in humans: severe acute respiratory syndrome coronavirus (SARS-CoV), which emerged between 2003 and 2004 [22,23], severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which spread worldwide and caused a global pandemic from 2019 to 2023 [24], and Middle East respiratory syndrome coronavirus (MERS-CoV), which resulted in the epidemic in Middle East and part of Asia in 2012–2013 [25] and continuous zoonotic transmission. SARS-CoV and SARS-CoV2 belong to the subgenus *Sarbecovirus*, while MERS-CoV belongs to the subgenus *Merbecovirus*. Furthermore, there are seasonal human coronaviruses that cause relatively milder respiratory infections, such as HCoV-OC43 and HCoV-HKU1, which belong to the subgenus *Embecovirus* [21,26,27]. It is worth noting that SARS-CoV and other β-CoVs with high genomic similarities to SARS-CoV were found to be infecting palm civets [28,29]. MERS-CoV and β-CoVs-HKU23, which share high sequence identity with MERS-CoV, were detected in camels [30–32]. SARS-CoV2 can also infect ferrets, mink, as well as many other diverse mammalian hosts [33–36] while SARS-CoV-2-like and β-CoV-HKU4-like β-CoVs were found to be infecting pangolins [37,38]. Additionally, the host range of β-CoVs further extends to other mammals such as porcine hemagglutinating encephalomyelitis virus (PHEV) [39], canine respiratory coronavirus (CRCoV) [40], rodents with the HKU24 β-CoVs [41], erinaceus CoV [42,43], bovine coronavirus (BCoV) [44] and equine coronavirus (ECoV) [45] (Fig. 1).

γ-CoV and δ-CoV are mainly found in avian hosts which are believed to be the natural reservoirs of these CoVs. The infectious bronchitis virus (IBV), a γ-CoV, was one of the earliest identified coronaviruses in poultry, causing severe respiratory disease and nephritis in chickens [46]. Other diverse avian coronaviruses (AvCoVs) have been discovered, causing respiratory, enteric and genitourinary infection in poultry including ducks and turkeys [47,48]. AvCoVs have also been detected in wild birds such as swans, peacocks and pigeons. γ-CoV was also found in diseased cetaceans. The beluga whale coronavirus SW1 was first identified in whales in 2008 [49] and subsequently, a γ-CoV that shares a close evolutionary relationship was found in bottlenose dolphins [50]. To date, no γ-CoV has been found to infect terrestrial mammals.

δ-CoVs, in contrast, are most commonly found in wild birds. δ-CoVs with diverse genomic relationships were detected in different species of wild birds such as sparrows, falcons, pigeons and quails intermittently from 1999 to 2020 [51]. δ-CoVs have also been found to infect terrestrial mammals, but not marine mammals. Two δ-CoVs were discovered in leopard cats and ferret badgers in China in 2007 [52]. The porcine δ-CoV (PDCoV) HKU15 was first identified in pigs in 2007 and has been consistently detected in pigs in Asia and America until 2019 [53,54]. Recently, PDCoV was identified in plasma samples from children with febrile illness in Haiti [55], indicating the potential risk of zoonotic spillover of these avian-associated CoVs to humans.

2. Spike proteins of CoVs

The four CoV genera have demonstrated the capacity to infect a wide range of hosts, particularly α-, β- and δ-CoVs which can cause mild to severe respiratory disease in humans. Receptor-binding capability, which is crucial for cell entry mechanisms, is a key determinant in interspecies transmission [56]. Receptor-binding relies on the CoV S protein, forming a trimeric structure anchored on the virion surface membrane, which then binds to the surface receptor of the host cell and mediates virus entry into the cell [56]. The CoV S protein is composed of three segments: an ectodomain, a single-pass transmembrane anchor and a short intracellular tail [57]. The ectodomain can be further divided into two functional subunits – the S1 subunit which recognizes cell receptors through binding, and the S2 subunit which mediates virus-to-cell and cell-to-cell fusion [57]. The S1 subunit hence plays a vital role in determining the binding affinity of CoV S protein to host cell receptors. There are two receptor binding domains (RBD) within the S1 subunit, one located at the N-terminus known as S1-NTD which binds to cell surface carbohydrates, and another located at the C-terminus known as S1-CTD that binds to proteinaceous receptors [56–58]. Importantly, the S protein regulates the tissue and cell tropisms of CoVs, as the presence and abundance of receptors varies between different cell and tissue types depending on their expression level [56–58].

3. Use of host receptors by CoVs

CoVs can bind to various cell-surface carbohydrates and proteinaceous receptors. Structural discrepancy of the S1-NTD and S1-CTD has been observed in different CoVs, driving the use of different receptors for cell entry. The structures of human-infecting CoVs (SARS-CoV, SARS-CoV-2, MERS-CoV, HCoV-OC43, HCoV-NL63, HCoV-229E and HCoV-HKU1) S proteins are demonstrated in Fig. 2.

Research on the receptor binding affinities of CoVs has primarily focused on α-CoVs and β-CoVs that infect humans (or CoVs with close evolutionary relationship) as well as domestic animals [56–58]. Most α-CoVs exploit aminopeptidase N (APN) as their receptors (Fig. 3) [56]. The S1-CTD of the S proteins belonging to TGEV, PRCV, CCoV and FCoV were shown to bind to their corresponding host APNs: swine in the case of TGEV and PRCV, canids for CCoV, and felids for FCoV [56]. The use of human APN for cell entry by HCoV-229E has been confirmed through a cell infection experiment [59]. Immunofluorescence assay and cell infection experiments have also shown that PDCoV, a δ-CoV, can use pig, chicken, feline, and human APNs as the receptor for cell entry [60], despite the high genomic difference between the S protein of PDCoV and α-CoVs [61]. This highlights the risk of human infection by PDCoV, which has been recently confirmed [55].

While most α-CoVs use APN as their receptor, β-CoVs use dipeptidyl-peptidase 4 (DPP4), also known as CD26, and angiotensin-converting enzyme 2 (ACE2) as a receptor for cell entry (Fig. 3) [62–65]. Most CoVs belonging to the subgenus *Merbecovirus* utilize the DPP4 protein as their receptor. This includes the epidemic MERS-CoV which can perform cell entry using DPP4 from different animals, including humans, camels and bats [63]. Bat *Merbecovirus* HKU4-CoV and the related Manis javanica HKU4-related CoV (MjHKU4r-CoV), isolated from pangolins, were also found to use DPP4, though not bat *Merbecovirus* HKU5-CoV [63]. CoVs of the *Sarbecovirus* subgenus use ACE2 as a receptor to initiate cell entry [56]. These include the two most transmissible CoVs in humans, SARS-CoV and SARS-CoV-2 [64,65]. The S protein of other *Sarbecoviruses* such as SARS-related CoVs found in bats and civet cats, and SARS-CoV-2-related CoVs found in bats and pangolins, were found to bind to ACE2 in different animal species. Surprisingly, HCoV-NL63, an α-CoV, also utilizes hACE2 as its receptor, but with a different receptor binding domain than β-CoVs [66]. Other β-CoVs use different protein and sugar receptors. For example, HCoV-OC43 and HCoV-HKU1 recognize sialic acid, whereas BCoV and CRCoV use the HLA-1 protein [56].

APN, DPP4, and ACE2 are the three major receptors for human-infecting CoVs [59,64–66]. ACE2 and DPP4 are expressed in the upper and lower respiratory tracts as well as in lung tissue cells, while APN is only expressed in the upper respiratory tract [67–69]. This suggests that CoVs using ACE2 and DPP4, including SARS-CoV, SARS-CoV-2, and MERS-CoV are more strongly associated with severe lower respiratory tract infections [70–72], while those using APN such as HCoV-229E are associated with milder respiratory syndromes [73]. Among the three pandemic, human-infecting CoVs, SARS-CoV and SARS-CoV-2 have a higher transmissibility than MERS-CoV [74,75]. This suggests that CoVs capable of using the hACE2 receptor may pose heightened pandemic threat with potential of high transmissibility, wider range of transmission, and risk of severe respiratory infection. Hence, establishing an understanding for the S protein binding affinity of diverse CoVs to the

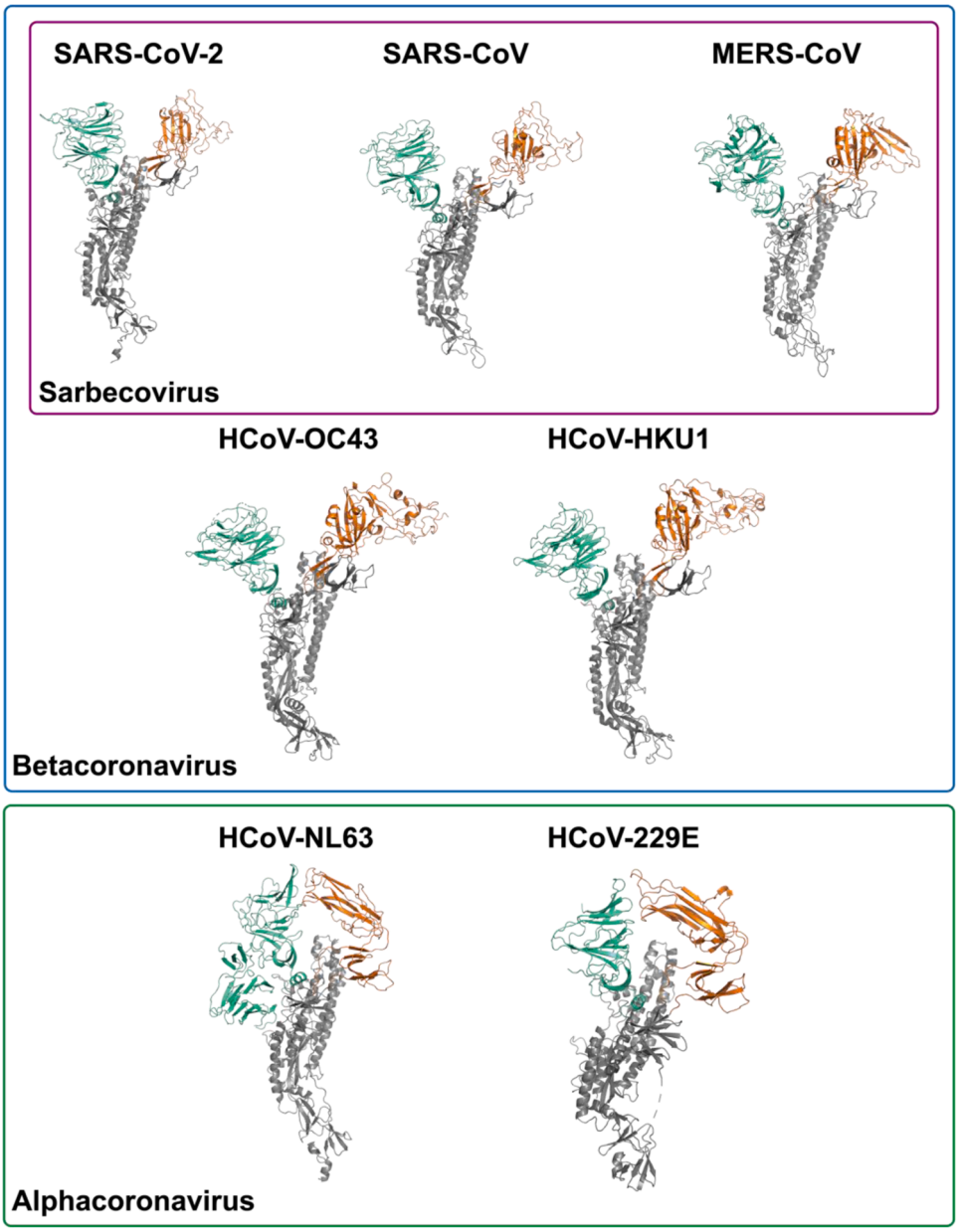

Fig. 2. Structure of monomers of the spike proteins from 7 human-infecting coronaviruses including SARS-CoV-2 (PDB: 7DDD), SARS-CoV (PDB: 5X58), MERS-CoV (PDB: 5X5C), HCoV-OC43 (PDB: 7SBV), HCoV-HKU1 (PDB: 8OHN), HCoV-NL63 (PDB: 5SZS) and HCoV-229E (PDB: 6U7H). The S1-NTD domain is highlighted in greencyan and the S1-CTD domain is highlighted in orange.

hACE2 receptor is a crucial step in assessing the risk of CoV zoonotic spillover from animals to humans, and may therefore help prevent future pandemics.

In this review, we have summarized and critically evaluated current approaches used to study the binding affinity between CoV S proteins and the hACE2 receptor (Fig. 4). The advantages and limitations of each approach are discussed (Table 1). Current knowledge on the binding affinity between the S protein of different CoVs and the hACE2 receptor are also reviewed.

4. Solid-phase enzyme immunoassays

Solid-phase enzyme immunoassays represent a conventional method for identifying binding between antigens and antibodies, and have later been used to study protein-protein interactions [76]. Enzyme-linked immunosorbent assays (ELISA) and dot-blot assays have been used to identify whether the S protein of a CoV can bind to hACE2 [77,78]. In brief, hACE2 is first prepared in solid phase (either precoated on the surface of each well on an ELISA plate for ELISA, or dotted onto nitrocellulose membranes for a dot-blot assay). The S1-NTD of the CoVs to be tested are then tagged, usually with a GST-tag or HIS-tag, and then incubated with the solid-phase hACE2. Enzyme-labeled antibodies such as horseradish peroxidase-conjugated goat anti-mouse IgG are then added and incubated. This is followed by the final step; an enzymatic reaction which is carried out using either ELISA substrate for ELISA or a chemiluminescent reagent for dot-blot assay [78,79]. For an ELISA, whether the S1-NTD region of the S protein has successfully bound to the hACE2 is determined by detecting changes in color intensity caused by absorbance at variable wavelengths. For a dot-blot assay, a positive result is instead signified by the presence of dots on the hACE2 coated membrane. Solid-phase enzyme immunoassays provide a cell-free approach in assessing the binding feasibility of the CoV S proteins and hACE2 with high specificity [79,80]. However, performing these cell-free assays requires professional laboratory expertise and is highly

Fig. 3. Structure of monomers of human ACE2 (PDB: 1R42), human DPP4 (PDB: 2ONC) and human CAN (PDB: 4FYQ).

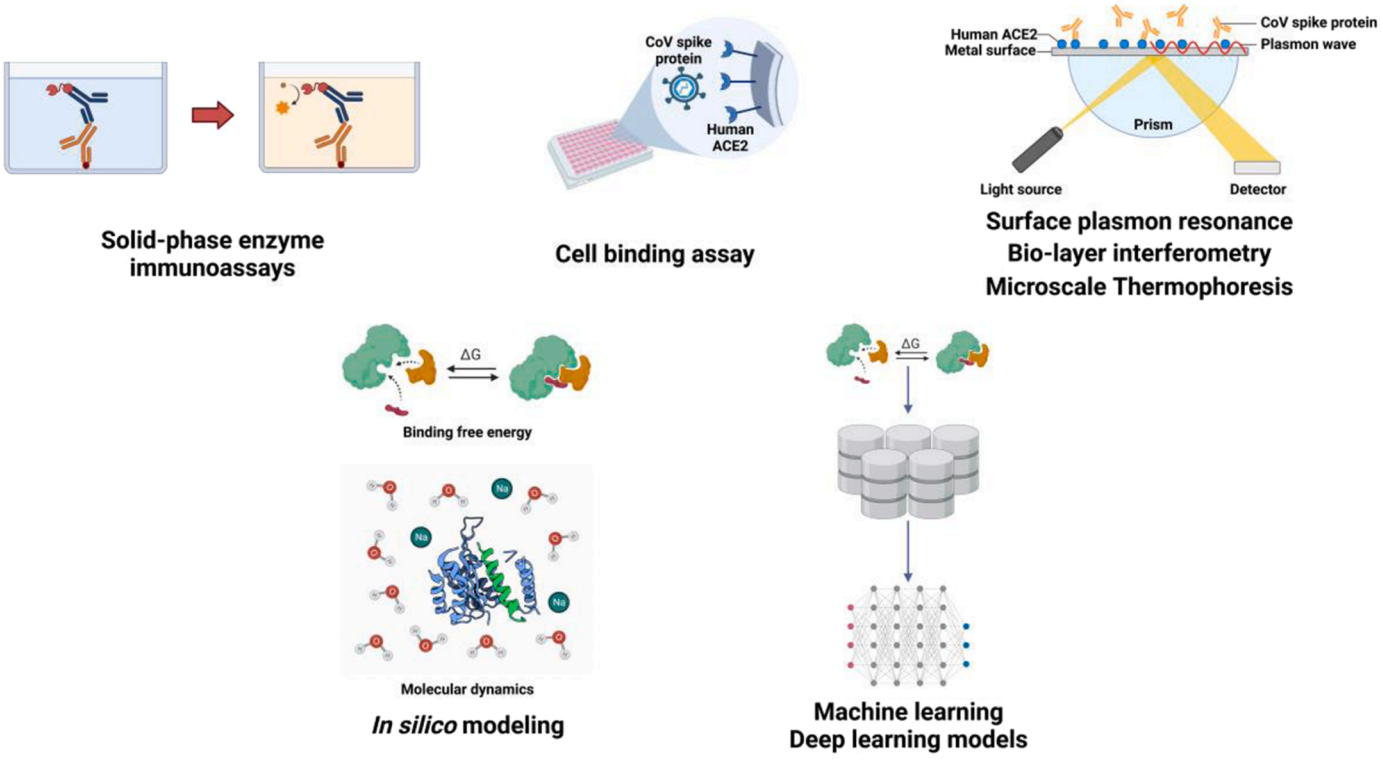

Fig. 4. Ideograph of the testing principal for different approaches in investigating the binding affinity between spike protein of coronavirus and human ACE2. Created with BioRender.com.

labor intensive. Furthermore, the absorbance value (in ELISA) and size of dots (in dot-blot assays) reflect the concentration of antigen (the concentration of CoV S1-NTD in this case), but not the binding affinity. Hence, these assays cannot be used to compare the binding affinities of different CoV S proteins to hACE2; instead, these are more frequently used to detect the presence and concentration of CoV in samples [81, 82].

5. Cell binding assay

Cell binding assays are widely used to study the binding affinity of different CoVs to hACE2 [83–91]. In such assays, the hACE2 protein is expressed in a variety of cells through transformation of plasmids carrying the hACE2 gene. These cells include yeast cells, e.g. *Saccharomyces cerevisiae* yeast strain EBY100 [92], mammalian cells, such as epithelial cells from different animals [88], as well as human cell lines such as HeLa cells, HEK 293 T cells, and Vero E6 cells [86,89,91]. A light-based detection method, similarly adopted in the solid-phase enzyme immunoassays, can be used to detect binding between cells expressing hACE2 and the S protein of diverse CoVs. They can also be employed to measure CoV infectivity, reflecting the entry efficiency of the virus, thereby partially revealing the binding affinity to hACE2 [81].

The pseudovirus system is another strategy commonly used in cell binding assays to investigate whether the hACE2 receptor can be targeted by different strains of SARS-CoV-2. One of the major reasons for the application of pseudovirus systems is due to the stringent biological safety requirements in handling live SARS-CoV-2 virus in biological experiments [93,94]. Within the pseudovirus system, the S protein of SARS-CoV-2 is expressed by modified viruses which are only capable of a single replication cycle upon entering susceptible cells. Consequently,

Table 1
Summary of the advantages and disadvantages of different approaches in investigating the binding affinity between spike protein of coronavirus and human ACE2.

Method	Advantage	Limitation
Solid-phase enzyme immunoassays	High specificity and sensitivity Provide a qualitative measure of whether the S protein can bind to human ACE2 or not	Unable to generate quantitative measurement of binging affinity for comparison
Cell binding assay	Reflecting the actual cell entry event Provide a qualitative measure of whether the S protein can bind to human ACE2 or not Application of pseudovirus method enable the study of virus with stringent biological safety requirement in laboratory experiment	Inconsistent result may be observed due to different human ACE2 expressing cells used.
Surface Plasmon Resonance, Bio-layer interferometry and Microscale Thermophoresis	Provide a quantitative measure of the binding affinity in terms of equilibrium dissociation constant (K_D) Compare a more accurate measurement of binding affinity	High purity requirement of the protein samples
In silico modeling	Provide a quantitative measure of the binding affinity in terms of binding energy computed using different mathematical approaches Do not require experimental validation	Inconsistent result may be observed when compared to wet lab experiment Results in different literatures are not comparable due to different parameters involved in the *in silico* modeling
Machine learning and deep learning models	Make use of the vast amount of experimental and simulated protein-protein interaction (PPI) data generated from *in silico* simulations Perform real-time analysis utilizing the graphics processing units	Comparatively low accuracy

the risk of an ongoing active infection is greatly minimized [93]. Another advantage is that these pseudoviruses possess surface protein conformations which closely resemble those of the native virus. These enable the pseudovirus to perform cell entry in a manner which accurately emulates the behavior of the actual virus [93]. Pseudoviruses previously used to study the binding affinity of SARS-CoV-2 to hACE2 include lentiviruses, vesicular stomatitis virus (VSV) and retrovirus (RV) [93,95]. Using a luciferase reporting system, cell entry can be observed and quantified through light-based detection methods [96].

While cell binding assays allow for the study of binding affinity of CoV to hACE2 under more realistic conditions (cell-based as opposed to solely protein-based in solid-phase enzyme immunoassays), they sometimes yield inconsistent results when different type of hACE2-expressing cells are used. For instance, Menachery et al. showed that pseudo-lentivirus expressing the S protein of CoV-SHC014 failed to enter HeLa cells expressing hACE2 [86]. However, Letko et al. demonstrated that pseudo-VSV expressing the CoV-SHC014 S protein could enter HEK 293 T cells expressing hACE2 [91]. One potential explanation for these incongruent results when using different hACE2 expressing cells could be the presence of other secondary receptors such as CD147 and L-SIGN protein [97]. Although secondary receptors have lower binding affinity to CoV S protein compared to the primary receptor, they have been shown to assist cell entry of CoV [56,57]. To further validate the results of pseudovirus-based infection experiments, conventional authentic viral infection assays should be conducted [98].

6. Surface plasmon resonance (SPR), Bio-layer interferometry (BLI) and Microscale Thermophoresis (MST)

Surface plasmon resonance (SPR) is a technique that can be used to measure interactions between two biomolecules, including binding affinity and specificity [99,100]. It has been recently applied to study binding affinity between different CoVs and hACE2 [89,101–107]. Similar to solid-phase enzyme immunoassays, SPR involves immobilizing the hACE2 on a metal film. However, instead of requiring a label in solid-phase enzyme immunoassays, SPR measures the change in the reflection angle of light when the S protein, or more specifically the RBD, binds to the immobilized hACE2, triggering the surface plasmon resonance phenomenon and altering the refractive index of the metal plate. The binding affinity between the S protein and hACE2 is measured by the equilibrium dissociation constant (K_D), which is calculated as the ratio of the rate of association (k_{on}, when the two proteins bind) to the rate of dissociation (k_{off}, when the two proteins detach) [99,100].

SPR is considered to provide a more accurate measurement of binding affinity compared to the methods mentioned in the previous sections [100]. It utilizes both the physical property of the metal chips and biophysical property when two proteins bind together. Therefore, there are potentially fewer factors that could lead to inconsistent results. In fact, various studies adopting SPR in studying the binding affinity of identical CoV S proteins to hACE2 have shown similar K_D values [96,98–100]. However, the high purity of CoV S protein required to conduct SPR currently represents a major obstacle to its routine application [99,108].

Like SPR, bio-layer interferometry (BLI) is a label-free technique for measuring the binding kinetics and affinity of protein-protein interactions, including the interaction between the CoV RBDs and its receptor. In BLI measurement, one protein is immobilized on a biosensor tip and its binding to the other protein in a static solution changes the optical interference pattern in a dose-dependent manner. This system allows the binding interactions to be monitored in real time. BLI is simpler, easier to operate, and higher in throughput than SPR. Furthermore, it is suitable for measuring binding affinities in the range of 10 pM to 1 mM [109]. BLI has been used to determine that the binding of SARS-CoV-2-RBD and SARS-CoV RBD to hACE2 has a K_D equivalent to 5.09 nM and 1.46 nM, respectively [110]. In a comparative study, BLI and SPR were found to afford similar binding affinity K_D but resulted in rather different association and dissociation rate constants (k_{on} and k_{off}) [111]. BLI has also been used to determine the binding affinities of RBD variants such as E484K for hACE2 [112] and screen for viral entry inhibitors [113,114].

To avoid the slow mass transfer and immobilization effects in both SPR and BLI, another technique called microscale thermophoresis (MST) may be more suitable to determine the binding affinities of the spike protein RBD for its hACE2 receptor. MST measures the motion of fluorescent molecules in a homogenous solution along a microscale temperature gradient, which reflects changes in molecular size, charge, and hydration shell. To perform a measurement, RBD is tagged with a fluorescent probe and mixed with a varied concentration of the receptor. This mixture is then injected into a glass capillary where changes in fluorescence can be measured over time. The resulting curves are then fitted with various kinetic protein-protein binding models to determine K_D. In a previous study, K_D for the binding of hACE2 with wild-type SARS-CoV-2 spike RBD was found to be 27.5 ± 4.8 nM [115]. Combined with steered molecular dynamics (SMD) simulations of the interactions with the RBD variants of concern (α-, β-, δ-CoV, and other variants), this MST study provides a deeper understanding of these interactions at the molecular-level.

7. *In silico* modeling

With the availability of the binding structure between the S protein of CoVs and the hACE2 receptor (such as SARS-CoV [64] and SARS-CoV-2 [105]), it is possible to harness recent advances in computational power for the *in silico* prediction of the binding affinity between different CoV S proteins and hACE2. Computational methods include binding structure modeling, protein-protein docking simulations, and binding energy calculations [116–125].

Binding structure modeling uses homology modeling to compute the heavy atom positions of each amino acid of a given CoV S protein based on a given reference structure using programs such as SWISS-MODEL [126] and Modeller [127]. While most studies employing *in silico* methods for the prediction of binding affinity between CoVs and hACE2 adopt homology modeling, reference-free structure construction of the S protein is now possible due to the introduction of deep learning (DL) models. AlphaFold [128], built with a deep neural network, is trained with an existing protein structure database. By aligning the query protein primary sequences to the database and constructing the embedded multiple sequence alignment as the input, AlphaFold can predict the 3D coordinates of all heavy atoms in the given query protein sequence's folded structure with high accuracy. ESM-Fold, on the other hand, utilizes a language learning model and relies on the token embedding of the large pre-trained protein sequences [129]. This allows for a reference-free prediction of the folded protein structure using ESM-Fold.

Homology modeling provides an initial prediction of the binding structure between the S protein (specifically, the S1-NTD) and hACE2. The coordinates of the light atoms, such as hydrogen, contribute heavily to binding affinity, and need to be fine-tuned using protein-protein docking. Upon homology modeling, the contact surface between the S protein and hACE2 is well-defined. Hence, local protein-protein docking is used where the computation of atom-atom interaction is restricted only to atoms located at the contact surface, resulting in shorter computational time. However, if the initial structure of the S protein is constructed using a reference-free approach, global protein-protein docking is necessary to identify the optimal orientation for the contact surface of the two proteins to interact. Programs used for protein-protein docking in CoVs binding affinity to hACE2 includes Rosetta [130] and HADDOCK [131].

The binding affinity between the S protein and hACE2 can be represented by estimating the binding free energy. Two simulation methods are commonly used to calculate binding free energy. These are Molecular mechanics Poisson-Boltzmann surface area (MM/PBSA) and molecular mechanics generalized Born surface area (MM/GBSA) combined with molecular dynamics (MD) [132]. The calculation of free energy using MM/PB(GB)SA includes the following steps: first, an MD simulation of the binding complex is performed using an explicit solvent model (implicit solvent simulations are not recommended as they have been shown to generate less accurate results [133]); second, solvent molecules and charged ions are removed from the MD snapshots; finally, the solvation energy of the ligand, protein receptor and the complex form are evaluated using the MD snapshots. The final binding free energies are obtained by summing the individual energy components. Programs such as AMBER [134] and GROMACS [135] can perform the MM/PB(GB)SA calculation. Although MM/PB(GB)SA can be combined with MD simulations to provide an accurate prediction of the binding free energy, it is comparatively less efficient and flexible due its computational requirements [136,137]. Alternatively, the binding free energy can be calculated directly from molecular mechanics-based empirical force fields. This approach is more efficient and yields a less accurate, yet comparable result to the MM/PB(GB)SA. The Rosetta software suite for macromolecular modeling and HADDOCK have been shown to produce results comparable with MM/PB(GB)SA-computed free energy estimates. Rosetta computes the Rosetta Energy Unit (REU) [138], whereas HADDOCK [131] computes the HADDOCK score. Both programs have been used to calculate the binding free energy between S protein and hACE2 in coronavirus [117,121,139–141].

The binding free energy can be used to compare binding affinities. The lower the binding energy, the stronger the binding affinity between two proteins [142]. Several different programs have shown that the SARS-CoV-2 S protein and hACE2 complex has a lower free binding energy than the SARS-CoV S protein and hACE2 complex, therefore suggesting that the former has a stronger binding affinity for the ACE2 receptor [117–119,122,124,125] (Fig. 5). The *in silico* binding free energy results are commensurate with the *in vitro* SPR results which also show a higher K_D for SARS-CoV-2 S protein with hACE2 than SARS-CoV S protein [103,105,107] (Fig. 5). The binding affinity of other CoVs to hACE2 can also be evaluated. Na et al. identified two *sarbecovirus* clade 2 CoVs from bats in Korea (KB-CoV) [118]. PRODIGY and MM/GBSA approaches were used to compute and reveal a high binding free energy between the KB-CoV-hACE2 complex, suggesting a low binding affinity to hACE2 for KB-CoV [118] (Fig. 5). This is consistent with the cell infection experiments conducted by Guo et al. and Starr et al., which showed that the *sarbecovirus* clade 2 CoVs cannot use hACE2 for cell entry due to the deletion of two loops in the receptor binding domain [143,144] (Fig. 5). Wu et al. performed MD simulation using GROMACS program for the RBDs-ACE2 complexes of the SARS-CoV-2 Delta and Omicron variants and used the MM/GBSA method to compute their binding free energy (ΔG) in 2022. Their analysis of three trajectories for each complex revealed that the Omicron variant has a similar binding affinity for hACE2 as the wild-type SARS-CoV-2 but a significantly weaker binding affinity than the Delta variant. These simulation results were further validated with ELISA bioassay [145]. Moreover, it is important to note that *in silico* results may not always be aligned with *in vitro* experiment results [106,107,117,124] (Fig. 5, annotated in orange). Chowdhury et al. computed the REU for the bat CoV-RaTG13, which is a CoV that shares a very close genetic relationship with SARS-CoV-2. The study reported that bat CoV RaTG13 had a lower REU, i.e. stronger binding, to hACE2 compared to SARS-CoV and a comparable REU to SARS-CoV-2 (RaTG13: -43.168 ± 2.1 kcal/mol, SARS-CoV: -37.308 ± 2.3 kcal/mol, SARS-CoV-2: -48.312 ± 3.4 kcal/mol [117]), suggesting a stronger binding affinity to hACE2 than SARS-CoV as well as the utilization of hACE2 as receptor. Comparatively, Zhang et al. measured the binding affinity of CoV-RaTG13 to hACE2 by conducting SPR and cell entry experiments. Their results indicated that the binding affinity between CoV-RaTG13 and hACE2 was five times weaker compared to SARS-CoV-2 and hACE2, and showed minimal evidence of CoV-RaTG13 entry in HEK 293 T cells expressing hACE2 [106]. Jamir et al. computed the free binding energy of the NL63-S-protein-to-hACE2 complex using Prodigy [124]. They found a lower free binding energy compared to SARS-CoV. In contrast, Lan et al. used the SPR approach to compare the binding affinity of hACE2 with SARS-CoV, SARS-CoV-2 and NL63 [107]. Their results showed that SARS-CoV had a stronger binding affinity to hACE2 compared to NL63, conflicting with the result based on the free binding energy computed from Prodigy.

Note that amino acid substitutions on CoV S proteins or hACE2 may affect the binding affinity between them. For instance, Dos Santos et al. identified that the E484K and N501Y mutations on the spike protein of SARS-CoV-2 increase the binding affinity between the RBD and hACE2 by introducing additional attractive interactions and a hydrogen bond with hACE2 [146]. Binding free energy can also be used to identify the contribution of individual amino acids to the binding affinity between CoV S protein and hACE2. Wierbowski et al. used PyRosetta [123] and Ren et al. used mCSM-PPI2 [102] to compute the changes in binding free energy for missense mutations found in the hACE2 protein. Their studies aimed to investigate the polymorphisms on hACE2 that contribute to the change in binding affinity to SARS2-CoV. Ren et al. also conducted SPR experiments to validate their findings. Additionally, Cavani et al. calculated the change in binding free energy based on the amino acid mutations between the wild type SARS-CoV-2 and the Delta lineage [120]. Xue et al. utilized Flex ΔΔG, a ΔΔG estimation method developed

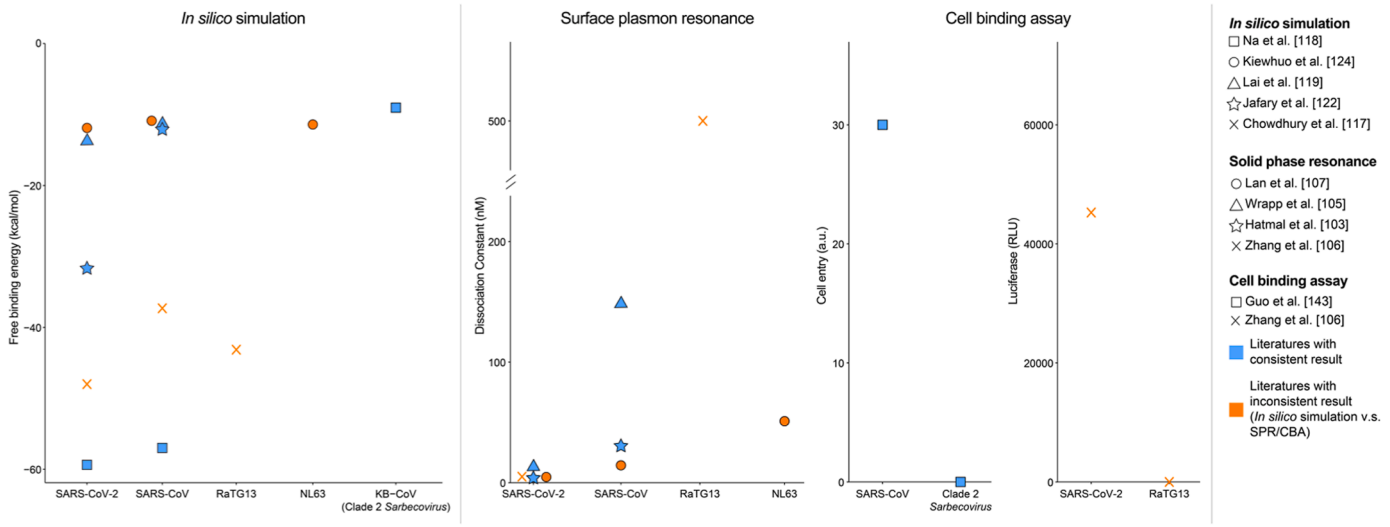

Fig. 5. *In silico* and *in vitro* assessments of binding between coronavirus spike proteins and ACE2 receptors reported in literature. Distribution of free binding energy computed from *in silico* simulation, dissociation constant measured from solid plasmon resonance (SPR), cell entry and luciferase RLU measured in cell binding assay (CBA) are shown in the left, middle and right panels. The lower the free binding energy and dissociation constant, the stronger the binding affinity. In contrast, the lower the cell entry and luciferase RLU, the lower the binding affinity. Inconsistent results between *in silico* simulation and SPR/CBA *in vitro* experiments are shown in orange symbols; consistent results are shown in blue symbols. Data sources are indicated at the figure legend on the right.

within the Rosetta software suite, to predict the binding affinity change by point mutation(s) on the RBD binding surface. Based on the predicted ΔΔG scores, candidate mutants with a significantly negative impact on binding affinities were selected for further experimental validation. Out of the nine recommended mutants, six showed improved affinity for hACE2 in SPR analysis [147]. Calcagnile et al. used HDOCK and FireDock to compute the global energy score (GES) of the interaction between wild type hACE2 or ACE2 missense variants and SARS-CoV-2 S protein [148]. Their results showed that multiple sites were associated with increased or reduced binding affinity. These sites include I21T and K26R which were found to result in higher binding affinity and I21V and K26D which were found to be reducing the binding affinity [148,149]. Glycosylation on CoV S proteins has also been associated with changes in their binding affinity to hACE2 [150–152]. Huang et al. performed an MD simulation and computed the binding free energy (ΔG) for glycosylated-RBDs-hACE2 complexes using GROMACS programs [150]. Their results showed that glycosylated RBDs had lower binding affinities to hACE2 compared with the wild-type RBD, no matter which types of glycan were attached. The presence of Man 5, NA2F, NA3F, A2G and A3F has led to significant steric effects hindering the binding between the RBDs and hACE2. Moreover, when the RBD is glycosylated with fully sialylated bi-antennary and tri-antennary A2F and A3F, Coulombic repulsion was observed. This phenomenon drives the RBD away from the hACE2 and hence reduces binding affinity. Huang et al. also conducted SPR assays confirming the findings in the MD simulations [150].

8. Machine learning and deep learning models

New methods have been developed to effectively utilize the increasing number of available MD snapshots and protein-protein interaction (PPI) data for real-time prediction of binding affinities and elucidating the key amino acid residues associated with CoV binding affinity to hACE2.

Machine learning (ML) classifiers use a training dataset to "learn" and identify important features which can then be used to classify test datasets [153]. ML classifiers have been trained to discover amino acid residues that significantly contribute to changes in SARS-CoV-2's binding affinity to hACE2. Pavlova et al. trained three ML models to distinguish between SARS-CoV-2 RBD from SARS-CoV RBD [154]. Specifically, this involved training a linear logistic regression model, a tree-based random forest model and a multilayer perceptron neural network, using over 4500 features generated from MD simulation snapshots of SARS-CoV RBD and SARS-CoV-2 bound to hACE2 By identifying the most important features which distinguish the two RBD, the ML classifier was able to select amino acid residues that contribute to an increased binding affinity between hACE2 and SARS-CoV-2 mutants such as N501 [155].

In recent years, machine learning has been used to increase the accuracy of data-driven modeling [156]. The large amount of data required can be generated by deep mutational scans (DMS), which reveal the functional consequence for all possible single mutations [157]. In 2020, Starr et al. conducted a comprehensive DMS assessment of the impact of various single-residue mutations in the SARS-CoV-2 RBD, and quantified the effects on RBD expression and hACE2 binding [158]. A large percentage of single residue mutations (84.5%) were shown to have negative effects on RBD expression and hACE2 binding, while roughly 7.5% have no impact, and only around 8% increase binding affinities.

In 2021, Chen et al. selected 27 variants with increased binding affinity and 54 variants with lowered binding affinity on the binding surface, to form an experimental DMS dataset together with 27 variants with enhanced binding affinity but do not directly interact with hACE2 [159]. The dataset was used to train a neural network regression model (NN_MM-GBSA) which used the decomposed MM/GBSA energy terms (Coulombic, covalent, van der Waals, lipophilic, generalized Born electrostatic solvation, hydrogen bonding, $\pi-\pi$ packing, and self-contact correction terms) from MD simulation trajectories as ML features, and the ratio of experimental dissociation constant ($K_{D,app}$) of the variants relative to the wild-type as the target. This model achieved a correlation coefficient of 0.73 between the ML-predicted results and experimental values, and a validation accuracy of 82.8% for predicting whether a single-residue variant of RBD increases or decreases the binding affinity for hACE2. It provides a new framework for predicting $K_{D,app}$ from features of decomposed MM/GBSA energy terms.

In silico simulations provide an approach to evaluate the binding affinity between two proteins. However, it becomes inefficient when one needs to compute the binding affinity of multiple protein pairs as each simulation necessitates MD simulation, leading to significant time requirements. DL architectures, including deep reinforcement learning, convolution neural networks and transformers, enable real-time

classification or prediction after training by utilizing the power of graphics processing units (GPUs) [160]. Xie et al. developed Emvirus, a DL model using the convolutional neural network and bi-directional long short-term memory, trained with over 27,000 protein-protein interaction data from different viruses including HIV, Herps, influenza, SARS-CoV and SARS-CoV-2 [161]. When trained with PPI data of only one type of virus (except SARS-CoV and SARS-CoV-2), Emvirus revealed a medium accuracy ranging from 46.4% to 54.0%.

9. Discussion and Conclusion

This review summarizes methods used in studying binding affinity between CoV S protein and hACE2. Among these methods, the solid-phase enzyme immunoassay is the most conventional technique for qualitatively assessing the potential for two proteins to bind. Cell binding assays allow the investigation of binding affinity between two proteins to be conducted in context of viral particles and host cells. However, inconsistent results can arise when different types of cells are used to express the target receptors. SPR, BLI and MST provide a more accurate and consistent measurement of binding affinity but stringent protein purity requirements limit its applicability. Also, the use of different detection buffers in these binding assays may potentially affect the result. *In silico* simulation of binding affinity has become feasible with continuing improvements in computational power and advances in molecular biophysics. They provide a convenient preliminary understanding of the binding affinity of two proteins with minimal laboratory requirements.

Taking SARS-CoV and SARS-CoV-2 as examples, results from *in silico* simulations of CoV S protein to hACE2 binding affinities appear to show greater inconsistency among themselves when compared to the *in vitro* SPR results [117–119,122,124] (Fig. 5). It is notable that these studies use inconsistent data (sequences, reference structure, etc.) and model parameters which can lead to incomparable results. Chowdhury et al. retrieved the SARS-CoV-2-hACE2 complex structure from PDB (ID: 6LZG) [117] where others used another one (ID: 6M0J) [118,119,122,124]. These studies used different binding free energy computation programs. Three estimated the binding free energy directly from molecular mechanics-based empirical force fields using Rosetta [117] and Prodigy [118,124]. The other two performs MD simulation and perform the MM/PB(GB)SA calculation using GROMACS [119] and AMBER [122], but with different factors in MD simulation. For instance, Lai et al. performed the MD simulation at a temperature of 310 K where the hydrogen bonds were constrained by the LINCS method and other electrostatic interaction was treated by the Particle Mesh Ewald method [119] whereas Jafary et al. conducted at a temperature of 300 K with all electrostatic interaction including hydrogen bonds using the SHAKE algorithm [122]. Therefore, it is crucial to establish and design a standardized approach for protein structure homology modeling, as well as the uniform selection of structures, solvents, ions, environment factors (such as temperature) and electrostatic interaction constrain method for MD simulation, perhaps one that resembles the human internal environment. This would allow the generation of comparable results when studying different CoV S proteins in various investigations.

Despite its low amino acid sequence identity and structural similarity to SARS-CoV and SARS-CoV-2, HCoV-NL63 also uses hACE2 as its host receptor. This is achieved via an RBD on its S protein that is completely different to those of the other coronaviruses, and binds to a completely different epitope on hACE2 [66]. The SARS-CoV-hACE2, SARS-CoV-2-hACE2 and MERS-CoV-hACE2 complex structures have been most commonly used as homology modeling references in *in silico* simulation studies [117,121,139–141]. While these complexes serve as the reference structures for the binding between β-CoV and hACE2, the binding structure of HCoV-NL63 and hACE2 [66] may serve as a reference structure for the binding between α-CoVs and hACE2. It can be used for *in silico* simulation analysis and has the potential to reveal other α-CoVs that also exploit hACE2 as their receptors. This could help identify the unknown intermediate host involved in the transmission of HCoV-NL63 from bats to humans.

ML and DL techniques offer the opportunity to leverage the vast amount of experimental and simulated PPI data generated by *in silico* simulations. They were introduced as a novel approach to elucidate key features contributing to the binding affinity to hACE2. Recently, DL models such as transformer and large language models, originally used in graphic generation [162] and language learning [163], have been successfully applied in protein structure prediction [128], protein structure alignment [164] and molecular docking [165]. Application of these DL models trained on extensive experimental PPI data may further optimize *in silico* simulation results and potentially explain the inconsistencies between computational simulations and biological experimental findings.

Author contributions

MHH Shum, Z Guo and TTY Lam contributed to the conception of this work. MHH Shum, Y Lee, L Tam collected the data. MHH Shum wrote the manuscript. MHH Shum, Y Lee, HX, OLW Chung, Z Guo and TTY Lam revised the manuscript. All authors read and approved the final manuscript.

Declaration of Competing Interest

We declare that all authors have read and approved the final version of the manuscript. We declare that the manuscript has not been published before and is not under consideration for publication elsewhere. We declare no conflicts of interest.

Acknowledgement

This project is mainly supported by Collaborative Research Fund (C7144-20GF) from Research Grants Council. It is also supported by InnoHK funding from Innovation and Technology Commission of the Hong Kong Special Administrative Region (HKSAR), China. We also thank Dr. Alexander Williams and Dr. Brian Worthington for editing the language of the manuscript, and Dr. Yun Lan for assistance in creating the figure.

Declaration of generative AI in scientific writing

The authors hereby declare that generative AI was NOT used in writing the manuscript nor figure generation.

References

[1] Su S, et al. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. Trends Microbiol 2016;24(6):490–502.
[2] King AML, Adams E, Carstens MJ, Virus EB. Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses, Vol. 9. Amsterdam, The Netherlands: Elsevier; 2011.
[3] Walker PJ, et al. Changes to virus taxonomy and the Statutes ratified by the International Committee on Taxonomy of Viruses (2020). Arch Virol 2020;165 (11):2737–48.
[4] Su S, et al. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. Trends Microbiol 2016;24(6):490–502.
[5] Cui J, Li F, Shi Z-L. Origin and evolution of pathogenic coronaviruses. Nat Rev Microbiol 2019;17(3):181–92.
[6] Cox E, et al. Intestinal replication of a porcine respiratory coronavirus closely related antigenically to the enteric transmissible gastroenteritis virus. Vet Microbiol 1990;23(1):237–43.
[7] Laude H, Van Reeth K, Pensaert M. Porcine respiratory coronavirus: molecular features and virus-host interactions. Vet Res 1993;24(2):125–50.
[8] Jackwood DJ, et al. Transmissible gastroenteritis virus and porcine respiratory coronavirus: molecular characterization of the S gene using cDNA probes and nucleotide sequence analysis. Adv Exp Med Biol 1993;342:43–8.
[9] Jung K, Saif LJ. Porcine epidemic diarrhea virus infection: Etiology, epidemiology, pathogenesis and immunoprophylaxis. Vet J 2015;204(2):134–43.
[10] Yang Y-L, Yu J-Q, Huang Y-W. Swine enteric alphacoronavirus (swine acute diarrhea syndrome coronavirus): An update three years after its discovery. Virus Res 2020;285:198024.

[11] Tennant BJ, et al. Studies on the epizootiology of canine coronavirus. Vet Rec 1993;132(1):7–11.
[12] Kipar A, et al. Sites of feline coronavirus persistence in healthy cats. J Gen Virol 2010;91(Pt 7):1698–707.
[13] Haake C, et al. Coronavirus Infections in Companion Animals: Virology, Epidemiology, Clinical and Pathologic Features. Viruses 2020;12. https://doi.org/10.3390/v12091023.
[14] Ghai RR, et al. Animal Reservoirs and Hosts for Emerging Alphacoronaviruses and Betacoronaviruses. Emerg Infect Dis 2021;27(4):1015–22.
[15] Decaro N. Alphacoronavirus(‡): Coronaviridae. Springe Index Virus 2011:371–83. https://doi.org/10.1007/978-0-387-95919-1_56.
[16] He WT, et al. Virome characterization of game animals in China reveals a spectrum of emerging pathogens. Cell 2022;185(7):1117–1129.e8.
[17] Wang W, et al. Discovery of a highly divergent coronavirus in the Asian house shrew from china illuminates the origin of the alphacoronaviruses. J Virol 2017; 91(17).
[18] Wang W, et al. Discovery, diversity and evolution of novel coronaviruses sampled from rodents in China. Virology 2015;474:19–27.
[19] van der Hoek L, et al. Identification of a new human coronavirus. Nat Med 2004; 10(4):368–73.
[20] Hamre D, Procknow JJ. A new virus isolated from the human respiratory tract. Proc Soc Exp Biol Med 1966;121(1):190–3.
[21] Myint SH. Human coronaviruses: a brief review. Rev Med Virol 1994;4(1):35–46. https://doi.org/10.1002/rmv.1980040108. Epub 2005 Oct 18.
[22] Tsang KW, et al. A cluster of cases of severe acute respiratory syndrome in Hong Kong. N Engl J Med 2003;348(20):1977–85.
[23] Lee J-W, McKibbin WJ. Globalization and disease: the case of SARS*. Asian Econ Pap 2004;3(1):113–31.
[24] Zhu N, et al. A novel coronavirus from patients with pneumonia in China, 2019. New Engl J Med 2020;382(8):727–33.
[25] Zaki AM, et al. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. New Engl J Med 2012;367(19):1814–20.
[26] Monto AS, Lim SK. The tecumseh study of respiratory illness. VI. Frequency of and relationship between outbreaks of coronavims infection. J Infect Dis 1974; 129(3):271–6.
[27] Woo PC, et al. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J Virol 2005;79 (2):884–95.
[28] Wang M, et al. SARS-CoV infection in a restaurant from palm civet. Emerg Infect Dis 2005;11(12):1860–5.
[29] Shi Z, Hu Z. A review of studies on animal reservoirs of the SARS coronavirus. Virus Res 2008;133(1):74–87.
[30] Azhar EI, et al. Evidence for camel-to-human transmission of MERS coronavirus. New Engl J Med 2014;370(26):2499–505.
[31] Raj VS, et al. Isolation of MERS coronavirus from a dromedary camel, Qatar, 2014. Emerg Infect Dis 2014;20(8):1339–42.
[32] Woo PC, et al. Novel betacoronavirus in dromedaries of the Middle East, 2013. Emerg Infect Dis 2014;20(4):560–72.
[33] Kim Y-I, et al. Infection and rapid transmission of SARS-CoV-2 in ferrets. Cell Host Microbe 2020;27(5):704–9. e2.
[34] Oreshkova N, et al. SARS-CoV-2 infection in farmed minks, the Netherlands, April and May 2020. Eur Surveill 2020;25(23).
[35] Palmer MV, et al. Susceptibility of white-tailed deer (Odocoileus virginianus) to SARS-CoV-2. J Virol 2021;95(11).
[36] Italiya J, et al. First detection of SARS-CoV-2 in white rhinoceros during a small-scale coronavirus surveillance in the Bandia Reserve, Senegal. Animals 2023;13 (16).
[37] Lam TT-Y, et al. Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. Nature 2020;583(7815):282–5.
[38] Shi W, et al. Trafficked Malayan pangolins contain viral pathogens of humans. Nat Microbiol 2022;7(8):1259–69.
[39] Mora-Díaz JC, et al. Porcine hemagglutinating encephalomyelitis virus: a review. Front Vet Sci 2019;6.
[40] Erles K, Brownlie J. Canine respiratory coronavirus: an emerging pathogen in the canine infectious respiratory disease complex. Vet Clin North Am: Small Anim Pract 2008;38(4):815–25.
[41] Lau Susanna KP, et al. Discovery of a novel coronavirus, china rattus coronavirus HKU24, from Norway rats supports the murine origin of betacoronavirus 1 and has implications for the ancestor of betacoronavirus lineage A. J Virol 2015;89 (6):3076–92.
[42] Corman Victor M, et al. Characterization of a novel betacoronavirus related to middle east respiratory syndrome coronavirus in European hedgehogs. J Virol 2014;88(1):717–24.
[43] Lau SKP, et al. Identification of a novel betacoronavirus (Merbecovirus) in amur hedgehogs from China. Viruses 2019;11. https://doi.org/10.3390/v11110980.
[44] Clark MA. Bovine coronavirus. Br Vet J 1993;149(1):51–70.
[45] Zhang J, et al. Genomic characterization of equine coronavirus. Virology 2007; 369(1):92–104.
[46] Raj GD, Jones RC. Infectious bronchitis virus: immunopathogenesis of infection in the chicken. Avian Pathol 1997;26(4):677–706.
[47] Chen G-Q, et al. Identification and survey of a novel avian coronavirus in ducks. PLoS One 2013;8(8):e72918.
[48] Guy JS. Turkey coronavirus is more closely related to avian infectious bronchitis virus than to mammalian coronaviruses: a review. Avian Pathol 2000;29(3): 207–12.
[49] Mihindukulasuriya Kathie A, et al. Identification of a novel coronavirus from a beluga whale by using a panviral microarray. J Virol 2008;82(10):5084–8.
[50] Woo Patrick CY, et al. Discovery of a novel bottlenose dolphin coronavirus reveals a distinct species of marine mammal coronavirus in gammacoronavirus. J Virol 2014;88(2):1318–31.
[51] Vlasova AN, et al. Deltacoronavirus evolution and transmission: current scenario and evolutionary perspectives. Front Vet Sci 2020;7:626785.
[52] Dong BQ, et al. Detection of a novel and highly divergent coronavirus from asian leopard cats and Chinese ferret badgers in Southern China. J Virol 2007;81(13): 6920–6.
[53] Zhang J. Porcine deltacoronavirus: overview of infection dynamics, diagnostic methods, prevalence and genetic evolution. Virus Res 2016;226:71–84.
[54] Boley PA, et al. Porcine deltacoronavirus infection and transmission in poultry, United States(1). Emerg Infect Dis 2020;26(2):255–65.
[55] Lednicky JA, et al. Independent infections of porcine deltacoronavirus among Haitian children. Nature 2021;600(7887):133–7.
[56] Everest H, et al. Known cellular and receptor interactions of animal and human coronaviruses: a review. Viruses 2022;14. https://doi.org/10.3390/v14020351.
[57] Belouzard S, et al. Mechanisms of coronavirus cell entry mediated by the viral spike protein. Viruses 2012;4(6):1011–33.
[58] Hulswit RJG, de Haan CAM, Bosch BJ. Chapter two - coronavirus spike protein and tropism changes. In: Ziebuhr J, editor. Advances in Virus Research. Academic Press; 2016. p. 29–57.
[59] Delmas B, et al. Aminopeptidase N is a major receptor for the entero-pathogenic coronavirus TGEV. Nature 1992;357(6377):417–20.
[60] Li W, et al. Broad receptor engagement of an emerging global coronavirus may potentiate its diverse cross-species transmissibility. Proc Natl Acad Sci USA 2018; 115(22). p. E5135-e5143.
[61] Pan Y, et al. Discovery of a novel swine enteric alphacoronavirus (SeACoV) in southern China. Vet Microbiol 2017;211:15–21.
[62] Raj VS, et al. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. Nature 2013;495(7440):251–4.
[63] Yang Y, et al. Receptor usage and cell entry of bat coronavirus HKU4 provide insight into bat-to-human transmission of MERS coronavirus. Proc Natl Acad Sci USA 2014;111(34):12516–21.
[64] Song W, et al. Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. PLoS Pathog 2018;14(8):e1007236.
[65] Wang Q, et al. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. Cell 2020;181(4):894–904.e9.
[66] Wu K, et al. Crystal structure of NL63 respiratory coronavirus receptor-binding domain complexed with its human receptor. Proc Natl Acad Sci USA 2009;106 (47):19970–4.
[67] Salamanna F, et al. Body localization of ACE-2: on the trail of the keyhole of SARS-CoV-2. Front Med 2020;7:594495.
[68] Meyerholz DK, Lambertz AM, McCray Jr PB. Dipeptidyl peptidase 4 distribution in the human respiratory tract: implications for the middle east respiratory syndrome. Am J Pathol 2016;186(1):78–86.
[69] van der Velden VH, et al. Expression of aminopeptidase N and dipeptidyl peptidase IV in the healthy and asthmatic bronchus. Clin Exp Allergy 1998;28(1): 110–20.
[70] Chan RW, et al. Tropism of and innate immune responses to the novel human betacoronavirus lineage C virus in human ex vivo respiratory organ cultures. J Virol 2013;87(12):6604–14.
[71] Petrovsky N. SARS coronavirus infections of the lower respiratory tract and their prevention. Microbiol Respir Syst Infect 2016:45–53. https://doi.org/10.1016/B978-0-12-804543-5.00003-8. Epub 2016 Aug 5.
[72] Murphy K. SARS CoV-2 detection from upper and lower respiratory tract specimens: diagnostic and infection control implications. Chest 2020;158(5): 1804–5.
[73] Liu DX, Liang JQ, Fung TS. Human coronavirus-229E, -OC43, -NL63, and -HKU1 (Coronaviridae). Encycl Virol 2021:428–40. https://doi.org/10.1016/B978-0-12-809633-8.21501-X. Epub 2021 Mar 1.
[74] Petrosillo N, et al. COVID-19, SARS and MERS: are they closely related? Clin Microbiol Infect 2020;26(6):729–34.
[75] Abdelrahman Z, Li M, Wang X. Comparative review of SARS-CoV-2, SARS-CoV, MERS-CoV, and influenza a respiratory viruses. Front Immunol 2020;11:552909.
[76] Guesdon J-L, Avrameas S. Solid phase enzyme immunoassays. In: Wingard LB, Katchalski-katzir E, Goldstein L, editors. in *Applied Biochemistry and Bioengineering*. Elsevier; 1981. p. 207–32.
[77] Shang J, et al. Cryo-electron microscopy structure of porcine deltacoronavirus spike protein in the prefusion state. J Virol 2018;92(4).
[78] Xiong Q, et al. Close relatives of MERS-CoV in bats use ACE2 as their functional receptors. Nature 2022;612(7941):748–57.
[79] Tabatabaei MS, Ahmed M. Enzyme-linked immunosorbent assay (ELISA). Methods Mol Biol 2022;2508:115–34.
[80] Halat MJ, Gravens E. Screening for antibody specificity and sensitivity with ELISA. Methods Mol Biol 2023;2612:45–58.
[81] Hayrapetyan H, et al. Enzyme-linked immunosorbent assay: types and applications. Methods Mol Biol 2023;2612:1–17.
[82] Pradhan M, et al. COVID-19: clinical presentation and detection methods. J Immunoass Immunochem 2022;43(1):1951291.
[83] Wong SK, et al. A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. J Biol Chem 2004;279(5): 3197–201.

[84] Yang ZY, et al. Evasion of antibody neutralization in emerging severe acute respiratory syndrome coronaviruses. Proc Natl Acad Sci USA 2005;102(3):797–801.
[85] Li JY, et al. The 442th amino acid residue of the spike protein is critical for the adaptation to bat hosts for SARS-related coronaviruses. Virus Res 2021;295:198307.
[86] Menachery VD, et al. A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. Nat Med 2015;21(12):1508–13.
[87] Hu B, et al. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. PLoS Pathog 2017;13(11):e1006698.
[88] Yang XL, et al. Isolation and Characterization of a Novel Bat Coronavirus Closely Related to the Direct Progenitor of Severe Acute Respiratory Syndrome Coronavirus. J Virol 2015;90(6):3253–6.
[89] Li F. Receptor recognition and cross-species infections of SARS coronavirus. Antivir Res 2013;100(1):246–54.
[90] Zech F, et al. Spike residue 403 affects binding of coronavirus spikes to human ACE2. Nat Commun 2021;12(1):6855.
[91] Letko M, Marzi A, Munster V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. Nat Microbiol 2020;5(4):562–9.
[92] Hunter SA, Cochran JR. Cell-binding assays for determining the affinity of protein-protein interactions. *Technol Consid* Methods Enzymol 2016;580:21–44.
[93] Chen M, Zhang XE. Construction and applications of SARS-CoV-2 pseudoviruses: a mini review. Int J Biol Sci 2021;17(6):1574–80.
[94] Nie J, et al. Establishment and validation of a pseudovirus neutralization assay for SARS-CoV-2. Emerg Microbes Infect 2020;9(1):680–6.
[95] Donofrio G, et al. A Simplified SARS-CoV-2 Pseudovirus Neutralization Assay. Vaccines 2021;9. https://doi.org/10.3390/vaccines9040389.
[96] Lin HH, Huang LM, Wu SC. A quantitative luciferase-based cell-cell fusion assay to measure four-serotype dengue virus E protein-triggered membrane fusion. Hum Vaccin Immunother 2020;16(9):2176–82.
[97] Bian J, Li Z. Angiotensin-converting enzyme 2 (ACE2): SARS-CoV-2 receptor and RAS modulator. Acta Pharm Sin B 2021;11(1):1–12.
[98] Nie J, et al. Establishment and validation of a pseudovirus neutralization assay for SARS-CoV-2. Emerg Microbes Infect 2020;9(1):680–6.
[99] Douzi B. Protein-protein interactions: surface plasmon resonance. Methods Mol Biol 2017;1615:257–75.
[100] Sparks RP, Jenkins JL, Fratti R. Use of Surface Plasmon Resonance (SPR) to Determine Binding Affinities and Kinetic Parameters Between Components Important in Fusion Machinery. Methods Mol Biol 2019;1860:199–210.
[101] Li W, et al. Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. Embo J 2005;24(8):1634–43.
[102] Ren W, et al. Susceptibilities of Human ACE2 Genetic Variants in Coronavirus Infection. J Virol 2022;96(1):e0149221.
[103] Hatmal MMM, et al. Comprehensive Structural and Molecular Comparison of Spike Proteins of SARS-CoV-2, SARS-CoV and MERS-CoV, and Their Interactions with ACE2. Cells 2020;9. https://doi.org/10.3390/cells9122638.
[104] Lu G, et al. Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26. Nature 2013;500(7461):227–31.
[105] Wrapp D, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science 2020;367(6483):1260–3.
[106] Zhang S, et al. Bat and pangolin coronavirus spike glycoprotein structures provide insights into SARS-CoV-2 evolution. Nat Commun 2021;12(1):1607.
[107] Lan J, et al. Structural insights into the binding of SARS-CoV-2, SARS-CoV, and hCoV-NL63 spike receptor-binding domain to horse ACE2. Structure 2022;30(10):1432–1442.e4.
[108] Drescher DG, Ramakrishnan NA, Drescher MJ. Surface plasmon resonance (SPR) analysis of binding interactions of proteins in inner-ear sensory epithelia. Methods Mol Biol 2009;493:323–43.
[109] Soltermann F, Struwe WB, Kukura P. Label-free methods for optical in vitro characterization of protein–protein interactions. Phys Chem Chem Phys 2021;23(31):16488–500.
[110] Yi C, et al. Key residues of the receptor binding motif in the spike protein of SARS-CoV-2 that interact with ACE2 and neutralizing antibodies. Cell Mol Immunol 2020;17(6):621–30.
[111] Forssén P, et al. Advanced Analysis of Biosensor Data for SARS-CoV-2 RBD and ACE2 Interactions. Anal Chem 2020;92(17):11520–4.
[112] Augusto G, et al. In vitro data suggest that Indian delta variant B.1.617 of SARS-CoV-2 escapes neutralization by both receptor affinity and immune evasion. Allergy 2022;77(1):111–7.
[113] Zhang D, et al. Identification of natural compounds as SARS-CoV-2 entry inhibitors by molecular docking-based virtual screening with bio-layer interferometry. Pharm Res 2021;172:105820.
[114] Coghi P, et al. Exploring SARS-CoV-2 Delta variant spike protein receptor-binding domain (RBD) as a target for tanshinones and antimalarials. Nat Prod Res 2022;36(23):6150–5.
[115] Kim S, et al. Differential Interactions between Human ACE2 and Spike RBD of SARS-CoV-2 Variants of Concern. J Chem Theory Comput 2021;17(12):7972–9.
[116] Brintnell E, Gupta M, Anderson DW. Phylogenetic and Ancestral Sequence Reconstruction of SARS-CoV-2 Reveals Latent Capacity to Bind Human ACE2 Receptor. J Mol Evol 2021;89(9-10):656–64.
[117] Chowdhury R, Boorla VS, Maranas CD. Computational biophysical characterization of the SARS-CoV-2 spike protein binding with the ACE2 receptor and implications for infectivity. Comput Struct Biotechnol J 2020;18:2573–82.
[118] Na EJ, et al. Comparative genetic analyses of Korean bat coronaviruses with SARS-CoV and the newly emerged SARS-CoV-2. J Vet Sci 2021;22(1):e12.
[119] Lai HTT, et al. A comparative study of receptor interactions between SARS-CoV and SARS-CoV-2 from molecular modeling. J Mol Model 2022;28(10):305.
[120] Cavani M, Riofrío WA, Arciniega M. Molecular Dynamics and MM-PBSA Analysis of the SARS-CoV-2 Gamma Variant in Complex with the hACE-2 Receptor. Molecules 2022;27(7).
[121] Sun H, et al. Systematic Tracing of Susceptible Animals to SARS-CoV-2 by a Bioinformatics Framework. Front Microbiol 2022;13:781770.
[122] Jafary F, Jafari S, Ganjalikhany MR. In silico investigation of critical binding pattern in SARS-CoV-2 spike protein with angiotensin-converting enzyme 2. Sci Rep 2021;11(1):6927.
[123] Wierbowski SD, et al. A 3D structural SARS-CoV-2–human interactome to explore genetic and drug perturbations. Nat Methods 2021;18(12):1477–88.
[124] Kiewhuo K, et al. Structure-function relationships among selected human coronaviruses. Indian J Biochem Biophys 2022;59:536–51.
[125] Singh AV, et al. Interfacial water in the SARS spike protein: investigating the interaction with human ACE2 receptor and in vitro uptake in A549 cells. Langmuir 2022;38(26):7976–88.
[126] Waterhouse A, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 2018;46(W1). W296-w303.
[127] Eswar, N., et al., Comparative protein structure modeling using Modeller. Curr Protoc Bioinformatics, 2006. Chapter 5: p. Unit-5.6.
[128] Jumper J, et al. Highly accurate protein structure prediction with AlphaFold. Nature 2021;596(7873):583–9.
[129] Rives A, et al. Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. Proc Natl Acad Sci USA 2021;118(15).e2016239118.
[130] Raveh B, et al. Rosetta FlexPepDock ab-initio: simultaneous folding, docking and refinement of peptides onto their receptors. PLoS One 2011;6(4):e18934.
[131] Dominguez C, Boelens R, Bonvin AMJJ. HADDOCK: A Protein−Protein Docking Approach Based on Biochemical or Biophysical Information. J Am Chem Soc 2003;125(7):1731–7.
[132] Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opin Drug Discov 2015;10(5):449–61.
[133] Anandakrishnan R, et al. Speed of Conformational Change: Comparing Explicit and Implicit Solvent Molecular Dynamics Simulations. Biophys J 2015;108(5):1153–64.
[134] Case DA, et al. The Amber biomolecular simulation programs. J Comput Chem 2005;26(16):1668–88.
[135] Páll S, et al. Heterogeneous parallelization and acceleration of molecular dynamics simulations in GROMACS. J Chem Phys 2020;153(13):134110.
[136] Bernardi RC, Melo MCR, Schulten K. Enhanced sampling techniques in molecular dynamics simulations of biological systems. Biochim Et Biophys Acta (BBA) - Gen Subj 2015;1850(5):872–7.
[137] Plimpton S. Computational limits of classical molecular dynamics simulations. Comput Mater Sci 1995;4(4):361–4.
[138] Alford RF, et al. The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. J Chem Theory Comput 2017;13(6):3031–48.
[139] Murdocca M, et al. Peptide Platform as a Powerful Tool in the Fight against COVID-19. Viruses 2021;13(8).
[140] Bastos RS, et al. Design and identification of inhibitors for the spike-ACE2 target of SARS-CoV-2. Int J Mol Sci 2023;24(10).
[141] Chen P, et al. Molecular dynamic simulation analysis of SARS-CoV-2 spike mutations and evaluation of ACE2 from pets and wild animals for infection risk. Comput Biol Chem 2022;96:107613.
[142] Pantsar T, Poso A. Binding affinity via docking: fact and fiction. Molecules 2018;23(8).
[143] Guo H, et al. ACE2-independent bat sarbecovirus entry and replication in human and bat cells. mBio 2022;13(6):e0256622.
[144] Starr TN, et al. ACE2 binding is an ancestral and evolvable trait of sarbecoviruses. Nature 2022;603(7903):913–8.
[145] Wu L, et al. SARS-CoV-2 Omicron RBD shows weaker binding affinity than the currently dominant Delta variant to human ACE2. Signal Transduct Target Ther 2022;7(1):8.
[146] Dos Santos VP, et al. E-Volve: understanding the impact of mutations in SARS-CoV-2 variants spike protein on antibodies and ACE2 affinity through patterns of chemical interactions at protein interfaces. PeerJ 2022;10:e13099.
[147] Xue T, et al. Single point mutations can potentially enhance infectivity of SARS-CoV-2 revealed by in silico affinity maturation and SPR assay. RSC Adv 2021;11(24):14737–45.
[148] Calcagnile M, et al. Molecular docking simulation reveals ACE2 polymorphisms that may increase the affinity of ACE2 with the SARS-CoV-2 Spike protein. Biochimie 2021;180:143–8.
[149] Bakhshandeh B, et al. Variants in ACE2; potential influences on virus infection and COVID-19 severity. Infect Genet Evol 2021;90:104773.
[150] Huang C, et al. The effect of N-glycosylation of SARS-CoV-2 spike protein on the virus interaction with the host cell ACE2 receptor. iScience 2021;24(11):103272.
[151] Casalino L, et al. Beyond shielding: the roles of glycans in the SARS-CoV-2 spike protein. ACS Cent Sci 2020;6(10):1722–34.
[152] Aloor A, et al. Glycosylation in SARS-CoV-2 variants: a path to infection and recovery. Biochem Pharm 2022;206:115335.
[153] Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. Science 2015;349(6245):255–60.
[154] Pavlova A, et al. Machine learning reveals the critical interactions for SARS-CoV-2 spike protein binding to ACE2. J Phys Chem Lett 2021;12(23):5494–502.

[155] Ali F, Kasry A, Amin M. The new SARS-CoV-2 strain shows a stronger binding affinity to ACE2 due to N501Y mutant. Med Drug Discov 2021;10:100086.
[156] Mowbray M, et al. Machine learning for biochemical engineering: a review. Biochem Eng J 2021;172:108054.
[157] Wei H, Li X. Deep mutational scanning: a versatile tool in systematically mapping genotypes to phenotypes. Front Genet 2023;14:1087267.
[158] Starr TN, et al. Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. Cell 2020;182(5):1295–310. e20.
[159] Chen C, et al. Computational prediction of the effect of amino acid changes on the binding affinity between SARS-CoV-2 spike RBD and human ACE2. Proc Natl Acad Sci USA 2021;118(42). e2106480118.
[160] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521(7553):436–44.
[161] Xie P, et al. Emvirus: An embedding-based neural framework for human-virus protein-protein interactions prediction. Biosaf Health 2023;5(3):152–8.
[162] Croitoru FA, et al. Diffusion models in vision: a survey. IEEE Trans Pattern Anal Mach Intell 2023;45(9):10850–69.
[163] Ofer D, Brandes N, Linial M. The language of proteins: NLP, machine learning & protein sequences. Comput Struct Biotechnol J 2021;19:1750–8.
[164] Hamamsy T, et al. Protein remote homology detection and structural alignment using deep learning. Nat Biotechnol 2023.
[165] McNutt AT, et al. GNINA 1.0: molecular docking with deep learning. J Chemin-2021;13(1):43.