Preprints.org

Article

Identification of Some Potential Cellular Receptors and Host Enzymes That Could Potentially Refine the Bovine Coronavirus (BCoV) Replication and Tissue Tropism. A Molecular Docking Study.

Mohd Yasir Khan , Nithyadevi Duraisamy , Abid Ullah Shah , Reda Nacif ElAlaoui , Mohammed Cherkaoui , Maged Gomaa Hemida *

Article

Identification of Some Potential Cellular Receptors and Host Enzymes That Could Potentially Refine the Bovine Coronavirus (BCoV) Replication and Tissue Tropism: A Molecular Docking Study

Mohd Yasir Khan [1], Nithyadevi Duraisamy [1], Abid Ullah Shah [2], Reda Nacif ElAlaoui [1], Mohammed Cherkaoui [1] and Maged Gomaa Hemida [2,*]

[1] Department of Computer Science, College of Digital Engineering and Artificial Intelligence, Long Island University, Brooklyn; mohd.yasirkhan@liu.edu (M.Y.K.), Nithyadevi.Duraisamy@liu.edu (N.D.), Reda.NacifElAlaoui@liu.edu (R.N.E.), mohammed.cherkaoui@liu.edu (M.C.)

[2] Department of Veterinary Biomedical Sciences, College of Veterinary Medicine, Long Island University, 720 Northern Boulevard, Brookville, NY, 11548, USA; abidullah.shah@liu.edu (A.S.)

* Correspondence: maged.hemida@liu.edu; Tel.: +1516-2993650

Abstract: Bovine coronavirus (BCoV) is one of the most common pathogens affecting cattle of all ages. BCoV possesses multiple tissue tropism in cattle, particularly in the respiratory and enteric tracts of the affected animals. Viral entry is a crucial step in the replication cycle of all viruses, including BCoV. The process of viral entry requires the orchestration between several proteins from the virus and the host side. The coronavirus spike glycoproteins contributed substantially to other viral infections and the molecular pathogenesis of coronaviruses, including BCoV. Meanwhile, among all coronaviruses, the hemagglutinin esterase enzyme is found in only BCoV and the human coronavirus-OC43 (HCoV-OC43). The viral receptors are among the key players in the viral entry and could explain the viral tropism, at least in part. The presence of some host cell proteases also enhances the virus replication from the entry until the release of the virus from the host cells. Previous studies have shown that the 9- O-acetylated sialic acids could potentially act as receptors for the BCoV. However, little is known about the BCoV receptors and if other co/receptors might be needed to ensure viral entry to the host cells. In addition to that, the roles of other host cell proteases, such as Furin and TMPRSS2, in BCoV infection have not yet been well studied. The main objectives of the current study are to identify some novel receptors for BCoV and some host cell proteases that might enhance virus replication. We used some in silico docking tools to predict some new BCoV receptors out of some known other coronavirus receptors (ACE-2, NRP1, DPP4, APN, and CEACAM1). Our results postulated the potential actions of the host Furin and TMPRRS2 enzymes in the activation of BCoV-S glycoprotein, which potentially enhances BCoV replication. Our modeling confirms the high binding affinities of both the 9- O-acetylated sialic and the NRP1 with the BCoV-S and BCoV-HE compared to the ACE-2, DPP4, and CEACAM-1, suggesting their roles as receptors for BCoV. Our prediction models show the potential bindings of the Furin and TMPRSS2 to the BCoV-S polybasic cleavage site (RRSRR↓A). To our knowledge, this is the first study to investigate the bovine coronavirus interaction with some other potential host receptors, particularly ACE-2, NRP1, DPP4, APN and CEACAM1, TMPRSS2, and the bovine Furin protease.

Keywords: BCoV; spike glycoprotein; Hemagglutinin esterase; ACE-2; NPR-1; Furin; TMPRRS2; docking; Neu5,9Ac2; homology modeling; virus/host interaction

1. Introduction

The virus replication occurs in many consecutive steps to generate many mature viral particles upon the completion of the viral replication in the permissive cells. The virus always hijacks the cellular machinery and directs them to synthesize their viral proteins instead of the cellular proteins [1,2]. The viral entry is a crucial step in any viral infection [3]. The viral entry requires the interaction between many proteins from the viral and host sides [1]. The presence of viral-specific receptors is considered an important factor in the process of viral entry into the host cells. The host cell is called permissive for certain viruses if they express the specific viral receptors and provide the suitable environments for the virus replication, including any auxiliary receptors, transcription, and translation factors. Coronaviruses are enveloped viruses containing positive sense RNA genomes and belong to the order Nidovirales and are classified into four genera (α, β, λ, and δ) [4]. The genus β-coronavirus includes five important human coronaviruses the sever acute respiratory syndrome coronavirus-1 (SARS-CoV-1), the Middel East respiratory syndrome coronavirus (MERS-CoV) the SARS-CoV-2, and the human coronavirus-OC43 (HCoV-OC-43). This genus includes some other important viruses affecting animals, particularly the BCoV and the equine coronavirus (ECoV) [4]. The coronavirus's genome size ranges from 27-31 kb in length and has a unique organization. The full-length genome is flanked with two untranslated regions at the 5' and 3' ends. Coronaviruses are characterized by the production of a set of sub-genomic messenger RNA (mRNA) at their 3' end [5]. The 5' end of the genome of most coronaviruses contains a large gene called gene-1, which consists of two overlapping open reading frames (ORFs) with a ribosomal frameshifting between those two ORFs. However, the 3' end of the genome is mainly occupied by the common structural proteins interspersed with some small accessory proteins. There are four major structural proteins in most coronaviruses, including the spike glycoprotein (S), the envelope (E), the membrane (M), and the nucleocapsid protein (N). Some members of the genus β-coronavirus, including BCoV and HCoV-OC43, have an additional structural protein called hemagglutinin esterase (HE); thus, their genome is a little larger in size compared to other coronaviruses (31 Kb) [6]. The S glycoprotein is a key player in all coronavirus replication. There are several proteins, including some cellular receptors, co-receptors, and cellular enzymes are involved in the BCoV/host interaction. The BCoV spike (BCoV/S) and BCoV-HE proteins are important in virus replication and pathogenesis (6). BCoV-S has potential bidding to the 5-N-acetyl-9-O-acetylneuraminic acid, suggesting their possible roles as BCoV receptors (7). On the other hand, the BCoV-HE acts as a receptor-destroying enzyme during BCoV replication [7]. However, there is a lack of comprehensive understanding of the interplay of the BCoV-S/BCoV-HE and the cellular receptors during BCoV replication. The availability of specific receptors is one of the main factors that make the target cells permissive to coronavirus (CoVs) infection. Each group of CoVs recognizes certain types of receptors and may require the presence of additional auxiliary receptors to facilitate virus attachment and downstream replication. SARS-CoV-2 uses the angiotensinogen-converting enzyme-2 (ACE-2) as the main receptor, and the chaperone GRP78 acts as an auxiliary receptor [8,9]. It was also shown that MERS-CoV utilize the dipeptidyl peptidase-4 (DPP4) as receptors in humans and dromedary camels [10,11]. The amino-peptidase N (APN) also called cluster of differentiation -13 (CD13) act as a major receptor for the transmissible gastroenteritis virus that mainly cause enteric infections in pigs [12]. The carcinoembryonic antigen cell adhesion molecule 1 (CEACAM-1) also acts as a major receptor to another coronavirus called murine hepatitis virus (MHV) [13]. Although the presence of the CoVs receptors and co/receptors is important for the success of viral replication, most coronaviruses require the presence of host cell enzymes that help in the cleavage of the CoV-S and CoV-HE proteins to initiate the process of viral infection. Usually, these host cell enzymes are enriched at the portal of entry of most coronaviruses, particularly the mucosal surfaces of the respiratory and enteric tracts of the affected hosts. It has been recently shown that TMPRSS2 (transmembrane protease serine 2) is a serine protease that plays an important role in SARS-CoV-2 replication, particularly during viral entry to the host cells. TMPRSS2 usually cleaves the viral spike glycoprotein, which activates the virus and facilitates its entrance to the host cells [14]. Furin is another host cell enzyme that is considered to be a subtilisin-like proprotein convertase. Furin usually cleaves the target proteins at a polybasic amino acid sequence R-X-(K/R)-R

(where R is arginine, K is lysine, and X can be any amino acid). The Furin cleavage to the SARS-CoV-2-S protein enhances the pathogenicity transmissibility and increases the virus infectivity in the target host [15,16]. Although the roles of the above-mentioned receptors and enzymes were intensively studied in SARS-CoV2, there is a lack of knowledge about the roles of these receptors and enzymes in BCoV infection and replication. The main goals of the current study are to use the in-silico prediction and docking tools to study the roles of these proteins and enzymes in BCoV replication. This study shed light on some unknown aspects of BCoV tissue tropism and pathogenesis. It will also pave the way for the development of some novel vaccines and antiviral therapies for BCoV infection in cattle.

Homology modeling is a well-established method that has been shown to produce quite accurate models for a protein sequence if an X-ray structure of a protein with a sufficient degree of sequence similarity is available [17]. The method is based on the fact that the structural conformation of a protein is more highly conserved than its amino acid sequence and that small or medium changes in sequence normally result in little variation in the 3D structures [18]. The quality of the model is directly linked to the identity between template and target sequences. As a rule, models built with over 50% sequence similarities are accurate enough for drug discovery applications [19].

2. Materials and Methods

2.1. BCoV Spike and Host Receptor Protein Structure Prediction

Homology modeling is a method to build predicted homology structures of proteins based on 3D confirmations of most identical template proteins. MODELLER tool of Biovia Discovery Studio v22.1.021297 can be used to predict comparative modeling with the help of the build homology model method, and it was employed for the 3D model structure confirmation prediction of our query protein sequences. The query protein sequences, including BCoV-Spike protein (BCoV-S) and the bovine ACE2, Furin, TMPRSS2, NRP1, DPP4, APN, and CEACAM-1 can be retrieved and downloaded in FASTA format. The FASTA file formats contain protein sequences, which can be retrieved and downloaded from NCBI (SciENcv - Home (nih.gov) and Uniprot databases (https://www.uniprot.org/).

The protein sequence of BCoV-spike (BCoV-S) glycoprotein (references strain; Mebus strain (ID: P15777) was downloaded from the UniProt database. The query sequences of other bovine host proteins including the ACE2 (ID: Q58DD0), Furin (ID: Q28193), the TMPRSS2 (ID: A2VDV7), the NRP1 (ID: E1BMX5), DPP4 (ID: P81425), the APN (ID: P79098) and the CEACAM-1 (ID: Q6VAN8) were retrieved from UniProt database. The homology modeling procedure requires the alignment of the query protein sequences with template protein sequences from the BLAST method (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC =blasthome). The query protein sequence aligned with 100 most identical template sequences with a maximum identity of 99% and minimum 30% in BLAST with BLOSSUM62 algorithm protocol. Five identical template sequences (identity: maximum 99-minimum 65%) were chosen to be aligned by load structure and alignment method. Further, the 3D-build homology model protocol was run to align template structures and predict query sequence 3D confirmation based on template proteins.

About (1-10) different homology models of query sequence can be generated with the help of this protocol. It gives you the liberty to choose the best model predicted. In this study, we predicted five different homology models for each query sequence. We used the **Discrete Optimized Protein Energy (DOPE) score** to assess the quality of the generated five different 3D protein models. The DOPE Score and PDF total energy score are for evaluating protein model accuracy. The more negative values of DOPE score and PDF total energy, the more stable and accurate the predicted protein model. The best model having the lowest DOPE score was selected for further *in silico* computational studies [20]. The correspondence homology model is then energy minimized (add hydrogen and CharMm force field) using the energy minimization method (full minimization). All the generated 3D structures of the all-atom models were verified through the Ramachandran plot and Verify 3D model tool of Biovia Discovery Studio v22.1.021297. Ramachandran plot suggested the stability of structure on the basis of amino acid residues located in the most favored, highly allowed, and allowed

region or site of the plot. If the verified score result from the Verify 3D method of the model protein is higher than the verified expected low Score value, then the model is of acceptable quality. The closer the verified score result is to the verified expected high Score value, the better the quality of the model

2.2. Mapping the BCoV-Spike Protein Domains

The InterPro is an international initiative that was conceived to streamline the efforts of the signature database providers (http://www.ebi.ac.uk/interpro/) and contains signatures diagnostic for protein families, domains, or functional sites. It has important tools for the computational functional classification of the newly determined sequences that lack biochemical characterization [21]. We used InterPro's family, domain, and functional site definitions in the annotation and functional Classification of uncharacterized sequences. The sequence-based searches are done using InterProScan (https://www.ebi.ac.uk/interpro/about/interproscan/) and the Prosite server (https://prosite.expasy.org/, Release 2024_03 of 29-May-2024), which combines the search methods from the member databases. The Web interface allows text-based and sequence-based searches using a sequence retrieval system (SRS) [22]. InterPro is accessible for interactive use via the EBI Web server (http://www.ebi.ac.uk/interpro), which can also be reached via each of the member databases [23].

2.3. Molecular Docking of Neu5,9Ac2 with BCoV/Spike Glycoprotein and the Hemagglutinin Esterase (HE).

The protein structure of the BCoV-S N-terminal domain (NTD) (residues 15–298), the BCoV-HE protein structure file was downloaded in the PDB file format from the Protein Data Bank (PDB) (https://www.rcsb.org/). In the molecular docking process of BCoV protein with the ligand, the N-terminal domain (NTD) of BCoV spike protein (PDB ID:4h14) was used. To confirm the structure similarity of NTD-from BCoV homology modeled and BCoV/NTD, we analyzed/compared the amino acid sequences from the PDB database (PDB ID: 4H14). First, protein sequences were aligned, and then the homology model as template and 4H14 as target sequence's projected 3D-structure with the alignment calculated using PyMOL molecular viewer yielded an RMSD value of 0.17. The ligand molecule, sugar molecule N-acetyl-9-O-acetylneuraminic acid (Neu5,9Ac2), the most common type of sialic acid, generally acts as the terminal sugar in cell surface glycans and polysaccharides, was taken in 3D-SDF file format from PubChem database (https://pubchem.ncbi.nlm.nih.gov/).

The proteins and ligand molecules were prepared using the built-in protein and ligand preparation methods using the Biovia Discovery studio v22.1.021297. This process typically involves removing the water molecules, adding hydrogens, and energy minimization to ensure accurate molecular docking. Both protein structures and ligand structures are manually examined and reverted to pre-reaction form. Grids with a grid space of 0.5 Å represent the receptor structure implicitly. In the CDOCKER tool from Biovia Discovery Studio docking protocol, we typically use ten docking poses. These computational tools enable the visualization of the ligand-target interaction (molecular docking) and the Identification of the compounds that bind more efficiently with the target (Binding energy) [24]. The docking output files were generated and analyzed to identify the most energetically favorable binding poses of the Neu5,9Ac2 with the BCoV/S glycoprotein. This analysis typically involves the examination of the docking scores, ligand-protein interactions, and the visualization of the docked complexes. To observe the exact binding affinity of the ligand to the protein, the calculation of binding energy is likely an internal step. The binding energy calculations were performed using the Biovia Discovery Studio to define the binding energy of ligands and proteins [25].

2.4. Protein-Protein Interaction Analysis-Molecular Docking

The protein-protein molecular docking was performed using the ZDOCK docking tool available in BIOVIA, Discovery studio v22.1.021297. The ZDOCK method uses a fast Fourier transform to perform a 3D search of the spatial degrees of freedom between two proteins. The molecular docking approaches, which take two (or more) structures as input and predict the structure of their complex, are increasingly being used for this purpose [26]. In each run of the ZDOCK, we used the default

angular sampling (3,600 ligand rotations) and a single 2.8 GHz 64-bit Opteron processor with 64 GB available system RAM. The protocol ran for 2000 different interaction poses of protein-protein interaction with 60 clusters. The best pose from ZDock output files can be selected based on your binding site and largest cluster choice. The resulting poses from the ZDOCK docking method must be further refined with the Energy refinement method (E_RDock). Therefore, the ZDOCK docking result poses were further refined for the best pose with minimum or lowest docking interaction energy/binding affinity with the refinement method, E_R Dock. This server refined protein-protein docking out of 2000 poses of ZDOCK with multiple conformations and calculates the weighted energies with the scoring function of E_RDock [E_RDock = E_Sol + beta + E_elec2] for every binding conformation, where E_RDock denotes the scoring function of RDOCK, E_vdw1 and E_vdw2 denotes van der Waals non-bond interaction energy of the protein complex after the first and second CHARMm minimization, E_elec1 and E_elec2 is an electrostatic energy term of the protein complex after the first and second CHARMm minimization, and E_Sol represents desolvation energy of the protein complex calculated by ACE method [26]. For protein-protein complex, the total desolvation score is simply the sum of the ACE scores of all receptor-ligand atom pairs within a distance cutoff of 6 Å. The protein-protein interactions interface was visualized by using the Biovia Discovery studio.

3. Results

3.1. The Homology Modelling for the BCoV-Spike and Receptor Proteins

To perform the homology modeling, we searched for the experimentally determined template reference sequences closely related to the query sequence and shared at least 65% sequence identity. The BCoV spike protein sequence (Strain: Mebus) of 1363 amino acids (a.a.) in length. The sequences of the target molecules from bovine as a receptor are ACE2 (804 a.a.), the host cell Furin (797 a.a.), the TMPRSS2 (490 a.a.), the NRP1 (924 a.a.), the DPP4 (765 a.a.), the APN (965 a.a.) and the CEACAM-1 (436 a.a.). Generally, the homology models, with lower DOPE scores and lower PDF total energy values, are considered more structurally stable and reliable. However, if the total energy values for multiple models are very similar, the DOPE score can be used as the deciding factor based on its statistical potential. Therefore, out of five, we selected the two best models with the PDF total energy values and lowest DOPE score for each protein homology model (Table S1). Out of these five models, based on the Lowest DOPE score and PDF total energy, one of the best models that have been further considered for other parameters, such as Ramachandran plot [27] and VERIFY 3D (Table 1), were used to assess the protein 3D mode**l'**s structural integrity. The best BCoV spike protein model has a DOPE score of −137057.54, and a PDF total energy value of 51589.75 (model 1) was chosen in the case of the BCoV. The spike protein Ramachandran plot shows that the maximum of the total amino acid residues is in the most favored regions and then is in the additional most allowed areas. A small proportion of residues are in the generously allowed regions. Based on Ramachandran plot results, all the predicted models, including BCoV/S, ACE2, NRP1, TMPRSS2, Furin, CEACAM-1, DPP4, and APN, indicate a highly accurate homology model structures (Figure 1). The modeled proteins verify that the score is closer to the expected high score, showing that the predicted homology models are of good quality (Table 1).

Table 1. Homology Model Verification Using Verify-3D Server in BIOVIA Discovery Studio.

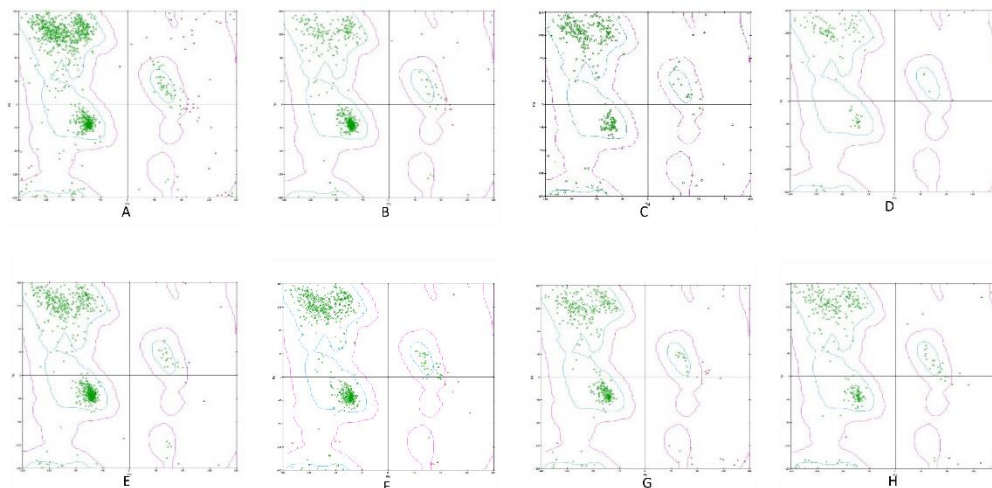| Homology model (low- DOPE Score)	Verify Score	Verify Expected High Score	Verify Expected Low Score
BCoV-Spike	476.141	561.414	252.636
ACE2	308.35	358.11	161.153
Furin	227.99	216.48	97.4159
TMPRSS2	142.08	152.573	68.6579
NRP1	125.32	159.001	71.5506
DPP4	308.11	333.627	150.132
APN	385.36	415.465	186.959
CEACAM-1	32.41	58.3052	26.2373

Figure 1. The validity and stability of the Homology modeled protein structure using the Ramachandran plot (A) BCoV-Spike (B) Bovine ACE2 (C) Bovine NRP1 (D) Bovine CEACAM-1 (E) Bovine APN (F) Bovine DPP4 (G) Bovine Furin (H) Bovine TMPRSS2.

3.2. BCoV Spike Protein Structure Domains

In the BCoV-spike-protein sequence, Identification of the conserved domain region was carried out through the PROSITE server and Interpro server. Based on the results of the PROSITE server and Interpro server, the BCoV spike glycoprotein has two chains: spike protein S1 and S2 chains. The chain S1 contains two domains: an N-terminal galectin-like domain (NTD) and a receptor-binding domain (S1 RBD), also referred to as the RBD site present in between the S1 C-terminal domain (CTD) (Figure 2). The NTD of S1 is of 15-298 amino acids. However, the S1-RBD site has 343-491 amino acid residues, which are present between the CTD of the S1 protein, and it is 329-617 amino acids long [28]. The amino acids from 629 to 1297 contain the S1/S2 cleavage region, the S2 fusion subunit of the spike (S) glycoprotein, and the HR1 and HR2 region from betacoronaviruses (BCoV). The protease action related to the S1/S2 cleavage region is between 758-797 amino acid residues of the BCoV spike protein structure (Figure 2).
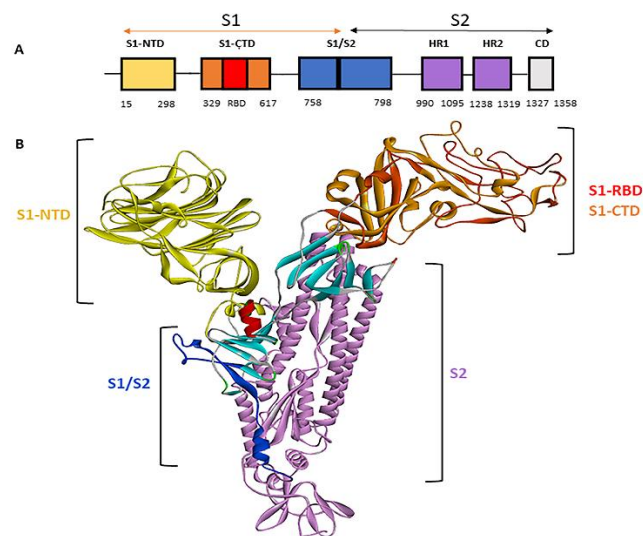

Figure 2. Schematic representation of the BCoV S Protein: S1 and S2 Chains (A) The listed domain boundaries are mostly defined as S-NTD, N-terminal Domain; RBD, Receptor Binding Domain; S1-CTD, C-Terminal domain; S2-HR1, Heptat Repeat 1; S2-HR2, Heptad Repeat 2; S2-CD, Cytoplasm

Domain. **(B) Schematic drawing of the three-dimensional Structure of BCoV Spike protein showing different domains on S1 and S2 Chains.**

3.3. Docking of the Bovine Neu5,9Ac2 with the BCoV-S Protein

It was previously reported that the Neu5,9Ac2 is a receptor for BCoV/S NTD [29,30]. The BCoV/NTD (4h14) shows sequence and structure similarity with the BCoV-isolate BCoV/NTD region (Figure 1S). To investigate the molecular interaction of the BCoV spike with the bovine cell surface receptor molecule, we performed molecular docking of the BCoV-NTD (PDB ID: 4H14) with molecule Neu5,9Ac2. The docking result gives ten poses of ligand-protein interaction complexes. Further, ligand-protein interaction binding energy (BE) calculation was done to calculate the binding energy of complexes. ΔG determines the binding energy and stability of any given protein-ligand complex or the binding affinity of a ligand to a protein. The binding energy calculation result of the ten best ligand-protein complexes indicated a strong binding affinity between the Neu5,9Ac2 molecule and the BCoV/S-NTD. The interaction between Neu5,9Ac2 and BCoV/S NTD with negative binding energy (-79.55 kcal/mol) is shown as the best ligand-protein interaction pose in Figure 3.

This interaction involves some key amino acid residues, including Asp-187, Gly-189, His-185, and Lys-196 (Figure 3). The docking results revealed that Neu5,9Ac2 interacts with a specific pocket above the β-sandwich core in the BCoV NTD.
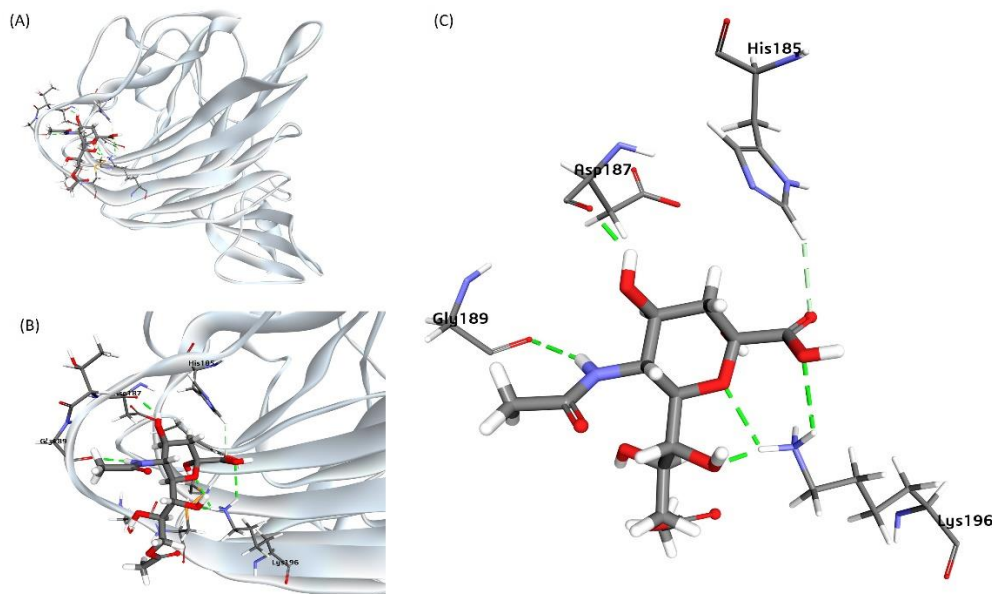

Figure 3. Sialic acid (Neu5,9Ac2) as the receptor for BCoV Spike protein at the NTD site. (A) Neu5,9Ac2 binding to the N-terminal domain (NTD) of the BCoV Spike protein is illustrated. (B) and (C) Detailed views of the interactions between Neu5,9Ac2 and the specific amino acid residues at the NTD site of the BCoV Spike protein: His185, Asp187, Gly189, and Lys196.

3.4. Interaction Interface between the BCoV Spike Glycoprotein and the Bovine ACE2

The energy minimization for the bovine ACE2 protein, the BCoV spike protein, and protein-protein dockings were performed on the ZDOCK server. The ACE2 interacted putatively with the BCoV-spike N terminal domain of the S1 protein based on the homology-modeled BCoV spike protein. The ACE2/BCoV spike protein-protein interaction interface residues are given in Table S2 and Figure 4. The total interaction energy between contact residues of bovine ACE2 and the BCoV-Spike were computed and reported in units of kcal/mol. The best docking pose shows a ZDOCK interaction score of 34.52 and an E_RDock score of -6.26 kcal/mol. The molecular interaction of the ACE2 shows interaction with BCoV/s at the receptor binding domain (RBD) of the C-terminal domain (CTD) of the S1 protein chain. ACE2 interaction residues are SER19, THR20, THR21, GLU23, GLN24,

THR27, GLU30, LYS81, THR 82 and other residues (Table S2). Spike S protein CTD-RBD site residues involved in interaction are TYR661, ASP662, SER663, GLY665, ASN666, SER611, THR425, ARG419, ALA527 and other residues (Table S2). The net interaction between ACE2 and the spike protein is attractive, as indicated by the negative binding energy of their interaction energy (RDOCK interaction energy) (Table 2). Specifically, the receptor-protein interaction analysis showed differences in the key residues at the interface between ACE2 and S1 protein CTD (Figure 4). This might be due to the local sequence difference of BCoV spike protein among different species, suggesting that bovine ACE2 is a putative receptor for BCoV-spike due to low binding energy (-6.26 kcal/mol) towards RBD or domain B of BCoV spike protein.
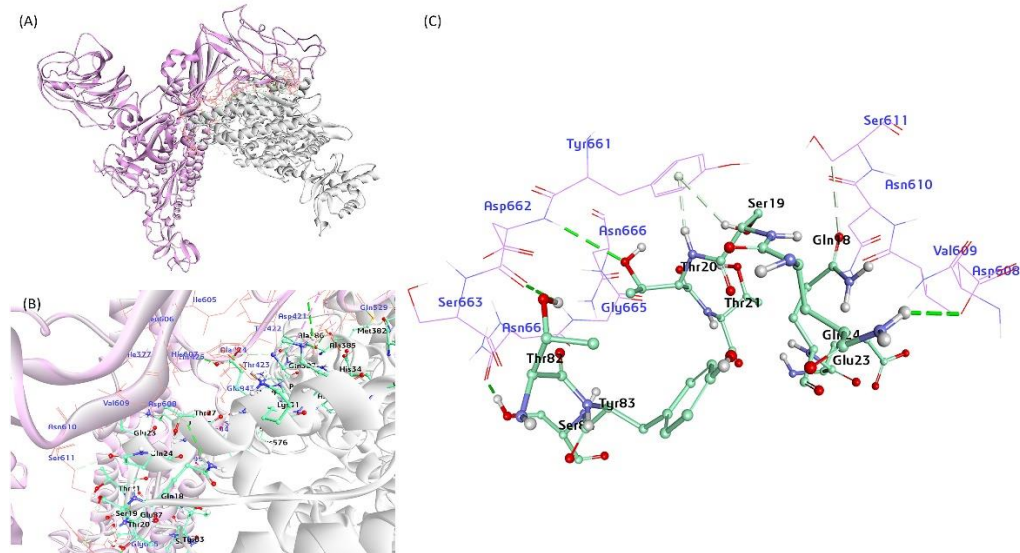

Figure 4. A proposed Model for the BCoV/S glycoprotein interaction with the bovine ACE2 protein. (A) The BCoV Spike protein interacting with ACE2, showing possible binding interactions at the receptor-binding domain (RBD) of the C-terminal domain (CTD) of the S1 protein chain. (B) and (C) Detailed views of the protein-protein interaction interface, highlighting the amino acid residues involved in the interaction between ACE2 and the N-terminal domain (NTD) of the BCoV Spike protein. ACE2 interaction residues: Ser19, Thr20, Thr21, Glu23, Gln24, Thr27, Glu30, Lys81, Thr82, and other residues. BCoV Spike NTD interaction residues: Tyr661, Asp662, Ser663, Gly665, Asn666, Ser611, Thr425, Arg419, Ala527, and other residues.

Table 2. Energetics and bonding interaction of the BCoV/ S glycoprotein with bovine receptor proteins.

Complex (Best Pose)	ZDock score	E_R Dock Score	Total hydrogen bonds	Total Pi bonds	Salt Bridge
BCoV/S-ACE2	22.58	-6.26	7	28	0
BCoV/S-NRP1	26	-21.22	29	9	0
BCoV/S-CEACAM-1	16.8	-6.29	12	9	1
BCoV/S-APN	17.6	-2.95	19	2	1
BCoV/S-DPP4	22.44	-6.23	20	9	1
BCoV/S-Furin	16.6	-13.14	30	1	0
BCoV/S-TMPRSS2	20	0.641	15	6	0

3.5. The Proposed Model for the Interactions between the Bovine NRP1 and the BCoV/S Glycoproteins

To test the potential roles of the bovine NRP1-as novel receptors for BCoV, the ZDock docking method of protein-protein interaction was performed to identify the interaction sites of the NRP1 with BCoV spike protein. The NRP1 potentially bids with both the NTD and RBD sites of BCoV-S glycoprotein. The interacting interface amino acid residues for both proteins are given in Table S2

and Figure 5. The NRP1 protein shows binding specificity with NTD and CTD site of BCoV/spike glycoprotein with residues Leu322, Leu333, val269, Arg307, Tyr138, Gln 266, and other amino acid residues (Table 4). The residues from BCoV spike NTD that interacted with NRP1 are Thr134, Leu194, Arg197, Phe232, Leu238, and other amino acid residues given in Table S2. The protein-protein docking between BCoV/S with the bovine NRP1 shows greater interaction, as suggested by the E_RDock score (-22.99 kcal/mol). The interaction affinity between both proteins is due to multiple non-covalent interactions (15 conventional hydrogen and 6 Pi bonds) (Figure 5 and Table 2). The binding of the NRP with the BCoV/S glycoprotein is highly stable, as suggested by the negative E_RDOCK binding energy values (-22.99 kcal/mol). Thus, the bovine NRP1 exhibits a higher binding affinity to the NTD and CTD or domain B of the BCoV/S glycoprotein.
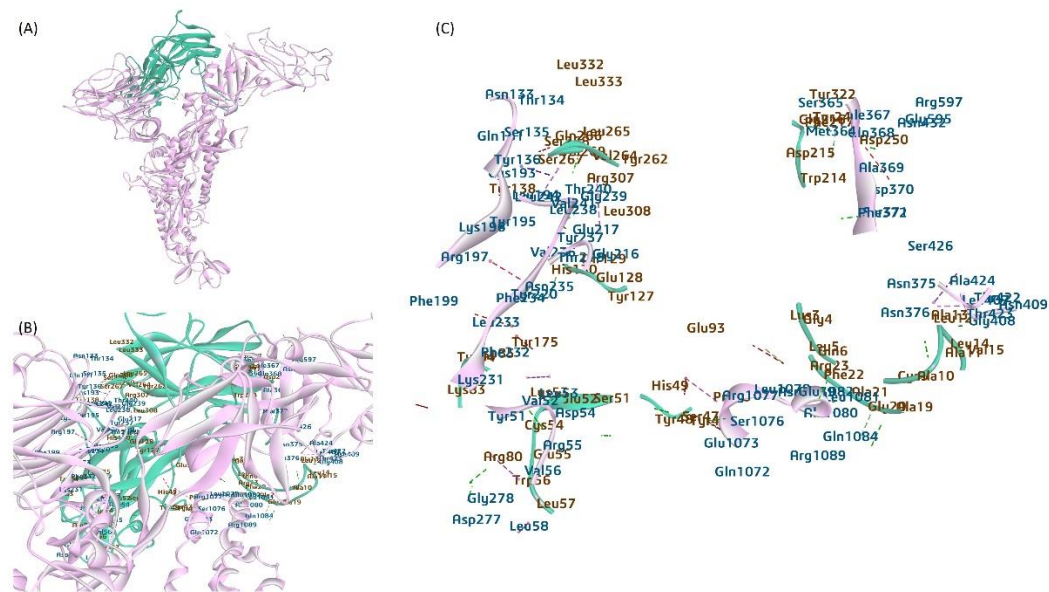

Figure 5. The proposed model for the interaction between the bovine NRP1 and the BCoV/S glycoprotein. (A) NRP1 interactions with the C-terminal domain (CTD) and N-terminal domain (NTD) of the S1 protein, illustrating possible binding conformations. **(B) and (C)** Specific NRP1 residues (Leu322, Leu333, Val269, Arg307, Tyr138, Gln266) and their binding affinities to amino acid residues in the NTD and CTD of the S1 protein.

3.6. BCoV Spike Interaction with the Bovine Type I Membrane Protein Receptor Carcinoembryonic Antigen-Related Cell Adhesion Molecule 1 (CEACAM-1)

To test the potential use of CEACAM-1 as a receptor for the BCoV, the ZDock docking method of protein-protein interaction was performed to map the interaction sites of the CEACAM-1 with BCoV/S glycoprotein. Our results show that the bovine CEACAM-1 protein binds with the BCOV-S/N-terminal domain (NTD). The interacting interface amino acid residues for both proteins are shown in (Figure 8 and Table 5). The CEACAM-1 protein shows a high binding specificity with the BCoV/S1-NTD site of Spike glycoprotein with residues ARG143, ASN170, THR171, ASN178, LYS196, LYS196, and other amino acid residues given in table. The residues from BCoV spike S1-NTD that interacted with CEACAM-1 are LEU233, PHE232, VAL229, Phe232, THR145, ASP148, and other amino acid residues shown in Table S2. The best dock pose shows a ZDOCK interaction score of 16.8, and E_R Dock interaction energy is -6.29 kcal/mol due to multiple non-covalent interactions, including nine conventional hydrogens bonds, 12 Pi bond, and one salt bridge interaction between the two proteins (Figure 6 and Table 2). Based on the E_R Dock Score of the interaction energy of the best pose result, the low binding affinity of BCOV/S glycoprotein with CEACAM-1 was suggested compared to the interaction of BCoV/NTD known receptors Neu5,9Ac2.
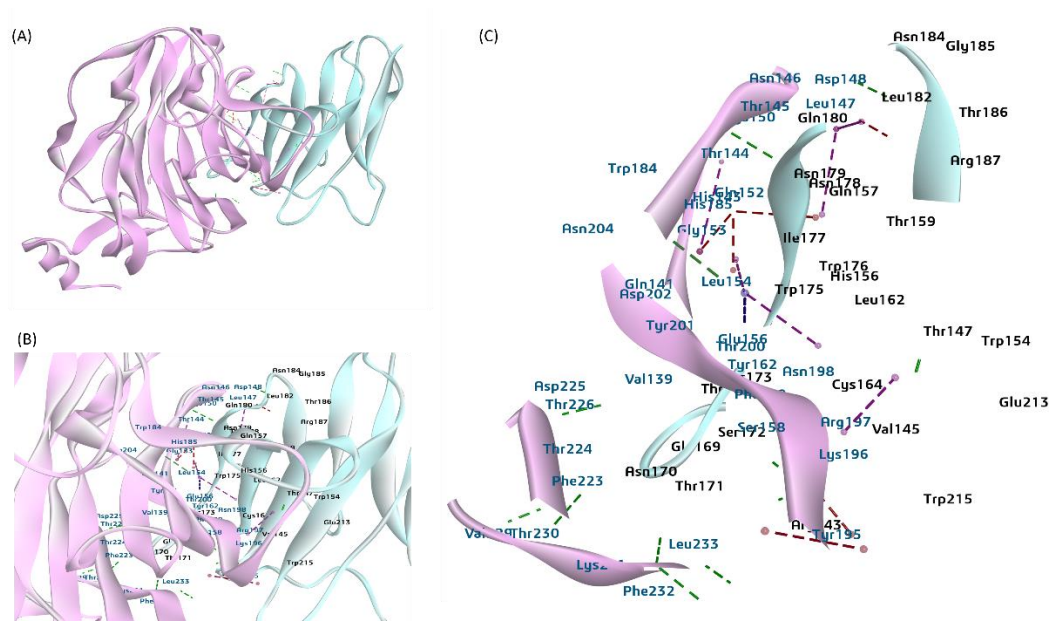
Figure 6. A putative model for the interaction between the bovine CEACAM-1 protein and the BCoV-S glycoprotein and mapping the interactive amino acid residues. (A) The bovine CEACAM-1 protein interacts with the BCoV/S glycoprotein NTD of the BCoV spike. (B) and (C) CEACAM-1 binds through LEU233, PHE232, VAL229, Phe232, THR145, ASP148 and other amino acids with N-terminal domain (NTD) of Spike glycoprotein with residues ARG143, ASN170, THR171, ASN178, LYS196, LYS196 with hydrogen and Pi bond interactions.

3.7. A Putative Model for the BCoV/S Glycoprotein Interaction with the bovine Aminopeptidase N (APN)

To explore the potential roles of the bovine APN as a receptor for the BCoV, the ZDock with RDOCK docking method of the protein-protein interaction was performed to map the BCoV spike protein interaction sites with the bovine APN protein. The interaction between the two proteins suggested that the APN showed binding with the RBD of the BCoV/S glycoprotein. The interacting interface amino acid residues for both proteins are shown in Figure 7. The APN protein is showing a binding specificity with RBD present in the BCoV/S1-CTD site of BCoV/S glycoprotein through the residues ASN397, TYR451, ARG514, LYS579, ASN583, CYS524, and other amino acid residues given in Table S2. The residues of BCoV/ S1-CTD interacted with the APN are CYS826, ASP849, GLU843, SER839, GLU843, SER839, ILE512, GLY511, PRO516, ASN495 and other amino acid residues given in Table 6. The best docking pose with low-energy binding conformation shows a ZDOCK interaction score 17.6 and an E_RDock score of -2.96 kcal/mol. The interaction between proteins is due to multiple non-covalent bonds, including 19 conventional hydrogen bonds, two Pi bonds, and one salt bridge interaction between the two proteins (Table 2).
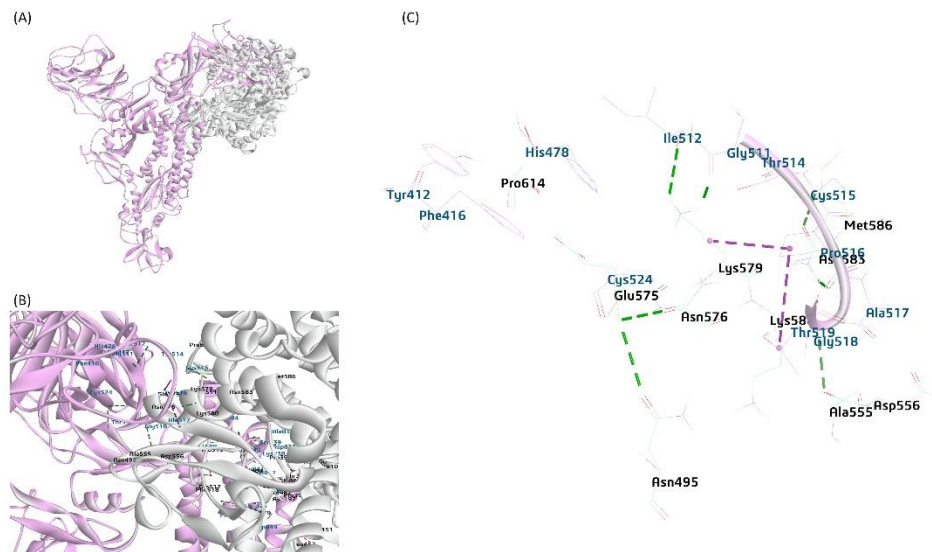
Figure 7. A putative model representing the interaction of BCoV/S glycoprotein and the bovine aminopeptidase N (APN) protein. (A) The APN interactions with BCoV/S1 RBD and CTD (B) (C) BCoV/S1 chain CTD residues binding to the bovine APN through the residues (TYR412, PHE416, HIS478, CYS524, ILE512, GLY511, THR514, CYS515, PRO516, THR518, ALA517 and GLY518 to amino acid residues ASN495, ALA555, ASP565, GLY575, ASP576, ARG614 and other amino acid residues of APN by hydrogen bond, Pi bond and salt bridge interactions).

3.8. A Putative Model for the Interaction between the BCoV/S Glycoprotein with The bovine Dipeptidyl Peptidase 4 (DPP4)

We used the ZDock docking method of the protein-protein interaction to map the interaction sites of DPP4 with BCoV/S glycoprotein. The bovine DPP4 protein binds to the N-terminal domain (NTD) of the BCoV-S1 protein. The interacting interface amino acid residues for both proteins are shown in the Table S2 and Figure 8. The DPP4 protein showed binding specificity with RBD in the BCoV/S1-CTD site with the residues (TRP184, HIS185, TRP186, LYS196, TRP401, LYS422, and other interacted amino acid residues presented in Table S2. The residues of the BCoV/S1-CTD interacted with the bovine DPP4 prptein are (ASP451, ASP500, LEU448, ASN146, THR144, GLU182, HIS185, THR188, GLY189 and other interacted amino acid residues given in the Table S2. The best docking pose shows a ZDOCK interaction score of 22.44 and an E_R Dock score of -6.23 kcal/mol. These parameters might be due to multiple non-covalent interactions, including 20 conventional hydrogen bonds, nine Pi bonds, and one salt bridge interaction between the two proteins (Table 2).

Figure 8. A proposed model for the interaction between the bovine DPP4 protein and the BCoV-S glycoprotein. (A) The DPP4 protein interacts with the BCoV/S1 protein NTD. (B) and (C) The BCoV/S1 chain N-terminal domain (NTD) residues bind to the DPP4 through TRP184, HIS185, TRP186, ASP187, THR188, GLY189, LYS196 which is core site for Neu5, 9Ac2 binding to BCoV-NTD. The binding amino acid residues from DPP4 to the NTD site are TYR43, VAL441 ARG552 GLY475, LEU476, LYS-512, PHE515, LYS520, HIS532, PHE533, LYS535, MET615, and other amino acid residues by hydrogen bond, Pi bond and salt bridge interactions. .

3.9. The Putative Interaction between the Bovine Host Cell Furin Enzyme and the BCoV/S Glycoprotein

The online ProP 1.0 Server software prediction result showed that there is one pro-peptide cleavage site Arg764, Arg765, Ser766, Arg767, Arg768, and Ala 769 (RRSRR|A) amino acids present across the full-length sequence of the BCoV/S (Uniprot ID: P15777) (Figure 2S). The multiple sequence alignments of the BCoV/S (Mebus) sequence with other betacoronaviruses spike protein sequences suggested protease-specific residues are Arg764, Arg765, Ser766, Arg767, Arg768 and Ala 769 (RRSRR/A) (Figures 9 and 3S). This finding suggests that host bovine cell furin enzyme could bind to and cleave at this polybasic amino acid residue site. The binding of Furin was determined by performing the *in-silico* protein-protein docking (ZDock) method. The protein-protein interaction of the bovine host cell furin and the BCoV/S glycoprotein suggested the potential binding of Furin at the cleavage site identified by the ZDOCK Score 16.72 and E_RDock score -13.14 kcal/mol. Among the ten refined poses from E_RDock, the best pose of the Furin binding to the S1/S2 region mapped near protease cleaving site of polybasic amino acid residues (Arg764, Arg765, Ser766, Arg767, Arg768 and Ala 769 (RRSRR|A) at BCoVS1/S2 cleavage sites (Figure 10). The protease furin interacted with spike protein at the amino acid residues (Glu230, Leu227, Asp228, Asp174, Asp177, Asn192, Met189), to the spike protein S1/S2 cleavage RRSRR site amino acid residues (Val758, Asp759, Tyr760 Ser761, Thr762 Lys763, Arg764, Arg765, Ser766, and Arg767) (Figure 10). The protein-protein interaction interface of BCoV/S-Furin residues is shown in Table S3. The protein-protein interface interacted residues, Lys736 from BCoV/S, form an electrostatic interaction, forming a salt bridge interaction with ASP174, ASP177, and Asp228 residues of Furin. The Arg767 and Ser766 of the BCoV/S RRSRR cleavage site show an electrostatic interaction and hydrogen bonding to the Glu230 of the Furin (Table 2).
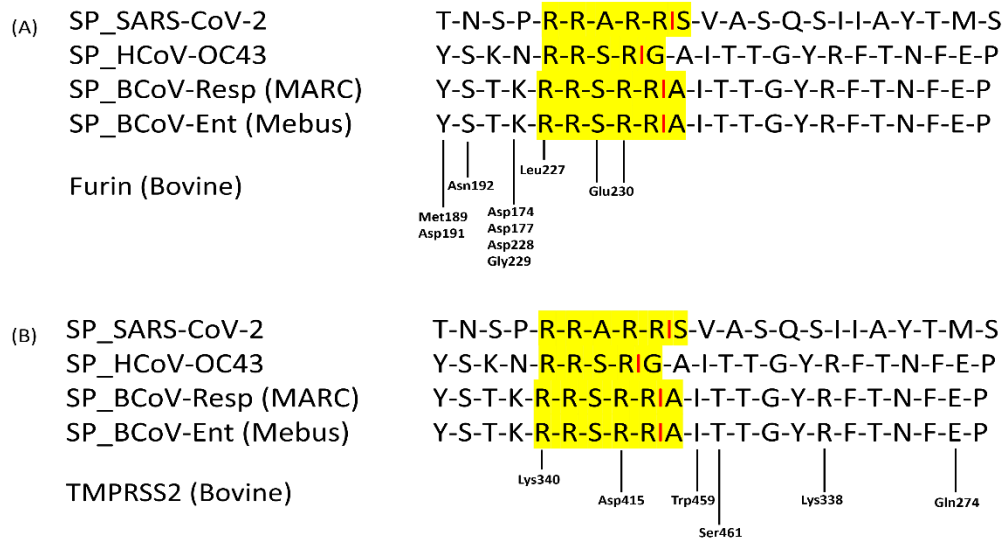
(A)
SP_SARS-CoV-2   T-N-S-P-R-R-A-R-R-I-S-V-A-S-Q-S-I-I-A-Y-T-M-S
SP_HCoV-OC43   Y-S-K-N-R-R-S-R-I-G-A-I-T-T-G-Y-R-F-T-N-F-E-P
SP_BCoV-Resp (MARC)   Y-S-T-K-R-R-S-R-R-I-A-I-T-T-G-Y-R-F-T-N-F-E-P
SP_BCoV-Ent (Mebus)   Y-S-T-K-R-R-S-R-R-I-A-I-T-T-G-Y-R-F-T-N-F-E-P

Furin (Bovine)

(B)
SP_SARS-CoV-2   T-N-S-P-R-R-A-R-R-I-S-V-A-S-Q-S-I-I-A-Y-T-M-S
SP_HCoV-OC43   Y-S-K-N-R-R-S-R-I-G-A-I-T-T-G-Y-R-F-T-N-F-E-P
SP_BCoV-Resp (MARC)   Y-S-T-K-R-R-S-R-R-I-A-I-T-T-G-Y-R-F-T-N-F-E-P
SP_BCoV-Ent (Mebus)   Y-S-T-K-R-R-S-R-R-I-A-I-T-T-G-Y-R-F-T-N-F-E-P

TMPRSS2 (Bovine)

Figure 9. Sequence alignment of the BCoV/S enteric (Mebus) strain with spike protein sequences of BCoV-Respiratory (MARC), Human Coronavirus (HCoV)-OC43 and SARS-CoV-2 strains. The yellow highlighted region contains polybasic residues and its specific cleavage site for furin and TMPRSS2 proteases. (A) showing the furin protease binding residues near its cleavage (RRSRR) at the S1/S2 site of S protein. (B) showing the TMPRSS2 protease binding residues near its cleavage (RRSRR) at the S1/S2 site of S protein.
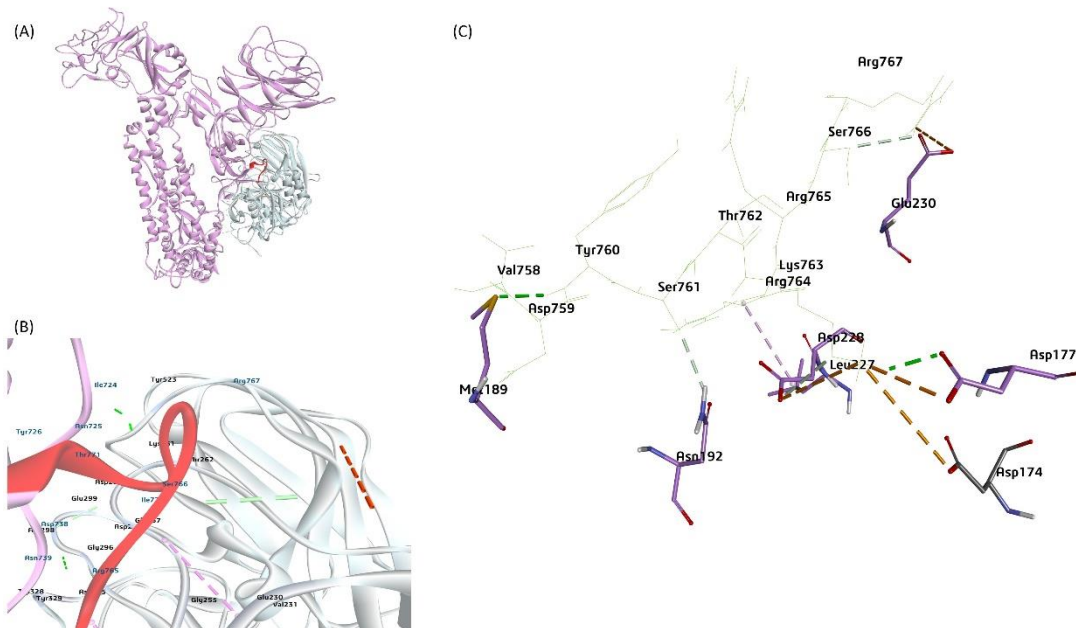

Figure 10. The proposed model of furin protease interactions with BCoV/S cleavage site (A) Furin protease interaction at the furin-specific recognition site (RRSRR region) at the BCoV/S1/S2 junction. (B) and (C) Conformation of the furin residues Glu230, Leu227, Asp228, Asp174, Asp177, Asn192, Met189 with binding affinity to BCoV/S1 protein amino acid residues including Val758, Asp759, Tyr760 Ser761, Thr762 Lys763, Arg764, Arg765, Ser766, Arg767 at RRSRR|A (S1/S2) cleavage site.

3.10. A Proposed Model for theTMPRSS2 Docking with the BCoV/Spike Glycoprotein

The multiple sequence alignments of the BCoV/S (Mebus) sequence with other betacoronaviruses spike protein sequences suggested protease-specific residues are Arg764, Arg765, Ser766, Arg767, Arg768 and Ala 769 (RRSRR/A) (Figure 9). The protein-protein interaction takes place between the bovine TMPRSS2 amino acid residues Val278, His294, Thr339, Asp415, Asn416, Trp459,

Gly460, Ser461, Gly462 near the RRSRR site of BCoV/S protein residues (Asp738, Ser740, Thr741, Ser742, Ser743, Ser761, Arg764, Thr762, Lys763, Arg767, Thr771 and Ile779). The protein-protein docking between the BCoV/S/ and TMPRSS2 in the ZDOCKER shows E_RDock energy by -0.62 kcal/mol due to multiple non-covalent interactions (conventional hydrogen and Pi bond) between the two proteins (Table 2 and Table S2). Among the top 10 refined models, the best pose of the TMPRSS2 binding at the active sites (substrate binding site and catalytic site) of the BCoV/S2 protein is shown in Figure 11.

Figure 11. A proposed model representing the TMPRSS2 serin protease binding upstream and downstream the BCoV/S protein polybasic cleavage site RRSRR|A. (B) and (C) Mapping the TMPRSS2 residues (Val278, His294, Thr339, Asp415, Asn416, Trp459, Gly460, Ser461, Gly46)2 binding affinity to the BCoV/S1/S2 site (Asp738, Ser740, Thr741, Ser742, Ser743, Ser761, Arg764, Thr762, Lys763, Arg767, Thr771 and Ile779).

3.11. The Proposed Model of the Neu5,9Ac2 Docking with the BCoV Haemagglutinin Esterase (BCoV-HE) Protein

The X-ray crystal structure of the BCoV/HE complex with the 4, 9-O'-diacetyl sialic acid retrieved from the RCSB PDB database (PDB ID:3CL5; 1.80 Å resolution) and was chosen as the receptor for ligand Neu5,9Ac2 [31]. The docking of Neu5,9Ac2 with HE indicated a strong binding affinity between the sugar molecule and the BCoV/HE, with the lowest binding energy (-122.73 Kcal/mol). The docking of Neu5,9Ac2 results suggested its interaction with a specific pocket where 4,9-O'-diacetyl sialic acid interacts with the HE. This interaction involves critical amino acid residues, including Asn-264, Glu-265, Ser-221, and Asn-236 (Figure 12). Our docking (molecular interaction) result of the BCoV/S and the BCoV/HE proteins with the known BCoV receptors, Neu5,9Ac2, showed high binding affinity towards both the proteins of BCoV. However, based on binding energy, HE interaction with receptor Neu5,9Ac2 (-122.73Kcal/mol) is more stable and more significant compared to the interaction of the BCoV/S protein with the Neu5,9Ac2 (-79.55 kcal/mol). This suggests that HE may help in the interaction of BCoV with or without spike protein to bind to its specific cellular receptor Neu5,9Ac2. The BCoV-HE could interact with its cellular receptors and may cause changes to the cell surface, which could lead to infection of the host cells.
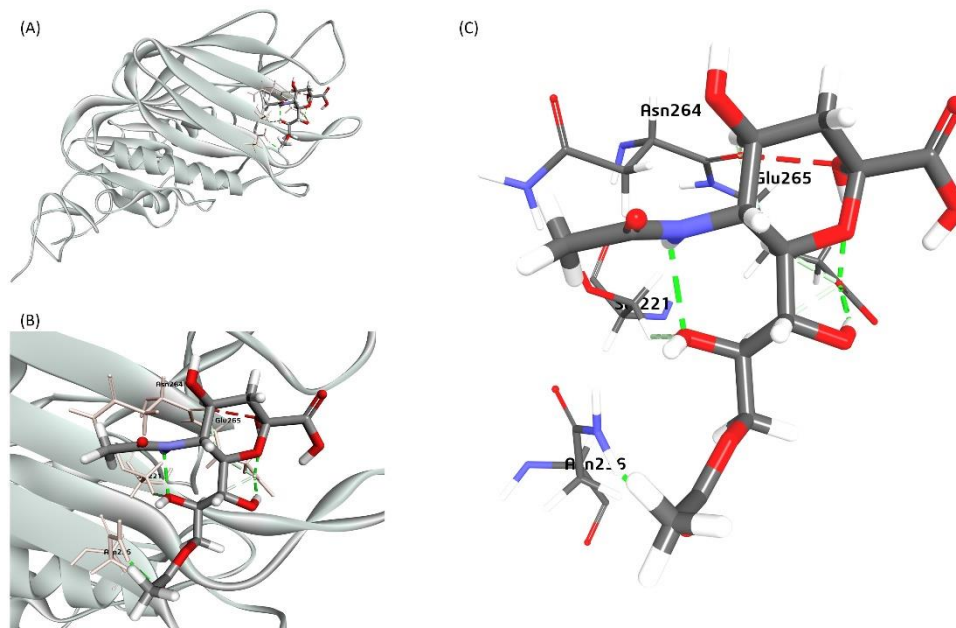
Figure 12. A proposed model for the bovine sialic acid (Neu5,9Ac2) interaction with the BCoV/HE protein (A) shows the binding of Neu5,9Ac2 as a Ligand for the BCoV/HE protein receptor. (B) and (C) showing ligand binding to the interacted amino acid residues- (Asn-264, Glu-265, Ser-221, and Asn-236) of the BCoV/HE protein.

4. Discussion

The tropism of coronaviruses is a complicated process that requires the availability of some factors from the viral side, including some attachment proteins, particularly the spike glycoprotein. This process also involves some cellular factors, including the receptors and other transcription and translation factors. It also requires some factors from the infected host, particularly the availability of some host enzymes that help activate some essential proteins [32–34]. BCoV possesses multiple tissue tropism in cattle. The virus mainly affects the digestive and respiratory tract of the affected animals; this pattern is called pneumoenteritis [35]. The viral tropism primarily depends on the availability of specific viral receptors, some other transcription translation factors, and some host cell enzymes [36]. Members of the family coronaviridae utilize many host cell receptors to attach to their target cells [37].

Meanwhile, most coronaviruses are inert outside the host. Coronaviruses require activation by some host cell proteases to initiate the viral infection inside the host. Usually, the coronaviruses spike glycoproteins cleaved by some host cell proteases of different classes to initiate the viral infection [32–34]. Although BCoV was discovered long ago, little is still known about viral tropism, especially the roles of the host cell receptors and the host enzymes in fine-tuning the viral tissue tropism [35]. The main aims of the current study are to explore the possibilities of identifying some novel receptors of BCoV and to predict the interaction of some key viral proteins (S and HE) with some host proteases, particularly the Furin and the TMPRSS2. The BCoV/S glycoprotein is the main viral protein involved in the process of viral attachment to the host cells [38]. The S protein comprises two subunits (S1 and S2): the N-terminal domain (NTD) and BCoV_S1_CTD are in the S1 subunit, whereas the fusion peptide (FP) and heptad repeat (HR.) domains 1 and 2 are located in the S2 subunit of the BCoV/S glycoprotein. The BCoV/ S1 protein usually attaches to the cell membrane by interacting with viral receptors on the surface of the target cells, initiating the viral infection. Spike protein S2 mediates the fusion of the virion and cellular membranes by acting as a class I viral fusion protein. Also, it acts as a viral fusion peptide, which is unmasked following the S2 cleavage site occurring upon virus endocytosis [39,40]. The distal S1 subunit of the coronavirus spike protein is responsible for receptor binding. Either the S1-NTD or the S1-RBD at C-terminal domain of the BCoV-S1 protein chain, or occasionally both, are involved in the binding to the host receptors [41,42]. The observed binding

between Neu5,9Ac2 and the BCoV NTD is particularly interesting because BCoV haemagglutinin esterase (HE) also utilizes this O-acylated sugar molecule as a substrate. While our study identified interacting residues like Asp-187, Gly-189, His-185, and Lys-196, another report [43] proposed a slightly similar set of critical residues for Neu5,9Ac2 binding (Tyr-162, Glu-182, Trp-184, and His-185).

BCoV/S and BCoV/HE proteins act synergistically and harmoniously to orchestrate the BCoV infection in the target cell [7,44]. The BCoV/S is mainly involved in the initial attachment of the virus to the host cells, while the BCoV/HE destroys the sialic acid in the cell's surface, promoting the viral release from the cell [7].

Most Betacoroanviruses, including the BCoV isolates, use the 9-O-acetyl-Sas as receptors; however, during the evolution of these viruses, some isolates started to recognize other forms of sialic acid (the 4-O-acetyl-SA isoform) [45]. It was recently shown that the HE gene of the SARS-CoV-2 recognizes and binds to the 9-O-acetyl-Sas in contrast to the type-II HE protein of SARS-CoV-2 uses the other isoform; the 4-O-acetyl-SA [45]. The ligand-interaction sites of the BCoV HE and the subset of coronavirus S glycoproteins evolved to recognize 9-O-Ac-Sia via hydrogen bonding [46] specifically.

In the case of most betacoronaviruses, the RBD of CTD or domain B of the spike glycoproteins showed the highest variability within S1 subunits across various members of the coronaviruses, including betacoronaviruses. This phenomenon allows coronaviruses to bind to various types of host cell receptors [46]. It has been proved that SARS-CoV-2 uses the ACE-2 as a valid receptor for the viral entry into the target host cells [47]. However, there are no records about the potential roles of ACE2 as receptors for BCoV. In our study, based on the interaction between bovine ACE2 and BCoV /S interaction with its CTD, the bovine cell surface ACE2 could act as a putative receptor for BCoV/S (Figure 13). This finding is based on the BCoV spike binding to the ACE2 receptor's tendency to show lower interaction energy [48]. This claim is also supported by the nature of the interaction energy, which is consistent with the structure of the virus receptor interface [49].
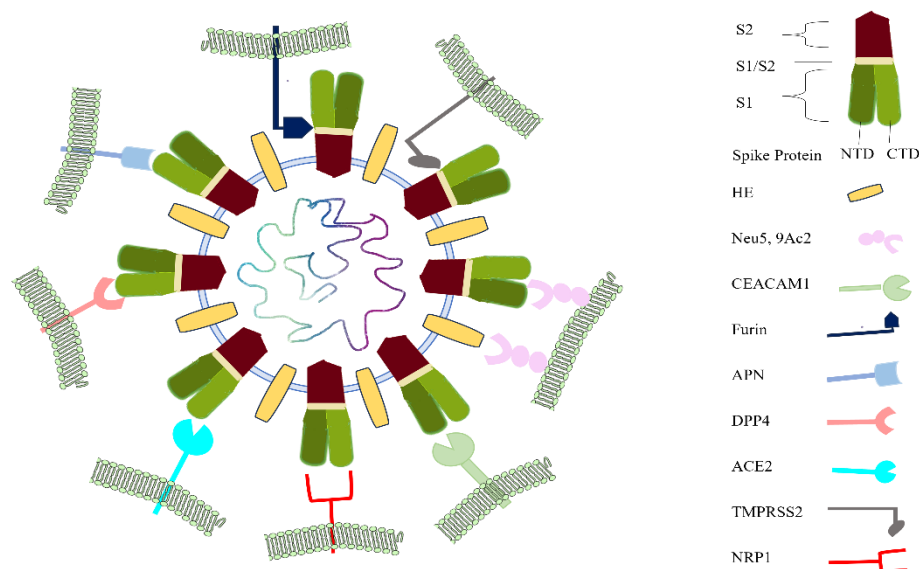

Figure 13. Schematic representation showing the interaction between the BCoV/S and BCoV/HE proteins with the potential cellular receptors, including host cell proteases (Furin and TMPRSS2). With most of the receptors, BCoV/S interacted with the NTD and CTD. The NTD is an N-terminal domain present in the BCoV/S1 protein. The CTD is the C-terminal domain of the BCoV/S1 chain.

We also investigated the potential use of the Neuropilin-1 (NRP1) as a receptor for the BCoV. Our ZDock method result suggested the high interaction affinity (-22.99 kcal/mol) of NRP1 with BCoV/S-NTD and BCoV/S-CTD region (Figures 5 and 13 and Table 2). This raises the possibility that

BCoV could potentially use the NRP receptors for viral entry into the host cell. According to some previous reports, the higher expression of NRP1 will facilitate virus-host cell interactions, especially in cells that do not express other potential BCoV receptors [50,51]. Previous research indicates that Neuropilin-1 (NRP-1) exhibits stronger binding affinity to the CTD region of S1, and this interaction stabilizes the folded conformation of the S protein [52]. Conversely, in the absence of NRP-1, the SARS-CoV-2 S protein tends to stretch and unfold.

It is week known that MHV utilizes the CEACAM1 protein for a dual function. The CEACAM-1 acts as a functional receptor for MHV in addition to enhancing the activation of the MHV-S glycoprotein by inducing some conformational changes, allowing the fusion of the virus with the target cells [53]. However, there is no data about the potential roles of CEACAM-1 in the BCoV replication. Our data showing low binding affinity interaction between BCoV/S and CEACAM-1 suggested that CEACAM-1 interacted with low affinity to the NTD of BCoV/S protein. The variations in binding of coronaviruses to CEACAM-1 can be directly attributed to the structural differences between their N-terminal domains, particularly those of BCoV and MHV. Unlike MHVs sugar-binding ancestors, contemporary MHVs utilize their NTDs for specific interaction with the host cell protein CEACAM-1 (Figure 13) [54].

The aminopeptidase N (APN), a widely found enzyme on cell surfaces, is involved in various cellular processes like survival, migration, blood pressure regulation, and even virus uptake [55]. PDCoV can utilize APN as a receptor to enter cells, highlighting AP**N**'s potential role as a viral infection gateway from different host species [56]. Aminopeptidase N from porcine functions as a receptor for the enveloped RNA virus TGEV [12]. This underscores the wide variety of membrane-bound proteins viruses exploit to infiltrate cells. The APN is a common receptor for the members of the alpha coronaviruses, particularly the HCoV-229E and the TGEV in pigs [57–59]. Based on our protein-protein interaction (E_RDock), the interaction energy of the best pose of APN interaction with BCoV/S shows low-affinity interaction (-2.96 kcal/mol). This result confirms that the binding interaction of BCOV/S is not strong, and it shows a putative binding with the APN. The RBD of CTD from BCoV/S is involved in interaction with Bovine APN (Figure 13).

The DPP4 is a cell surface protease that exhibits exopeptidase activity and is expressed on the surface of various cell types, including those found in the human airways [60]. Researchers have compared the potential binding interactions between the receptor binding domain (RBD) of different spike variants of SARS-CoV-2 and DPP4 with the interactions observed in the experimentally determined structure of the MERS-CoV complex with DPP4 [61]. Members of the Betacoroanviruses, especially the SARS-CoV-2 and the MERS-CoV, utilize the ACE2 and DPP4 as receptors, respectively [62]. The pathogenicity of MERS-CoV is caused by the specific binding of its S1B domain or S1-CTD to the human DPP4 receptor [63]. In our study, docking of the BCoV spike with the DPP4 shows low interaction affinity of DPP4 towards BCoV/S (E_R Dock score -6.23 kcal/mol). However, the binding of DPP4 with BCoV/S1-CTD region or S1B domain is correlated with MERS-CoV binding of its S1B domain or S1-CTD to the human DPP4 (Figure 13).

For virus entry and infections, proteolytic cleavage is widely used to activate the fusion machinery of viral glycoproteins. Furin binding to RRSRR (Arginine-Arginine-Serine-Arginine-Arginine) site of BCoV spike protein, which is related to the conserved cleavage site of SARS-CoV-2 cleavable PRRAR|S residues at receptor-binding (S1) and fusion (S2) domains of the spike protein [64]. Such a motif may allow Spikes to be cut into S1 and S2 by Furin and TMPRSS2-like proteases before maturity, which provides S1 with the flexibility to change the conformation to better fit the host receptor. The arginine residues at the SARS-CoV-2 spike protein catalytic site are popped out of the closed state of the S protein to form multiple non-covalent interactions with the Furin [65]. In our study, the BCoV/S interaction with the Furin, shows that Furin binds near the spike protein S1/S2 cleavage RRSRR site specific for Furin and TMPRSS2 proteases (Figure 13).

TMPRSS2 has been shown to proteolytically activate the S glycoprotein of many coronaviruses, including SARS-CoV-2, SARS-CoV-1, MERS-CoV, 229E, as well as influenza virus [66–68]. TMPRSS2 triggers HKU1-mediated cell-cell fusion and viral entry and binds with high affinity to both HKU1A and HKU1B RBDs.[69].

The protease furin binds firmly with the S protein RRSRR site, which is conserved in BCoV as RRSRR at the S1/S2 site, whereas the TMPRSS2 cleaves the S protein in the lungs of SARS-CoV-2 infected person and promotes pathogenicity [70]. SARS-CoV-2 S protein S1/S2 is cleaved by furin protease, and subsequently, TMPRSS2 mediates the cleavage and activation of the S2 region on the S2 protein [71]. TMPRSS2, which colocalizes with ACE2 at the cell membrane, has been identified as the dominant proteolytic driver of S protein activation and SARS-CoV-2 infection of the aerodigestive tract [72]. TMPRSS2 protease can be opted as a potential therapeutic target for bovine host-specific viruses. In this study, TMPRSS2 binds near the S1/S2 junction RRSRR protease cleavage-specific region for furin and TMPRSS2 proteases. Therefore, there is limited knowledge about the bovine-specific virus-host interactions that determine cellular entry of SARS-CoV-2. Viruses display considerable redundancy and flexibility because they can exploit weak multivalent interactions to enhance affinity. To date, studies of BCoV entry have focused almost entirely on ACE2 and Neu5,9Ac2. In this study, we aimed to see the interaction of bovine cell surface molecules as receptors for BCoV/S, including Nue5,9Ac2, NRP1, DPP4, CEACAM-1, APN ACE2, TMPRSS2, and Furin shown in Figure 13. The results from our study suggested that NRP1 showed greater binding affinity towards BCoV/S other than receptor Neu5,9Ac2. However, bovine HE also showed a stronger binding affinity with Neu5,9Ac2. Our in silico structural interaction data provide a blueprint for understanding the BCoV specificity for the different bovine receptors.

5. Conclusion

Most of the studies suggested different receptor interactions with varying virus strains from animal to human coronaviruses. Our proposed in silico modeling and docking show the potential interaction between the BCoV/S glycoprotein and some known coronaviruses receptors (ACE2, DPP4, NPR1, APN, and CEACAM-1) at various degrees. The NRP1/BCoV/S interaction shows the highest refined dock score (E_RDock score). NRP1 interacts with BCoV/S with both of its ends (NTD and CTD). Therefore, based on the docking result from our study, NRP-1 might play the role of receptor for BCoV entry into the host cell. However, regarding cell surface proteases, our data confirms that the host cellular proteases (Furin and TMPRSS2) have conserved recognition sites within the BCoV/S glycoprotein. The interaction of BCoV-HE shows high-affinity interaction towards the Neu5,9Ac2 receptor from bovine. The BCoV/HE protein may augment the interaction between the BCoV and Neu5,9Ac2, enhancing the virus's attachment to the cellular receptors and the downstream BCoV replication steps. Our in-silico study predicted the bovine NRP1 as a potential receptor for Bovine coronaviruses, and it also appears that the use of NRP1 inhibitors may considered as antiviral therapy. The finding of this in silico study could pave the way for investigating the entry receptor for bovine coronavirus in vivo and understanding the virus replication mechanism in the host through the BCoV/S-Receptor interaction.

References

1. Masters, P.S., *The molecular biology of coronaviruses.* Adv Virus Res, 2006. **66**: p. 193-292.
2. Wild, J.R., et al., *Neuropilins: expression and roles in the epithelium.* Int J Exp Pathol, 2012. **93**(2): p. 81-103.

3. Nassar, A., et al., A Review of Human Coronaviruses' Receptors: The Host-Cell Targets for the Crown Bearing Viruses. Molecules, 2021. **26**(21).

4. Liao, Y., et al., Classification, replication, and transcription of Nidovirales. Front Microbiol, 2023. **14**: p. 1291761.

5. Cavanagh, D., Nidovirales: a new order comprising Coronaviridae and Arteriviridae. Arch Virol, 1997. **142**(3): p. 629-33.

6. Brian, D.A. and R.S. Baric, *Coronavirus genome structure and replication.* Curr Top Microbiol Immunol, 2005. **287**: p. 1-30.

7. Lang, Y., et al., Coronavirus hemagglutinin-esterase and spike proteins coevolve for functional balance and optimal virion avidity. Proc Natl Acad Sci U S A, 2020. **117**(41): p. 25759-25770.

8. Carlos, A.J., et al., The chaperone GRP78 is a host auxiliary factor for SARS-CoV-2 and GRP78 depleting antibody blocks viral entry and infection. J Biol Chem, 2021. **296**: p. 100759.

9. Jackson, C.B., et al., *Mechanisms of SARS-CoV-2 entry into cells.* Nat Rev Mol Cell Biol, 2022. **23**(1): p. 3-20.

10. Alnaeem, A., et al., The dipeptidyl peptidase-4 expression in some MERS-CoV naturally infected dromedary camels in Saudi Arabia 2018-2019. Virusdiseae, 2020. **31**(2): p. 200-203.

11. Widagdo, W., et al., Host Determinants of MERS-CoV Transmission and Pathogenesis. Viruses, 2019. **11**(3).

12. Delmas, B., et al., Aminopeptidase N is a major receptor for the entero-pathogenic coronavirus TGEV. Nature, 1992. **357**(6377): p. 417-20.

13. Nakagaki, K., K. Nakagaki, and F. Taguchi, Receptor-independent spread of a highly neurotropic murine coronavirus JHMV strain from initially infected microglial cells in mixed neural cultures. J Virol, 2005. **79**(10): p. 6102-10.

14. Rossi, A.D., et al., Association between ACE2 and TMPRSS2 nasopharyngeal expression and COVID-19 respiratory distress. Sci Rep, 2021. **11**(1): p. 9658.

15. Lubinski, B. and G.R. Whittaker, *The SARS-CoV-2 furin cleavage site: natural selection or smoking gun?* Lancet Microbe, 2023. **4**(8): p. e570.

16. Peacock, T.P., et al., The furin cleavage site in the SARS-CoV-2 spike protein is required for transmission in ferrets. Nat Microbiol, 2021. **6**(7): p. 899-909.

17. Kemmish, H., M. Fasnacht, and L. Yan, Fully automated antibody structure prediction using BIOVIA tools: Validation study. PLoS One, 2017. **12**(5): p. e0177923.

18. Vyas, V.K., et al., Homology modeling a fast tool for drug discovery: current perspectives. Indian J Pharm Sci, 2012. **74**(1): p. 1-17.

19. Takeda-Shitaka, M., et al., *Protein structure prediction in structure based drug design.* Curr Med Chem, 2004. **11**(5): p. 551-8.

20. Takeda-Shitaka, M., et al., Evaluation of homology modeling of the severe acute respiratory syndrome (SARS) coronavirus main protease for structure based drug design. Chem Pharm Bull (Tokyo), 2004. **52**(5): p. 643-5.

21. Blum, M., et al., *The InterPro protein families and domains database: 20 years on.* Nucleic Acids Res, 2021. **49**(D1): p. D344-D354.

22. Etzold, T. and P. Argos, *SRS--an indexing and retrieval tool for flat file data libraries.* Comput Appl Biosci, 1993. **9**(1): p. 49-57.

23. Apweiler, R., et al., The InterPro database, an integrated documentation resource for protein families, domains and functional sites. Nucleic Acids Res, 2001. **29**(1): p. 37-40.

24. Ding, X., et al., Accelerated CDOCKER with GPUs, Parallel Simulated Annealing, and Fast Fourier Transforms. J Chem Theory Comput, 2020. **16**(6): p. 3910-3919.

25. Gagnon, J.K., S.M. Law, and C.L. Brooks, 3rd, Flexible CDOCKER: Development and application of a pseudo-explicit structure-based docking method within CHARMM. J Comput Chem, 2016. **37**(8): p. 753-62.

26. Pierce, B.G., Y. Hourai, and Z. Weng, Accelerating protein docking in ZDOCK using an advanced 3D convolution library. PLoS One, 2011. **6**(9): p. e24657.

27. Ramachandran, G.N., C. Ramakrishnan, and V. Sasisekharan, *Stereochemistry of polypeptide chain configurations.* J Mol Biol, 1963. **7**: p. 95-9.

28. Qian, Z., et al., Identification of the Receptor-Binding Domain of the Spike Glycoprotein of Human Betacoronavirus HKU1. J Virol, 2015. **89**(17): p. 8816-27.

29. Schultze, B., et al., The S protein of bovine coronavirus is a hemagglutinin recognizing 9-O-acetylated sialic acid as a receptor determinant. J Virol, 1991. **65**(11): p. 6232-7.

30. Kunkel, F. and G. Herrler, Structural and functional analysis of the surface protein of human coronavirus OC43. Virology, 1993. **195**(1): p. 195-202.

31. Kumar, A., et al., Identification of phytochemical inhibitors against main protease of COVID-19 using molecular modeling approaches. J Biomol Struct Dyn, 2021. **39**(10): p. 3760-3770.

32. Cheng, Y.R., et al., *Cell Entry of Animal Coronaviruses.* Viruses, 2021. **13**(10).

33. Hoffmann, M., H. Hofmann-Winkler, and S. Pöhlmann, *Priming time: how cellular proteases arm coronavirus spike proteins.* Activation of viruses by host proteases, 2018: p. 71-98.

34. Millet, J.K. and G.R. Whittaker, Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis. Virus Res, 2015. **202**: p. 120-34.

35. Shah, A.U. and M.G. Hemida, The Potential Roles of Host Cell miRNAs in Fine-Tuning Bovine Coronavirus (BCoV) Molecular Pathogenesis, Tissue Tropism, and Immune Regulation. Microorganisms, 2024. **12**(5).

36. Susi, P., Special Issue: Virus Receptors and Viral Tropism. Viruses, 2021. **14**(1).

37. Everest, H., et al., Known Cellular and Receptor Interactions of Animal and Human Coronaviruses: A Review. Viruses, 2022. **14**(2).

38. Malik, Y.A., *Properties of Coronavirus and SARS-CoV-2.* Malays J Pathol, 2020. **42**(1): p. 3-11.

39. Belouzard, S., V.C. Chu, and G.R. Whittaker, *Activation of the SARS coronavirus spike protein* via *sequential proteolytic cleavage at two distinct sites.* Proc Natl Acad Sci U S A, 2009. **106**(14): p. 5871-6.

40. Walls, A.C., et al., Tectonic conformational changes of a coronavirus spike glycoprotein promote membrane fusion. Proc Natl Acad Sci U S A, 2017. **114**(42): p. 11157-11162.

41. Xia, S., et al., Inhibition of SARS-CoV-2 (previously 2019-nCoV) infection by a highly potent pan-coronavirus fusion inhibitor targeting its spike protein that harbors a high capacity to mediate membrane fusion. Cell Res, 2020. **30**(4): p. 343-355.

42. Xia, S., et al., The role of furin cleavage site in SARS-CoV-2 spike protein-mediated membrane fusion in the presence or absence of trypsin. Signal Transduct Target Ther, 2020. **5**(1): p. 92.

43. Peng, G., et al., Crystal structure of bovine coronavirus spike protein lectin domain. J Biol Chem, 2012. **287**(50): p. 41931-8.

44. Bahoussi, A.N., et al., Evolutionary adaptation of bovine coronavirus (BCoV): Screening of natural recombinations across the complete genomes. J Basic Microbiol, 2023. **63**(5): p. 519-529.

45. Kim, C.H., SARS-CoV-2 Evolutionary Adaptation toward Host Entry and Recognition of Receptor O-Acetyl Sialylation in Virus-Host Interaction. Int J Mol Sci, 2020. **21**(12).

46. Tortorici, M.A., et al., Structural basis for human coronavirus attachment to sialic acid receptors. Nat Struct Mol Biol, 2019. **26**(6): p. 481-489.

47. Scialo, F., et al., *ACE2: The Major Cell Entry Receptor for SARS-CoV-2.* Lung, 2020. **198**(6): p. 867-877.

48. Gupta, A., et al., Recent Developments and Future Perspectives of Vaccines and Therapeutic Agents against SARS-CoV2 Using the BCOV_S1_CTD of the S Protein. Viruses, 2023. **15**(6).

49. Rodriguez, J.H. and A. Gupta, Contact residue contributions to interaction energies between SARS-CoV-1 spike proteins and human ACE2 receptors. Sci Rep, 2021. **11**(1): p. 1156.

50. Davies, J., et al., Neuropilin-1 as a new potential SARS-CoV-2 infection mediator implicated in the neurologic features and central nervous system involvement of COVID-19. Mol Med Rep, 2020. **22**(5): p. 4221-4226.

51. Cantuti-Castelvetri, L., et al., *Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity.* Science, 2020. **370**(6518): p. 856-860.

52. Pal, D., et al., Mutating novel interaction sites in NRP1 reduces SARS-CoV-2 spike protein internalization. iScience, 2023. **26**(4): p. 106274.

53. Miura, H.S., K. Nakagaki, and F. Taguchi, N-terminal domain of the murine coronavirus receptor CEACAM1 is responsible for fusogenic activation and conformational changes of the spike protein. J Virol, 2004. **78**(1): p. 216-23.

54. Taylor, M.E. and K. Drickamer, Mammalian sugar-binding receptors: known functions and unexplored roles. Febs j, 2019. **286**(10): p. 1800-1814.

55. Amin, S.A., N. Adhikari, and T. Jha, *Design of Aminopeptidase N Inhibitors as Anti-cancer Agents.* J Med Chem, 2018. **61**(15): p. 6468-6490.

56. Yang, Y.L., et al., Aminopeptidase N Is an Entry Co-factor Triggering Porcine Deltacoronavirus Entry via an Endocytotic Pathway. J Virol, 2021. **95**(21): p. e0094421.

57. Liu, Y., et al., Roles of Two Major Domains of the Porcine Deltacoronavirus S1 Subunit in Receptor Binding and Neutralization. J Virol, 2021. **95**(24): p. e0111821.

58. Wang, B., et al., Porcine Deltacoronavirus Engages the Transmissible Gastroenteritis Virus Functional Receptor Porcine Aminopeptidase N for Infectious Cellular Entry. J Virol, 2018. **92**(12).

59. Yin, L., et al., Aminopeptidase N Expression, Not Interferon Responses, Determines the Intestinal Segmental Tropism of Porcine Deltacoronavirus. J Virol, 2020. **94**(14).

60. Lambeir, A.M., et al., Dipeptidyl-peptidase IV from bench to bedside: an update on structural properties, functions, and clinical aspects of the enzyme DPP IV. Crit Rev Clin Lab Sci, 2003. **40**(3): p. 209-94.

61. Roy, A.N., et al., Unraveling DPP4 Receptor Interactions with SARS-CoV-2 Variants and MERS-CoV: Insights into Pulmonary Disorders via Immunoinformatics and Molecular Dynamics. Viruses, 2023. **15**(10).

62. Wang, Q., L. Sun, and S. Jiang, Potential recombination between SARS-CoV-2 and MERS-CoV: calls for the development of Pan-CoV vaccines. Signal Transduct Target Ther, 2023. **8**(1): p. 122.

63. Lu, G., et al., Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26. Nature, 2013. **500**(7461): p. 227-31.

64. Whittaker, G.R., *SARS-CoV-2 spike and its adaptable furin cleavage site.* Lancet Microbe, 2021. **2**(10): p. e488-e489.

65. Vardhan, S. and S.K. Sahoo, Virtual screening by targeting proteolytic sites of Furin and TMPRSS2 to propose potential compounds obstructing the entry of SARS-CoV-2 virus into human host cells. J Tradit Complement Med, 2022. **12**(1): p. 6-15.

66. Bertram, S., et al., TMPRSS2 activates the human coronavirus 229E for cathepsin-independent host cell entry and is expressed in viral target cells in the respiratory epithelium. J Virol, 2013. **87**(11): p. 6150-60.

67. Shirato, K., M. Kawase, and S. Matsuyama, Middle East respiratory syndrome coronavirus infection mediated by the transmembrane serine protease TMPRSS2. J Virol, 2013. **87**(23): p. 12552-61.

68. McCallum, M., et al., Human coronavirus HKU1 recognition of the TMPRSS2 host receptor. bioRxiv, 2024.

69. Saunders, N., et al., TMPRSS2 is a functional receptor for human coronavirus HKU1. Nature, 2023. **624**(7990): p. 207-214.

70. Zhang, Y., et al., Transmembrane serine protease TMPRSS2 implicated in SARS-CoV-2 infection is autoactivated intracellularly and requires N-glycosylation for regulation. J Biol Chem, 2022. **298**(12): p. 102643.

71. Fraser, B.J., et al., structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation. Nat Chem Biol, 2022. **18**(9): p. 963-971.

72. Bestle, D., et al., TMPRSS2 and Furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. Life Sci Alliance, 2020. **3**(9).