**Identification of some potential novel host enzymes and cellular receptors that could refines the bovine coronavirus (BCoV) replication and the viral tissue tropism. A molecular docking study.**

Mohd Yasir Khan1, Nithyadevi Duraisamy1, Abid Ullah Shah2, Reda Nacif ElAlaoui1, Mohammed Cherkaoui1, Maged Gomaa Hemida2\*

1 Affiliation 1; Department of Computer Scicnes, College of Digital engineering and artificial intelligence, Long Island University, Brooklyn, [mohd.yasirkhan@liu.edu](mailto:mohd.yasirkhan@liu.edu) (MYK), [Nithyadevi.Duraisamy@liu.edu](mailto:Nithyadevi.Duraisamy@liu.edu) (ND), [Reda.NacifElAlaoui@liu.edu](mailto:Reda.NacifElAlaoui@liu.edu) (RNE), [mohammed.cherkaoui@liu.edu](mailto:mohammed.cherkaoui@liu.edu) (MC).

2 Affiliation 2; Department of Veterinary Biomedical Sciences, College of Veterinary Medicine, Long Island University, 720 Northern Boluvered, Brookville, NY, 11548, USA, [abidullah.shah@liu.edu](mailto:abidullah.shah@liu.edu) (AS), [maged.hemida@liu.edu](mailto:maged.hemida@liu.edu) (MGH)

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

**Supplementary data**

**3.1 The Homology Modelling for the BCoV-Spike glycoprotein**

**Table 1S. Top Homology Models by Sequence ID Based on DOPE Score and PDF Total Energy**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sequence UniProt ID | Protein | Homology Model | PDF Total Energy | DOPE Score |
| **P15777** | **BCoV-Spike** | Model 2 | 51989.78 | -135770.54 |
| Model 1 | 53052.81 | -133727.64 |
| **Q58DD0** | **ACE2** | Model 1 | 34827.88 | -94345.04 |
| Model 2 | 34765.74 | -94330.17 |
| **Q28193** | **Furin** | Model 2 | 16666.25 | -53696.58 |
| Model 1 | 16719.49 | -53651.78 |
| **A2VDV7** | **TMPRSS2** | Model 1 | 11,639.4 | -34,427 |
| Model 2 | 11,659.4 | -34,507 |
| **E1BMX5** | **NRP1** | Model 2 | 15313.95 | -25423.16 |
| Model 1 | 15491.51 | -25920.80 |
| **P81425** | **DPP4** | Model 1 | 33378.53 | -90256.60 |
| Model 2 | 33326.76 | -89901.53 |
| **P79098** | **APN** | Model 1 | 35879.33 | -118318.42 |
| Model 2 | 35532.13 | -118316.10 |
| **Q6VAN8** | **CEACAM-1** | Model 5 | 5361.93 | -10424.35 |
| Model 2 | 5319.78 | -9986.85 |

**Structural alignment of the BCoV-Spike protein (mebus strain) NTD with BCoV-NTD (PDB: 4h14)**

The structural alignment results of BCoV NTD (PDB ID: 4h14) structure and NTD-of homology model BCoV (UniProt ID: p157777) was done through CLUSTAL Omega and 3D structure superimposition by PyMol. The results of the CLUSTAL Omega and 3D structure superimposition shows structure similarity and the sequence similarity and between BCoV spike NTD region (Figure 1S).

A close-up of a document

Description automatically generated

**Figure 1S:** Structural Alignment of N-terminal domain of BCoVSpike(A) Sequence alignment of the BCoV spike protein whole sequence (UniProt ID: P15777) and the sequence N-terminal domain (NTD) of the BCoV spike protein (PDB:ID 4h14), performed using CLUSTAL Omega. Structural alignment is indicated by dots underlying the two sequences. (B) and (C) Structural alignment of the N-terminal domain of BCoV/S (UniProt ID: P15777) and BCoV/S-NTD (PDB:ID 4h14), visualized using PyMol software.

**Table 2S**: The interacting residues between the BCoV-Spike protein and bovine ACE2, NRP1, CEACAM, APN and DPP4 protein predicted by ZDOCK method of Biovia, Discovery studio.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **BCoV/S-ACE2** | **BCoV/S-NRP1** | **BCoV/S-CEACAM1** | **BCoV/S-APN** | **BCoV/S-DPP4** |
| :ARG419 - :GLU37  :GLN18 - :ASP608​  :HIS34 - :CYS530  :THR82 - :ASP662  :SER84 - :ASN664  :GLN324 - :ILE512  :ASN329 - :THR519  :THR333 - :ALA517  :ASP354 - :ASN526  :ARG419 - :ALA385  :THR425 - :GLU30  :THR514 - :ASP328  :CYS515 - :GLU304 ​  :ASN520 - :ASP328  :THR523 - :ASN329  :CYS524 - :ASN329  :ASP662 - :THR20  :SER846 - :GLY574  :GLU30 - :THR423​  :HIS478 - :GLN324  :THR519 - :ASP328  :HIS607 - :GLU30  :HIS607 - :HIS34:  :SER611 - :GLN24​  :SER846 - :GLU570  :LYS31 - :HIS607​  :GLU570 - :PHE847​  :SER19 - :TYR661​  :THR20 - :TYR661​  :ASN526 - :TYR41​  :ALA385 - :LEU410​  :ALA528 - :LYS352​  :HIS34 - :CYS530​ | :THR422 - :LEU12  :ASN432 - :ASP250  :ARG1077 - :SER47  :ARG1077 - :TYR48  :LEU1081 - :ALA21  :GLN1084 - :ALA19  :ARG1089 - :GLN6:  :ARG80 - :GLY278  :ARG80 - :ASP277  :HIS130 - :ASN235  :TRP214 - :ASP370  :ARG307 - :TYR237  :SER1076 - :TYR44:  :LYS3 - :GLU1082  :GLY20 - :GLN1084  :HIS130 - :ASN235  :GLY216 - :ILE367  :ARG197 - :TYR175  :ARG1077 - :HIS49 ​  :LYS53 - :PHE232  :GLN368 - :PHE217  :VAL269 - :TYR136 ​  :HIS130 - :TYR195 ​  :TYR237 - :VAL264 ​  :HIS49 - :ARG1077​ | :ARG143 - A:LEU233  :ARG143 - A:PHE232  :ASN170 - A:VAL229  :THR171 - A:PHE232  :ASN178 - A:THR145  :LYS196 - :TRP154  :ARG197 - :ARG143​  :ASP225 -: TYR173  :LYS231 - :ASN170  :HIS156 - :GLN141​  :LYS150 - :TRP175​  :LYS150 - :TRP176​  :ARG197 - :TRP215​  :GLU156: - :HIS156​  :TRP215 - :TYR195​  :TRP154 - :LYS196​  :TRP176 - :LEU147​  :ARG187 - :ASP148​ | :ASN397 - :CYS826  :TYR451 - :ASP849  :ARG514 - :GLU843​  :ARG514 - :SER839​​​  :LYS579 - :ILE512​  :LYS579 - :GLY511  :ASN583 - :PRO516​  :CYS524 - :ASN495​  :CYS524 - :ASN576​  :SER827 - :ASN397​  :LEU841 - :GLU398​  :SER846 - :GLN517​  :PHE847 - :GLN517​  :PRO515 -: GLN840​  :CYS515 - :ASN583​  :GLY518 - :ALA555​  :SER516 - :TYR844 | :TRP184 - :ASP451  :HIS185 - :ASP451  :TRP186 - :ASP500​  :LYS196 - :LEU448  :TRP401 - :ASN146  :LYS422 - :THR144  :LYS422 - :GLU182​  :ARG452 - :HIS185  :ARG452 - :THR188  :ARG452 - :GLY189  :LYS512 - :LYS179​  :LYS522 - :ASN175  :LYS522 - :PRO174​  :LYS522 - :ASN175  :PHE533 - :THR23  :LEU183 - :ASP451  :GLY189 - :ASN495  :SER510 - :CYS21  :VAL24 - :PHE533​  :GLY475 - :TRP186​  :HIS143 - :PRO450​  :TRP184 - :PRO450​  :HIS185 - :ARG452​  :TRP186 - :LEU476​  :TRP524 - :PRO174​  :LYS179 - :ASP514 |

**Table 3S**: The interacting residues between the BCoV-Spike protein and bovine furin and TMPRSS2 protease protein predicted by ZDOCK method of Biovia, Discovery studio.

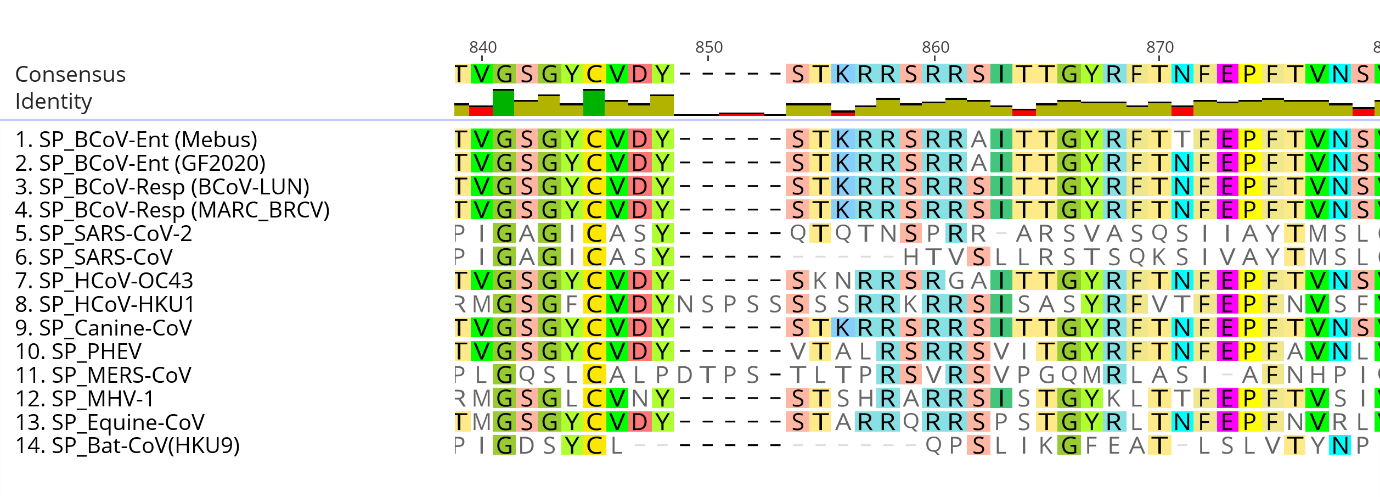
|  |  |
| --- | --- |
| BCoV/S-Furin | BCoV/S-TMPRSS2 |
| :LYS763 - :ASP174​  :LYS763 - :ASP177​  :LYS763 - :ASP228​  :ARG767 - :GLU230​  :TYR760 - :MET189​  :TYR760 - :ASP191  :LYS763 - :ASP177​  :GLY229 - :LYS763  :SER761 - :ASN192​  :LYS763 - :ASP228  :SER766 - :GLU230  :ARG764 - :LEU227 | :ASN739 - :GLY462  :ASN739 - :CYS463  :GLN746 - :PRO299  :ARG767 - :ASP415  :ARG775 - :LYS338  :GLN274 - :GLU780  :GLU297 - :SER743  :LYS304 - :SER786  :ARG325 - :ASP789  :ARG325 - :GLU792  :LYS340 - :ARG764  :SER461 - :THR771  :ARG468 - :ILE724  :ARG468 - :ILE724  :SER742 - :VAL278  :LYS390 - :TYR730 ​  :LYS465 - :TYR726, ​  :TRP459 - :ILE770 |

A graph with numbers and lines

Description automatically generated

Figure 2S. The ProP 1.0 Server software prediction of the Propeptide cleavage site (RRSRR|A) amino acids residues present in the full-length sequences of the BCoV/S protein

(A)



(B)



**Figure S3**: The sequence alignment of the (A) Bovine coronavirus (BCoV) enteric and respiratory strain spike protein sequence with spike protein sequences of other betacoronaviruses. (B) The red highlighted residues are Furin and TMPRSS2 specific cleavage site present at S1/S2 junction of spike protein. This polybasic residue (RRXRR) site present in all the spike protein sequences (SP) of betacoronaviruses except the SARS-COV and Bat-coronavirus (HKU9) strain.