



|                                     |  |
|-------------------------------------|--|
| <b>Title</b>                        | Network proteins of angiotensin-converting enzyme 2 but not angiotensin-converting enzyme 2 itself are host cell receptors for SARS-Coronavirus-2 attachment   |
| <b>Authors(s)</b>                   | Kumar, Arun H.S.   |
| <b>Publication date</b>             | 2020-05-07   |
| <b>Publication information</b>      | Kumar, Arun H.S. "Network Proteins of Angiotensin-Converting Enzyme 2 but Not Angiotensin-Converting Enzyme 2 Itself Are Host Cell Receptors for SARS-Coronavirus-2 Attachment." Research Square, May 7, 2020. <a href="https://doi.org/10.21203/rs.3.rs-27149/v1">https://doi.org/10.21203/rs.3.rs-27149/v1</a> . |
| <b>Publisher</b>                    | Research Square  |
| <b>Item record/more information</b> | <a href="http://hdl.handle.net/10197/11397">http://hdl.handle.net/10197/11397</a>  |
| <b>Publisher's statement</b>        | This work is licensed under a Creative Commons Attribution 4.0 International License.  |
| <b>Publisher's version (DOI)</b>    | 10.21203/rs.3.rs-27149/v1  |

Downloaded 2026-05-02 00:27:20

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd\_oa)



© Some rights reserved. For more information

# **Network proteins of angiotensin-converting enzyme 2 are host cell receptors for SARS-Coronavirus-2 attachment**

**Short title:** Cell surface receptors for SARS-CoV-2 entry

Arun HS Kumar

Corresponding Address

Arun HS Kumar, DVM, PhD.

Room 216, School of Veterinary Medicine, University College Dublin, Belfield, Dublin-04, Ireland. Phone: 0035317166230, Fax: 00353017166104. Email: arun.kumar@ucd.ie

## **Abstract**

**Coronaviruses causing severe acute respiratory syndrome (SARS-CoV) are known to enter the host cells by attaching to the membrane bound angiotensin-converting enzyme 2 (ACE2)<sup>1</sup>. A low binding affinity was observed between SARS-CoV-2 spike proteins (protein S, M and 6YLA) and ACE2. Coronaviruses are also reported to bind to dipeptidyl peptidase 4 (DPP4),<sup>2,3</sup> which is a network protein of ACE2. Network analysis showed five membrane proteins associated with ACE2. The ACE2 network proteins were assessed for their binding affinity with all known SARS-CoV-2 surface proteins. The SARS-CoV-2 surface proteins showed preferential binding to network proteins such as DPP4 and Meprin A alpha but not ACE2. The binding efficacy (affinity (-5.86 to -7.10 Kcal/mol), Ki (6.32 – 22.04  $\mu$ M) and IC50 (12.63 – 113.71  $\mu$ M) values) of DPP4 inhibitors (saxagliptin and sitagliptin) against SARS-CoV-2 surface proteins, was observed to be at a therapeutically feasible concentration to prevent SARS-CoV-2 attachment and entry into host cells.**

The recent pandemic caused by a new strain of coronavirus (SARS-CoV-2) has resulted in serious health, social and economic setbacks<sup>4,5</sup>. Efforts to identify effective therapeutics or vaccine against SARS-CoV-2 illness (Covid-19) is currently being extensively explored globally<sup>6-9</sup>. Understanding the mechanisms by which

SARS-CoV-2 enters to host cell and replicates to cause Covid-19 is necessary for our efforts to develop effective therapeutics or vaccine. Coronavirus are reported to enter the host cells by attaching to the membrane bound ACE2 using their spike protein<sup>1,10</sup>. The spike protein (protein S) of SARS-CoV-2 was hence assessed for its binding with ACE2 using molecular docking analysis in Chimera software<sup>11</sup>. The crystal structure of human ACE2 (hACE2, PDB ID: 1r42) was downloaded from the protein data bank (PDB) and the 3D structure of SARS-CoV-2 protein S (Uniport ID P59594) was constructed by homology modelling using SWISS-MODEL server<sup>12</sup>. In contrast to the published reports, SARS-CoV-2 protein S wasn't observed to be binding with hACE2 and no hydrogen bonds were observed between the two molecules. This wasn't surprising because the previous reports showing the interaction of ACE2 with coronavirus spike protein were performed in cell based assay systems using Vero E6 cells isolated from African green monkeys<sup>1,10</sup>. While structural similarities of proteins are evident across different species, the functional molecular interactions may differ between different species and as well as between different individuals of same species. The Insilco molecular interaction analysis allows to overcome this limitations by facilitating species specific molecular interaction analysis. Further the short peptide sequences used to identify protein targets based on antibody-antigen interactions in cell based assay system may often cross react due to similarities between different protein sequence, resulting in false positive outcomes. Due to the poor interaction of SARS-CoV-2 protein S with hACE2, seventeen other SARS-CoV-2 target proteins (see online methods) were further assessed for their binding potential with hACE2. Only the SARS-CoV-2 protein M (a membrane associated matrix glycoprotein) and 6YLA (a receptor binding domain on spike) were observed very weakly bound to hACE2 (Fig 1B). ACE2 exists in soluble<sup>13</sup> as well as membrane bound<sup>10,14</sup> forms. It is likely that the reported affinity of SARS-CoV to soluble form of ACE2<sup>1</sup> may differ with the membrane bound form of ACE2, this together with the species specific differences highlighted above may explain the weak interactions observed between ACE2 and SARS-CoV-2. Considering the weak interactions between ACE2 and SARS-CoV-2, it is unlikely that approaches to upregulate ACE2 or using soluble forms of ACE2<sup>10,14</sup> will be therapeutically beneficial in the clinical management of COVID-19. Coronavirus are also reported to bind to DPP4 receptors on cell surface<sup>2,3</sup>. Both protein M and 6YLA but not protein S, were observed to strongly bind with DPP4

(Figure 1B). Interestingly DPP4 is a network protein of hACE2. Hence a network analysis<sup>15,16</sup> was performed using the STRING database to identify all the primary ACE2 network proteins. Six membrane associated network host cell proteins including ACE2 were identified (Fig 1A) showing significant ( $p\text{-value} = 1.3 \times 10^{-09}$ ) interactions among themselves. Each of the network proteins were screened for molecular interactions with the seventeen SARS-CoV-2 target proteins (Fig 2). The SARS-CoV-2 protein S did not bind to hACE2 or any of its other five network proteins, hence it is unlikely that this spike protein is involved in attachment and entry of the virus into the host cells. Further the stronger binding of SARS-CoV-2 protein M and 6YLA to DPP4 compared to hACE2 (Fig 1B), suggests that SARS-CoV-2 may preferentially attach to DPP4 rather than hACE2 for attaching and entering into the host cell. Consistent with this study, another strain of human coronavirus was reported to preferentially use DPP4 rather than ACE2 as the host cell entry receptor<sup>2,3</sup>. DPP4 is highly expressed in apical surfaces of human bronchiolar epithelial cell and lung tissue,<sup>2,3</sup> hence may be involved in increasing the susceptibility of the respiratory system to human coronavirus infection.

The SARS-CoV-2 surface proteins were observed to effectively bind with all the network proteins of ACE2 with better efficiency (Fig 2A). Among the network proteins DPP4 and meprin A alpha were observed to bind to eight different SARS-CoV-2 surface proteins (Fig 2B). While the meprin A beta, MME, PRCP and XPNPEP2 were binding with at least 3-4 SARS-CoV-2 target proteins (Fig 2B). This observations is in contrast to current knowledge on specific strains of coronavirus binding to selective host cell receptors<sup>6,9</sup>. It is likely that SARS-CoV-2 can either sequentially or preferentially interact with ACE2 network proteins (but not ACE2) for its attachment and entry into host cells. Currently the following three; ACE2, aminopeptidase N (APN or CD13) and DPP4 are reported to be receptor for coronaviruses<sup>1,3</sup>. However this is the first report of a single coronavirus strain potentially utilizing multiple host cell receptors for attachment and cell entry. Such wider choices in host cell receptors for SARS-CoV-2 attachment and entry may enhance the infectivity potential of this virus and probably also lead to poor development of immunity by the host. Although APN was not among the network protein of ACE2, it was evaluated for its interaction with the seventeen SARS-CoV-2 target proteins. APN did not bind to any of the SARS-CoV-2 target proteins, but was observed to bind with SARS-CoV-2 non-structural protein 10/16 (nsp10/16), which is

a RNA methyltransferase involved in virus replication rather than virus attachment to host cell (data not shown). Unlike the previously known strains of coronavirus, SARS-CoV-2 seem to specifically bind with ACE2 network proteins (preferentially DPP4 and meprin A alpha) for attachment and entry into host cells. This knowledge of SARS-CoV-2 using the ACE2 network proteins will be helpful in the development of novel therapeutics and repurposing existing therapeutics for the clinical management of COVID-19.

Among the host cell receptors for SARS-CoV-2 identified in this study, DPP4 was the only receptors for which FDA approved selective inhibitors (sitagliptin, vildagliptin, saxagliptin) are available. Hence the DPP4 inhibitors, saxagliptin and sitagliptin were selected to assess their pharmacological efficacy against SARS-CoV-2. To assess the efficacy of DPP4 inhibitors, saxagliptin and sitagliptin were docked against the selected SARS-CoV-2 surface proteins (protein M, 6YLA and 6M0J) in the AutoDock Vina software<sup>17-19</sup>. Molecular docking with DPP4 and 6Y2E (SARS-CoV-2 main protease) were used as positive and negative controls respectively. The dose response effect of saxagliptin and sitagliptin against their selected targets (Fig 3A) were simulated using the one standard deviation variation of the IC<sub>50</sub> values<sup>20</sup>. This new approach to model the dose response curves of ligand based on the IC<sub>50</sub> values estimated from the data of molecular docking will be a valuable tool in the Insilco and network pharmacology<sup>15,16</sup> for drug evaluation or repurposing. Saxagliptin and sitagliptin showed high binding affinity against all the three SARS-CoV-2 surface proteins (Fig 3A, B). The Ki (6.32 – 22.04 μM) and IC<sub>50</sub> (12.63 – 113.71 μM) values of sitagliptin and saxagliptin against SARS-CoV-2 surface proteins were within the therapeutically acceptable range (Fig 3 C, D). The effectiveness of interaction between saxagliptin and sitagliptin with SARS-CoV-2 surface proteins was also evident by presence of number of hydrogen bonds between the ligand and the target molecule (supplementary figure 1). The DPP4 inhibitors also showed binding to 6Y2E (SARS-CoV-2 main protease), however this was at a very high concentration (IC<sub>50</sub> 318.97 – 455.75 μM). Although DPP4 was reported as a selective receptor for hCoV-EMC,<sup>2,3</sup> its inhibitors were unlikely to block hCoV-EMC infection due to irrelevance of the DPP4 mediated proteolytic activity in the host cell entry of the virus<sup>3</sup>. However in this study DPP4 inhibitors were observed to have similar binding efficacy to both DPP4 and SARS-CoV-2 surface proteins (Fig 3, supplementary

figure 1). Hence highlighting the potential of DPP4 inhibitors in preventing SARS-CoV-2 attachment and entry into the host cells. Further the analysis of the binding regions of the SARS-CoV-2 surface protein with ACE2 network proteins may be valuable in development of effective therapeutic, vaccine or diagnostic molecules.

## Methods

**Protein network analysis:** The ACE2 protein network was analysed using the STRING database (<https://string-db.org/cgi/info.pl>)<sup>16</sup>. The STRING database was searched using the ACE2 as protein name and Homo sapiens as organism. The first search was enhanced once to include additional nodes and 6 proteins were identified in the ACE2 network. All these 6 proteins (ACE2, DDP4, MEP1B, MEP1A, MME, PRCP and XPNPEP2) were membrane bound.

**SARS-CoV-2 surface proteins:** The reported SARS-CoV-2 target proteins were searched in the protein data bank (<https://www.rcsb.org/>) and uniprot database (<https://www.uniprot.org/peptidesearch/>)<sup>8</sup>. The following SARS-CoV-2 proteins were identified for binding analysis:

| PDB/Protein ID     | Brief Description                             |
|--------------------|---|
| P59596 (Protein M) | Membrane protein (Matrix glycoprotein)        |
| P59594 (Protein S) | Spike glycoprotein (S glycoprotein)           |
| P59637 (Protein E) | Envelope small membrane protein               |
| 6YLA               | Receptor binding domain on spike              |
| 6W41               | Spike glycoprotein in receptor binding domain |
| 6W75               | NSP10 - NSP16 Complex                         |
| 6W6Y               | ADP ribose phosphatase of NSP3                |
| 6LXT               | Post fusion core of S2 subunit                |
| 6LVN               | HR2 Domain                                    |
| 6VW1               | Receptor binding domain                       |
| 6M1D               | ACE2-B0AT1 complex, open conformation         |
| 6M17               | Receptor binding domain /ACE2-B0AT1 complex   |
| 6VXX               | Spike glycoprotein (closed state)             |
| 6VYB               | Spike glycoprotein (open state)               |
| 6M0J               | Spike receptor-binding domain                 |
| 6LZG               | Spike receptor-binding domain                 |

**Protein 3D structure and molecular docking:** The 3D structure of SARS-CoV-2 targets listed above were downloaded as PDB files from the protein data bank (<https://www.rcsb.org/>) and were optimized for molecular docking in the Chimera software<sup>11</sup>. The targets for which 3D structure was not available, the protein sequence (FASTA format) was downloaded from the uniprot database and the 3D

structure was constructed by homology modelling using the SWISS-MODEL server (<https://swissmodel.expasy.org/>).<sup>12</sup> For generating the 3D structures in the SWISS-MODEL server, each of the protein sequence in FASTA format was loaded into the SWISS-MODEL server<sup>12</sup> and the molecular modelling was initiated to generate PDB format of the protein. The interfaces of all the structures generated were refined with the Galaxy Refine Complex program installed in a computer with 2.66 GHz Intel Core 2 Duo processor, 4 GB 1067 MHz DDR3 RAM and with Mac OS X as the operating system. The 3D structural model generated was validated by measuring the probability of amino acids in the interface of the models. VADAR server was also used for the validation of 3D structure modelled by plotting Ramachandran plot<sup>8</sup>. The protein-protein interaction (between host cell receptor and SARS-CoV-2 targets) were assessed using the Chimera software<sup>8,12</sup>.

The structures of sitagliptin and saxagliptin were accessed from PubChem database and were processed into PDB file format and minimised for molecular docking using the Chimera software<sup>8,12</sup>. Molecular docking was performed to evaluate the binding efficacy of these compounds against the SARS-CoV-2 targets using AutoDock Vina (version 1.5.4) and the docked protein-ligand complex were visualised using the Chimera and PyMOL v 1.8.2.0 software<sup>11,17-19</sup>. AutoDock-MGLTools was employed to visualize and modify the receptor and ligand structures to PDBQT file formats. The PDBQT file formats of the ligand and receptor were used for molecular docking using the AutoDock Vina program. Specific grid size (x, y, and z) points, and the grid centre (x, y, and z dimensions) were set for each receptor type with a grid spacing of 1Å. An exhaustiveness setting of 30 was used in the AutoDock Vina for achieving best molecular docking result possible, with the centre of the matrix grid and the box was set at 27 Å on each side. The ligand was in torsion free condition for interacting with the receptor under fixed configuration. Polar hydrogen atoms and charges were included in the receptor structure. Ligands were docked individually to the receptor with grid coordinates and grid boxes of specific sizes for each receptor centralised in the AutoDock-MGLTools. The output file was saved in the PDBQT format and the ligand-receptor binding affinity estimated as negative Gibbs free energy ( $\Delta G$ ) scores (Kcal/mol), were documented based on AutoDock Vina scoring function. Post-docking analyses were visualized using PyMOL and Chimera, giving details of the sizes, locations of binding sites, hydrogen-bond interactions of the docked ligand in various confirmations.

**Simulation of dose response curves:** Dose-response curves were modelled based on nonlinear regression analysis approach by specifying IC<sub>50</sub> values as independent variable and response (% inhibition) as dependent variable. IC<sub>50</sub> values were estimated from the binding affinity values using the following formula  $IC_{50} = \exp(\Delta G/RT) (1 + ([S]/K_m))$ . Where  $\Delta G$  = Binding affinity (Kcal/mol),  $RT = 298K$ ,  $S$  = substrate concentration,  $K_m$ =Michaelis constant. The increase in the ligand dose increases results in sequential changes to the response (% receptor inhibition) eventually achieving minimum or maximum response limits. As the 20 to 80% response is linear, this was modelled using four-point logistic increments of IC<sub>50</sub> values (0.5x, 1x, 1.5x and 2x) and averaged. The resulting equation from this was  $y = 26.362x - 158.85$  ( $R^2 = 0.9967$ ). This one standard deviation increase and decrease of the IC<sub>50</sub> values was estimated from this equation. The resulted three IC<sub>50</sub> values estimated were used to calculate the IC 5, 20, 60, 80 and 100 values by employing a five-parameter logistic equation<sup>20</sup>. The means of the IC 5, 20, 60, 80 and 100 values obtained (in x axis) were plotted against the % inhibition response (in y axis) to obtain the simulated dose response curves.

**Conflict of interest:** none

**Acknowledgement:** Research support from University College Dublin-Seed funding/Output Based Research Support Scheme (AHSK), Royal Society-UK (AHSK) and Stemcology (AHSK) is acknowledged.

## References

1. Li, W., *et al.* Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450-454 (2003).
2. Muller, M.A., *et al.* Human coronavirus EMC does not require the SARS-coronavirus receptor and maintains broad replicative capability in mammalian cell lines. *mBio* **3**(2012).
3. Raj, V.S., *et al.* Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature* **495**, 251-254 (2013).
4. Cucinotta, D. & Vanelli, M. WHO Declares COVID-19 a Pandemic. *Acta Biomed* **91**, 157-160 (2020).
5. Fauci, A.S., Lane, H.C. & Redfield, R.R. Covid-19 - Navigating the Uncharted. *N Engl J Med* **382**, 1268-1269 (2020).
6. Beck, B.R., Shin, B., Choi, Y., Park, S. & Kang, K. Predicting commercially available antiviral drugs that may act on the novel coronavirus (SARS-CoV-2) through a drug-target interaction deep learning model. *Comput Struct Biotechnol J* **18**, 784-790 (2020).

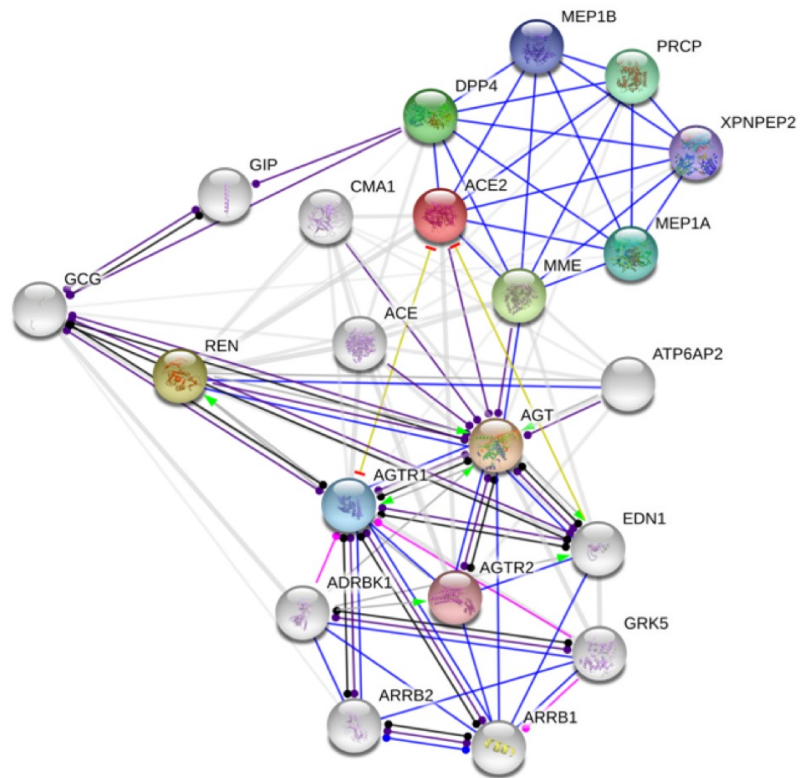
7. Choy, K.T., *et al.* Remdesivir, lopinavir, emetine, and homoharringtonine inhibit SARS-CoV-2 replication in vitro. *Antiviral Res* **178**, 104786 (2020).
8. Kumar, A.H.S. & Sharma, V. Acetamido-Propanoic Acid Derived Compounds as Protease Inhibitors to Target Coronaviruses. *BEMS Reports* **5**, 1-3 (2019).
9. Stower, H. Virological assessment of SARS-CoV-2. *Nat Med* **26**, 465 (2020).
10. Imai, Y., *et al.* Angiotensin-converting enzyme 2 protects from severe acute lung failure. *Nature* **436**, 112-116 (2005).
11. Yang, Z., *et al.* UCSF Chimera, MODELLER, and IMP: an integrated modeling system. *J Struct Biol* **179**, 269-278 (2012).
12. Bordoli, L., *et al.* Protein structure homology modeling using SWISS-MODEL workspace. *Nat Protoc* **4**, 1-13 (2009).
13. Huentelman, M.J., Zubcevic, J., Katovich, M.J. & Raizada, M.K. Cloning and characterization of a secreted form of angiotensin-converting enzyme 2. *Regul Pept* **122**, 61-67 (2004).
14. Kuba, K., *et al.* A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. *Nat Med* **11**, 875-879 (2005).
15. Guney, E., Menche, J., Vidal, M. & Barabasi, A.L. Network-based in silico drug efficacy screening. *Nat Commun* **7**, 10331 (2016).
16. Sharma, A., *et al.* A disease module in the interactome explains disease heterogeneity, drug response and captures novel pathways and genes in asthma. *Hum Mol Genet* **24**, 3005-3020 (2015).
17. O'Donoghue, S.I., *et al.* Visualization of macromolecular structures. *Nat Methods* **7**, S42-55 (2010).
18. Seeliger, D. & de Groot, B.L. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J Comput Aided Mol Des* **24**, 417-422 (2010).
19. Trott, O. & Olson, A.J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**, 455-461 (2010).
20. Ricketts, J.H. & Head, G.A. A five-parameter logistic equation for investigating asymmetry of curvature in baroreflex studies. *Am J Physiol* **277**, R441-454 (1999).

## Figure legends

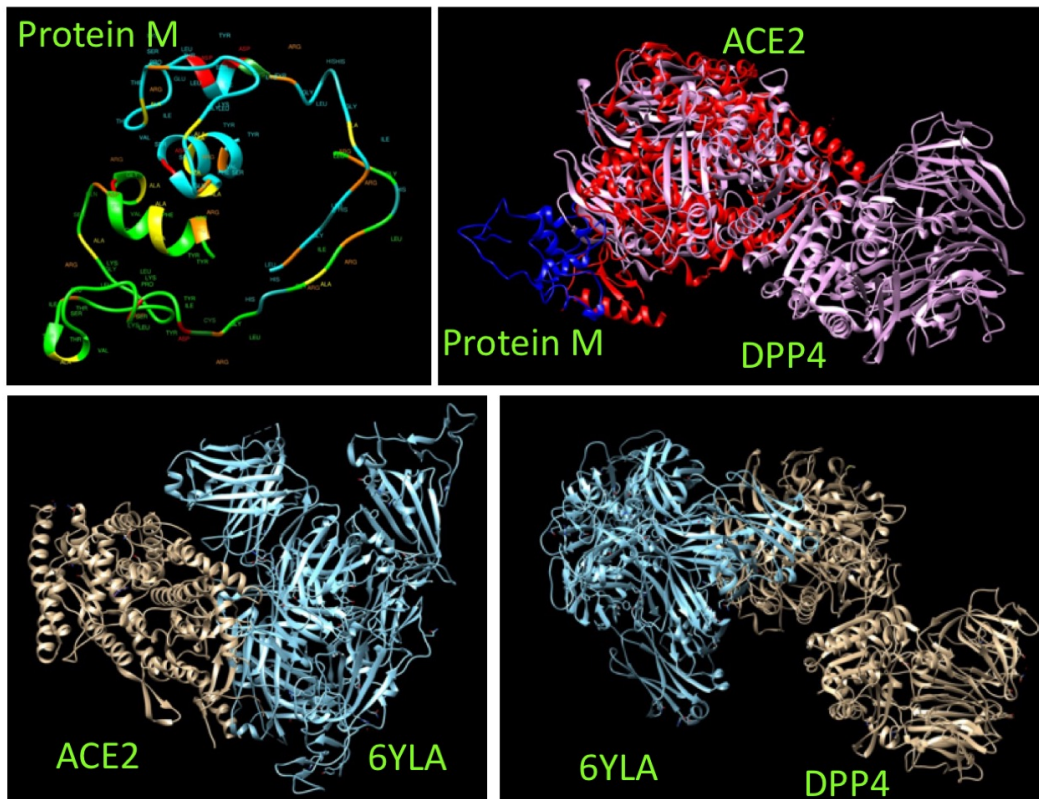
**Figure 1: ACE2 Network protein analysis and its interaction with SARS-COV-2 surface proteins.** **a**, six membrane associated proteins were identified in the human ACE2 network in the STRING database. All the associated proteins showed interaction score greater than 0.9, average node degree of 6.73, average local clustering coefficient of 0.844 and the PPI enrichment p-value of  $1.3 \times 10^{-09}$  suggesting significant interactions. **b**, Membrane bound Matrix glycoprotein (protein M) and Receptor binding domain on spike (6YLA) of SARS-CoV-2 were the only two virus surface proteins binding to human ACE2 (PDB ID 1r42). The binding interaction were analysed using the chimera software. The 3D structure of protein M was constructed by homology modelling using the SWISS-MODEL server and the amino acid labelled structure of protein M is presented. Both Protein M and 6YLA bound very weakly to the human ACE2 but were see binding very effectively (more region of interaction) to DPP4.

Figure 1

a

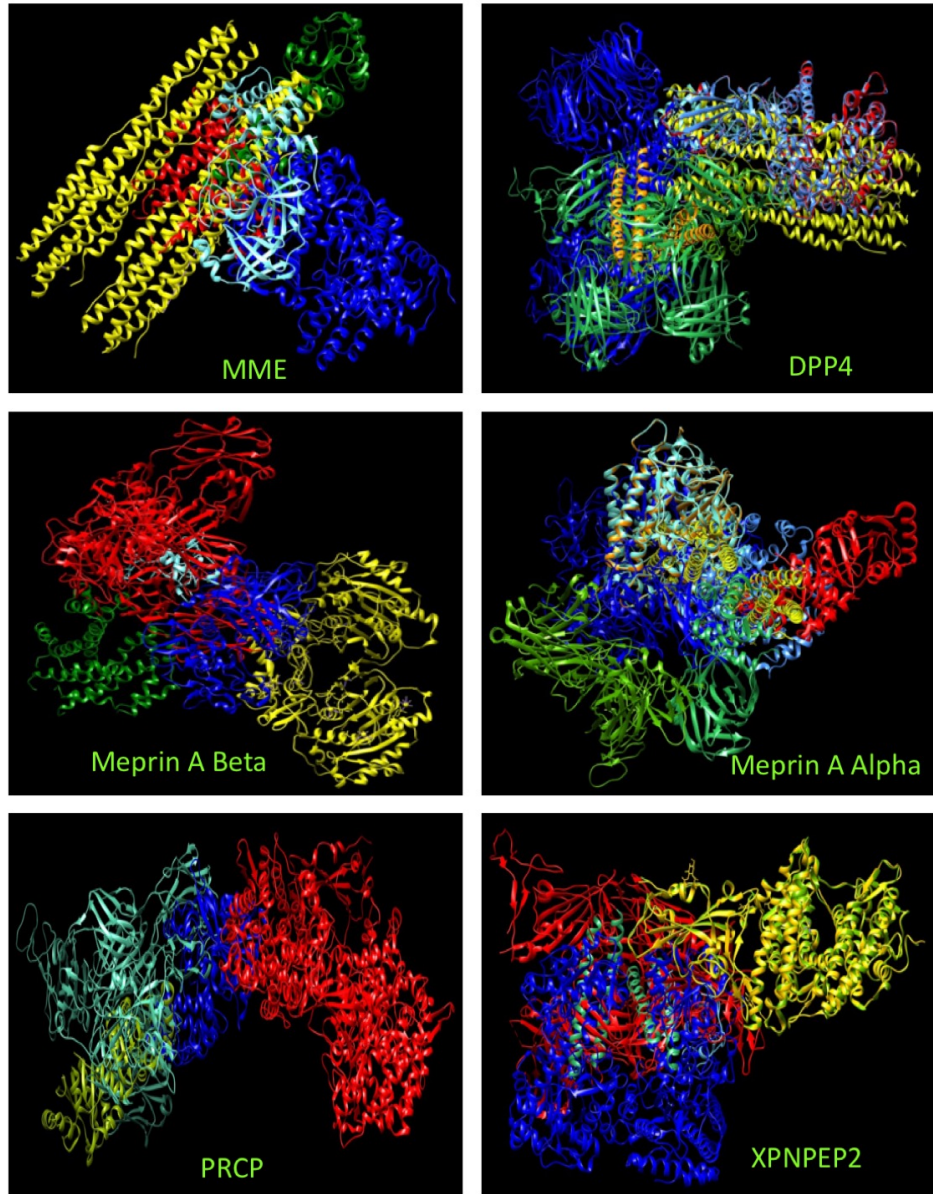


b

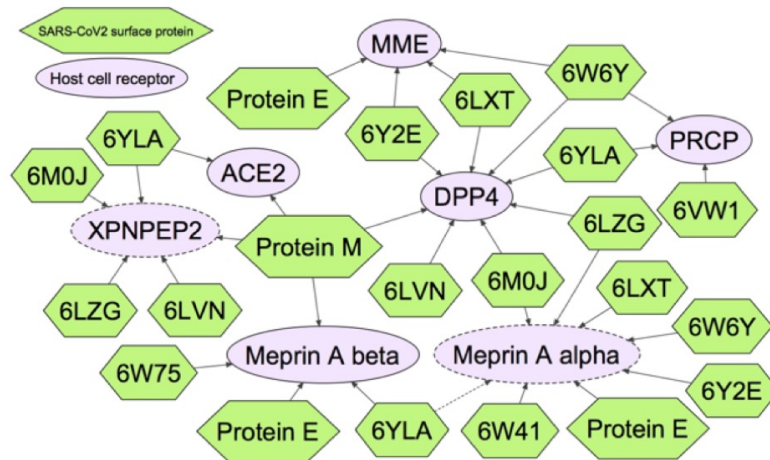


**Figure 2: Binding interaction of ACE2 network proteins with SARS-CoV-2 surface/target proteins.** **a**, Binding of 3D structures of ACE2 network proteins (ACE2, DDP4, MEP1B, MEP1A, MME, PRCP and XPNPEP2) with surface and key target proteins of SARS-CoV-2. The ACE2 network proteins are coloured in dark blue in all the panels, while the SARS-CoV-2 proteins are given different colour. The binding interaction were analysed using the chimera software. **b**, Summary of the binding interactions of SARS-CoV-2 proteins (green hexagons) with the ACE2 network proteins (light purple ovals). The ACE2 network proteins, whose 3D were constructed by homology modelling using the SWISS-MODEL server are indicated by dotted borders (MEP1A and XPNPEP2). The binding interaction is indicated by an arrow connecting the SARS-CoV-2 proteins (green hexagons) to the ACE2 network proteins (light purple ovals).

Figure 2 a

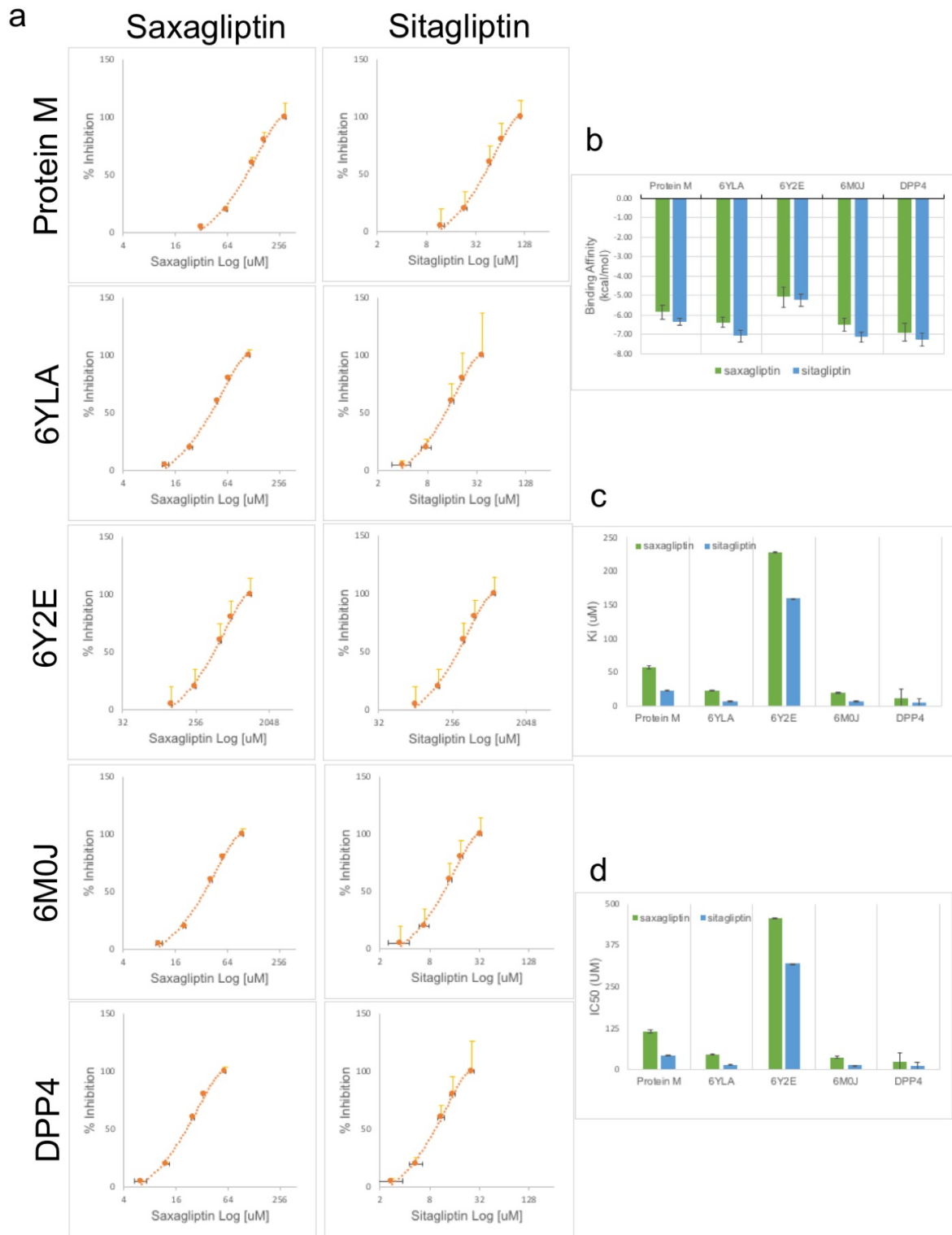


b



**Figure 3: Insilco pharmacology of DPP4 inhibitors on selected SARS-CoV-2 targets.** **a**, Dose response curves of DPP4 inhibitors (saxagliptin and sitagliptin) showing their percent inhibition of selected SARS-CoV-2 targets modelled by Insilco pharmacology. DPP4 and 6Y2E (SARS-CoV-2 main protease) were used as positive and negative controls respectively. **b, c, and d**, bar graphs showing the binding affinity (kcal/mol), inhibitor constant ( $K_i$ ,  $\mu\text{M}$ ) and  $\text{IC}_{50}$  values ( $\mu\text{M}$ ) respectively of DPP4 inhibitors (saxagliptin and sitagliptin) against selected SARS-CoV-2 targets. Data is represented as mean $\pm$ SD of nine different chemical confirmations of saxagliptin and sitagliptin.

Figure 3



**Supplementary figure 1: Binding of DPP4 inhibitors to selected SARS-CoV-2 targets.** Binding of DPP4 inhibitors (saxagliptin and sitagliptin) to selected SARS-CoV-2 targets and DPP4 analysed by molecular docking using AutoDock Vina. The nine different structural confirmations (red spheres or sticks) of the saxagliptin and sitagliptin are shown bound to the SARS-CoV-2 targets (cyan). The molecular docking with the protein M shows only one structural confirmation of saxagliptin and sitagliptin as all the confirmations were observed to bind to a single location on the protein M. The hydrogen bonds involved in the molecular interactions are also shown as yellow lines.

Supplementary figure 1

