doi.org/10.36721/PJPS.2022.35.6.REG.1531-1538.1

Structure-based virtual screening and molecular docking of drugs against the SARS-CoV-2 spike protein-ACE2 receptor complex

Irshad Ahmad[1], Muhammad Ali[1], Roshan Ali[1], Nighat Nawaz[2]* and Simon G Patching[3]
[1]Institute of Basic Medical Sciences, Khyber Medical University, Peshawar, Pakistan
[2]Department of Chemistry, Islamia College Peshawar, Peshawar, Pakistan
[3]School of Biomedical Sciences, University of Leeds, Leeds, UK

Abstract: The agent responsible for the COVID-19 pandemic was the newly discovered coronavirus SARS-CoV-2. A trimeric spike protein on the SARS-CoV-2 virion binds to the ACE2 receptor on host cells. In this study we performed a structure-based virtual screening and molecular docking of existing drugs against a high-resolution structure of the SARS-CoV-2 spike protein-ACE2 receptor complex. The 2.5-Å crystal structure of the C-terminal domain of the SARS-CoV-2 spike protein (residues 319-541) in complex with human ACE2 (SARS-CoV-2-S-CTD/hACE2) (PDB ID: 6LZG) was used as the target for screening 4,374 FDA-approved drugs from the ZINC15 database using PyRx software. Molecular docking was performed using BIOVIA Discovery Studio Visualizer. The top twenty highest affinity drugs had binding energies of -7.0 to -8.8 kcal/mol. The highest affinity drug was the selective vasopressin V2-receptor antagonist Tolvaptan, for which molecular docking identified drug-amino acid residue interactions with ACE2. Other drugs displaying binding energies better than -8.0 kcal/mol were Nizoral, Amaryl, Accolate, Sorafenib, Glipizide and Azelastine. The predicted interactions of these highest affinity drugs with residues in ACE2 were at positions that could disrupt the spike protein-ACE2 complex, so these drugs have the potential to be repurposed as inhibitors of the SARS-CoV-2 virus.

Keywords: Coronavirus, drug discovery, infections, molecular docking simulation, therapeutics.

INTRODUCTION

During the month of December 2019 a pneumonia-type infection of unknown origin was noticed in the city of Wuhan, China (Zhu *et al*., 2020). In early January 2020 the local Center for Disease Control launched an emergency response to this potentially lethal infection (Li *et al*., 2020). The agent causing this infection was identified as a novel β-coronavirus, first named 2019-nCoV and then SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) (Coronaviridae Study Group of the International Committee on Taxonomy of Viruses, 2020). The SARS-CoV-2 virus causes the disease referred to as COVID-19. The genome sequence of SARS-CoV-2 (GenBank: MN908947) was initially made available from China on 10th January 2020 (Guo *et al*., 2020; Lu *et al*., 2020). On 30th January 2020 the World Health Organisation (WHO) announced a "public health emergency of international concern" (Contini *et al*., 2020). The SARS-CoV-2 virus subsequently spread to create a global pandemic with devastating health and socio-economic consequences. As of 1st February 2022, 223 countries and territories had reported a total of 380,947,112 confirmed cases with 5,699,059 related deaths (https://www.worldometers.info/coronavirus/countries-where-coronavirus-has-spread/).

The discovery of SARS-CoV-2 triggered a remarkable response from scientists to develop testing strategies and vaccines for the virus (Golob *et al*., 2021; Ling *et al*., 2021; Fiolet *et al*., 2022; Hadj Hassine, 2022). The development of novel vaccines and other therapeutics to control the spread and minimise the effects of SARS-CoV-2 infection requires a rigorous understanding of the virus structure and how it interacts and functions at the cellular, molecular and atomic level. This led to the elucidation of hundreds of protein structures from the SARS-CoV-2 virus using the techniques of cryo-electron microscopy, cryo-electron tomography, X-ray crystallography and NMR spectroscopy (Hardenbrook and Zhang, 2021). The SARS-CoV-2 virion has a roughly spherical or ellipsoidal shape with an average diameter of 108 ± 8 nm. In common with other coronaviruses, the SARS-CoV-2 virion comprises four main structural proteins: spike protein (S), small envelope protein (E), membrane protein (M), nucleocapsid protein (N). It also has the non-structural proteins Nsp1-16. The S, E and M proteins together form the viral envelope and the N protein is found in the virion matrix (fig. 1).

The S protein (1273 residues, 141.2 kDa) (https://www.uniprot.org/uniprot/P0DTC2) forms homotrimers that appear as a flexible head on a stalk that protrude from the virus surface and facilitate binding of the virus to host cells by binding to the angiotensin-converting enzyme 2 (ACE2) receptor (Lan *et al*., 2020; Shan *et al*., 2020). The S protein contains an N-terminal S1 subunit (14-685 residues) and a C-terminal S2 subunit (686-1273 residues) (fig. 1). An individual SARS-CoV-2 virion contains an estimated 24 ± 9 S protein trimers (Ke *et al*., 2020). The S1 subunit comprises an N-terminal domain (NTD), a

*Corresponding author: e-mail: nnawaz@mail.com

receptor-binding domain (RBD) and C-terminal domains (CTD1 and CTD2). The head of S1 is the part involved in receptor binding. The S2 subunit comprises an internal membrane fusion peptide, two heptapeptide repeat sequences (HR1 and HR2), a membrane-proximal external region and a transmembrane domain (TM) that constitutes a single transmembrane helix. S2 assists transfer of the genome into host cells by fusion of the host and viral membranes. The S protein contains heavy glycosylation where each protomer possesses 22 N-linked glycosylation sites (Zhang *et al*., 2021).

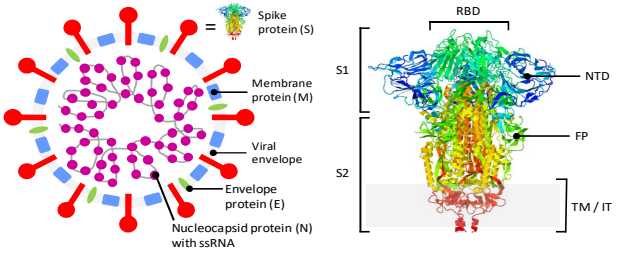

Fig. 1: Structural components of the SARS-CoV-2 virion and structure of the spike protein. Left. Cartoon representation of the major structural components of the SARS-CoV-2 virion. Right. High-resolution crystal structure of the SARS-CoV-2 spike protein (PDB ID: 7DF3) (Xu *et al*., 2021) with rainbow colouring (N-terminus blue, C-terminus red). The protein comprises two subunits (S1 and S2). S1 contains the receptor binding domain (RBD) that binds to the ACE2 receptor and contains the N-terminal domain (NTD). S2 contains the fusion protein (F2), the transmembrane (TM) segment that anchors the protein to the viral envelope and the intracellular tail (IT). The spike protein image was produced from the PDB entry using Jmol: an open-source Java viewer for chemical structures in 3D (http://www.jmol.org/).

The M protein (222 residues, 25.1 kDa) (https://www.uniprot.org/uniprot/P0DTC5) is the most abundant of the structural proteins, it contains a three-helix transmembrane bundle that resembles the sugar transporter semiSWEET (Thomas, 2020) and it functions as a homodimer. The E protein (75 residues, 8.4 kDa) (https://www.uniprot.org/ uniprot/P0DTC4) has a single transmembrane helix that functions as a structural protein in the viral capsid and it also self-assembles in host membranes to form pentameric protein-lipid pores that allow ion transport (Mandala *et al*., 2020). The N protein (419 residues, 45.6 kDa) (https://www.uniprot.org/uniprot/P0DTC9) packages the genomic RNA of the virus into a helical ribonucleocapsid and participates in assembling the virion through interacting with the viral genome and the M protein (Bai *et al*., 2021).

Because of the critical role that the S protein has in virus-host receptor interaction, the S protein has been the major molecular target for developing vaccines and therapeutics against the SARS-CoV-2 virus (Huang *et al*., 2020; Wang *et al*., 2020; Dai and Gao, 2021; Xia, 2021). Some of the high-resolution structures that have been determined for SARS-CoV-2 are in complex with human ACE2 (Lan *et al*., 2020; Wang *et al*., 2020; Yan *et al*., 2020; Xu *et al*., 2021), thus providing a base for interrogation of the key virus-host recognition and binding interaction. Indeed, various drug screening strategies have been conducted to identify inhibitors of the S protein-ACE2 interaction, especially through the repurposing of existing drugs (Acharya *et al*., 2021; Farouk *et al*., 2021; Mohebbi *et al*., 2021; Tsegay *et al*., 2021; Chen *et al*., 2022; Yang *et al*., 2022).

The objective of the present study was to perform a structure-based virtual screening of 4,374 FDA-approved drugs from the ZINC15 database against a high-resolution crystal structure of the C-terminal domain of the SARS-CoV-2 S protein in complex with human ACE2. Then to perform a molecular docking analysis for the highest affinity drugs in order to identify potential inhibitors of the S protein-ACE2 interaction.

MATERIALS AND METHODS

Target structure selection and analysis

The high-resolution three-dimensional structure of the C-terminal domain of the SARS-CoV-2 spike protein (residues 319-541) in complex with human ACE2 (2.5 Å crystal structure) (SARS-CoV-2-S-CTD/hACE2 complex) was downloaded from the RCSB Protein Data Bank (PDB) under the accession code of 6LZG (https://www.rcsb.org/structure/6LZG) (Wang *et al*., 2020). In this structure, distinct electron densities are observed for 596 hACE2 residues (S19 to A614) of the N-terminal peptidase domain and 195 SARS-CoV-2-CTD residues spanning T333 to P527. The structure illustrates how the SARS-CoV-2-CTD external subdomain recognises subdomain I in the hACE2 N-terminal domain.

The active binding site residues of the SARS-CoV-2-S-CTD/hACE2 complex were identified from the 6LZG structure using BIOVIA Discovery Studio Visualizer (https://discover.3ds.com/discovery-studio-visualizer-download), the PDB entry, ProBis (http://probis.cmm.ki.si) (Konc and Janezic, 2010) and from literature study. From the common identified binding site residues the most important and functional residues that were revealed in the literature were selected as target residues for virtual screening. Analysis of the binding site of the SARS-CoV-2-S-CTD/hACE2 complex was performed using BIOVIA Discovery Studio Visualizer.

Virtual drug screening and analysis

The FDA approved drug library of 4,374 molecules was downloaded from the ZINC15 online resource (https://zinc15.docking.org) (Sterling and Irwin, 2015).

These molecules were screened against the crystal structure of the SARS-CoV-2-S-CTD/hACE2 complex (PDB ID: 6LZG) by employing the PyRx drug screening program (https://pyrx.sourceforge.io) (Dallakyan and Olson, 2015). The interface between the docked protein complex and ligand molecules was analysed by BIOVIA Discovery Studio Visualizer

RESULTS

Out of the 4,374 FDA-approved drugs from the ZINC15 database that were screened against the SARS-CoV-2-S-CTD/hACE2 complex, the top twenty highest affinity drugs are shown in table 1 and the chemical structures of these are shown in fig. 2. These had binding energies in the range -7.0 to -8.8 kcal/mol compared with the reference molecule Pavetannin C1, which had a binding energy of -7.7 kcal/mol. The highest affinity drug was the selective vasopressin V2-receptor antagonist Tolvaptan.

The model of the SARS-CoV-2-S-CTD/hACE2 complex with docked Tolvaptan is shown in fig. 3A. This identified a number of drug-amino acid residue interactions with ACE2, including hydrogen bonds with His345, His374, His378, Ala348 and Asp350 and pi-anion interactions with His378, Asp382 and Glu402 (fig. 3B). Residue interactions in the SARS-CoV-2-CTD/hACE2 complex and its interactions with Tolvaptan are shown in fig. 4.

Table 1: Top twenty highest affinity drugs binding to the SARS-CoV-2-CTD/hACE2 complex from virtual screening. These include the reference molecule Pavetannin C1 (No. 11).

No.	Drug name Ligand code	Binding energy (kcal/mol)	Mol. formula	Heavy atoms	Mol. weight	LogP
1	Tolvaptan ZINC538658	-8.8	$C_{26}H_{25}ClN_2O_3$	32	449.0	5.683
2	Nizoral ZINC643153	-8.7	$C_{26}H_{28}Cl_2N_4O_4$	36	531.4	4.206
3	Amaryl ZINC537791	-8.6	$C_{24}H_{34}N_4O_5S$	34	490.6	3.278
4	Accolate ZINC896717	-8.4	$C_{31}H_{33}N_3O_6S$	41	575.7	6.271
5	Sorafenib ZINC1493878	-8.2	$C_{21}H_{16}ClF_3N_4O_3$	32	464.8	5.550
6	Glipizide ZINC537795	-8.2	$C_{21}H_{27}N_5O_4S$	31	445.5	2.283
7	Azelastine ZINC601229	-8.1	$C_{22}H_{24}ClN_3O$	27	381.9	4.298
8	Haloperidol ZINC537822	-7.8	$C_{21}H_{23}ClFNO_2$	26	375.9	4.426
9	Thalitone ZINC57255	-7.8	$C_{14}H_{11}ClN_2O_4S$	22	338.8	0.924
10	Anzemet ZINC2688	-7.8	$C_{19}H_{20}N_2O_3$	24	324.4	2.519
11	Pavetannin C1 PubChem CID 16165472	-7.7	$C_{60}H_{48}O_{24}$	84	1153.0	4.300
12	Methotrexate ZINC1529323	-7.7	$C_{20}H_{22}N_8O_5$	33	454.4	0.268
13	Isavuconazole ZINC1485935	-7.7	$C_{22}H_{17}F_2N_5OS$	31	437.5	4.243
14	Thalitone ZINC20253	-7.6	$C_{14}H_{11}ClN_2O_4S$	22	338.8	0.924
15	Labetalol ZINC403010	-7.5	$C_{19}H_{24}N_2O_3$	24	328.4	2.135
16	Aerius ZINC1261	-7.4	$C_{19}H_{19}ClN_2$	22	310.8	4.019
17	Ofloxacin ZINC537891	-7.3	$C_{18}H_{20}FN_3O_4$	26	361.4	1.544
18	Ofloxacin ZINC538273	-7.2	$C_{18}H_{20}FN_3O_4$	26	361.4	1.544
19	Lodine ZINC57313	-7.1	$C_{17}H_{21}NO_3$	21	287.4	3.383
20	Triamteril ZINC120286	-7.0	$C_{12}H_{11}N_7$	19	253.3	0.833

In addition to Tolvaptan, another six drugs (Nizoral, Amaryl, Accolate, Sorafenib, Glipizide, Azelastine) had binding energies better than -8.0 kcal/mol (table 1). Interactions between these virtually docked drugs and the structure of the SARS-CoV-2-CTD/hACE2 complex are shown in fig. 5. Residue interactions in the SARS-CoV-2-CTD/hACE2 complex and its interactions with these six docked drugs are shown in fig. 6.

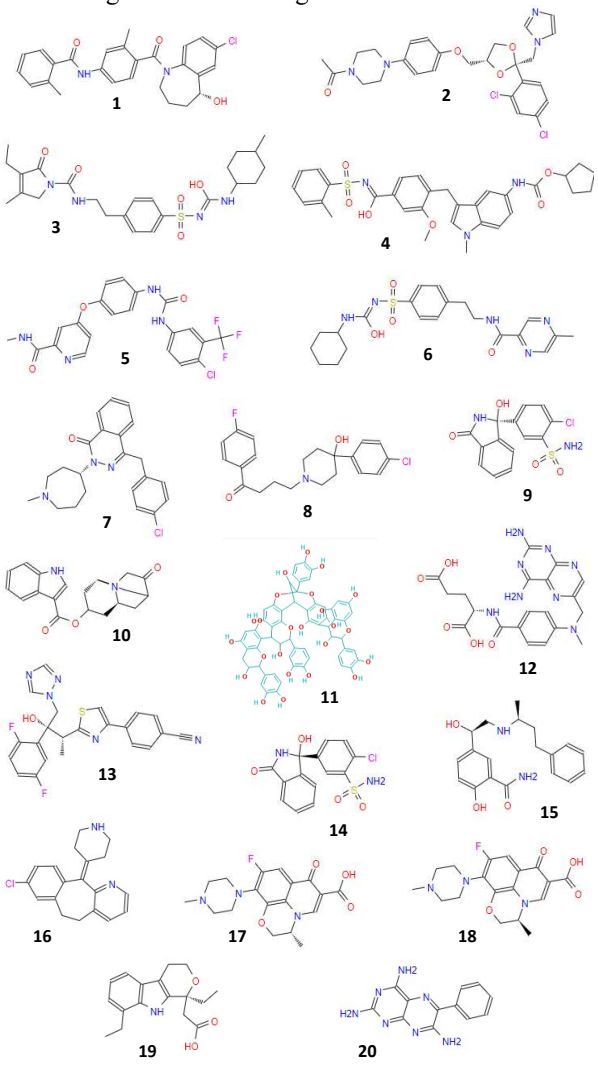

Fig. 2: Chemical structures of the top twenty highest affinity drugs binding to the SARS-CoV-2-CTD/hACE2 complex from virtual screening.

DISCUSSION

The high-resolution crystal structure of the SARS-CoV-2-S-CTD/hACE2 complex (PDB ID: 6LZG) was chosen as the target for structure-based virtual drug screening so that it focuses on the key virus-host receptor interactions. The residues that show clear electron densities in this crystal structure are 195 SARS-CoV-2 spike protein residues spanning T333 to P527 and 596 hACE2 residues spanning S19 to A614 of the N-terminal peptidase domain. The structure illustrates how the SARS-CoV-2-S-CTD external subdomain recognises subdomain I in the ACE2 N-terminal domain.

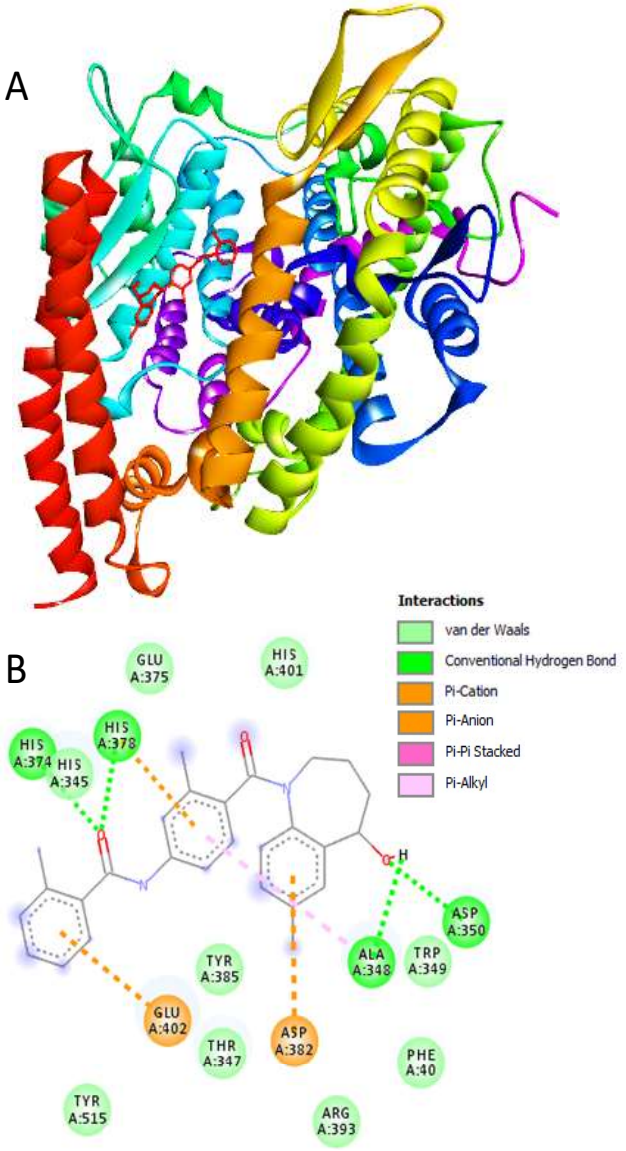

Fig. 3: Interactions between the virtually docked drug Tolvaptan (ZINC538658) and the structure of the SARS-CoV-2-CTD/hACE2 complex. A. Three-dimensional model. B. Two-dimensional drug-amino acid residue interactions. The images were produced using BIOVIA Discovery Studio Visualizer.

From the drugs screened against the SARS-CoV-2-S-CTD/hACE2 complex, the highest affinity drug Tolvaptan (brand name Samsca) is used as a treatment for hyponatremia related to congestive heart failure, cirrhosis and the syndrome of inappropriate antidiuretic hormone (SIADH) (Brunton *et al*., 2018; Blair, 2019). Tolvaptan (IUPAC name: N-[4-(7-chloro-5-hydroxy-2,3,4,5-tetrahydro-1-benzazepine-1-carbonyl)-3-methylphenyl]-2 methylbenzamide) is benzazepine derivative with a

benzenedicarboxamide function that is administered as a racemate. The model of the SARS-CoV-2-S-CTD/hACE2 complex with docked Tolvaptan identified drug-amino acid residue interactions with ACE2, including hydrogen bonds with His345, His374, His378, Ala348 and Asp350 and pi-anion interactions with His378, Asp382 and Glu402 (fig. 3B).

SARS-CoV-2 spike protein (P0DTC2)

```
LLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK 300
CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISN 360
CVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIAD 420
YNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPC 480
NGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVN 540
FNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITP 600
```

Human ACE2 receptor (Q9BYF1)

```
MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQ  60
NMNNAGDKWSAFLKEQSTLAQMYPLQEIQNLTVKLQLQALQQNGSSVLSEDKSKRLNTIL 120
NTMSTIYSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNERLWAWESWRSEVGKQLRPLY 180
EEYVVLKNEMARANHYEDYGDYWRGDYEVNGVDGYDYSRGQLIEDVEHTFEEIKPLYEHL 240
HAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWTNLYSLTVPFGQKPNIDVTDAMVDQ 300
AWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILM 360
CTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS 420
IGLLSPDFQEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMVFKGEIPKDQWMKKWWEM 480
KREIVGVVEPVPHDETYCDPASLFHVSNDYSFIRYYTRTLYQFQFQEALCQAAKHEGPLH 540
KCDISNSTEAGQKLFNMLRLGKSEPWTLALENVVGAKNMNVRPLLNYFEPLFTWLKDQNK 600
NSFVGWSTDWSPYADQSIKVRISLKSALGDKAYEWNDNEMYLFRSSVAYAMRQYFLKVKN 660
QMILFGEEDVRVANLKPRISFNFFVTAPKNVSDIIPRTEVEKAIRMSRSRINDAFRLNDN 720
SLEFLGIQPTLGPPNQPPVSIWLIVFGVVMGVIVVGIVILIFTGIRDRKKKNKARSGENP 780
YASIDISKGENNPGFQNTDDVQTSF 805
```

Fig. 4: Residue interactions in the SARS-CoV-2-CTD/hACE2 complex and its interactions with the docked drug Tolvaptan. In the sequences of the SARS-CoV-2 spike protein (P0DTC2) (top) and the human ACE2 receptor (Q9BYF1) (bottom), residues that show clear electron density in the crystal structure of the SARS-CoV-2-CTD/hACE2 complex (PDB ID: 6LZG) (www.rcsb.org/structure/6LZG) (Wang *et al*., 2020) are highlighted (grey). These residues are T333 to P527 in the SARS-CoV-2 spike protein and S19 to A614 in hACE2. Residues that are involved in strong polar contacts (cyan) and hydrophobic interactions (green) between the SARS-CoV-2 spike protein CTD and hACE2 are highlighted. Also highlighted are residues involved in interactions with the virtually docked drug Tolvaptan (ZINC538658) (red).

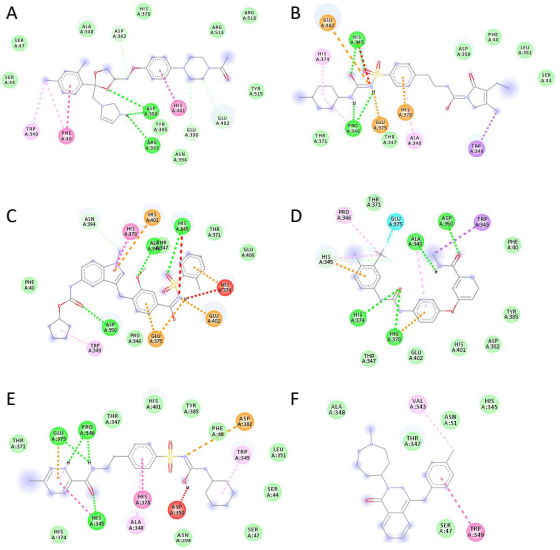

Fig. 5: Interactions between virtually docked drugs and the structure of the SARS-CoV-2-CTD/hACE2 complex. Two-dimensional drug-amino acid residue interactions for A. Nizoral (ZINC643153), B. Amaryl (ZINC537791), C. Accolate (ZINC896717), D. Sorafenib (ZINC1493878), E. Glipizide (ZINC537795) and F. Azelastine (ZINC601229). Images produced by BIOVIA Discovery Studio Visualizer.

Nizoral (ZINC643153)

```
MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQ  60
NMNNAGDKWSAFLKEQSTLAQMYPLQEIQNLTVKLQLQALQQNGSSVLSEDKSKRLNTIL 120
NTMSTIYSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNERLWAWESWRSEVGKQLRPLY 180
EEYVVLKNEMARANHYEDYGDYWRGDYEVNGVDGYDYSRGQLIEDVEHTFEEIKPLYEHL 240
HAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWTNLYSLTVPFGQKPNIDVTDAMVDQ 300
AWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILM 360
CTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS 420
IGLLSPDFQEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMVFKGEIPKDQWMKKWWEM 480
KREIVGVVEPVPHDETYCDPASLFHVSNDYSFIRYYTRTLYQFQFQEALCQAAKHEGPLH 540
KCDISNSTEAGQKLFNMLRLGKSEPWTLALENVVGAKNMNVRPLLNYFEPLFTWLKDQNK 600
NSFVGWSTDWSPYADQSIKVRISLKSALGDKAYEWNDNEMYLFRSSVAYAMRQYFLKVKN 660
```

Amaryl (ZINC537791)

```
MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQ  60
NMNNAGDKWSAFLKEQSTLAQMYPLQEIQNLTVKLQLQALQQNGSSVLSEDKSKRLNTIL 120
NTMSTIYSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNERLWAWESWRSEVGKQLRPLY 180
EEYVVLKNEMARANHYEDYGDYWRGDYEVNGVDGYDYSRGQLIEDVEHTFEEIKPLYEHL 240
HAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWTNLYSLTVPFGQKPNIDVTDAMVDQ 300
AWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILM 360
CTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS 420
IGLLSPDFQEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMVFKGEIPKDQWMKKWWEM 480
KREIVGVVEPVPHDETYCDPASLFHVSNDYSFIRYYTRTLYQFQFQEALCQAAKHEGPLH 540
KCDISNSTEAGQKLFNMLRLGKSEPWTLALENVVGAKNMNVRPLLNYFEPLFTWLKDQNK 600
NSFVGWSTDWSPYADQSIKVRISLKSALGDKAYEWNDNEMYLFRSSVAYAMRQYFLKVKN 660
```

Accolate (ZINC896717)

```
MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQ  60
NMNNAGDKWSAFLKEQSTLAQMYPLQEIQNLTVKLQLQALQQNGSSVLSEDKSKRLNTIL 120
NTMSTIYSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNERLWAWESWRSEVGKQLRPLY 180
EEYVVLKNEMARANHYEDYGDYWRGDYEVNGVDGYDYSRGQLIEDVEHTFEEIKPLYEHL 240
HAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWTNLYSLTVPFGQKPNIDVTDAMVDQ 300
AWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILM 360
CTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS 420
IGLLSPDFQEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMVFKGEIPKDQWMKKWWEM 480
KREIVGVVEPVPHDETYCDPASLFHVSNDYSFIRYYTRTLYQFQFQEALCQAAKHEGPLH 540
KCDISNSTEAGQKLFNMLRLGKSEPWTLALENVVGAKNMNVRPLLNYFEPLFTWLKDQNK 600
NSFVGWSTDWSPYADQSIKVRISLKSALGDKAYEWNDNEMYLFRSSVAYAMRQYFLKVKN 660
```

Sorafenib (ZINC1493878)

```
MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQ  60
NMNNAGDKWSAFLKEQSTLAQMYPLQEIQNLTVKLQLQALQQNGSSVLSEDKSKRLNTIL 120
NTMSTIYSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNERLWAWESWRSEVGKQLRPLY 180
EEYVVLKNEMARANHYEDYGDYWRGDYEVNGVDGYDYSRGQLIEDVEHTFEEIKPLYEHL 240
HAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWTNLYSLTVPFGQKPNIDVTDAMVDQ 300
AWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILM 360
CTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS 420
IGLLSPDFQEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMVFKGEIPKDQWMKKWWEM 480
KREIVGVVEPVPHDETYCDPASLFHVSNDYSFIRYYTRTLYQFQFQEALCQAAKHEGPLH 540
KCDISNSTEAGQKLFNMLRLGKSEPWTLALENVVGAKNMNVRPLLNYFEPLFTWLKDQNK 600
NSFVGWSTDWSPYADQSIKVRISLKSALGDKAYEWNDNEMYLFRSSVAYAMRQYFLKVKN 660
```

Glipizide (ZINC537795)

```
MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQ  60
NMNNAGDKWSAFLKEQSTLAQMYPLQEIQNLTVKLQLQALQQNGSSVLSEDKSKRLNTIL 120
NTMSTIYSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNERLWAWESWRSEVGKQLRPLY 180
EEYVVLKNEMARANHYEDYGDYWRGDYEVNGVDGYDYSRGQLIEDVEHTFEEIKPLYEHL 240
HAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWTNLYSLTVPFGQKPNIDVTDAMVDQ 300
AWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILM 360
CTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS 420
IGLLSPDFQEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMVFKGEIPKDQWMKKWWEM 480
KREIVGVVEPVPHDETYCDPASLFHVSNDYSFIRYYTRTLYQFQFQEALCQAAKHEGPLH 540
KCDISNSTEAGQKLFNMLRLGKSEPWTLALENVVGAKNMNVRPLLNYFEPLFTWLKDQNK 600
NSFVGWSTDWSPYADQSIKVRISLKSALGDKAYEWNDNEMYLFRSSVAYAMRQYFLKVKN 660
```

Azelastine (ZINC601229)

```
MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQ  60
NMNNAGDKWSAFLKEQSTLAQMYPLQEIQNLTVKLQLQALQQNGSSVLSEDKSKRLNTIL 120
NTMSTIYSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNERLWAWESWRSEVGKQLRPLY 180
EEYVVLKNEMARANHYEDYGDYWRGDYEVNGVDGYDYSRGQLIEDVEHTFEEIKPLYEHL 240
HAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWTNLYSLTVPFGQKPNIDVTDAMVDQ 300
AWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILM 360
CTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS 420
IGLLSPDFQEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMVFKGEIPKDQWMKKWWEM 480
KREIVGVVEPVPHDETYCDPASLFHVSNDYSFIRYYTRTLYQFQFQEALCQAAKHEGPLH 540
KCDISNSTEAGQKLFNMLRLGKSEPWTLALENVVGAKNMNVRPLLNYFEPLFTWLKDQNK 600
NSFVGWSTDWSPYADQSIKVRISLKSALGDKAYEWNDNEMYLFRSSVAYAMRQYFLKVKN 660
```

Fig. 6: Residue interactions in the SARS-CoV-2-CTD/hACE2 complex and its interactions with docked

drugs. In the sequence of the human ACE2 receptor (Q9BYF1), residues that show clear electron density in the crystal structure of the SARS-CoV-2-CTD/hACE2 complex (PDB 6LZG) (www.rcsb.org/structure/6LZG) (Wang *et al*., 2020) are highlighted (grey). Residues that are involved in strong polar contacts (cyan) and hydrophobic interactions (green) with the SARS-CoV-2 spike protein CTD are highlighted. Also highlighted are residues involved in interactions with the virtually docked drugs Nizoral (ZINC643153), Amaryl (ZINC537791), Accolate (ZINC896717), Sorafenib (ZINC1493878), Glipizide (ZINC537795) and Azelastine (ZINC601229) (red).

If the interactions between the drug and ACE2 are at locations close to the interactions that bind the SARS-CoV-2-S-CTD/hACE2 complex, then the drug has the potential to be an inhibitor of the SARS-CoV-2 virus. According to the 6LZG structure, the key interactions that stabilise the SARS-CoV-2-S-CTD/hACE2 complex involve strong polar contacts between eleven residues in SARS-CoV-2 (Ala475, Asn487, Glu484, Tyr453, Lys417, Gly446, Tyr449, Gly496, Gln498, Thr500, Gly502) and ten residues in ACE2 (Ser19, Gln24, Lys31, His34, Asp30, Asp38, Tyr41, Gln42, Lys353, Asp355) and hydrophobic contacts between two residues in the SARS-CoV-2 spike protein (Tyr489, Phe486) and four residues in ACE2 (Phe28, Lue79, Met82, Tyr83) (fig. 4). The residues in the SARS-CoV-2-CTD/hACE2 complex involved in interactions with the virtually docked drug Tolvaptan are the following residues in ACE2: Phe40, His345, Thr347, Ala348, Trp349, Asp350, His374, Glu375, His378, Asp382, Tyr385, Arg393, His401, Glu402, Tyr515 (fig. 3B and fig. 4). Some of the residues interacting with Tolvaptan are close to residues that bind the SARS-CoV-2-S-CTD/hACE2 complex. Tolvaptan could therefore disrupt the SARS-CoV-2-S-CTD/hACE2 complex and therefore serve as an inhibitor of the SARS-CoV-2 virus. Similarly, all of the other highest affinity drugs with binding energies better than -8.0 kcal/mol (Nizoral, Amaryl, Accolate, Sorafenib, Glipizide, Azelastine) (table 1), are involved in interactions with ACE2 that are close to residues that bind the SARS-CoV-2-S-CTD/hACE2 complex (fig. 5 and fig. 6). These drugs also have potential to serve as inhibitors of the SARS-CoV-2 virus.

CONCLUSION

The present virtual screening study has identified some FDA-approved drugs that could be repurposed as inhibitors of the SARS-CoV-2 virus by targeting the spike protein-ACE2 interaction, the highest affinity ones being Tolvaptan, Nizoral, Amaryl, Accolate, Sorafenib, Glipizide and Azelastine. Future work to realise such drugs as therapeutics would need to involve *In vitro* testing of the drugs against infected cell lines to find out their activity against the virus and *In vivo* testing of the drugs for their ability to prevent infection or reduce the effects of infection of the virus on animals. Further investigation of the interaction between the drug molecule and the SARS-CoV-2-S-CTD/hACE2 complex should also be performed. The FDA-approved drugs from the ZINC15 database that were used in the present study could be screened against high-resolution structures of the SARS-CoV-2 envelope, membrane and nucleocapsid proteins.

REFERENCES

Acharya A, Pandey K, Thurman M, Klug E, Trivedi J, Sharma K, Lorson CL, Singh K and Byrareddy SN (2021). Discovery and evaluation of entry inhibitors for SARS-CoV-2 and its emerging variants. *J. Virol*., **95**(24): e0143721.

Bai Z, Cao Y, Liu W and Li J (2021). The SARS-CoV-2 nucleocapsid protein and its role in viral structure, biological functions and a potential target for drug or vaccine mitigation. *Viruses*, **13**(6): 1115.

Blair HA (2019). Tolvaptan: A review in autosomal dominant polycystic kidney disease. *Drugs*, **79**(3): 303-313.

Brunton LL, Hilal-Dandan R and Knollmann BC (2018). Goodman & Gilman's the pharmacological basis of therapeutics, 13th Edition. McGraw-Hill Education, New York, USA, pp.527-546.

Chen GY, Pan YC, Wu TY, Yao TY, Wang WJ, Shen WJ, Ahmed A, Chan ST, Tang CH, Huang WC, Hung MC, Yang JC and Wu YC (2022). Potential natural products that target the SARS-CoV-2 spike protein identified by structure-based virtual screening, isothermal titration calorimetry and lentivirus particles pseudotyped (Vpp) infection assay. *J. Tradit. Complement*., **12**(1): 73-89.

Contini C, Di Nuzzo M, Barp N, Bonazza A, De Giorgio R, Tognon M and Rubino S (2020). The novel zoonotic COVID-19 pandemic: An expected global health concern. *J. Infect. Dev. Ctries.*, **14**(3): 254-264.

Coronaviridae Study Group of the International Committee on Taxonomy of Viruses (2020). The species severe acute respiratory syndrome-related coronavirus: Classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.*, **5**(4): 536-544.

Dai La and Gao GF (2021). Viral targets for vaccines against COVID-19. *Nat. Rev. Immunol.*, **21**(2): 73-82.

Dallakyan S, Olson AJ (2015). Small-molecule library screening by docking with PyRx. *Methods Mol. Biol.*, **1263**: 243-250.

Farouk AE, Baig MH, Khan MI, Park T, Alotaibi SS and Dong JJ (2021). Screening of inhibitors against SARS-CoV-2 spike protein and their capability to block the viral entry mechanism: A viroinformatics study. *Saudi J. Biol. Sci.*, **28**(6): 3262-3269.

Fiolet T, Kherabi Y, MacDonald CJ, Ghosn J and Peiffer-Smadja N (2022). Comparing COVID-19 vaccines for

their characteristics, efficacy and effectiveness against SARS-CoV-2 and variants of concern: A narrative review. *Clin. Microbiol. Infect.*, **28**(2): 202-221.

Golob JL, Lugogo N, Lauring AS and Lok AS (2021). SARS-CoV-2 vaccines: A triumph of science and collaboration. *JCI Insight*, **6**(9): e149187.

Guo YR, Cao QD, Hong ZS, Tan YY, Chen SD, Jin HJ, Tan KS, Wang DY and Yan Y (2020). The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak - an update on the status. *Mil. Med. Res.*, **7**(1): 11.

Hadj Hassine I (2022). Covid-19 vaccines and variants of concern: A review. *Rev. Med. Virol.* **32**(4): e2313.

Hardenbrook NJ and Zhang P (2020). A structural view of the SARS-CoV-2 virus and its assembly. *Curr. Opin. Virol.*, **52**: 123-134.

Huang Y, Yang C, Xu XF, Xu W and Liu SW (2020). Structural and functional properties of SARS-CoV-2 spike protein: Potential antivirus drug development for COVID-19. *Acta Pharmacol. Sin*., **41**(9): 1141-1149.

Ke Z, Oton J, Qu K, Cortese M, Zila V, McKeane L, Nakane T, Zivanov J, Neufeldt CJ, Cerikan B, Lu JM, Peukes J, Xiong X, Kräusslich HG, Scheres SHW, Bartenschlager R and Briggs JAG (2020). Structures and distributions of SARS-CoV-2 spike proteins on intact virions. *Nature*, **588**(7838): 498-502.

Konc J and Janezic D (2010). ProBiS: A web server for detection of structurally similar protein binding sites. *Nucleic Acids Res.*, **38**: W436-W440.

Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, Zhang Q, Shi X, Wang Q, Zhang L and Wang X (2020). Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*, **581**(7807): 215-220.

Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KSM, Lau EHY, Wong JY, Xing X, Xiang N, Wu Y, Li C, Chen Q, Li D, Liu T, Zhao J, Liu M, Tu W, Chen C, Jin L, Yang R, Wang Q, Zhou S, Wang R, Liu H, Luo Y, Liu Y, Shao G, Li H, Tao Z, Yang Y, Deng Z, Liu B, Ma Z, Zhang Y, Shi G, Lam TTY, Wu JT, Gao GF, Cowling BJ, Yang B, Leung GM and Feng Z (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.*, **382**(13): 1199-1207.

Ling Y, Zhong J and Luo J (2021). Safety and effectiveness of SARS-CoV-2 vaccines: A systematic review and meta-analysis. *J. Med. Virol*., **93**(12): 6486-6495.

Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, Wang W, Song H, Huang B, Zhu N, Bi Y, Ma X, Zhan F, Wang L, Hu T, Zhou H, Hu Z, Zhou W, Zhao L, Chen J, Meng Y, Wang J, Lin Y, Yuan J, Xie Z, Ma J, Liu WJ, Wang D, Xu W, Holmes EC, Gao GF, Wu G, Chen W, Shi W and Tan W (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding. *Lancet*, **395**(10224): 565-574.

Mandala VS, McKay MJ, Shcherbakov AA, Dregni AJ, Kolocouris A and Hong M (2020). Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers. *Nat. Struct. Mol. Biol.*, **27**(12): 1202-1208.

Mohebbi A, Askari FS, Sammak AS, Ebrahimi M and Najafimemar Z (2021). Druggability of cavity pockets within SARS-CoV-2 spike glycoprotein and pharmacophore-based drug discovery. *Future Virol.*, doi: 10.2217/fvl-2020-0394.

Shang J, Ye G, Shi K, Wan Y, Luo C, Aihara H, Geng Q, Auerbach A and Li F (2020). Structural basis of receptor recognition by SARS-CoV-2. *Nature*, **581**(7807): 221-224.

Sterling, T and Irwin JJ (2015). ZINC 15 – Ligand discovery for everyone. *J. Chem. Inf. Model.*, **55**(11): 2324-2337.

Thomas S (2020). The structure of the membrane protein of SARS-CoV-2 resembles the sugar transporter SemiSWEET. *Pathog Immun.*, **5**(1): 342-363.

Tsegay KB, Adeyemi CM, Gniffke EP, Sather DN, Walker JK and Smith SEP (2021). A repurposed drug screen identifies compounds that inhibit the binding of the COVID-19 spike protein to ACE2. *Front. Pharmacol.*, **12**: 685308.

Wang MY, Zhao R, Gao LJ, Gao XF, Wang DP and Cao JM (2020). SARS-CoV-2: Structure, biology and structure-based therapeutics development. *Front. Cell. Infect. Microbiol*., **10**: 587269.

Wang Q, Zhang Y, Wu L, Niu S, Song C, Zhang Z, Lu G, Qiao C, Hu Y, Yuen KY, Wang Q, Zhou H, Yan J and Qi J (2020). Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell*, **181**(4): 894-904.e9

Xia X (2021). Domains and functions of spike protein in Sars-Cov-2 in the context of vaccine design. *Viruses*, **13**(1): 109.

Xu C, Wang Y, Liu C, Zhang C, Han W, Hong X, Wang Y, Hong Q, Wang S, Zhao Q, Wang Y, Yang Y, Chen K, Zheng W, Kong L, Wang F, Zuo Q, Huang Z and Cong Y (2021). Conformational dynamics of SARS-CoV-2 trimeric spike glycoprotein in complex with receptor ACE2 revealed by cryo-EM. *Sci Adv.*, **7**(1): eabe5575.

Yan R, Zhang Y, Li Y, Xia L, Guo Y and Zhou Q (2020). Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science*, **367**(6485): 1444-1448.

Yang Q, Jian X, Syed AAS, Fahira A, Zheng C, Zhu Z, Wang K, Zhang J, Wen Y, Li Z, Pan D, Lu T, Wang Z and Shi Y (2022). Structural comparison and drug screening of spike proteins of ten SARS-CoV-2 variants. *Research (Wash D C)*, **2022**: 9781758.

Zhang J, Xiao T, Cai Y and Chen B (2021). Structure of SARS-CoV-2 spike protein. *Curr. Opin. Virol*. **50**: 173-182.

Zhou S, Wang R, Liu H, Luo Y, Liu Y, Shao G, Li H, Tao Z, Yang Y, Deng Z, Liu B, Ma Z, Zhang Y, Shi G, Lam TTY, Wu JT, Gao GF, Cowling BJ, Yang B, Leung GM and Feng Z (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.*, **382**(13): 1199-1207.

Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P, Zhan F, Ma X, Wang D, Xu W, Wu G, Gao GF and Tan W (2020). China novel coronavirus investigating and research team. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.*, **382**(8): 727-733.