Molecular docking analysis of azithromycin and hydroxychloroquine with spike surface glycoprotein of SARS-CoV-2

Jitender Singh, Deepti Malik & Ashvinder Raina*

Post Graduate Institute of Medical Education and Research, Sector-12, Chandigarh, India. Pin-160012; Jitender Singh - jitenderpgi@gmail.com and Deepti Malik - deepti180@gmail.com; Ashvinder Raina - Email: drashvinder@gmail.com, *Corresponding author

Received October 22, 2020; Revised November 6, 2020; Accepted November 6, 2020; Published January 31, 2021

DOI: 10.6026/9732063001701

The authors are responsible for the content of this article. The Editorial and the publisher has taken reasonable steps to check the content of the article in accordance to publishing ethics with adequate peer reviews deposited at PUBLONS.

Declaration on official E-mail:
The corresponding author declares that official e-mail from their institution is not available for all authors

Declaration on Publication Ethics:
The authors state that they adhere with COPE guidelines on publishing ethics as described elsewhere at https://publicationethics.org/. The authors also undertake that they are not associated with any other third party (governmental or non-governmental agencies) linking with any form of unethical issues connecting to this publication. The authors also declare that they are not withholding any information that is misleading to the publisher in regard to this article.

Abstract:
Millions of people are affected by COVID-19 since the last quarter of 2019. Treatment using hydroxychloroquine (HCQ) as monotherapy in combination with azithromycin (HCQ-AZ) were administered at several clinical centres to patients tested positive to the virus across continents. Therefore, it is of interest to document the molecular docking analysis data of azithromycin and hydroxychloroquine drug with the spike surface glycoprotein of novel COVID-19. Thus, we report the molecular modelling docking based structural binding features of HCQ-AZ with the spike surface glycoprotein of COVID-19 for further evaluation in this regard.

Keywords: Drugs; pharmacokinetics; binding pocket; molecular docking; structure activity relationship.

Background:
COVID-19 in humans has emerged in Wuhan, China. Its genome has been sequenced and the genomic information promptly released. Despite a high similarity with the genome sequence of SARS-CoV and SARS-like CoVs, we identified a peculiar furin-like cleavage site in the Spike protein of the 2019-nCoV, lacking in the
other SARS-like CoVs [1]. In this research article, we discuss the possible functional impact of HCQ-AZ on the spike protein, pathogenicity and its potential implication in the development of COVID-19 therapy. Human coronaviruses (CoV) are enveloped positive-stranded RNA viruses belonging to the order Nidovirales and are mostly responsible for upper respiratory and digestive tract infections. Among them SARS-CoV and MERS-CoV that spread in 2002 and 2013 respectively, have been associated with severe human illnesses, such as severe pneumonia and bronchiolitis, and even meningitis in more vulnerable populations [2]. In December 2019, a new CoV (2019-nCoV) has been detected in the city of Wuhan, and this emerging viral infection was associated with severe human respiratory disease with a ~2-3% fatality rate [3]. In this article, we focus on a specific furin-like protease recognition pattern present in the vicinity of one of the maturation sites of the S protein and the impact of HCQ-AZ inhibiting this spike. Coronavirus spike (S) glycoproteins promote entry into cells and are the main target of antibodies. SARS-CoV-2 S uses ACE2 to enter cells and that the receptor-binding domains of SARS-CoV-2 S and SARS-CoV S bind with similar affinities to human ACE2, correlating with the efficient spread of SARS-CoV-2 among humans [4]. The coronavirus spike (S) glycoprotein initiates infection by promoting fusion of the viral and cellular membranes through conformational changes that remains largely uncharacterized. Coronavirus entry is mediated by the trimeric transmembrane spike (S) glycoprotein, which is responsible for receptor binding and fusion of the viral and host membranes. S is a class I viral fusion protein that is synthesized as a single-chain precursor of ~1, 300 amino acids and trimerizes upon folding. It forms an extensive crown decorating the virus surface and is the main target of neutralizing antibodies upon infection [5]. Therefore, it is of interest to document the molecular docking analysis data of azithromycin and hydroxychloroquine drug with the spike surface glycoprotein of novel COVID-19.

Materials and Methods:

Protein preparation:
Primary sequence of surface glycoprotein of severe acute respiratory syndrome coronavirus 2 with protein NCBI ID: YP_009724390.1 was retrieved in FASTA format from National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) database which has been recently updated in NCBI database. The 3D structure of surface glycoprotein was constructed using SWISS MODEL server (https://swissmodel.expasy.org/). The 3D model of the target protein generated by us was also available in the Swiss model database. The best model was selected from the Swiss model database. Plotting the Ramachandran plot did further validation of the 3D structure model and MolProbity Score1.50 was taken from SWISS MODEL database.

Binding site prediction:
Protein–Ligand-binding sites are the active sites on protein surface that perform protein functions. Thus, the identification of these binding sites is often the first step to study protein functions and structure-based drug design. Active sites (ligand binding sites) of the receptor protein were analysed by Metapocket 2.0 online tool (https://projects.biotec.tu-dresden.de/metapocket/). Binding sites are the distribution of surrounding residues in the active sites and act as the catalytic residues [6].

Antigenic peptides prediction:
Antigenicity is the capacity of an antigen to bind specifically with a group of certain products that have adaptive immunity: T cell receptors or antibodies. Antigenic determinants were analysed using ImmunoMedicine Group Server Antigenic Peptide Prediction (http://imed.med.ucm.es/Tools/index.html).

Ligand preparation:
The ligands azithromycin and hydroxychloroquine were selected for binding with the target protein (spike surface glycoprotein) of COVID-19. The structure of these drug compounds was downloaded from PubChem database (http://pubchem.ncbi.nlm.nih.gov/). The drug files was downloaded in SDF file format and were converted into PDB files using open PyMOL software [7] (and Discovery Visualize Software [8]. The ligands preparation included 2D-3D conversion, verifying and optimizing the structures. We have taken two FDA approved drug molecule structures for docking analysis with target surface glycoprotein.

Docking:
AutoDock vina (PyRx) software was used to dock protein and ligand molecules. Molecular docking of selected drug molecules at the binding site of target protein of COVID-19 was docked by automated docking program AutoDock vina (PyRx) [9]. Preparation of required input files for docking was completed using AutoDock Vina (PyRx) bioinformatics software. The target receptor file and all ligand files were uploaded to the software for the docking with individual drug molecules. An autogrid map dimensions was set to find the best binding site in the coordinating centre. The results were analysed after the docking process was complete.
Figure 1: Analysis of Primary sequence characterization of spike surface glycoprotein of COVID-19: Snapshot of 1273 amino acid single chain long primary sequence of spike protein of COVID-19 retrieved from NCBI database having NCBI ID: YP_009724390.1

Table 1: Actives site residues prediction of target surface glycoprotein: table showing three different metapocket having no. of amino acid residues positions within these metaPockets were analysed using Metapocket2.0 online tool (https://projects.biotec.tu-dresden.de/metapocket/)
**Table 2:** Antigenic peptide prediction of target surface glycoprotein: Table representing antigenicity of target surface glycoprotein using online server ImmunoMedicine group tool Antigenic Peptide Prediction (http://imed.med.ucm.es/Tools/index.html). Sixty three antigenic determinants were found in the target sequence of surface glycoprotein with antigenic propensity score 1.0416 considered to be a valid score.

| S. No. | Start Position | Sequence | End Position |
|--------|----------------|----------|--------------|
| 1      | 4              | FLVLLPLVSQCVNL | 18 |
| 2      | 22             | TQLPPAY    | 28 |
| 3      | 33             | TRGVYYPDKVFRSSVLHSTQD LFLPFFSNVTWFHAIHV | 70 |
| 4      | 80             | DKNFLPFNDDGYFA | 93 |
| 5      | 114            | TQSLLIV    | 120 |
| 6      | 122            | NATNVVVKCEFQFCNDPFLGYY | 145 |
| 7      | 156            | EFRVYSS    | 162 |
| 8      | 167            | TFHEYVSQPFML | 177 |
| 9      | 199            | GYFKIYSK   | 206 |
| 10     | 208            | TPINLRDLPLQFSALEPLVDLPIG | 232 |
| 11     | 237            | RFQTLALHRSYLT | 250 |
| 12     | 261            | GAAAYYVGYLQPRTFELL | 277 |
| 13     | 285            | ITDANDLCPD | 295 |
| 14     | 297            | SETKCTLSETVVEK | 310 |
| 15     | 316            | SNFRVQPTESIVRF | 329 |
| 16     | 331            | NITNLCPFGE | 340 |
| 17     | 346            | RFASYA     | 352 |
| 18     | 357            | RSNCVADYSVLYNS | 371 |
| 19     | 373            | SPSFKCYGVSP | 384 |
| 20     | 386            | KLDLCFTNYADSFVIR | 403 |
| 21     | 428            | DFTGCVIA   | 435 |
| 22     | 449            | YNYLYRL    | 455 |
| 23     | 469            | STEYQAGS   | 477 |
| 24     | 484            | EQFCNYFPLOQSYG | 496 |
| 25     | 503            | VGYQPRYVTVLSFELLHAPATVCGP | 527 |
| 26     | 534            | VKNKCVN    | 540 |
| 27     | 547            | TGTCVLT    | 553 |
| 28     | 557            | KKFLPQO    | 563 |
| 29     | 580            | QTLEILDITPCSFQGVSVIT | 599 |
| 30     | 606            | NQVAVLQVDVNCETVPVAIHDQ | 628 |
| 31     | 631            | PTWRYVSTSCNSNVFQTRACCLIGA | 653 |
| 32     | 660            | YEDCIPAGICASY | 674 |
| 33     | 683            | RARSVASQHIAYTM | 697 |
| 34     | 701            | AENSVAY    | 707 |
| 35     | 718            | FTISVTTEILPVS | 730 |
| 36     | 734            | TSVDCTMYICGD | 745 |
| 37     | 748            | ECSNLILLQYGGSPCTQLN | 764 |
| 38     | 779            | QEFAQVRQVYKT | 791 |
| 39     | 801            | NFQGILPQD | 809 |
| 40     | 817            | FIEDLFLNVKTLADA | 831 |
| 41     | 833            | FKQYGDCLGD | 843 |
| 42     | 845            | AARDLCAQ   | 853 |
| 43     | 856            | NGLTVLPPILLT | 866 |
| 44     | 871            | AQYTSALLAG | 880 |
| 45     | 890            | AGAALQIPF | 896 |
| 46     | 911            | VTQVNLVENQ | 920 |
| 47     | 931            | IGXQIQSDSLSSTASALKLQDQVNVQ | 954 |
| 48     | 958            | ALNITLQKQLSS | 968 |
| 49     | 972            | AISSVLNDSRDLKVEAEVQIDR | 995 |
| 50     | 997            | ITGRILSQLQTYVQLLIRA | 1015 |
| 51     | 1028           | KMSECVLGQ | 1056 |
Table 3: Molecular docking binding affinity energy and interacting residues. Table showing both the drug molecules showing different binding energy, no of interacting residues, no of hydrogen bond and bond length.

| Name of compound       | Binding affinity (kcal/mol) | Interacting residues               | No of H-bonds | Bond length (Å) |
|------------------------|-----------------------------|-----------------------------------|---------------|-----------------|
| Azithromycin           | -6.9                        | ASN-953, GLN-957, SER-943, TYR-313 | 3-H bonds     | 2.56, 2.16      |
| Hydroxy chloroquine    | -5.7                        | VAL-42, ALA-222, PRO-225, TYR-38  | 3-H-bonds     | 2.13, 2.34      |

Figure 2: Model quality and Ramachandran plot of spike surface glycoprotein of COVID-19: a) Showing global quality estimation of target protein which has -3.72 QMEAN value. b) Local quality estimation of S protein of COVID-19, having maximum number of residues showing scores more than 0.8. C) z-score of modelled structure which comparison with non-redundant set of PDB structure. d) Ramachandran plot retrieved from SWISS-MODEL database of target protein representing 97.50% of the favoured residues in the plot area with good validate score thus can be considered as a target molecule.
Figure 3: Functional and structural characterization of spike surface glycoprotein of COVID-19: a) Representative 3D modelled structure of target protein (spike surface glycoprotein) generated through SWISS-MODEL database. b) Representation of top three metaPocket regions within 3D carton structure marked as metaPocket-1, 2, 3 with good pocket z-score: 20.28 of 3 pocket sites. The amino acid positions within these meta-Pockets were analysed using Metapocket2.0 online tool (https://projects.biotec.tu-dresden.de/metapocket/).

Figure 4: Structural and chemical properties of selected drug molecules: Representative 3D structure of a) azithromycin b) hydroxychloroquine. SDF structure retrieved from PUBCHEM database was converted to PDB structure using Discovery Studio Visualizer free software.
Figure 5: Molecular Modeling and docking studies of both selected drugs with surface glycoprotein of COVID-19: Representative 3D structure of complex docked target protein molecule (surface glycoprotein) with a) azithromycin, b) hydroxychloroquine.

Figure 6: 3D docking binding interactions of both selected drugs with surface glycoprotein of COVID-19. Bonding interaction(s) between drug molecules and target protein receptor; (a) target protein + azithromycin; (b) target protein + hydroxychloroquine.
Figure 7: Representation of 2D complex structure of both drugs and target protein indicating different types of bonds interacting with various amino acid residues: a) Target protein + azithromycin. b) Target protein + hydroxychloroquine. All the 2D structures were generated from the Discovery Studio Visualize software and PYMOL Visualize.

Figure 8: Graphical representation of binding energy of both the drugs with surface target glycoprotein from minimum to maximum energy score in kcal/mol.
Results and Discussion:
The 1273 amino acids residue long primary sequence of spike surface glycoprotein of severe acute respiratory syndrome coronavirus 2 with protein NCBI ID: YP_009724390.1 was retrieved in FASTA format from National Centre for Biotechnology Information database (Figure 1). 3D structures of target protein spike surface glycoprotein of COVID-19 were built with the help of SWISS MODEL server. The modelled structure is also available on the SWISS MODEL database. The best model was selected from the SWISS MODEL database which having 99.26% sequence identity with the template sequence of PDB ID: 6VSB. After the generation of 3D structure of target protein, the model quality was checked with SWISS MODEL online program. The SWISS MODEL online program showed that the ‘target’ protein model, which was generated, was the best model structure, with good global quality and having best Z score for quality of the structures. The GNMQE value was 0.73 and QMEAN score was -3.72 (Figure 2a-c). For the prediction of the active binding site(s) of ‘target’ protein structure, first we validated the 3D structure of target protein using Ramachandran plot and we found 90.49% residues in the favoured region (Figure 2d). The final predicted 3D structure model of the target proteins is as shown in (Figure 3a). After validation of this predicted 3D structure of target protein, next we wanted to explore ligand binding site residues in surface glycoprotein using Meta Pocket 2.0 Web online tool. On analysis the 3D structure of target protein of COVID-19, we found there were top three active sites (meta-pockets), considered as ligand binding sites in this protein (Figure 3b). Analysing the number of important amino acid residues in the catalytic region followed site prediction. Important amino acid residues present in these top three metaPocket sites are shown in Table 1. These three metaPocket(s) were studied for further analysis as they had important amino acid residues. Next we predicted the antigenicity of target surface glycoprotein using online server Immuno Medicine group tool Antigenic Peptide Prediction. On analysis, we found there were 63 antigenic determinants found in the target sequence of surface glycoprotein with average antigenic propensity score is 1.0416 which is considered to be a valid score as shown in Table 2. After the confirmation of Meta Pocket site(s) and antigenic peptide prediction, two drug compounds were selected as ligands for docking with the selected spike surface glycoprotein of COVID-19 protein. 3D structure of both the chemical compounds was downloaded from the NCBI PubChem database (Figures 4a-b). These two compounds azithromycin and hydroxychloroquine, which are already FDA, established drugs and therefore follow the Lipinski rule of five to evaluate drug-likeness (www.scbio-ittd.res.in) [10]. These two drug compounds were then checked for their binding with target protein and docking score was evaluated using AutoDock vina Docking (PyRx) software. An autogrid map dimension was set at grid centre X: 2000.63, Y: 227.356, Z: 217.7743. Number of point was X: 68, Y: 58, Z: 76 and spacing was 0.9873 Angstrom. For each compound, many docking poses were obtained for target protein with different binding affinity score. Out of many docking poses, only those were selected which had the best and highest docking score, good hydrogen bond interaction and highest binding energy. The binding interaction(s) of all the docked with complex structure of selected two drug molecules are shown in Figure 5 a-b. The 3D interactions of receptor target protein of COVID-19 with these two selected drugs are represented in Figure 6 a-b. Similarly the 2D poses of receptor target protein of COVID-19 with selected two drugs are represented in Figures 7a-b. Different types of bonding interactions were observed between the same and it is a universal acknowledged fact that hydrogen bonding is of considerable importance in the interaction between such molecules and from our results we also found that both the selected drug molecules showed highest binding energy with the target surface glycoprotein of COVID-19 (Table 3). We observed that the binding affinity score of Azithromycin is -6.9kcal/mol and that of hydroxychloroquine is -5.7 kcal/mol, which is considered as good docking score. Graphical representation of the same is shown in Figure 8. In addition to this both complex structure showed 3 Hydrogen bond interactions (Table 3). In coherence to our study reported the binding energies obtained from the docking of S protein with ligand, chloroquine and hydroxychloroquine were -7.1 and -6.8kcal/mol [11]. Another study reported the binding affinity of phyto compounds derived from Silybum marianum (Silybin), Withania somnifera (Withaferin A), Tinospora cordifolia (Cordioside) and Aloe barbadensis (Catechin and Quercetin) with SARS-CoV-2 target [12]. Other studies are also there which underline the significance of Hydroxychloroquine and Azithromycin in COVID-19 [13-17]. From these findings, the study concluded that these two selected drug molecules azithromycin and hydroxychloroquine may act as potential candidate drug molecules against the target spike surface glycoprotein of COVID-19.

Conclusion:
The focus of our study was molecular docking of azithromycin and hydroxychloroquine drug compounds, which potentially have therapeutic impact on spike surface glycoprotein of novel COVID-19. Docked result confirmed our hypothesis that both the selected drug molecules bind to surface glycoprotein protein of COVID-19 showing best binding affinity which attests to the good binding of these drugs to the target protein. On the basis of molecular docking energy score of both selected FDA approved drug molecules suggest that they may act as inhibitor that can potentially be used
for the treatment of severe acute respiratory syndrome coronavirus 2. Based on results we report the molecular modelling docking based structural binding features of HCQ-AZ with the spike surface glycoprotein of COVID-19 for further evaluation in this regard.

**Declaration of Competing Interest:**
The authors declare that they have no conflict of interest.

**References:**

[1] Coutar B et al. Antiviral Res. 2020 176:104742. [PMID: 32057769]
[2] De Wit E et al. Nat Rev Microbiol. 2016 14: 523. [PMID: 27344959]
[3] She J et al. Clin Transl Med. 2020 9:19. [PMID: 32078069]
[4] Walls AC et al. Cell 2020 181:281. [PMID: 32155444]
[5] Walls AC et al. Proc Natl Acad Sci U S A. 2017 114:11157. [PMID: 29073020]
[6] Zhang Z et al. Bioinformatics 2011 27:2083. [PMID: 21636590]
[7] PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC
[8] Dassault Systèmes BIOVIA, Discovery Studio Modeling Environment, Release 2017, San Diego: Dassault Systèmes, 2016.

[9] Trott O and Olson AJ Journal of Computational Chemistry. 2010 31:455. [PMID: 19499576]
[10] Lipinski CA 2004. Lead- and drug-like compounds: the rule-of-five revolution. Drug Discov Today Technol. 1, 337.
[11] Seshu Vardhan et al. Comput Biol Med 2020 124:103936. [PMID: 32738628]
[12] Nabajyoti Bailddy et al. J Mol Struct. 2020 1219: 128595. [PMID: 32834108]
[13] Gautret P et al. Int J Antimicrob Agents. 2020 56:105949. [PMID: 32205204].

Choudhary R and Sharma AK New Microbes New Infect. 2020 35:100684. PMID: 32322397.

Arshad S et al. Int J Infect Dis. 2020 97:396. [PMID: 32623082]

Rosenberg ES et al. JAMA. 2020 323:2493. [PMID: 32392282]

Gautret P et al. Travel Med Infect Dis. 2020 34:101663. [PMID: 32289548].

http://www.ncbi.nlm.nih.gov/

http://swissmodel.expasy.org/

https://projects.biotec.tu-dresden.de/metapocket/

http://imed.med.ucm.es/Tools/index.html

http://www.scfbio-iitd.res.in
Articles published in BIOINFORMATION are open for relevant post publication comments and criticisms, which will be published immediately linking to the original article for FREE of cost without open access charges. Comments should be concise, coherent and critical in less than 1000 words.
