SARS-CoV-2 spike protein 13-mer subdomain corresponds to the drug-binding domain of glutamyl-propyl-tRNA synthetase 1 and is targetable by halofuginone

Lora Benoit (lbenoit@chsu.edu)
College of Osteopathic Medicine, CHSU
Vinayak Shenoy
College of Pharmacy, CHSU
Simon Meykler
College of Osteopathic Medicine, CHSU
Ali Mohseni
PACAN

Research Article

Keywords: spike protein, transmission, infectivity, halofugenine, cladosporin, SARS-CoV-2, COVID19, treatment, emergency drug intervention

Posted Date: July 22nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-738132/v1

Research Reports

Title: SARS-CoV-2 spike protein 13-mer subdomain corresponds to the drug-binding domain of glutamyl-propyl-tRNA synthetase 1 and is targetable by halofuginone

Authors: Lora Benoit, PhD[1*], Vinayak Shenoy, PhD[2], Simon Meykler, DO[1], Ali Mohseni, PhD[3]

Key words: spike protein, transmission, infectivity, halofugenine, cladosporin, SARS-CoV-2, COVID19, treatment, emergency drug intervention

Affiliations:
[1]*Department of Biomedical Sciences, College of Osteopathic Medicine, California Health Sciences University, 2500 Alluvial Avenue, Clovis, CA 93611*
[2]*Department of Pharmaceutical and Biomedical Sciences, College of Pharmacy, California Health Sciences University, 120 N. Clovis Avenue, Clovis, CA, 93612*
[3]*Persian American Civic Action Network (PACAN)*

[*]**Materials and Correspondence:** Tel: +1-559-712-4153
E-mail: lbenoit@chsu.edu

Author contribution statement:
L.B. and A.M. undertook the preliminary sequence analysis leading to the concept and the first two figures. L.B., A.M., and S.M. conceptualized and contributed to the design of the experiments. L.B. and S.M. performed the docking experiments. V.K. designed and performed the in vitro experiment. All authors contributed to writing and preparation of the manuscript.

Corresponding Author identification: https://orcid.org/0000-0002-2131-213X

Acknowledgements: This work was partially funded by ART FUND.

Competing interests: The authors have no competing interests or conflict of interests.

ABSTRACT

Since its emergence, SARS-CoV-2 has been the subject of intense investigation. Early sequence analysis identified a unique 13 amino acid region (13-mer) nested within the receptor-binding domain (RBD) of the spike protein that directly interacts with the ACE2 receptor. Blasting with the 13-mer identified a highly conserved segment in propyl-tRNA synthetase enzymes. Comparison with the human analogue, glutamyl-propyl-tRNA synthetase 1, showed a high level of identity with its drug binding domain, which is targeted by halofuginone, a drug recently shown to block SARS-CoV-2 infection in vitro. In silico experiments predicted a high affinity interaction between halofuginone and the 13-mer. In vitro addition of halofuginone effectively inhibited binding of recombinant S1 monomer to ACE2. Accordingly, it appears that halofuginone inhibits viral infection by preventing correct interactions between spike protein and ACE2. These findings indicate that viral entry can potentially be drug-targeted and support the application of halofuginone in mitigation of COVID-19.

INTRODUCTION

Since the emergence of SARS CoV-2 in December 2019, significant morbidity and mortality has ensued worldwide prompting intense investigation into the viral adaptations that have allowed for its continued propagation. Of particular relevance to understanding the biology of SARS-CoV-2 is its spike protein which directly facilitates infection through binding to its receptor, ACE2, on the surface of the host cell (1, 2). For this reason, the spike protein critically defines the host tropism, transmissibility, and pathogenicity of the virus. Sequence comparison of the SARS-CoV-2 spike protein against spike protein derived from other *Coronaviridae* (CoV) family members that utilize ACE2 for cell entry demonstrates variable levels of sequence relatedness: 23.3% identity with endemic HCoV-NL63 spike, 76.2% identity with SARS-CoV spike, and 97.4% identity with bat/Yunnan/RaTG13/2013 CoV (bat RaTG13) spike. Of note, phylogenetic analysis has revealed that Bat RaTG13 CoV is the closest extant relative, supporting the hypothesis of a zoonotic origin in bats for SARS-CoV-2 (3). Direct comparison of SARS-CoV-2 spike protein with Bat RaTG13 CoV spike protein revealed a gain of 4 amino acid residues, PRRA, at position 681-684, which function as a furin cleavage site as well as 29 non-synonymous changes, 6 of which localize to the 13-mer region nested within the receptor-binding domain (RBD) of the spike protein (alignment data not shown). The consequence of this seemingly nominal variation is a striking 1000-fold enhancement in receptor binding by SARS-CoV-2 spike protein (4). In contrast, sequence comparison with SARS-CoV spike protein again reveals the gain of the furin cleavage site, but additionally shows the occurrence of several novel insertions as well as 303 non-synonymous changes, 7 of which cluster within the 13-mer region (alignment data not shown). Despite this increased sequence variation at the protein level, SARS-CoV-2 spike is reported to exhibit a comparatively modest 10-40-fold higher binding affinity for ACE2 compared to the SARS-CoV spike protein (5). This incongruent relationship between sequence variation and receptor binding affinity led us to pursue this current study.

While there are no reported cases of human infections attributed to Bat RaTG13 CoV, SAR-CoV was reported to have infected 8,098 people during a four month-long epidemic (6). In comparison, SARS-CoV-2 has resulted in ~ 200 million confirmed cases to date over an extended 18-month long pandemic. The apparent increase in virulence of SARS-CoV-2 compared to SARS-CoV has been attributed in part to the enhanced binding affinity of the spike protein to ACE2, the result of many distinct modifications that globally effect changes in the spike protein, including improvement of complementary electrostatic interactions (7, 8). The biological significance of the furin cleavage site is well established; it is targeted by host proteases, such as TMPRSS2, to form individual S1 and S2 subunits (1), which function co-operatively during infection leading to improved transmissibility and increased pathogenicity. Specifically, the S1 subunit contains the RBD that directly contacts ACE2. Following initial binding, the S1 subunit shifts to the up conformation causing the S2 subunit

to "jack-knife" upwards, enabling it to catalyze the fusion of the viral envelope with the host cell membrane (9). The significance of the 13-mer located at position 482-494 within the S1 subunit has been partially elucidated; it forms a projection that critically contacts the ACE2 receptor (2, 10, 11, 12, 13).

METHODS

In Silico Analyses. Blastp analysis was performed using https://blast.ncbi.nlm.nih.gov/Blast.cgi. The non-redundant protein sequence database was blasted using the 13-mer sequence GVEGFNCYFPLQS using the blastp algorithm and excluding Viruses (taxid:10239). Theoretical drug binding analysis was performed using the PBD 6vw1 (SARS-CoV-2) and PDB 2AJF (SARS-CoV) structures downloaded from https://www.rcsb.org. PBD 6vw1 represents a high-fidelity SARS-CoV-2 chimeric receptor-binding domain complexed with its receptor human ACE2 (10). The *in silico* molecular docking experiments were performed using the Autodock Vina application provided by the Scripps Research Institute (14) in complex with halofuginone, two structural analogues, deoxyhalofuginone and febrifugine, as well as cladosporin, a selective lysyl-tRNA-synthetase inhibitor. An energy range of 4 and an exhaustiveness of 8 were applied to configure the assay. Prior to the assay, molecules (ligands and proteins) were first edited in AutoDockTools, to remove hetatoms, repair atoms, add kollman charges, and remove unwanted chains from the analysis.

In vitro binding assay. The ACE2 activity assay is based on the use of the fluorogenic peptide substrate VI [7Mca-Y-V-A-D-A-PK(Knp)-OH]. ACE2 removes the c-terminal dinitrophenyl moiety of the substrate thereby quenching the inherent fluorescence activity of the 7-methoxycoumain (7Mca) group and resulting in an increase in fluorescence of the overall peptide substrate. Human recombinant ACE2 and its fluorogenic substrate were purchased from R&D systems, Minneapolis, MN. The enzymatic activity was measured in a total reaction volume of 100 µL using Spectra Max Gemini EM Florescence Plate Reader (Molecular Devices, San Jose, CA) at an excitation wavelength of 320 nm and emission wavelength of 405 nm. All assays were performed in triplicate in buffer containing 1 M NaCl, 75 mM Tris-HCl, and 0.5 µM ZnCl2, at pH 7.4, 0.5 ng ACE2 enzyme, 20 uM substrate, drug (various concentrations) and the receptor-binding domain (RBD) of SARS CoV-2 spike protein, consisting of residues 330-530 (1.8 ug). The drug, halofuginone, was purchased from Selleckchem, Houston, Texas and solubilized in DMSO. The RBD of spike protein was purchased from Genscript Biotech Corporation, Piscataway, NJ. To perform the experiment, the drug was pre-incubated at variable concentrations with the RBD of spike protein for 30 min, followed by addition of the ACE2 enzyme and the fluorogenic peptide substrate. The microplate was then read at 37°C every 36 seconds for 30 minutes to determine efficacy of enzyme activity in the context of the drug.

RESULTS

The positioning of the 13-mer within the landscape of the spike proprotein is schematically represented in Figure 1. To gain insight into the possible function of the 13-mer, Blastp analysis was performed against non-viral taxa, revealing significant identity and coverage with a segment derived from prolyl tRNA synthetase and its bifunctional analogue, glutamyl-prolyl tRNA synthetase (ligase) derived from diverse taxa including a bacterium, a protozoan parasite, and a boney fish (Figure 2). This enzyme catalyzes the covalent attachment of proline to its cognate tRNA molecule and therefore plays a key role in protein translation. These blastp hits in turn revealed highly

conserved identity with the key drug binding site of human glutamyl-prolyl tRNA synthetase1 (15) as delineated in Figure 2, implying that the drug might also bind to the 13-mer region of the spike protein. Structural analysis of the SARS-CoV-2 Spike trimer-ACE2 complex revealed that the 13-mer forms a projection from S1 RBD that allows for asymmetric pairing with ACE2, forming the primary docking site in the co-complex (Figure 3a), (10, 12). Comparatively, the 13-mer sequence from SARS-CoV-1, Bat RaTG13 CoV, and endemic hCoV NL63 demonstrated less relatedness to the SARS-CoV-2 13-mer, implying possible ectopic acquisition of this 13-mer sequence. Corresponding analysis of the 13-mer region of the spike protein from the South African variant of concern (VOC), Pango lineage B.1.351, revealed the presence of a single amino acid substitution, E484K, which is predicted to impair drug binding based on the conversion of a negatively charged residue to a positively charged residue at a critical drug contact surface (Figure 3.b). In comparison, the delta variant, which is currently circulating world-wide, retains a persevered 13-mer sequence relative to the reference strain (data not shown), indicating that this VOC has retained the genetic ability to theoretical respond to halofuginone treatment and can be considered as a potential candidate for intervention of severe COVID-19.

The human orthologue, glutamyl-prolyl tRNA synthetase 1 (EPRS), plays an additional regulatory role in Stat and Smad3 signal transduction in response to various cytokines. Specifically, phosphorylation of EPRS allows for the enzyme to combine and form the Gamma-interferon Activated Inhibitor of Translation (GAIT) complex that regulates translation in immune cells in response to IFN-γ stimulation (16) and can be pharmacologically targeted using the inhibitor halofuginone (17). In particular, halofuginone has been shown to exert pleotropic effects on the immune system (18) and has received orphan drug designation by the FDA in March 2020 for the treatment of scleroderma. More recently, halofuginone has been shown to be a potent inhibitor of SARS-CoV-2 infection in vitro, through two distinct putative mechanisms: down regulation of TMPRSS2 or through the effects of amino acid starvation on genomic biosynthesis and heparin sulfate decoration of proteoglycans on the cell surface (19, 20).

Due to the degree of shared similarity between the EPRS drug binding domain and the SARS-CoV-2 13-mer, we pursued in silico docking analysis of the drug on the RBD of spike protein. Halofuginone docked flawlessly in an internal cavity formed inside the 13-mer projection, with a theoretical affinity of interaction of -3.9 kcal/mol. We posit that the apparent effect of drug binding to the 13-mer cavity is to lock it in the down configuration such that the S1 subunit is unable to transition upwards and facilitate the jack-knife-like movement of the S2 subunit to facilitate the ensuing fusion event (Figure 3b). In comparison, cladosporin, a structural analogue with specificity for lysyl tRNA synthase bound slightly below and outside the cavity formed by the 13-mer region, with a theoretical affinity of -4.9 kcal/mol (Supplemental Figure 1). Structural analogues (+)-deoxyhalofuginone and febrifugine were also docked. Deoxyhalofuginone docked within the 13-mer cavity, although not as precisely as halofuginone, with an apparent affinity of -4.5 kcal/mol (Supplemental Figure 2), whereas febrifugine did not exhibit any binding proximal to the 13-mer (data not shown). The structure of these four drugs is represented in Supplemental Figure 3 and underscores the specificity of halofuginone docking to the 13-mer concavity. Lastly, the drug was docked against the spike protein of SARS-CoV (supplemental Figure 4). Halofuginone failed to bind within the concavity formed by the 13-mer, though it did bind below the 13-mer with a theoretical binding affinity of -4.8 kcal/mol, indicating that the drug was specific for the 13-mer of SARS-CoV-2.

We subsequently undertook in vitro enzymatic experiments to see if the drug could inhibit binding of the spike protein to the ACE2 receptor, thereby restoring the enzymatic activity of ACE2. In this analysis, the receptor binding domain (RBD) of spike protein (consisting of amino acids 330-530) was used rather than complete spike protein so that we could restrict our

findings as tightly as possible to a minimal structure containing the 13-mer region (amino acids 482-494) and therefore definitively conclude the mechanism by which the drug might be inhibiting infection as reported by Chen et al., 2020, and Sandoval et al., 2021. Indeed, at the 0.01- 0.1 nM range, halofuginone blocked the ability of the RBD to bind to ACE2, as evidenced by restoration of ACE2 activity in the context of spike protein RBD.

DISCUSSION

our data support the findings of Chen et al., and Sandoval et al. showing that halofuginone blocks infection of ACE2-expressing cells in vitro (19, 20), however, it extends their findings by providing clear mechanism. The concept of using a small molecule inhibitor to bind to spike protein such that it cannot effectively engage the host cell presents a novel mechanism by which to minimize infection. Although there are several monoclonal antibody-based therapies that are currently authorized by FDA that target epitopes within the 13-mer region (21), their efficacy and widespread use is somewhat stymied by the cost, delivery mechanism, and nature of antibodies. In particular, although the antibodies bind to and accordingly neutralize the virus, they paradoxically promote enhancement of infection by bridging the virus to Fc receptors on the host cell (22). Thus application of a small drug inhibitor such as halofuginone eliminates this concern. Collectively, these findings strongly support the use of halofuginone to treat patients with COVID19 by the ability of the drug to directly block infection of the host cell through its action on binding to the spike protein 13-mer region.

REFERENCES

1. Hoffmann, M. et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*. 181, 271-280.e8 (2020).
2. Wan, Y. et al. Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *J. Virol*. 94, e00127-20 (2020).
3. Li, X. et al. Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. *J. Med. Virol*. 92, 602-611 (2020).
4. Wrobel, A. G. et al. SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects. *Nat. Struct. Mol. Bio.l* 27, 763–767 (2020).
5. Wrapp, D. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*. 367:1260-1263 (2020).
6. Stadler, K. et al. SARS--beginning to understand a new virus. *Nat. Rev. Microbiol*. 1, 209-218 (2003).
7. Amin, M. et al. Comparing the Binding Interactions in the Receptor Binding Domains of SARS-CoV-2 and SARS-CoV. *J. Phys. Chem. Lett*. 11, 4897-4900 (2020).
8. Hatmal, M. M. et al. Comprehensive Structural and Molecular Comparison of Spike Proteins of SARS-CoV-2, SARS-CoV and MERS-CoV, and Their Interactions with ACE2. *Cells*. 9, 2638 (2020).
9. Cai, Y. et al. Distinct conformational states of SARS-CoV-2 spike protein. *Science*. 369, 1586-1592 (2020).
10. Shang, J. et al. Structural basis of receptor recognition by SARS-CoV-2. *Nature*. 581, 221-224 (2020).
11. Yan, R., et al. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science*. 367, 1444-1448 (2020).

12. Xu, C. et al. Conformational dynamics of SARS-CoV-2 trimeric spike glycoprotein in complex with receptor ACE2 revealed by cryo-EM. *Sci. Adv*. 7, eabe5575 (2021).
13. Lan, J. et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*. 581, 215-220 (2020).
14. Trott, O. and 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).
15. Son, J. et al. Conformational changes in human prolyl-tRNA synthetase upon binding of the substrates proline and ATP and the inhibitor halofuginone. *Acta Crystallogr. D. Biol. Crystallogr*. 69, 2136-2145 (2013).
16. Arif, A. et al. Two-site phosphorylation of EPRS coordinates multimodal regulation of noncanonical translational control activity. *Mol. Cell*. 35,164-180 (2009).
17. Keller, T. L. et al. Halofuginone and other febrifugine derivatives inhibit prolyl-tRNA synthetase. *Nat. Chem. Biol*. 8, 311-317 (2012).
18. Pines, M. and Spector, I. Halofuginone - the multifaceted molecule. *Molecules*. 20, 573-594 (2015).
19. Chen, Y. et al. A high throughput screen for TMPRSS2 expression identifies FDA-approved and clinically advanced compounds that can limit SARS-CoV-2 entry. *Res Sq.* rs.3.rs-48659 (2020).
20. Sandoval, D. R. et al. The Prolyl-tRNA Synthetase Inhibitor Halofuginone Inhibits SARS-CoV-2 Infection. *bioRxiv.* 2021;2021.03.22.436522 (2021).
21. Taylor, P. C. et al. Neutralizing monoclonal antibodies for treatment of COVID-19. *Nature reviews. Immunology*, *21*, 382–393 (2021).
22. Eroshenko, N. et al. Implications of antibody-dependent enhancement of infection for SARS-CoV-2 countermeasures. *Nature biotechnology*, *38*, 789–791 (2020).

FIGURE LEGENDS

Figure 1: Schematic representation of SARS-CoV-2 Spike protein: S1 domain (binds to Ace2), S2 domain (mediates fusion with host cell membrane), RBD or receptor-binding domain (portion of S1 that directly contacts ACE2) positioned between amino acids 438-508, and the nested 13-mer (unique 13 amino acid subdomain) positioned between amino acids 482-494.

Figure 2: SARS-Cov-2 13-mer insert sequence alignment: Reference strain spike protein 13-mer (NCBI NC_045512.2), South African variant spike protein 13-mer (GISAID EPI_ISL_736980), endemic CoV-NL63 coronavirus spike protein 13-mer (NCBI NC_005831.2), bat coronavirus RaTG13 spike protein 13-mer (NCBI MN996532), SARS-1 spike protein 13-mer (NCBI NC_004718.3), human glutamyl-prolyl-tRNA synthetase 1 halofuginone binding domain (NCBI XP_036448626.1), *Babesia microti* (protozoan) prolyl-tRNA synthetase (NCBI XP_021338015.1), *Novosphingobium aromaticivorans* (Gram-negative bacteria) prolyl-tRNA synthetase (NCBI WP_011445599.1). Grey = identical, blue = similar, and red = disparate amino acid residues.

Figure 3: Analysis of the 13-mer subdomain of spike protein: A) Structural conformation of SARS-CoV-2 Spike trimer-ACE2 co-complex resolved using cryo-EM at 3.80 Å (PBD 7DF4). The 13-mer is indicated in light green, ACE2 in dark green, and the remainder of ribbon diagram is the spike trimer. B) *In silico* docking of halofuginone with the monomeric SARS CoV-2 spike RBD complexed with ACE2 (ACE2 has been removed to facilitate viewing; PDB 6vw1).

Figure 4: Effect of halofuginone on the enzymatic activity of ACE2: Spike protein RBD domain and halofuginone at variable concentrations were co-incubated and then added to ACE2 and is fluorogenic peptide substrate 7Mca-Y-V-A-D-A-PK(Knp)-OH to determine the effect of halofuginone on RBD. Under normal conditions, binding of RBD to ACE2 inhibits the enzyme's activity as measured by a reduction of fluorescent emission.

Supplemental Figure 1: *In silico* docking of cladopsorin with the monomeric SARS CoV-2 spike RBD complexed with ACE2 (ACE2 has been removed to facilitate viewing; PDB 6vw1).

Supplemental Figure 2: *In silico* docking of (+)-deoxyhalofuginone with the monomeric SARS CoV-2 spike RBD complexed with ACE2 (ACE2 has been removed to facilitate viewing; PDB 6vw1).

Supplemental Figure 3: Comparison of 3-D ligand structures: A) Halofuginone, B) (+)-Deoxyhalofuginone, C) Cladosporin, D) Febrifugine

Supplemental Figure 4: *In silico* docking of halofuginone with the monomeric SARS CoV spike RBD complexed with ACE2 (ACE2 has been removed to facilitate viewing; PDB 2AJF).

FIGURES

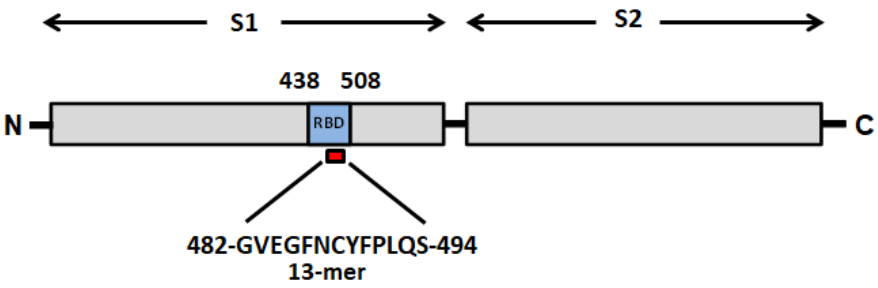


Figure 1

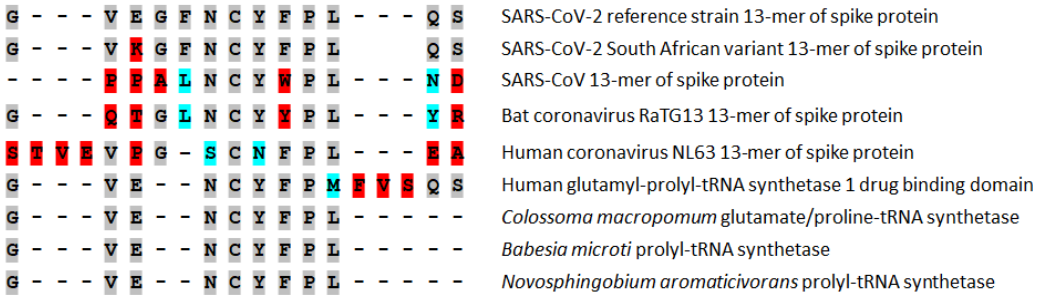


Figure 2

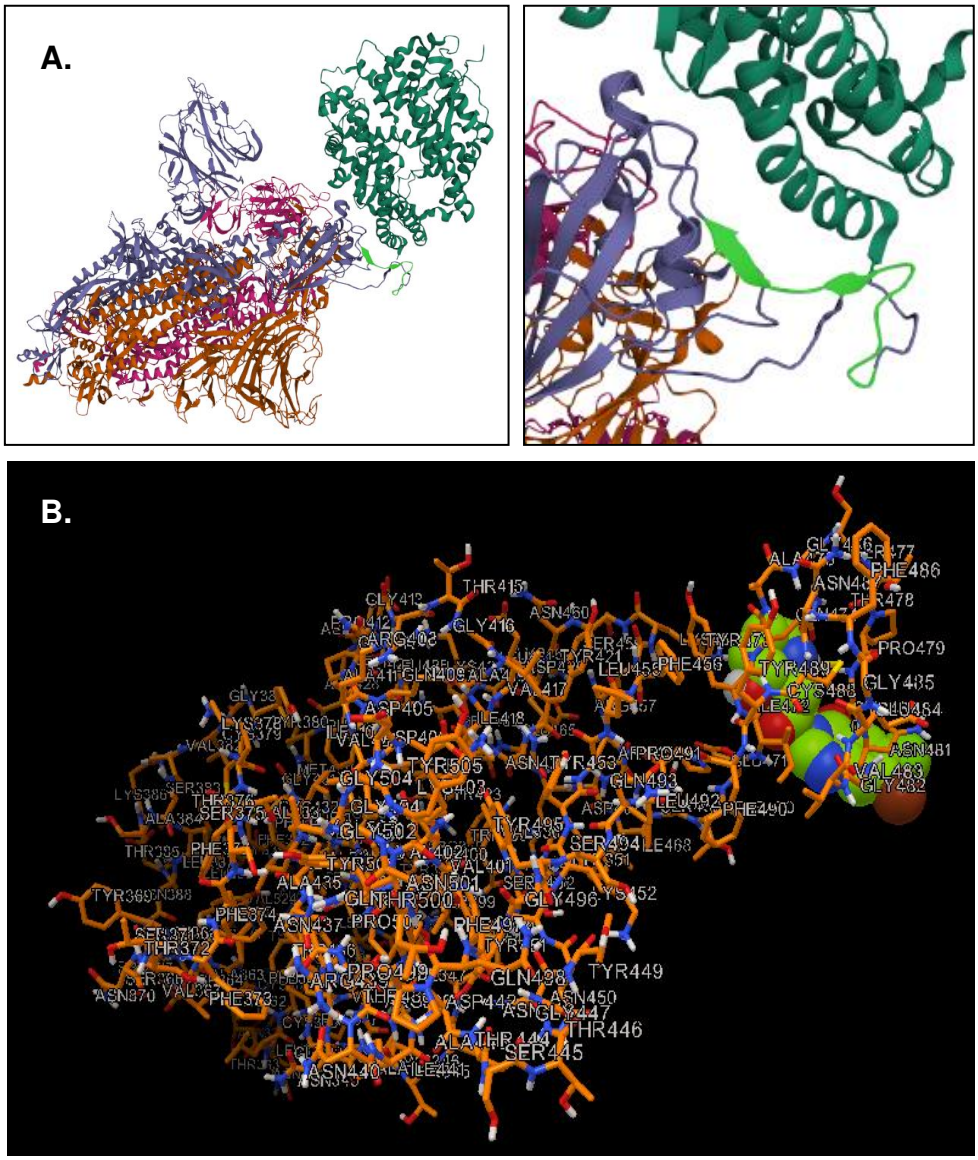

A.
B.
THR415
GLY412
ASN460
GLY416
ARG403
TYR421
LEU455
PHE456
TYR489
PHE486
THR478
PRO479
GLY485
CYS488
GLY496
ASP405
VAL417
LEU418
ASN481
VAL483
GLY482
TYR505
ASN487
GLN493
LEU492
PHE490
GLY504
SER494
TYR495
GLY502
LEU468
VAL401
ALA435
ASN501
GLN498
TYR449
ASN450
THR446
SER445
ASN440
THR444

Figure 3

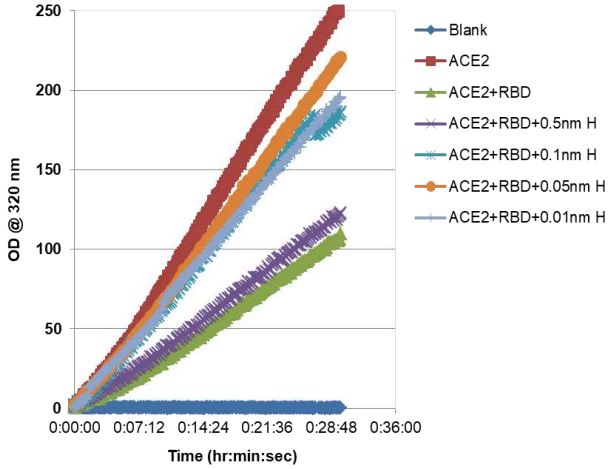
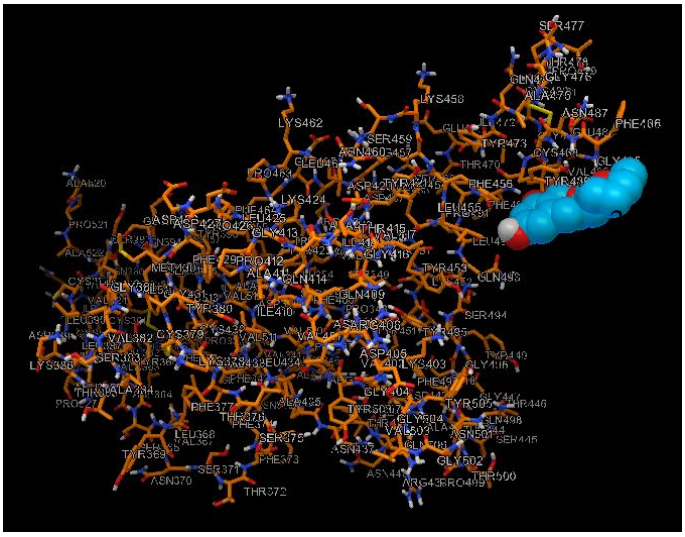
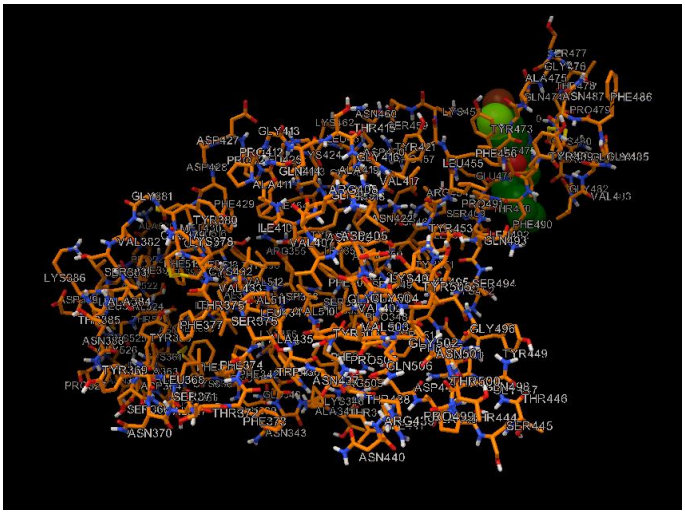


Figure 4



Supplemental Figure 1

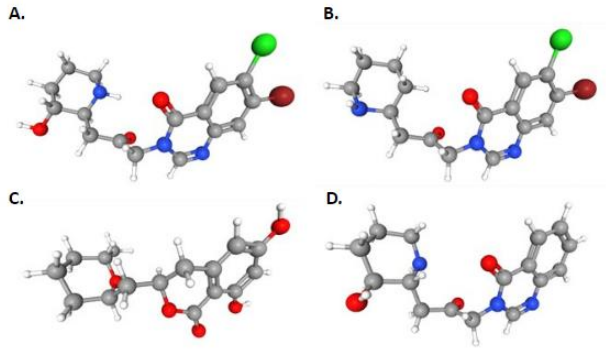


357

358

Supplemental Figure 2

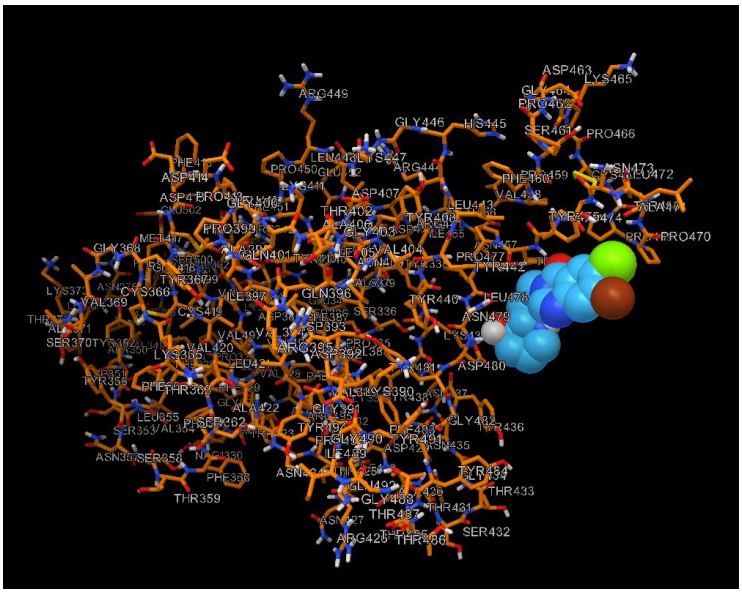
359



360

Supplemental Figure 3

361



362

Supplmental Figure 4

363