

Direct ACE2- Spike RBD binding disruption with small molecules: a strategy for COVID-19 treatment

Smieszek SP¹, Przychodzen B¹, Polymeropoulos V¹, Polymeropoulos C¹ Polymeropoulos MH¹

¹ Vanda Pharmaceuticals Inc., Washington, DC

Corresponding Author: Sandra.Smieszek@vandapharma.com

Abstract

ACE2 is a key receptor for SARS-CoV-2 cell entry. Binding of SARS-Cov-2 to ACE2 involves the viral Spike protein. The molecular interaction between ACE2 and Spike has been resolved. Interfering with this interaction might be used in treating patients with COVID-19. Inhibition of this interaction can be attained via multiple routes: here we focus on identifying small molecules that would prevent the interaction. Specifically we focus on small molecules and peptides that have the capacity to effectively bind the ACE2: RBD contact domain to prevent and reduce SARS-CoV-2 entry into the cell. We aim to identify molecules that prevent the docking of viral spike protein (mediated by RBD) onto cells expressing ACE2, without inhibiting the activity of ACE2. We utilize the most recent ACE2-RBD crystallography resolved model (PDB-ID:6LZG). Based on animal susceptibility data we narrowed down our interest to the location of amino acid 34 (Histidine) located on ACE2. We performed an *in silico* screen of a chemical library of compounds with several thousand small molecules including FDA approved compounds. All compounds were tested for binding to the proximal binding site located close to histidine 34 on ACE2. We report a list of four potential small molecules that potentially have the capacity to bind target residue: AY-NH₂, a selective PAR4 receptor agonist peptide (CAS number: 352017-71-1), NAD⁺ (CAS number: 53-84-9), Reproterol, a short-acting β_2 adrenoreceptor agonist used

in the treatment of asthma (CAS number: 54063-54-6), and Thymopentin, a synthetic immune-stimulant which enhances production of thymic T cells (CAS number: 69558-55-0). The focus is on a High Throughput Screen Assay (HTSA), or *in silico* screen, delineating small molecules that are selectively binding/masking the crucial interface residue on ACE2 at His34.

Consequently, inhibiting SARS-CoV-2 binding to host ACE2 and viral entry is a potent strategy to reduce cellular entry of the virus. We suggest that this anti-viral nature of this interaction is a viable strategy for COVID19 whereas the small molecules including peptides warrant further *in vitro* screens.

Introduction

Similar to SARS-CoV, SARS-CoV-2 enters the cell by means of binding to the Angiotensin-converting enzyme 2 (ACE2) cellular receptor¹. ACE2 is a type 1 integral membrane glycoprotein protein which functions as a carboxypeptidase and its main role is the degradation of angiotensin II². It is expressed widely across human tissues including lungs, heart, kidneys and intestine, and adipose tissue as shown in The Genotype-Tissue Expression (GTEx) project³.

Pulmonary ACE2 regulates the balance of circulating angiotensin II / angiotensin 1–7 levels².

Furthermore, in preclinical ARDS models, ACE2 knockout mice displayed more severe symptoms of this disease compared with wild-type². It has been shown that ACE2 is expressed primarily in alveolar epithelial type II cells, likely a viral reservoir⁴. ACE2 has been shown to be the viral entry receptor also for SARS-CoV⁵. Overexpression of human ACE2 enhanced the severity of disease in mice infected with SARS-CoV⁵ which would imply higher viral titer.

Decreased expression of ACE2 is associated with cardiovascular diseases. Rare genetic variation in ACE2 likely affects the propagation capacity of the virus ultimately affecting variation in susceptibility across individuals. *ACE2* genotype-tissue analysis points to a rather rare population frequency of potential variants that may constitute susceptibility or resilience to infection of SARS-CoV-2 in certain individuals. The capacity of viral binding to the target cell might be specifically altered by natural human variation that would affect the ACE2-RBD binding region.

The risk for serious disease and death from COVID-19 is higher in males, in older individuals and those with co-morbidities. Host genetic variation is essential in determining susceptibility to COVID-19 as well as in variation in course of disease and outcomes. It has been recently discovered that polymorphisms in the host IFNL region which control expression of IFNL3 and

IFNL4 modulate ACE2 expression and could affect varied response to COVID-19⁶. Furthermore, there are age related differences in expression of ACE2 and in ACE2 relative to ACE. The ACE2/ACE ratio is much higher among the young as compared to the older individuals. Furthermore, the fact that the human ACE2 protein is located on the X chromosome means that males who carry rare ACE2 coding variants will be hemizygous, expressing only those variants in all ACE2-expressing cells. Conversely, females will typically express those rare ACE2 variants in a mosaic distribution determined by early X-inactivation events⁷. In a recent study, site-directed mutagenesis analysis of ACE2 interface interactions identified polymorphisms expected to have an impact on the function of the resulting protein, but also grouped these variants based on their predicted effects on interactions with the SARS-CoV-2 spike protein⁸. Authors report variants that could alter the binding affinity to for example the p.Thr92Ile variant (rs763395248) amongst others⁸.

The structural basis for the ACE-2 RBD Spike protein recognition has been recently mapped out and the cryo-EM structure of the full length, viral spike protein that targets human ACE2 complex has been reported⁹. The viral spike glycoprotein (S protein) mediates receptor recognition¹⁰. The authors show that viral S protein binds ACE2 up to x20 times more tightly, when compared to the previous SARS-CoV strain. Recently the 3.5-angstrom-resolution structure of the S protein has been described¹⁰. The S protein is cleaved into two subunits: S1 and S2. This cleavage of S proteins by host proteases is critical for viral infection⁹. The structure of ACE2 as well as that of Spike protein is shown on **Figure1**. The interaction between the viral Spike protein and ACE2 is a critical step in the replication cycle. Currently there are two models describing the interaction^{11,12}. Both studies show the C-terminal domain (CTD, the part of the Spike protein where RBD is located) physically attaching to host ACE2 (the Spike protein's N-terminal domain, NTD, does not bind). Authors¹² show interaction of the CTD (but not NTD) of Spike by flow cytometry (expressed in mouse cells). Interaction was shown between partial structure SARS-CoV-2 Spike (RBD, residues Arg319–Phe541) and the N-terminal peptidase domain of ACE2 (residues Ser19–Asp615)¹¹. In the trimer structure, only one molecule of RBD is exposed (in the up conformation) to directly interact with ACE2^{11,12}. Additionally S binding is independent of ACE2 catalytic activity and occurs on the outer surface of ACE2, whereas angiotensin substrates bind within a deep cleft that houses the active site⁸.

In the present study, we first identified via crystal structure and genetics a target amino acid region, and second we identified potential small molecule inhibitors of RBD-ACE2 binding. Specifically we focus on small molecules and peptides that have the capacity to effectively bind to histidine 34 (His³⁴) of the ACE2 protein in order to modulate viral binding.

Results

Natural genetic variation in ACE2 likely affect the propagation capacity of the virus and variation in susceptibility across individuals. The capacity of viral binding to target cell might be altered by natural, human variation that is present within ACE2 RBD region. Looking at tolerance status of ACE2 in gnomAD¹³ database, there are on average 3 LOF variants in *ACE2* and the gene is predicted to be loss-of-function variant intolerant with a pLI of 1. Among reported inherited ACE2 missense variant list from gnomAD, we considered variants near the ACE2S protein interface. We inspected the region for nonsynonymous polymorphisms and the LOF variants with MAF>0.0005 are shown on **Figure1A** and **Figure1B**. One interesting variant closest to the region of interest is rs4646116, missense ACE2 locus (MAF_gnomAD_all = 0.003) encoding p.Lys26Arg. Rare variants (MAF_gnomAD_all≤0.0005) are not shown. All of these variants could plausibly affect risk for progression to COVID-19 after an initial exposure with SARS-CoV-2 which greatest effect coming from p.Lys26Arg as it is the more common variation.

We used the most recent inter-species susceptibility data to narrow down whichaa residues are crucial for maintaining ACE2 receptor entry¹⁴. SARS-CoV-2 replicates poorly in dogs, pigs, chickens and ducks, whereas cats and ferrets are more permissive to infection¹⁴. We aligned the ACE2 sequence across species based on a study by comparing susceptibility to infection¹⁴. We focused on the functional contact amino acid residues located on either ACE2 or the RBD domain (Spike protein)¹⁰. The authors show RBD residues Gln⁴⁹⁸, Thr⁵⁰⁰, and Asn⁵⁰¹ form a network of H-bonds with Tyr⁴¹, Gln⁴², Lys³⁵³, and Arg³⁵⁷ from ACE2. Additionally Lys⁴¹⁷ and Tyr⁴⁵³ of the RBD interact with Asp³⁰ and His³⁴ of ACE2⁹.

We aligned multiple amino acid sequences of ACE2 at the RBD-interacting domain across selected species, and surprisingly one amino acid was segregating with inter-species susceptibility (Histidine 34, region shown on **Figure1C**). We hypothesize that this one amino acid is essential for effective viral entry and we therefore sought to identify small molecules and

peptides that could block or mask this residue. We focused on a simultaneous *in silico* chemical library screen in order to find potential small molecule candidates that will selectively bind or mask that residue, effectively inhibiting SARS-CoV-2 binding to ACE2 and subsequent viral entry.

A glide docking protocol was applied followed by the determination of pharmacokinetically relevant molecular descriptors to narrow down the initial hits. Next, molecular dynamics simulations were conducted to validate the stability of docked binding modes. Altogether ~11,000 non-redundant molecules were tested (Tocris, TargetMol and FDA approved – downloaded from Zinc database¹⁵). We report a list of small molecules in Table 1 together with the docking Glide Score¹⁶. We used visual docking modeling as well as the docking score to guide and select candidates. Docking score (GlideScore) is an empirical scoring function designed to maximize separation of compounds with strong binding affinity from those with little to no binding ability. As an empirical scoring function it is comprised of terms that account for the physics of the binding process including a lipophilic-lipophilic term, hydrogen bond terms, a rotatable bond penalty, and contributions from protein-ligand coulomb-vdW energies. A lower docking score corresponds with a more optimal docking. Candidate molecules were assessed for actual interactions at the docking site, hbond score (hbond term is lower if the hydrogen bonds are more optimal, for example a closer distance) and ligand efficiency (normalized version of the glide score, gscore divided by a number of heavy atoms).

The top binding candidates are shown in **Table 1**. A full list is provided in Supplementary Material 1. AY-NH₂ is a selective PAR4 receptor agonist peptide (H-Ala-Tyr-Pro-Gly-Lys-Phe-NH₂) and yielded the most favorable docking score. **Figure 2A** is a graphical depiction of the AY-NH₂ compound annotated according to the determinations described above with respect to predicted interactions of the compound with the relevant amino acids of ACE2. These include predicted hydrogen bonding with the ACE2 Asp³⁰, Ala³⁸⁷, Gln³⁸⁸, and Glu⁵⁶⁴ amino acids. Figure 3B is a three-dimensional representation of AY-NH₂ bound to ACE2 as predicted above. Bound as such, binding by the SARS-CoV-2 S protein RBD to the His³⁴ (blue) or ASP³⁰ (red) of ACE2 is effectively blocked by NY-HN₂ (grey).

Table 1 Top hits from in silico small molecule docking experiments

DrugID	Drug	CAS ID	glide rotatable	docking score	glide ligand efficiency	glide hbond
ZINC000098052516	AY-NH₂	352017-71-1	24	-8.541	-0.174	-0.916
ZINC000008214766	Nicotinamide adenine dinucleotide NAD⁺	53-84-9	15	-7.979	-0.181	-0.852
ZINC000001542931	Reproterol	54063-54-6	10	-7.71	-0.275	-1.169
ZINC000245219534	Thymopentin	69558-55-0	28	-7.645	-0.159	-1.423
ZINC000009228252	CGS 21680 HCl	124431-80-7	13	-7.57	-0.21	-1.084
ZINC000008215403	Disodium NADH	606-68-8	15	-7.467	-0.17	-1.214
ZINC000098052511	Nociceptin (1-7)	178249-42-8	25	-7.249	-0.154	-0.83
ZINC000257482989	Mupirocin	12650-69-0	20	-7.078	-0.202	-1.124
ZINC000026468553	SLIGRL-NH ₂	171436-38-7	29	-6.973	-0.152	-0.949

ZINC000000001872	Oxiniacic Acid	2398-81-4	1	-6.948	-0.695	-0.16
------------------	----------------	-----------	---	--------	--------	-------

NAD⁺ (oxidized nicotinamide adenine dinucleotide), a coenzyme involved in many metabolic reactions, yielded the second most favorable docking score. It is an essential pyridine nucleotide that serves as an electron carrier in metabolism. It was reported that significant decline in the plasma levels of NAD⁺ occurs with age¹⁷. Interestingly recent studies show SARS-CoV-2 infection of cell lines significantly dysregulates the NAD⁺ pathway with regards to NAD⁺ synthesis and utilization¹⁸. Furthermore, gene expression and data suggest that increasing NAD⁺ with nicotinamide may restore antiviral PARP functions. Viral replication and host cell homeostasis are dependent on NAD coenzymes including NAD⁺. PARPs which are upregulated due to the virus are dependent on NAD⁺ in order to transfer the ADP-ribose moiety¹⁸. It has been shown that PARP10-mediated loss in NAD⁺ could be restored by boosting NAD⁺ levels¹⁸. Furthermore, in a small placebo-controlled clinical trial of Nicotinamide riboside (NR) in older men, it was discovered that 1 gram of Nicotinamide riboside per day decreases the inflammatory response (including levels of IL-6, IL-5, IL-2), ultimately reducing cytokine storm¹⁹. **Figure 2B** is a graphical depiction of the NAD⁺ compound similarly annotated according to the determinations described above with respect to predicted interactions of the compound with the relevant amino acids of ACE2. **Figure 2B** is a three-dimensional representation of NAD⁺ bound to ACE2 as predicted. Bound as such, binding by the SARS-CoV-2 S protein RBD to the His³⁴ (blue) or ASP³⁰ (red) of ACE2 is effectively blocked by NAD⁺ (grey).

Reproterol (7-[3-[[2-(3,5-dihydroxyphenyl)-2-hydroxyethyl]amino]propyl]-1,3-dimethylpurine-2,6-dione) is a short-acting β_2 adrenoreceptor agonist approved for use in the treatment of asthma. It yielded the third most favorable docking score. **Figure 2C** is a graphical depiction of the reproterol compound annotated according to the determinations described above with respect to predicted interactions of the compound with the relevant amino acids of ACE2. **Figure 2C** is a three-dimensional representation of reproterol bound to ACE2 as predicted. Bound as such, binding by the SARS-CoV-2 S protein RBD to the His³⁴ (blue) or ASP³⁰ (red) of ACE2 is effectively blocked by reproterol (grey).

Thymopentin (H-Arg-Lys-Asp-Val-Tyr-OH) is a synthetic pentapeptide used to enhance the production of thymic T cells. It yielded the fourth most favorable docking score. **Figure 2D** is a

graphical depiction of the thymopentin compound annotated according to the determinations described above with respect to predicted interactions of the compound with the relevant amino acids of ACE2. **We furthermore show** a three-dimensional representation of thymopentin bound to ACE2 as predicted. Bound as such, binding by the SARS-CoV-2 S protein RBD to the His³⁴ (blue) or ASP³⁰ (red) of ACE2 is effectively blocked by thymopentin (grey).

The other compounds of Table 1 yielded lower docking scores but were still believed to be capable of acting to inhibit binding at His34 and other of the ACE2 amino acids noted above, thereby having the potential to inhibit SARS-CoV-2 infection. We previously evaluated the natural variation in ACE2 and RBD (Nextstrain²⁰) to ensure docking was not within a polymorphic region. Positive controls of actual, known docking molecules were therefore run. We have confirmed docking of TRMPRSS2 with ligand 56677007, as described recently in the literature²¹. We have further tested a benchmark set of molecules²². We describe this set of true positive controls and the mapping strategy in Supplementary Table.

Discussion

There are several strategies to block SARS-CoV-2 entry into cell, and further replication. Small-molecule inhibitors of SARS-CoV replication that block viral entry have been described to block SARS-CoV²³. In previous reports SSAA09E2 interferes with the interaction of the RBD with ACE2²³. There is a therapeutic strategy whereby they administer ACE2 soluble as a trap for the virus. Furthermore, given recent data that human recombinant soluble ACE2 can block early stages of SARS-CoV-2²⁴. Authors showed that clinical grade human recombinant soluble ACE2 reduced SARS-CoV-2 recovery from Vero cells by a factor of 1,000-5,000²⁴. The described candidates from our screen call for further validation. One option would be to use models of CoV-2 infection on Caco-2 cells with overexpression of ACE2 differentiating only by this one amino acid – quick proof of concept. There are other ACE2 residues that are segregating with susceptibility across other species, but they were not confirmed RBD-contact residues. Those residues were initially ruled out in the light of extensive work that was done to elucidate RBD/ACE2 physical interaction using Electron Microscopy.

Conclusions

There is urgent need for the development of specific antiviral therapeutics to conquer SARS-CoV-2. In order to find novel inhibitors of ACE2-RBD binding, we computationally screened a compound libraries for binding using the solved crystal structure of ACE2. The focus on *in silico* screen delineating small molecules selectively binding/masking the histidine residue and consequently inhibiting SARS-CoV-2 binding to ACE2 and viral entry may be a potent strategy to reduce cellular entry of the virus. We suggest that this is a viable antiviral strategy whereas the small molecules, including peptides, warrant further validation.

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* **0**, (2020).
2. Tikellis, C. & Thomas, M. C. Angiotensin-Converting Enzyme 2 (ACE2) Is a Key Modulator of the Renin Angiotensin System in Health and Disease. *Int. J. Pept.* **2012**, 1–8 (2012).
3. Lonsdale, J. *et al.* The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
4. Zhao, Y. *et al.* Single-cell RNA expression profiling of ACE2, the putative receptor of Wuhan 2019-nCov. *bioRxiv* 2020.01.26.919985 (2020). doi:10.1101/2020.01.26.919985
5. Yang, X.-H. *et al.* Mice transgenic for human angiotensin-converting enzyme 2 provide a model for SARS coronavirus infection. *Comp. Med.* **57**, 450–9 (2007).
6. Ansari, M. A. *et al.* A gene locus that controls expression of ACE2 in virus infection. *medRxiv* 2020.04.26.20080408 (2020). doi:10.1101/2020.04.26.20080408
7. Gibson, W. T., Evans, D. M., An, J. & Jones, S. J. ACE 2 Coding Variants: A Potential X-linked Risk Factor for COVID-19 Disease. *bioRxiv* 2020.04.05.026633 (2020). doi:10.1101/2020.04.05.026633

8. Procko, E. The sequence of human ACE2 is suboptimal for binding the S spike protein of SARS coronavirus 2. *bioRxiv* 2020.03.16.994236 (2020). doi:10.1101/2020.03.16.994236
9. Yan, R. *et al.* Structural basis for the recognition of the SARS-CoV-2 by full-length human ACE2. *Science* (2020). doi:10.1126/science.abb2762
10. Wrapp, D. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020).
11. Lan, J. *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215–220 (2020).
12. Wang, Q. *et al.* Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* (2020). doi:10.1016/J.CELL.2020.03.045
13. Karczewski, K. J. *et al.* Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *bioRxiv* 531210 (2019). doi:10.1101/531210
14. Shi, J. *et al.* Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS-coronavirus 2. *Science* (2020). doi:10.1126/science.abb7015
15. Irwin, J. J. & Shoichet, B. K. ZINC – A Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model.* **45**, 177 (2005).
16. Friesner, R. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **47**, 1739–1749 (2004).
17. Clement, J., Wong, M., Poljak, A., Sachdev, P. & Braidy, N. The Plasma NAD⁺ Metabolome Is Dysregulated in “Normal” Aging. *Rejuvenation Res.* **22**, 121 (2019).
18. Heer, C. D. *et al.* Coronavirus Infection and PARP Expression Dysregulate the NAD Metabolome: A Potentially Actionable Component of Innate Immunity. *bioRxiv* 2020.04.17.047480 (2020). doi:10.1101/2020.04.17.047480
19. Elhassan, Y. S. *et al.* Nicotinamide Riboside Augments the Aged Human Skeletal Muscle NAD⁺ Metabolome and Induces Transcriptomic and Anti-inflammatory Signatures. *Cell Rep.* **28**, 1717–1728.e6 (2019).

20. Hadfield, J. *et al.* Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121–4123 (2018).
21. Rensi, S. *et al.* Homology Modeling of TMPRSS2 Yields Candidate Drugs That May Inhibit Entry of SARS-CoV-2 into Human Cells. (2020). doi:10.26434/CHEMRXIV.12009582.V1
22. Bohari, M. H. & Sastry, G. N. FDA approved drugs complexed to their targets: evaluating pose prediction accuracy of docking protocols. *J. Mol. Model.* **18**, 4263–4274 (2012).
23. Adedeji, A. O. *et al.* Novel Inhibitors of Severe Acute Respiratory Syndrome Coronavirus Entry That Act by Three Distinct Mechanisms. *J. Virol.* **87**, 8017–8028 (2013).
24. Monteil, V. *et al.* Inhibition of SARS-CoV-2 Infections in Engineered Human Tissues Using Clinical-Grade Soluble Human ACE2. *Cell* (2020). doi:10.1016/j.cell.2020.04.004