An *In silico* scientific basis for LL-37 as a therapeutic and Vitamin D as preventive for Covid-19.

Kiran Bharat Lokhande[1,2], Tanushree Banerjee[1,3], K. Venkateswara Swamy[4], Manisha Deshpande[1]

[1]Dr. D.Y. Patil Biotechnology and Bioinformatics Institute, [1]Dr. D.Y. Patil Vidyapeeth, [2]Bioinformatics Research Laboratory, [3]Molecular Neuroscience Research Laboratory, Bangalore-Mumbai Highway, Tathawade, Pune – 411 033, India.

[4]MIT School of Bioengineering Sciences and Research, MIT Art, Design and Technology University, Pune - 412 201, India.

Correspondence:-

Dr. Manisha Deshpande,

Dr. D.Y. Patil Biotechnology and Bioinformatics Institute,

Dr. D.Y. Patil Vidyapeeth, Bangalore-Mumbai Highway,

Tathawade, Pune – 411033, India.

Email: manisha.deshpande@dpu.edu.in

Phone: +91-20-67919444

Fax: +91 20 27420010

ABSTRACT

Even as clinical trials are underway for vaccines and therapeutics for Covid-19, establishment of modalities with a strong and complete foundation is still awaited and until then, the uncertainty remains associated. Thus, there is a requirement to research as many new and different types of approaches as possible to tackle the pandemic. In this report, *in silico* scientific findings are presented, which are indicative of the putative potential for the use of the LL-37 human anti-microbial peptide as a therapeutic or possibly even as a prophylactic against SARS-CoV-2. This indication is mainly based on the high structural similarity of LL-37 to the N-terminal helix of the receptor-binding domain of SARS-CoV-2, and the positive prediction of binding of LL-37 to the receptor-binding domain of SARS-CoV-2. Also, as Vitamin D is known to upregulate the expression of LL-37, the vitamin is a candidate preventive molecule. This report also provides the possible basis for why there is an inverse correlation between Vitamin D levels in the body and the severity of or susceptibility to Covid-19, as described in a large body of published literature. The path for development of LL-37 as a therapeutic could be of lesser duration, as LL-37 is native to the human body. With the scientific link put forth in this work, Vitamin D could be used at an effective, medically prescribed dose as a preventive measure. As Vitamin D is insoluble in water, it should be taken only in consultation with a medical practitioner to prevent adverse effects of its accumulation in the body. The information in this report would be valuable in bolstering the worldwide efforts to control the pandemic as early as possible.

KEYWORDS:- SARS-CoV-2, Receptor Binding Domain, LL-37, Structural analysis, Molecular Docking, Vitamin D.

INTRODUCTION

The world is plagued with the peculiar and daunting problem that a vast population of humans does not mount an effective innate immune response against SARS-CoV-2 (Vabret *et al*, 2020). World-wide efforts are ongoing for the development of vaccines and therapeutics. These efforts need to surmount different challenges in producing effective and safe vaccines, additionally confounded by the nature and behaviour of SARS-CoV-2 (Calina *et al*, 2020). Likewise, there are various problems associated with drugs and drug development for Covid-19 (El-Aziz & Stockand, 2020). Considering the difficulties well-reviewed in these articles, there should be research in the direction of as many *different* types of approaches that can be thought of, including non-typical approaches. In this work, a different possibility is presented, in the form of the human anti-microbial peptide LL-37 as a therapeutic, the reasoning for which is described below, after a brief introduction to (i) the binding of SARS-CoV-2 to its receptor Angiotensin-Converting Enzyme-2 (ACE2) on cells, (ii) the cathelicidin-derived LL-37 peptide, and (iii) the upregulation of LL-37 by Vitamin D.

SARS-CoV-2 attaches to the cell surface receptor ACE2 by the binding of the Receptor Binding Domain (RBD) of its Spike protein to ACE2. The N-terminal helix (NTH) of the peptidase domain of ACE2 is mainly responsible for the interaction with the RBD (Lan *et al*, 2020). The Receptor Binding Motif (RBM) within the RBD, forms a concavity into which the NTH fits in. Most of the residues of ACE2 that form interactions with the RBD are present in the N-terminal helix. A stretch of ACE2 from Ser19 to Asn53 comprises the NTH (Chowdhury & Maranas, 2020).

LL-37, a 37 amino acid peptide, produced from the cleavage of an 18 kDa polypeptide, hCAP18, is the only anti-microbial peptide (AMP) present in humans which belongs to the cathelicidin family (Agier *et al*, 2015). It is produced not only by multiple immune cells like NK cells, B cells, mast cells etc, but also by epithelial cells present in the skin and the respiratory tract (Mansour *et al*, 2014). LL-37, the peptide which has the anti-microbial activity, is produced by extracellular cleavage of hCap18 (Sorenson *et al*, 2001) and is majorly present extracellularly. Within cells, it is present in its proform in granules. LL-37 acts as a polycationic agent and binds to the anionic regions of the membranes of pathogens (Burton & Steel, 2009). There are many examples of internalization of active LL-37. To name a few, it can get internalized into human macrophages where it can act in clearance of intracellular pathogens (Tang *et al*, 2015). LL-37 is internalized by human mast cells which it induces to

release nucleic acids (Dahl *et al*, 2018). LL-37 also forms complexes with extracellular DNA and transports it into monocytes and also plasmacytoid dendritic cells (Chamilos *et al* 2012). Zhang *et al* (2010) have reported about LL-37 being a cell-penetrating peptide (CPP).

The physiological concentration of LL-37 is known to be in the range of 2-5 µg/ml. However, in case of infection its concentration could rise up to 20 µg/ml (Tripathi *et al*, 2015). LL-37 is reported to be cleaved even further in the epithelial cells into more potent anti-microbial peptides which can act as a chemoattractant for various immune cells like neutrophils, dendritic cells, T cells, etc (Bandurska *et al*, 2015). In broncho-alveolar fluid, nasal epithelium, nasal secretions, LL-37 concentrations are very high especially during infections (Tripathi *et al*, 2015). LL-37 also exhibits antiviral activity against several viruses by interacting with the viral envelopes (Ahmed, 2019). LL-37 has also been found to interfere with the replication of several single-stranded enveloped RNA viruses similar to SARS-CoV-2, such as Respiratory Syncytial Virus, Influenza A hepatitis C virus, Dengue virus, HIV-1 and Vaccinia Virus (Crane-Godreau *et al*, 2020).

A metabolite of Vitamin D3, 1α,25- dihydroxy vitamin D3 (1,25D3), is known for its role in transcriptional regulation of many genes of the immune system (Gurlek *et al*, 2002). Genes regulated by 1,25D3 not only code for pro-inflammatory mediators, cytokines, chemokines but also for anti-microbial peptides. Human antimicrobial peptide hCAP-18/LL-37 gene is one such gene which is upregulated by Vitamin D (Svensson *et al*, 2016).

Moving on to the crux of the matter, i.e., the logic of this study, a large collection of evidence has mounted that there is a correlation between Covid-19 and Vitamin D. In PubMed, there are numerous articles reporting or discussing such correlation. To give an example, in an Israeli population-based analysis involving 7807 people, 10% individuals who were Covid-19 positive had lower 25(OH) Vitamin D levels in the plasma than the individuals who were tested negative and the finding was statistically significant (Merzon *et al*, 2020). In another work, which is most relevant to our study, the binding of LL-37 to enveloped viruses and their consequent disablement has been described and it has been postulated that Vitamin D deficiency manifests as aggravated Covid-19 through consequent low expression of cathelicidin (Crane-Godreau *et al*, 2020). In our study, the question was asked - Could this association be something that could be tapped for developing a potential therapy? LL-37 is known to have a helical structure (Zhao *et al*, 2018) as depicted in Figure 1. The actual part of ACE2 that the RBD binds to is the N-terminal helix (NTH) (Figure 1) of ACE2. Both being

helical, the PDB structure of NTH was obtained by submitting its sequence to the structure prediction server Phyre2 (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index; Kelley *et al*, 2015) (Figure 1). This PDB structure was submitted as a query in the Protein structure comparison service, PDBeFold at European Bioinformatics Institute (http://www.ebi.ac.uk/msd-srv/ssm) (Krissinel & Henrick, 2005; Krissinel & Henrick, 2004). The second molecule that emerged in the results of this search was LL-37 (Figure 1).

Thus, a hypothesis was made that LL-37, could be binding to RBD and if true, this could be one of the most effective way in which LL-37 could interfere with SARS-CoV-2 by binding to the receptor-binding domain (RBD) of the virus. This, in fact, is the scientific basis for the development of vaccines that give rise to neutralizing antibodies (Tian *et al*, 2020), and also therapeutics which act by binding to the RBD, and hence preventing attachment of the virus to its receptor ACE2 (Huang *et al*, 2020). Therefore, in this work, *in silico* studies were undertaken, mainly to test the hypothesis put forth and to unearth the possibility, if any, of binding of LL-37 to RBD, just as RBD binds to NTH. A great value of this approach to develop a treatment would be that a molecule from the body itself could be used as a therapeutic.

The Spike protein RBD is amply glycosylated (Shajahan *et al*, 2020) and the glycan moieties are considered to be steric hindrances for neutralizing antibodies to bind – a mechanism of immune evasion by the virus. The N-glycans on the Spike protein affect its folding and consequently, the immune surveillance system is evaded and an antibody response averted (Sternberg & Naujokat, 2020) . The glycans have been reported to control the conformation of the RBD (Henderson *et al*, 2020). The RBD does get exposed when the virus is about to bind to ACE2; this is referred to the "UP" state of the RBD, as against the unavailable "DOWN" state, these states being alternated by the glycans which is also discussed by Cai *et al* (2020). Indeed neutralizing antibodies to this UP state of the RBD have been found in the plasma of convalescent Covid-19 patients (Barnes *et al*, 2020). Also, the development of neutralizing antibodies as a therapeutic depends on the "Open" or available state of the RBD as against its "closed" state (Mercurio *et al*, 2020). So, in the UP or Open state of the RBD, LL-37 could also bind to it.

METHODS

Retrieval of 3-D Structures and their Preparation

The NMR structure of human LL-37 and two X-ray crystal structures of SARS-CoV-2 Receptor-Binding Domain (RBD) in complex with its receptor i.e. human ACE2 (hACE2) were retrieved from the PDB database (https://www.rcsb.org/) with PDB ID: 2K6O (Wang, 2008), 6LZG (Wang *et al*, 2020) and 6M0J (Lan *et al*, 2020) respectively. In this report, we have used another two peptides as negative controls that are structurally almost similar to LL-37 (Figure 1). The 3-dimensional solution structures determined by solution spectroscopy of these two peptides viz. transmembrane domain of syndecan-2 (TDS-2) (PDB ID: 6ITH) (Li *et al*, 2019) and glucagon-like peptide-2 (GLP-2) (PDB ID: 2L63) (Venneti & Hewage, 2011) were retrieved from the PDB database.

The Schrödinger's Protein Preparation Wizard tool (Sastry *et al*, 2013**;** Schrödinger Release 2020-2**:** Protein Preparation Wizard, Schrödinger, LLC, New York, NY, 2020) was used to prepare the downloaded PDB structures for structural correctness with respect to optimization of H-bonds and energy minimization of heavy atoms by using OPLS-2005 force field, resulting into high-confidence structures for further molecular docking studies.

Quantitative Assessment of NTH Structure Similarity with LL-37, TDS-2, and GLP-2

To assess the structural similarity between N-terminal Helix (NTH) of hACE2 with human anti-microbial peptide LL-37, and two other unrelated helical peptides viz. TDS-2 and GLP-2, the quantitative assessment of protein structure similarity was carried out using TM-align online server (Zhang & Skolnick, 2005; https://zhanglab.ccmb.med.umich.edu/TM-align/). The TM-align program produces the optimized residual alignment between two structures depending on their structural similarity followed by structural superimposition and gives the TM-score which is a measure of the structural similarity. TM-score gives the values in between 0 to 1, where 0 indicates the given structures are structurally different while 1 indicates perfect alignment and implies that both the structures are identical to each other. If the TM-score higher than 0.5, then it signifies that the similar structural folds are present in both the structures as described in SCOP/CATH database. Further, a greater value for TM-score signifies stronger structural similarity (Xu & Zhang, 2010).

Macromolecular or Protein-Protein Docking

For the macromolecular docking of given structures, we used the HDOCK online server (http://hdock.phys.hust.edu.cn/) to predict their binding complexes with binding affinity. HDOCK predicts the intermolecular interactions at the interface of two proteins in binding

complexes through a hybrid algorithm (Yan *et al*, 2020). The HDOCK algorithm executes rigid docking by considering both the receptor and ligand as rigid molecules.

In this study, we performed macromolecular docking of LL-37 (PDB ID: 2K6O) with RBD of SARS-CoV-2 (from PDB ID: 6LZG). For this, we separated the RBD-ACE2 complex structures using Maestro software (Schrödinger Release 2020-2**:** *Maestro, Schrödinger, LLC, New York, NY, 2020*) and used only RBD structural coordinates for docking. Then, these docking results were compared with site-specific positive and negative control docking. Positive control docking i.e. docking of SARS-CoV-2 RBD with its receptor hACE2 was done by taking the 6LZG PDB complex, separating the two protein structures and then docking them. For negative control docking, we used two unrelated helical peptides (PDB IDs: 6ITH and 2L63) and docked them with SARS-CoV-2 RBD to evaluate binding affinities and interacting residues. Also, as a different type of negative control, LL-37 was docked with the Spike protein region from residue 96 to residue 318, which is adjacent (upstream) to RBD by extracting the structure of this region from the full spike protein structure of SARS-CoV-2 (PDB ID: 6Z97). All docked complexes were downloaded from the HDOCK server and their intermolecular interaction pattern was analyzed using Maestro.

RESULTS

Structural Alignment of LL-37 and the other peptides TDS-2 and GLP-2, with NTH

Initially, the validation of the TM-align server was done by giving the same hACE2 NTH structure as structure 1 and structure 2 inputs to the server, to find out whether the same structure produces a TM-score of 1 or not. The resulting TM-score of this alignment was found to be 1 with RMSD 0 Å and this confirmed the accuracy of the TM-align server. Next, the structural alignment of hACE2 NTH was carried out with LL-37, TDS-2 (35 amino acids long), and GLP-2 (33 amino acids long), resulting in TM-scores and RMSD values from the TM-align program as summarized in Table 1. The results of TM-alignment give insights into the structural similarities and reveals that the hACE2 NTH is structurally very similar to LL-37 than the other two peptides and is in similar structural folds as NTH.

The structural alignments generated from the TM-align server are shown in Figure 2. It can be seen that LL-37 even shares a similar bend in the helix as NTH, and this could help

in fitting of LL-37 into the concavity of RBD. Also, although the peptide TDS-2 has aligned with NTH with a score a little greater than 0.5, it can be seen from Figure 2, that there is a clear bifurcation of the two structures towards the N-terminus. This study suggests that the reported human anti-microbial peptide LL-37 is structurally very close to hACE2 NTH and this feature could facilitate its binding to RBD.

Intermolecular Interactions between hACE2 and RBD (Positive Control)

The already determined crystal structure of SARS-CoV-2 RBD complexed with hACE2 (PDB ID: 6LZG) was used for docking analysis to predict the *in silico* derived complex by HDOCK server. The complexed structure was split into separate SARS-CoV-2 RBD and separate hACE2 structure and subjected to macromolecular docking calculations. The docking results gave a structurally very similar complex as compared to the crystal structure (PDB ID: 6LZG) with -353.05 Kcal/mol docking energy. Intermolecular interactions between this docked complex are summarized in Table 2. Figure 3 shows the binding mode of RBD with hACE2 with a detailed interaction pattern. The N-Terminal Helix of hACE2 is exposed towards the Receptor Binging Motif (RBM) of SARS-CoV-2 RBD and this interface makes very stable complex by forming twelve different interactions comprising of four hydrogen bond interactions, six aromatic hydrogen bonds, and two salt bridges.

$Thr27_{hACE2}$ and $Asp30_{hACE2}$ are involved in aromatic hydrogen bonding with $Phe456_{RBD}$ at bond distances of 3.01 Å and 3.47 Å respectively. Amino acid residue $Asp30_{hACE2}$ also forms a hydrogen bond and a salt bridge with $Lys417_{RBD}$ at 1.94 Å and 2.87 Å bond distances respectively. $Lys31_{hACE2,}$ $GLu35_{hACE2}$ and $Asp38_{hACE2}$ interacted with $Glu484_{RBD}$, $Gln493_{RBD}$ and $Tyr449_{RBD}$ with salt bridge, hydrogen bond, and aromatic hydrogen bond interactions respectively. $Tyr83_{hACE2}$ forms one aromatic hydrogen bond and one hydrogen bond interaction with $Asn487_{RBD}$ at a bond distance of 2.76 Å and 1.78 Å respectively. Other residues from hACE2 viz. $Phe28_{hACE2}$ and $Met82_{hACE2}$ form aromatic hydrogen bond interactions with $Tyr489_{RBD}$ and $Phe486_{RBD}$. Also, $Gln42_{hACE2}$ interacted with $Gly446_{RBD}$ by forming one hydrogen bond with 2.74 Å bond distance. It was analyzed that the residues of the RBD found to interact with ACE2 in the docking, viz., residue numbers 417, 446, 449, 456, 486, 487, 489 and 493, were common with the RBD residues interacting with ACE2 in the crystal structure determined by Lan *et al* (2020). This showed the accuracy of the docking method used in this work.

Intermolecular Interactions between LL-37 and RBD (Test)

The docked complexes of LL-37-RBD_{6LZG} and LL-37-RBD_{6M0J} gave docking energies of -253.75 Kcal/mol and -238.18 Kcal/mol respectively (a more negative value indicates stronger binding and stable complex). The intermolecular interactions between LL-37 and SARS-CoV-2 RBD from 6LZG and 6M0J PDB are illustrated in Figure 4(A) and Figure 4(B) respectively.

LL-37-RBD (PDB ID: 6LZG) Complex: Eight interactions including three hydrogen bonds, three aromatic hydrogen bonds, and two salt bridges were detected at the interface of the LL-37-RBD complex, and the detailed interaction analysis is summarized in Table 3. The amino acid residue $Glu16_{LL-37}$ forms a hydrogen bond and a salt bridge with the amino acid residue $Lys417_{RBD}$ at a bond distance of 2.53 Å and 3.18 Å respectively. $Glu16_{LL-37}$ is also involved in the formation of an aromatic hydrogen bond with $Tyr421_{RBD}$. $Arg23_{LL-37}$ forms one hydrogen bond and one salt bridge with $Tyr505_{RBD}$ and $Glu406_{RBD}$ respectively. The amino acid residues viz. $Asp26_{LL-37}$ and $Phe27_{LL-37}$ form an aromatic hydrogen bond with residues $Tyr505_{RBD}$ and $Gly496_{RBD}$ respectively with bond distances of 2.73 Å and 2.75 Å. In addition to this, $Asn30_{LL-37}$ also forms a strong hydrogen bond with $Asn501_{RBD}$ at 1.67 Å bond distance.

LL-37-RBD (PDB ID: 6M0J) Complex: In this complex also, eight interactions were observed at the interface including two hydrogen bonds, four aromatic hydrogen bond interactions and two salt bridges (Table 3). Here also $Glu16_{LL-37}$ interacted with SARS-CoV-2 RBD by forming a salt bridge with $Lys417_{RBD}$, hydrogen bond, and aromatic hydrogen bond with $Tyr421_{RBD}$ at a bond distance of 3.09 Å, 1.94 Å and 2.2 Å respectively. $Glu16_{LL-37}$ also forms one aromatic hydrogen bond with $Phe456_{RBD}$ at 3.38 Å bond distance. Another amino acid residue $Arg23_{LL-37}$ forms a hydrogen bond and one salt bridge with $Tyr505_{RBD}$ and $Glu406_{RBD}$ at 2.74 Å and 3.25 Å bond distances respectively. $Asp26_{LL-37}$ and $Phe27_{LL-37}$ form two aromatic hydrogen bond interactions with $Tyr505_{RBD}$ and $Gly496_{RBD}$ with 2.96 and 2.82 bond distances respectively.

From this docking study, it is clear that the residues from LL-37, viz., Glu16, Arg23, Asp26 and Phe27, also residues from SARS-CoV-2 RBD, viz., Glu406, Lys417, Tyr421, Gly 496 and Tyr505 are very important residues for binding as these interactions were observed in both the Test complexes. The regions of RBD spanned by LL-37 in each case are shown diagrammatically in Figure 5. Thus, five residues of RBD contacted by LL-37 are common

between the two test dockings. This finding also demonstrates that LL-37 binding to RBD appears to be specific.

Comparing the docking of LL-37 with RBD and hACE2 with RBD, hACE2 forms a very stable complex with RBD (-353.05 Kcal/mol binding energy), as not only N-terminal helix but other secondary structure elements of hACE2 are exposed to the surface binding site of SARS-CoV-2 RBD. In the case of LL-37, which consists of only a helix for docking with RBD, therefore the binding energy for this complex was -253.75 Kcal/mol. Despite the weaker binding energy of the LL-37-RBD complex, the stability of this complex is very similar to the stability of the hACE2-RBD complex because the scoring function of any docking program not only considers the number of interactions but also takes into consideration the short-ranged Van der Waals' and electrostatic interactions, the loss of entropy which occurs upon ligand binding, the hydrogen bonds and solvation factor. As seen in Table 2, not only the NTH of hACE2 contributed in the binding but other residues, viz., $Met82_{hACE2}$ and $Tyr83_{hACE2}$ are also involved in the binding with SARS-CoV-2 RBD resulting in a lower binding energy as compared to LL-37-RBD complex. The presence of salt bridges in both the Test complexes predicts thermostability of the binding of LL-37 to RBD.

Intermolecular Interactions between other peptides and RBD (Negative Control)

The HDOCK server gave the following docking energies for TDS-2-RBD and GLP-2-RBD complexes: -241.83 Kcal/mol and -202.59 Kcal/mol respectively. The detailed interaction analysis is summarized in Table 4. The interface between TDS-2 and SARS-CoV-2 RBD complex has six interactions including two hydrogen bonds, three aromatic hydrogen bonds, and one Pi-Pi stacking as shown in Figure 6(A). The amino acid residue $Val7_{TDS-2}$ interacted with $Tyr473_{RBD}$ by forming hydrogen bonding at a distance of 2.47 Å. $Ala10_{TDS-2}$ and $Val11_{TDS-2}$ interacted with the $Phe456_{RBD}$ by forming three aromatic hydrogen bonds at 3.14 Å, 3.26 Å and 3.35 Å bond distances respectively. Other residues, viz., $Phe24_{TDS-2}$ and $Arg32_{TDS-2}$ from TDS-2 form one Pi-Pi stacking with $Tyr505_{RBD}$ and one hydrogen bonding with $Thr500_{RBD}$ respectively. TDS-2 was found to interact with only four residues of RBD as compared to six residues contacted by LL-37, and the number of interactions was also lesser for TDS-2 than LL-37. No salt bridges are formed.

The other peptide GLP-2 forms only three interactions with RBD of SARS-CoV-2 and gives docking energy of -202.59 Kcal/mol. Two aromatic hydrogen bonds and one hydrogen

bond was observed between GLP-2-RBD binding interface as shown in Figure 6(B). The amino acid residue $His1_{GLP-2}$ forms one hydrogen bond with $Asn343_{RBD}$ and one aromatic hydrogen bond with $Glu340_{RBD}$ at a bond distance of 2.23 Å and 3.28 Å respectively. Also, $Phe22_{GLP-2}$ is involved in aromatic hydrogen bond interaction with Asn450 at 2.85 Å bond distance. No salt bridges are formed. Figure 5 shows the alignments of the regions of RBD spanned by the negative control peptides.

We also explored the binding stability of LL-37 peptide towards the region which is immediately adjacent and upstream to SARS-CoV-2 RBD by taking the secondary structure of RBD which includes the amino acid sequence from 96 to 318 (PDB ID: (6Z97) (Huo *et al*, 2020). This secondary structure was extracted from the full spike protein structure by using Maestro software and submitted to the HDOCK server for binding with LL-37. This docking result gives docking energy of -209.74 Kcal/mol by forming two interactions at the interface of this complex. Figure 6(C) shows the amino acid residue $Arg7_{LL-37}$ forms one salt bridge with Asp138 at a bond distance of 3.13 Å. $Phe17_{LL-37}$ is also involved in the Pi-Pi stacking with Phe238 Å at 2.80 Å bond distance. This also is indicative of the specificity of LL-37 binding to RBD.

From these docking studies, which are to be further refined, we can conclude that *in silico*, LL-37 shows a binding affinity and relative stability of binding towards the RBD of SARS-CoV-2.

Sequence variation in geographical isolates of SAR-CoV2:- For analyzing the frequency of variation in the sequence of SARS-CoV2 isolated from different geographical regions we used the 'Latest Global Analysis' tool of 'nextstrain.org' database wherein we downloaded the analyzed sequences of 3868 genomes of different geographical isolates of SARS-CoV-2 reported between Dec. 2019-June 2020 (Hadfield *et al*, 2018). Subsequently we analyzed the frequency of variation of amino acid residues which are observed to participate in the interaction with LL-37. As shown in Figure 5, the frequency of variation for these residues is less than 0.002 and hence seem to be well conserved.

Cell-penetrating peptide activity of LL-37:- Two prediction servers were used. The SkipCPP-Pred tool (http://server.malab.cn/SkipCPP-Pred/Index.html) (Wei *et al*, 2017) predicted LL-37 to belong to the cell-penetrating class with a prediction confidence of 0.806. The MLCPP algorithm (http://www.thegleelab.org/MLCPP/) (Manavalan *et al*, 2018) predicted LL-37 to be a CPP with a probability score of 0.68. The uptake efficiency was predicted as Low with a

probability score of 0.45. Additionally, LL-37 is found deposited as a CPP in the database CPPsite 2.0 (Agrawal *et al*, 2015; Gautam *et al*, 2012).

Safety analyses of LL-37:- It is recognized that peptides, even including LL-37, could have adverse effects when used as therapeutics (Chen & Lu, 2019). For possible delivery of LL-37 through the inhalatory route, LL-37 was assessed for allergenicity. Taking an example of a human polypeptide which is used by inhalation in clinical practice, namely insulin, it has been reported that allergic reactions can occur although insulin is an innate protein (Gatto *et al*, 2019; Mastrandrea, 2010). The algorithm AllergenFP v.1.0 (Dimitrov *et al*, 2014) predicted LL-37 to be a "probable non-allergen". Using the server AllerCatPro (Maurer-Stroh *et al*, 2019) LL-37 as a query gave the result "No Evidence" for allergenicity. AllergenFP v.1.0 predicted human Insulin (UniProt sequence ID P01308) to be a "probable allergen" and AllerCatPro gave the result "Strong Evidence" for insulin as an allergen.

The software program ToxinPred (Gupta *et al*, 2013) predicted LL-37 to be a Non-Toxin by all Support Vector Machine methods available on the server. For prediction of hemolytic activity, the tool HAPPENN, which employs neural networks method was used (Timmons & Hewage, 2020). LL-37 was predicted to have very low hemolytic scores of 0.073, 0.089, and 0.09 by the three methods available in the tool, degree of hemolytic activity increasing from 0 to 1.

DISCUSSION

In summary, based on our *in silico* findings, we putatively propose that LL-37 could be used as a therapeutic for Covid-19, and maybe as a prophylactic in the current state of emergency. Besides the known action of LL-37 against enveloped viruses, we provide *in silico* data for the binding of LL-37 to the RBD of SARS-CoV-2 itself. The good structural alignment of LL-37 with the NTH suggests that LL-37 could stop RBD from binding ACE2 even by simply occupying the space intended by the virus for NTH. The molecular docking studies predict strong binding interactions between LL-37 and RBD, further demonstrating the RBD-blocking potential of LL-37. Also, from analysis of the docking studies (Figure 5), it can be seen that LL-37 spans a considerable stretch of RBD, from Glu407 to Tyr506 in both the Test dockings which includes the entire Receptor Binding Motif. This again is indicative of the potential for effective blocking of the RBD. It may be noted that the interactions of the peptide TDS-2 with RBD span a much shorter region of RBD than LL-37. Also, although the

interactions of GLP-2 span a long region of RBD, the RBM is almost completely missed out by this peptide. This finding of LL-37 binding to RBD also provides a boost for the use of Vitamin D at an effective, medically prescribed dose for protection against SARS-CoV-2, especially as LL-37 is expressed in the respiratory epithelial cells also, and Vitamin D elevates LL-37 levels. To help control SARS-CoV-2 infection effectively, Vitamin D should be administered under consultation with medical practitioners. In the future, it would be helpful to design a water-soluble derivative of Vitamin D to avoid the adverse effect of hypervitaminosis.

Even when the RBD is in the UP state, it might be difficult for the immune surveillance mechanism of neutralizing antibodies to reach the RBD, and hence the efficient immune evasion by the virus. Also, it is reasonable to conjecture that a small peptide like LL-37 would more easily access the RBD in its UP state, than large molecules like neutralizing antibodies. Other studies using peptides based on the N-terminal helix itself, demonstrating *in vitro* binding to the RBD of SARS-CoV-2 (Romano *et al*, 2020) or *in vitro* anti-SARS-CoV activity (Han *et al*, 2006), provide support for our approach. In fact, using LL-37 as a mimic of NTH is closer to home than these studies, LL-37 being a molecule already present in the human circulation. Other peptides based on NTH would require to be evaluated for any effects of placing in circulation a part of a molecule (ACE2) which normally is a part of a cell-surface receptor and is not found in the blood. The prediction that LL-37 is a cell-penetrating peptide raises the interesting possibility of LL-37 intracellularly binding to the Spike protein of SARS-CoV-2 as the protein is produced in the infected cells and then, following viral particle assembly, release of virions from the cells which already have the RBD blocked by LL-37 and hence cannot infect new cells.

Regarding the safety of using LL-37 as a therapeutic, it has been cautioned that there could be adverse effects including cell membrane-destabilization above critical concentrations of LL-37 (Chen & Lu, 2019). A clinical trial testing the safety of LL-37 in the treatment of non-healing leg ulcers demonstrated that topical LL-37 application was safe (Grönberg *et al*, 2020). Now, LL-37 has been approved for Phase II clinical trials (https://clinicaltrials.gov; ClinicalTrials.gov Identifier: NCT04098562) for its anti-microbial action in the management of diabetic foot ulcers (Chen & Lu, 2019). There is also approval for a clinical study of intratumoral injection of LL-37 for the condition of melanoma (ClinicalTrials.gov Identifier: NCT02225366). At high concentrations *in vitro*, LL-37 has been found cytotoxic to different eukaryotic cells (Johansson *et al*, 1998), but in the presence of human serum, the cytotoxic

activity of LL-37 is inhibited, and this could be the reason why LL-37 is not cytotoxic in the body. Indeed, the cytotoxic effects of LL-37 secreted into the circulation are negated by the binding of LL-37 to plasma proteins such as apolipoprotein A-I (Ciornei *et al*, 2005). Data is also indicative that eukaryotic cells are more resistant (than prokaryotic cells) to the cytotoxicity of LL-37 due to the expression of heparan sulfate by eukaryotic cells and the structure and composition of their membranes (Zhang *et al*, 2010). It is also well elucidated by Xhindoli *et al* (2016) how the cytotoxic effects of LL-37 are counteracted by its binding to apolipoprotein A-I. The K_D of the binding is such that apolipoprotein A-I at its physiological concentration in plasma, can render inactive (including cytotoxic activity) 90% of the LL-37 molecules present at concentrations which would be toxic to the body's cells. This still leaves sufficient free LL-37 molecules for antimicrobial action. Now, it has been reported that excision of the N-terminal hydrophobic residues from LL-37 reduces its cytotoxicity (Ciornei *et al*, 2005). From Figure 2(E), it can be seen that the first three N-terminal amino acids (LLG) of LL-37 do not align with the NTH residues. Also, from the docking data, these 3 amino acids do not form interactions with RBD. This means that, for the development of LL-37 as a therapeutic, the option of removing 2-3 amino acids from the N-terminus is available.

The predicted non-allergenicity of LL-37 can mean that LL-37 could be administered to patients through the respiratory tract route, which would have more immediate efficacy in treating the disease. If administered intravenously, being a peptide, LL-37 could migrate from the circulation to the infected regions in the lungs. Peptide therapeutics enter the tissues from the vasculature by diffusion or convective extravasation, and transfer from the circulation to the tissues also depends on the properties of specific peptides (Diao & Meibohm, 2013). In relation to the lungs, as respiratory epithelial cells express hCap-18, LL-37 has been found to be present in the bronchial alveolar lavage fluid (Tjabringa *et al*, 2005), and increased levels have been detected in tracheal aspirates during infection. In the lungs, the normal role of LL-37 is to provide innate immunity against bacterial infections. In the case of Covid-19, SARS-CoV-2 infects the respiratory epithelial cells (Mason, 2020), the hCap-18 expression system of which could hence be adversely affected, which anyway is upregulated in response to bacterial infection. Thus, as we have shown that there could be a binding connection between LL-37 and RBD, after confirmatory experimental work, it *could* be advocated that external administration of LL-37 directly into the lungs would be beneficial in the treatment of Covid-19. Besides, the epithelial cells yet unaffected by the virus, can be stimulated to produce LL-37 by the use of Vitamin D. Caution would have to be exercised if LL-37 is to be used as a therapeutic, because

it is reported to contribute to inflammatory processes in the lungs. At the same time, LL-37 is also involved in wound repair involving skin epithelial cells (Tjabringa *et al*, 2005). This role of LL-37 *may* have (positive) implications for Covid-19. Grant et al (2020) have reviewed that Vitamin D could have a protective effect against Covid-19 severity through its upregulation of cathelicidin, which in turn results in decrease in concentrations of pro-inflammatory cytokines and increase in anti-inflammatory cytokines. Consequently, the cytokine storm associated with Covid-19 could be controlled. The development of LL-37 as a therapeutic for Covid-19 will have to be in the context of the various roles and effects of the molecule. But the fact that LL-37 as a therapeutic has received approval for clinical trials could help in paving the way for its use in the case of Covid-19.

The analysis of mutation frequencies of the residues of RBD that are bound by LL-37 in the docking studies, shows that these residues are not highly mutating ones, on the contrary, they are conserved. So, LL-37 could be effective against all present and future mutational variants of SARS-CoV-2. Refinement of the molecular docking interactions and prediction of other immune responses are part of upcoming studies. Further actual experimental studies would be required to evaluate the binding of LL-37 with the RBD. If such studies confirm the finding of our *in silico* studies, then LL-37 has a good potential to be developed as a therapeutic. As it is a molecule native to the human body, its development up the regulatory pathway would be of shorter duration than other vaccines and therapeutics being developed. A recent clinical study has reported that there is a decline in the levels of neutralizing antibodies in patients who recovered from Covid-19 within two to three months (Long *et al*, 2020). This finding, and the fact that the information landscape about the nature of SARS-CoV-2 is rather dynamic, paints a grim picture and also further highlights the need for developing different therapies and prophylactics at a rapid pace.

ACKNOWLEDGEMENTS

The authors are thankful for the valued support of Dr. D. Y. Patil Vidyapeeth, Pune, India, and of Prof. Jayanta K. Pal, Director, Dr. D. Y. Patil Biotechnology and Bioinformatics Institute, Pune, for carrying out this work. Senior Research Fellowship awarded to Mr. Kiran Bharat Lokhande (Project ID: 2019-3458; File No.: ISRM/11(54)/2019) by the Indian Council of Medical Research, New Delhi is also acknowledged.

COMPETING INTERESTS The authors declare no competing interests.

SUPPLEMENTARY INFORMATION

Due to the benefit that *could* putatively arise from this work towards tackling the pandemic, as a humanitarian gesture, no Intellectual Property Rights have been created for this work, and we are not carrying out further work beyond more *in silico* studies. Based on these studies, if any research laboratory across the world wishes to take up development of LL-37 as a therapeutic, the corresponding author may please be contacted.

REFERENCES

1. Abd El-Aziz, T. M., & Stockand, J. D. (2020). Recent progress and challenges in drug development against COVID-19 coronavirus (SARS-CoV-2) - an update on the status. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases*, *83*, 104327. https://doi.org/10.1016/j.meegid.2020.104327
2. Agier, J., Efenberger, M., & Brzezińska-Błaszczyk, E. (2015). Cathelicidin impact on inflammatory cells. *Central-European journal of immunology*, *40*(2), 225–235. https://doi.org/10.5114/ceji.2015.51359
3. Agrawal, P., Bhalla, S., Usmani, S. S., Singh, S., Chaudhary, K., Raghava, G. P., & Gautam, A. (2016). CPPsite 2.0: a repository of experimentally validated cell-penetrating peptides. *Nucleic acids research*, *44*(D1), D1098–D1103. https://doi.org/10.1093/nar/gkv1266
4. Ahmed, A., Siman-Tov, G., Hall, G., Bhalla, N., & Narayanan, A. (2019). Human Antimicrobial Peptides as Therapeutics for Viral Infections. *Viruses*, *11*(8), 704. https://doi.org/10.3390/v11080704
5. Bandurska, K., Berdowska, A., Barczyńska-Felusiak, R., & Krupa, P. (2015). Unique features of human cathelicidin LL-37. *BioFactors (Oxford, England)*, *41*(5), 289–300. https://doi.org/10.1002/biof.1225
6. Barnes, C. O., West, A. P., Jr, Huey-Tubman, K. E., Hoffmann, M., Sharaf, N. G., Hoffman, P. R., Koranda, N., Gristick, H. B., Gaebler, C., Muecksch, F., Lorenzi, J., Finkin, S., Hägglöf, T., Hurley, A., Millard, K. G., Weisblum, Y., Schmidt, F., Hatziioannou, T., Bieniasz, P. D., Caskey, M., … Bjorkman, P. J. (2020). Structures of Human Antibodies Bound to SARS-CoV-2 Spike Reveal Common Epitopes and Recurrent Features of Antibodies. *Cell*, S0092-8674(20)30757-1. Advance online publication. https://doi.org/10.1016/j.cell.2020.06.025

7. Burton, M. F., & Steel, P. G. (2009). The chemistry and biology of LL-37. *Natural product reports*, *26*(12), 1572–1584. https://doi.org/10.1039/b912533g
8. Cai, Y., Zhang, J., Xiao, T., Peng, H., Sterling, S. M., Walsh, R. M., Jr, Rawson, S., Rits-Volloch, S., & Chen, B. (2020). Distinct conformational states of SARS-CoV-2 spike protein. *Science (New York, N.Y.)*, eabd4251. Advance online publication. https://doi.org/10.1126/science.abd4251
9. Calina, D., Docea, A. O., Petrakis, D., Egorov, A. M., Ishmukhametov, A. A., Gabibov, A. G., Shtilman, M. I., Kostoff, R., Carvalho, F., Vinceti, M., Spandidos, D. A., & Tsatsakis, A. (2020). Towards effective COVID-19 vaccines: Updates, perspectives and challenges (Review). *International journal of molecular medicine*, *46*(1), 3–16. https://doi.org/10.3892/ijmm.2020.4596
10. Chamilos, G., Gregorio, J., Meller, S., Lande, R., Kontoyiannis, D. P., Modlin, R. L., & Gilliet, M. (2012). Cytosolic sensing of extracellular self-DNA transported into monocytes by the antimicrobial peptide LL37. *Blood*, *120*(18), 3699–3707. https://doi.org/10.1182/blood-2012-01-401364
11. Chen, C. H., & Lu, T. K. (2020). Development and Challenges of Antimicrobial Peptides for Therapeutic Applications. *Antibiotics (Basel, Switzerland)*, *9*(1), 24. https://doi.org/10.3390/antibiotics9010024
12. Chowdhury, R., & Maranas, C. D. (2020). Biophysical characterization of the SARS-CoV-2 spike protein 1 binding with the ACE2 receptor and implications for infectivity. bioRxiv, https://doi.org/10.1101/2020.03.30.015891.
13. Ciornei, C. D., Sigurdardóttir, T., Schmidtchen, A., & Bodelsson, M. (2005). Antimicrobial and chemoattractant activity, lipopolysaccharide neutralization, cytotoxicity, and inhibition by serum of analogs of human cathelicidin LL-37. *Antimicrobial agents and chemotherapy*, *49*(7), 2845–2850. https://doi.org/10.1128/AAC.49.7.2845-2850.2005
14. Crane-Godreau, M. A., Clem, K. J., Payne, P., & Fiering, S. (2020). Vitamin D Deficiency and Air Pollution Exacerbate COVID-19 Through Suppression of Antiviral Peptide LL37. *Frontiers in public health*, *8*, 232. https://doi.org/10.3389/fpubh.2020.00232
15. Dahl, S., Anders, E., Gidlöf, O., Svensson, D., & Nilsson, B. O. (2018). The host defense peptide LL-37 triggers release of nucleic acids from human mast cells. *Peptides*, *109*, 39–45. https://doi.org/10.1016/j.peptides.2018.10.001

16. Diao, L., & Meibohm, B. (2013). Pharmacokinetics and pharmacokinetic-pharmacodynamic correlations of therapeutic peptides. *Clinical pharmacokinetics*, *52*(10), 855–868. https://doi.org/10.1007/s40262-013-0079-0
17. Dimitrov, I., Naneva, L., Doytchinova, I., & Bangov, I. (2014). AllergenFP: allergenicity prediction by descriptor fingerprints. *Bioinformatics (Oxford, England)*, *30*(6), 846–851. https://doi.org/10.1093/bioinformatics/btt619
18. Gautam, A., Singh, H., Tyagi, A., Chaudhary, K., Kumar, R., Kapoor, P., & Raghava, G. P. (2012). CPPsite: a curated database of cell penetrating peptides. *Database : the journal of biological databases and curation*, *2012*, bas015. https://doi.org/10.1093/database/bas015
19. Gatto, N. M., Bracken, M. B., Kolitsopoulos, F., Duggan, W. T., Koch, G. G., Wise, R. A., & Jackson, N. C. (2019). Pulmonary and cardiovascular safety of inhaled insulin in routine practice: The Exubera Large Simple Trial (VOLUME). *Contemporary clinical trials communications*, *18*, 100427. https://doi.org/10.1016/j.conctc.2019.100427
20. Grant, W. B., Lahore, H., McDonnell, S. L., Baggerly, C. A., French, C. B., Aliano, J. L., & Bhattoa, H. P. (2020). Evidence that Vitamin D Supplementation Could Reduce Risk of Influenza and COVID-19 Infections and Deaths. *Nutrients*, *12*(4), 988. https://doi.org/10.3390/nu12040988
21. Grönberg, A., Mahlapuu, M., Ståhle, M., Whately-Smith, C., & Rollman, O. (2014). Treatment with LL-37 is safe and effective in enhancing healing of hard-to-heal venous leg ulcers: a randomized, placebo-controlled clinical trial. *Wound repair and regeneration : official publication of the Wound Healing Society [and] the European Tissue Repair Society*, *22*(5), 613–621. https://doi.org/10.1111/wrr.12211
22. Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., Open Source Drug Discovery Consortium, & Raghava, G. P. (2013). In silico approach for predicting toxicity of peptides and proteins. *PloS one*, *8*(9), e73957. https://doi.org/10.1371/journal.pone.0073957
23. Gurlek, A., Pittelkow, M. R., & Kumar, R. (2002). Modulation of growth factor/cytokine synthesis and signaling by 1alpha,25-dihydroxyvitamin D(3): implications in cell growth and differentiation. *Endocrine reviews*, *23*(6), 763–786. https://doi.org/10.1210/er.2001-0044
24. Hadfield, J., Megill, C., Bell, S. M., Huddleston, J., Potter, B., Callender, C., Sagulenko, P., Bedford, T., & Neher, R. A. (2018). Nextstrain: real-time tracking of pathogen

evolution. *Bioinformatics (Oxford, England)*, *34*(23), 4121–4123. https://doi.org/10.1093/bioinformatics/bty407

25. Han, D. P., Penn-Nicholson, A., & Cho, M. W. (2006). Identification of critical determinants on ACE2 for SARS-CoV entry and development of a potent entry inhibitor. *Virology*, *350*(1), 15–25. https://doi.org/10.1016/j.virol.2006.01.029
26. Henderson, R., Edwards, R. J., Mansouri, K., Janowska, K., Stalls, V., Gobeil, S., Kopp, M., Li, D., Parks, R., Hsu, A. L., Borgnia, M. J., Haynes, B. F., & Acharya, P. (2020). Controlling the SARS-CoV-2 spike glycoprotein conformation. *Nature structural & molecular biology*, 10.1038/s41594-020-0479-4. Advance online publication. https://doi.org/10.1038/s41594-020-0479-4
27. Huang, X., Pearce, R., & Zhang, Y. (2020). *De novo* design of protein peptides to block association of the SARS-CoV-2 spike protein with human ACE2. *Aging*, *12*(12), 11263–11276. https://doi.org/10.18632/aging.103416
28. Huo, J., Zhao, Y., Ren, J., Zhou, D., Duyvesteyn, H., Ginn, H. M., Carrique, L., Malinauskas, T., Ruza, R. R., Shah, P., Tan, T. K., Rijal, P., Coombes, N., Bewley, K. R., Tree, J. A., Radecke, J., Paterson, N. G., Supasa, P., Mongkolsapaya, J., Screaton, G. R., … Stuart, D. I. (2020). Neutralization of SARS-CoV-2 by Destruction of the Prefusion Spike. *Cell host & microbe*, S1931-3128(20)30351-6. Advance online publication. https://doi.org/10.1016/j.chom.2020.06.010
29. Johansson, J., Gudmundsson, G. H., Rottenberg, M. E., Berndt, K. D., & Agerberth, B. (1998). Conformation-dependent antibacterial activity of the naturally occurring human peptide LL-37. *The Journal of biological chemistry*, *273*(6), 3718–3724. https://doi.org/10.1074/jbc.273.6.3718
30. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols*, *10*(6), 845–858. https://doi.org/10.1038/nprot.2015.053
31. Krissinel, E., & Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta crystallographica. Section D, Biological crystallography*, *60*(Pt 12 Pt 1), 2256–2268. https://doi.org/10.1107/S0907444904026460
32. Krissinel, E., & Henrick, K. (2005). Multiple Alignment of Protein Structures in Three Dimensions. In: M.R. Berthold *et.al.* (Eds.): CompLife 2005, LNBI 3695, pp. 67--78. Springer-Verlag Berlin Heidelberg.

33. Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., Zhang, Q., Shi, X., Wang, Q., Zhang, L., & Wang, X. (2020). Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*, *581*(7807), 215–220. https://doi.org/10.1038/s41586-020-2180-5
34. Li, Q., Ng, H. Q., & Kang, C. (2019). Secondary structure and topology of the transmembrane domain of Syndecan-2 in detergent micelles. *FEBS letters*, *593*(5), 554–561. https://doi.org/10.1002/1873-3468.13335
35. Long, Q. X., Tang, X. J., Shi, Q. L., Li, Q., Deng, H. J., Yuan, J., Hu, J. L., Xu, W., Zhang, Y., Lv, F. J., Su, K., Zhang, F., Gong, J., Wu, B., Liu, X. M., Li, J. J., Qiu, J. F., Chen, J., & Huang, A. L. (2020). Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nature medicine*, 10.1038/s41591-020-0965-6. Advance online publication. https://doi.org/10.1038/s41591-020-0965-6
36. Manavalan, B., Subramaniyam, S., Shin, T. H., Kim, M. O., & Lee, G. (2018). Machine-Learning-Based Prediction of Cell-Penetrating Peptides and Their Uptake Efficiency with Improved Accuracy. *Journal of proteome research*, *17*(8), 2715–2726. https://doi.org/10.1021/acs.jproteome.8b00148
37. Mansour, S. C., Pena, O. M., & Hancock, R. E. (2014). Host defense peptides: front-line immunomodulators. *Trends in immunology*, *35*(9), 443–450. https://doi.org/10.1016/j.it.2014.07.004
38. Mason R. J. (2020). Pathogenesis of COVID-19 from a cell biology perspective. *The European respiratory journal*, *55*(4), 2000607. https://doi.org/10.1183/13993003.00607-2020
39. Mastrandrea L. D. (2010). Inhaled insulin: overview of a novel route of insulin administration. *Vascular health and risk management*, *6*, 47–58. https://doi.org/10.2147/vhrm.s6098
40. Maurer-Stroh, S., Krutz, N. L., Kern, P. S., Gunalan, V., Nguyen, M. N., Limviphuvadh, V., Eisenhaber, F., & Gerberick, G. F. (2019). AllerCatPro-prediction of protein allergenicity potential from the protein sequence. *Bioinformatics (Oxford, England)*, *35*(17), 3020–3027. https://doi.org/10.1093/bioinformatics/btz029
41. Mercurio, I., Tragni, V., Busto, F., De Grassi, A., & Pierri, C. L. (2020). Protein structure analysis of the interactions between SARS-CoV-2 spike protein and the human ACE2 receptor: from conformational changes to novel neutralizing antibodies. *Cellular and molecular life sciences : CMLS*, 1–22. Advance online publication. https://doi.org/10.1007/s00018-020-03580-1

42. Merzon, E., Tworowski, D., Gorohovski, A., Vinker, S., Golan Cohen, A., Green, I., & Frenkel Morgenstern, M. (2020). Low plasma 25(OH) vitamin D level is associated with increased risk of COVID-19 infection: an Israeli population-based study. *The FEBS journal*, 10.1111/febs.15495. Advance online publication. https://doi.org/10.1111/febs.15495

43. Romano, M., Ruggiero, A., Squeglia, F., & Berisio, R. (2020) An engineered stable mini-protein to plug SARS-Cov-2 Spikes. *bioRxiv.* https://doi.org/10.1101/2020.04.29.067728

44. Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R., & Sherman, W. (2013). Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of computer-aided molecular design*, *27*(3), 221–234. https://doi.org/10.1007/s10822-013-9644-8

45. Shajahan, A., Supekar, N. T., Gleinich, A. S., & Azadi, P. (2020). Deducing the N- and O- glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2. *Glycobiology*, cwaa042. Advance online publication. https://doi.org/10.1093/glycob/cwaa042.

46. Sørensen, O. E., Follin, P., Johnsen, A. H., Calafat, J., Tjabringa, G. S., Hiemstra, P. S., & Borregaard, N. (2001). Human cathelicidin, hCAP-18, is processed to the antimicrobial peptide LL-37 by extracellular cleavage with proteinase 3. *Blood*, *97*(12), 3951–3959. https://doi.org/10.1182/blood.v97.12.3951

47. Sternberg, A., & Naujokat, C. (2020). Structural features of coronavirus SARS-CoV-2 spike protein: Targets for vaccination. *Life sciences*, *257*, 118056. Advance online publication. https://doi.org/10.1016/j.lfs.2020.118056

48. Svensson, D., Nebel, D., & Nilsson, B. O. (2016). Vitamin D3 modulates the innate immune response through regulation of the hCAP-18/LL-37 gene expression and cytokine production. *Inflammation research : official journal of the European Histamine Research Society ... [et al.]*, *65*(1), 25–32. https://doi.org/10.1007/s00011-015-0884-z

49. Tang, X., Basavarajappa, D., Haeggström, J. Z., & Wan, M. (2015). P2X7 Receptor Regulates Internalization of Antimicrobial Peptide LL-37 by Human Macrophages That Promotes Intracellular Pathogen Clearance. *Journal of immunology (Baltimore, Md. : 1950)*, *195*(3), 1191–1201. https://doi.org/10.4049/jimmunol.1402845

50. Tian, X., Li, C., Huang, A., Xia, S., Lu, S., Shi, Z., Lu, L., Jiang, S., Yang, Z., Wu, Y., & Ying, T. (2020). Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. *Emerging microbes & infections*, *9*(1), 382–385. https://doi.org/10.1080/22221751.2020.1729069
51. Timmons, P. B., & Hewage, C. M. (2020). HAPPENN is a novel tool for hemolytic activity prediction for therapeutic peptides which employs neural networks. *Scientific reports*, *10*(1), 10869. https://doi.org/10.1038/s41598-020-67701-3
52. Tjabringa, G. S., Rabe, K. F., & Hiemstra, P. S. (2005). The human cathelicidin LL-37: a multifunctional peptide involved in infection and inflammation in the lung. *Pulmonary pharmacology & therapeutics*, *18*(5), 321–327. https://doi.org/10.1016/j.pupt.2005.01.001
53. Tripathi, S., Wang, G., White, M., Qi, L., Taubenberger, J., & Hartshorn, K. L. (2015). Antiviral Activity of the Human Cathelicidin, LL-37, and Derived Peptides on Seasonal and Pandemic Influenza A Viruses. *PloS one*, *10*(4), e0124706. https://doi.org/10.1371/journal.pone.0124706
54. Vabret, N., Britton, G. J., Gruber, C., Hegde, S., Kim, J., Kuksin, M., Levantovsky, R., Malle, L., Moreira, A., Park, M. D., Pia, L., Risson, E., Saffern, M., Salomé, B., Esai Selvan, M., Spindler, M. P., Tan, J., van der Heide, V., Gregory, J. K., Alexandropoulos, K., … Sinai Immunology Review Project (2020). Immunology of COVID-19: Current State of the Science. *Immunity*, *52*(6), 910–941. https://doi.org/10.1016/j.immuni.2020.05.002
55. Venneti, K. C., & Hewage, C. M. (2011). Conformational and molecular interaction studies of glucagon-like peptide-2 with its N-terminal extracellular receptor domain. *FEBS letters*, *585*(2), 346–352. https://doi.org/10.1016/j.febslet.2010.12.011
56. Wang G. (2008). Structures of human host defense cathelicidin LL-37 and its smallest antimicrobial peptide KR-12 in lipid micelles. *The Journal of biological chemistry*, *283*(47), 32637–32643. https://doi.org/10.1074/jbc.M805533200
57. Wang, Q., Zhang, Y., Wu, L., Niu, S., Song, C., Zhang, Z., Lu, G., Qiao, C., Hu, Y., Yuen, K. Y., Wang, Q., Zhou, H., Yan, J., & Qi, J. (2020). Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell*, *181*(4), 894–904.e9. https://doi.org/10.1016/j.cell.2020.03.045
58. Wei, L., Tang, J., & Zou, Q. (2017). SkipCPP-Pred: an improved and promising sequence-based predictor for predicting cell-penetrating peptides. *BMC genomics*, *18*(Suppl 7), 742. https://doi.org/10.1186/s12864-017-4128-1

59. Xhindoli, D., Pacor, S., Benincasa, M., Scocchi, M., Gennaro, R., & Tossi, A. (2016). The human cathelicidin LL-37--A pore-forming antibacterial peptide and host-cell modulator. *Biochimica et biophysica acta*, *1858*(3), 546–566. https://doi.org/10.1016/j.bbamem.2015.11.003
60. Xu, J., & Zhang, Y. (2010). How significant is a protein structure similarity with TM-score = 0.5?. *Bioinformatics (Oxford, England)*, *26*(7), 889–895. https://doi.org/10.1093/bioinformatics/btq066
61. Yan, Y., Tao, H., He, J., & Huang, S. Y. (2020). The HDOCK server for integrated protein-protein docking. *Nature protocols*, *15*(5), 1829–1852. https://doi.org/10.1038/s41596-020-0312-x
62. Zhang, X., Oglęcka, K., Sandgren, S., Belting, M., Esbjörner, E. K., Nordén, B., & Gräslund, A. (2010). Dual functions of the human antimicrobial peptide LL-37-target membrane perturbation and host cell cargo delivery. *Biochimica et biophysica acta*, *1798*(12), 2201–2208. https://doi.org/10.1016/j.bbamem.2009.12.011
63. Zhang, Y., & Skolnick, J. (2005). TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic acids research*, *33*(7), 2302–2309. https://doi.org/10.1093/nar/gki524
64. Zhao, L., Cao, Z., Bian, Y., Hu, G., Wang, J., & Zhou, Y. (2018). Molecular Dynamics Simulations of Human Antimicrobial Peptide LL-37 in Model POPC and POPG Lipid Bilayers. *International journal of molecular sciences*, *19*(4), 1186. https://doi.org/10.3390/ijms19041186

AUTHOR CONTRIBUTIONS

Manisha Deshpande conceived, designed and led the project and also carried out the Structural Alignment study, the Safety studies and CPP prediction. She has also written most of the Introduction and Discussion of the manuscript. Kiran Lokhande performed all the Molecular Docking studies and analyses, analysed the Structural Alignment results and has written all parts about the docking and alignment in the manuscript. Tanushree Banerjee gave the advice to carry out structural studies, performed the mutation analysis, and contributed to the writing. K.V. Swamy mentored Kiran Lokhande, did critical reading of the manuscript and gave valuable suggestions. All authors contributed by valuable discussions and/or suggestions. All authors have read and approved the final manuscript.

FIGURES AND TABLES

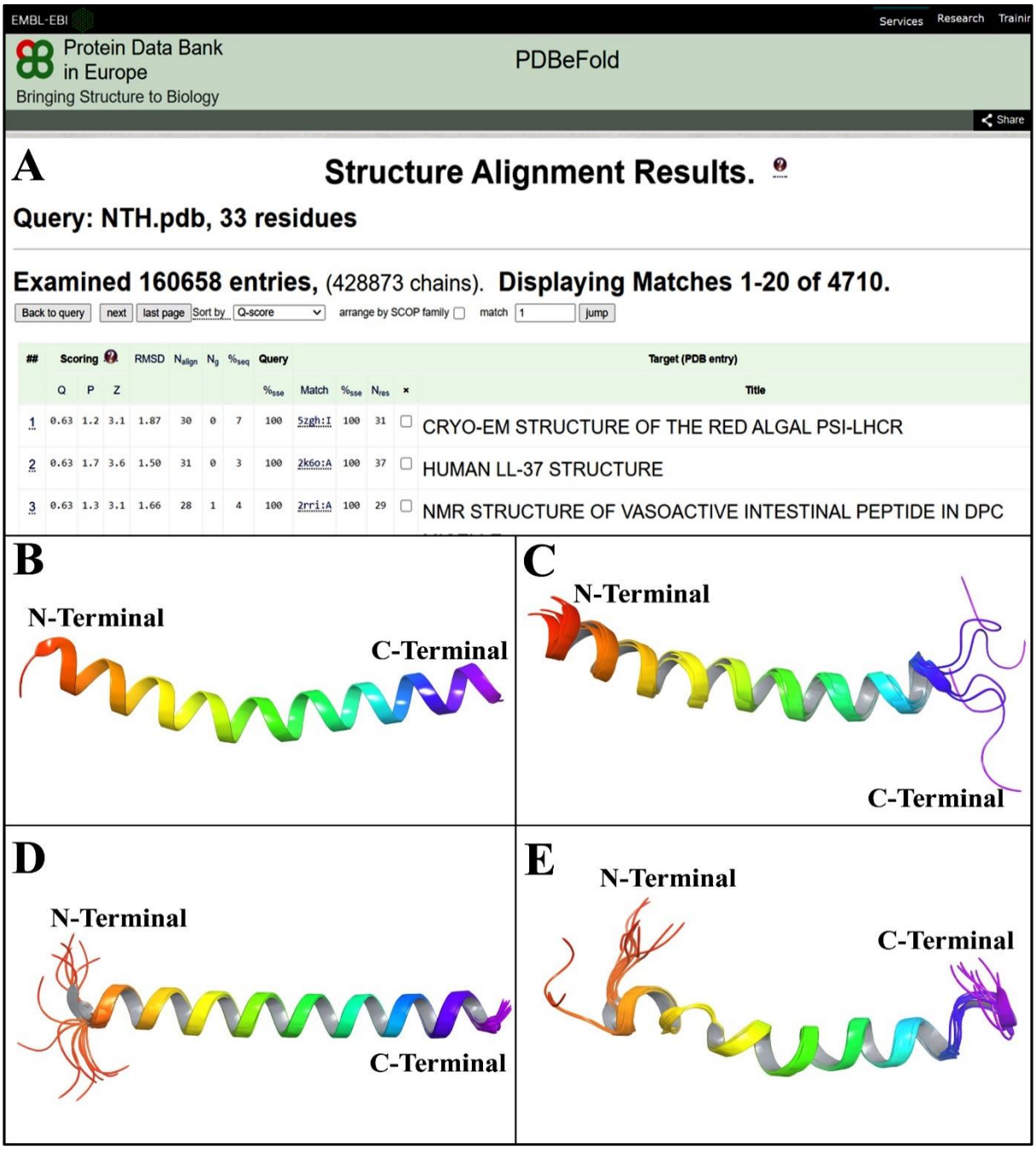

Figure 1. (A) The results of the search for structural similarity to the NTH of ACE2 in PDBeFold search engine. LL-37 is seen as the 2nd result. **(B)** The secondary structure of NTH was obtained by submitting its sequence to the structure prediction server Phyre2 (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index; Kelley *et al*, 2015) **(C)** The structure of the LL-37 peptide (2K6O) from https://www.rcsb.org/. **(D)** The structure of the TDS-2 peptide (6ITH) from https://www.rcsb.org/ **(E)** The structure of the GLP-2 peptide (2L63) from https://www.rcsb.org/.

Table 1. Quantitative assessment of hACE2 NTH Structural Similarity with LL-37, TDS-2, and GLP-2 using TM-align server. The reference protein was NTH of ACE2.

Structural Proteins	**hACE2 NTH**	**LL-37**	**TDS-2**	**GLP-2**
Sequence Length	35	37	35	33
Aligned length	35	32	32	23
RMSD (Å)	0.00	1.5	1.64	1.8
TM-Score	1.00000	0.68037	0.55154	0.44354

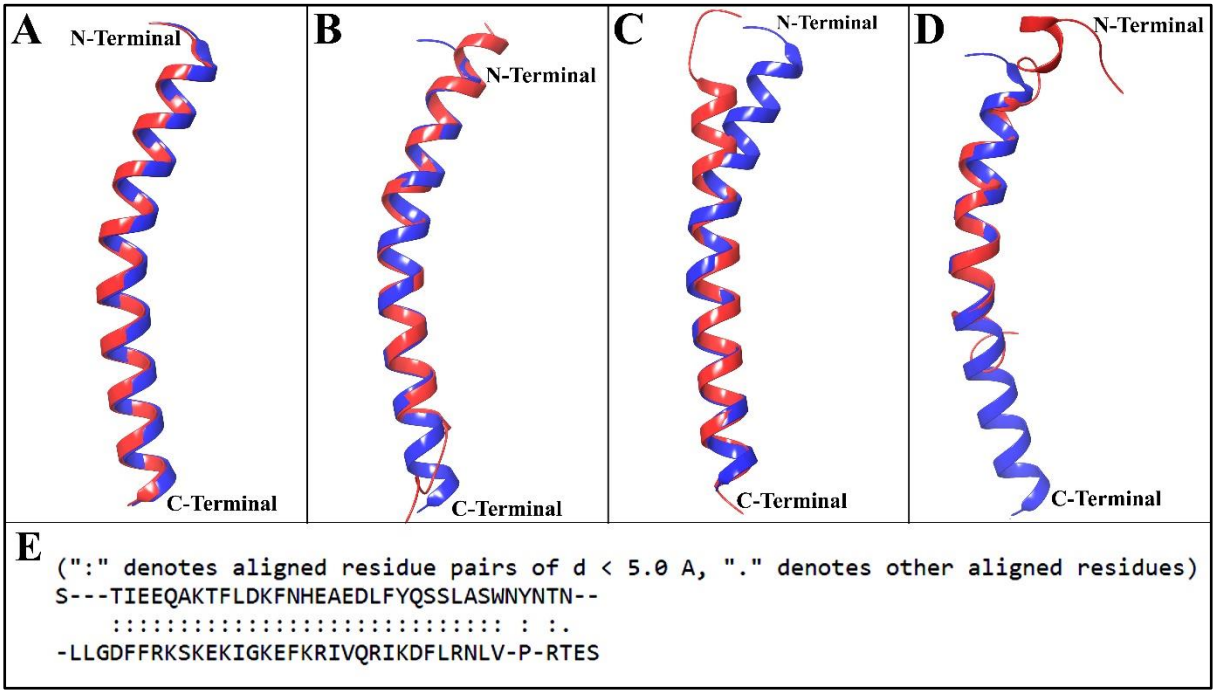

Figure 2. Structural alignment of hACE2 NTH with **(A)** hACE2 NTH, **(B)** LL-37, **(C)** TDS-2, and **(D)** GLP-2. The structure of hACE2 NTH represented in blue color and aligned structures are shown in red color. (E) The amino acids participating in the structural alignment of LL-37 with NTH. The top sequence is the NTH, and the bottom sequence is of LL-37.

Table 2. Detailed intermolecular interaction analysis of the hACE2-SARS-CoV-2 RBD complex.

Interacting Residues		**Bond Type**	**Bond Distance (Å)**
hACE2	**SARS-CoV-2 RBD**		
Met82	Phe486	Ar-HB	3.23
Tyr83	Asn487	Ar-HB	2.76
Tyr83	Asn487	HB	1.78
Phe28	Tyr489	Ar-HB	3.16
Thr27	Phe456	Ar-HB	3.01
Asp30	Phe456	Ar-HB	3.47
Asp30	Lys417	HB	1.94
Asp30	Lys417	Salt Bridge	2.87
Lys31	Glu484	Salt Bridge	3.50
Glu35	Gln493	HB	2.31
Asp38	Tyr449	Ar-HB	2.45
Gln42	Gly446	HB	2.74

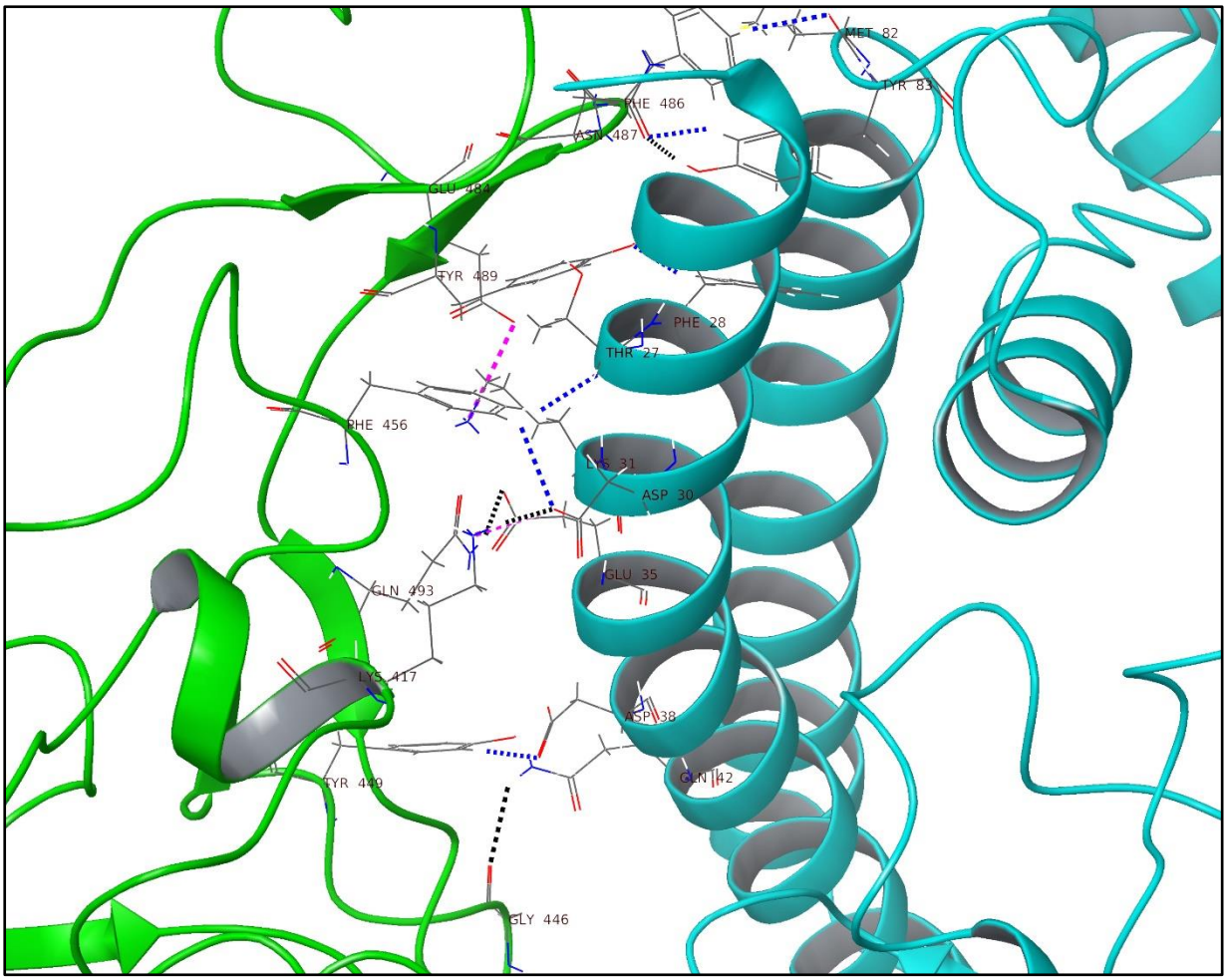

Figure 3. Intermolecular interactions and binding mode of hACE2 (cyan color) with SARS-CoV-2 RBD (green color). Hydrogen bonds, aromatic hydrogen bonds, and salt bridges are represented by dashed black, blue and pink colored lines respectively.

Table 3. Detailed intermolecular interaction analysis of the LL-37-SARS-CoV-2 RBD complex.

Interacting Residues		**Bond Type**	**Bond Distance (Å)**
LL-37 (PDB ID: 2K6O)	**SARS-CoV-2 RBD (PDB ID: 6LZG)**		
Asn30	Asn501	HB	1.67
Phe27	Gly496	Ar-HB	2.75
Asp26	Tyr505	Ar-HB	2.73
Arg23	Tyr505	HB	2.48
Arg23	Glu406	Salt Bridge	4.12
Glu16	Lys417	Salt Bridge	3.18
Glu16	Lys417	HB	2.53
Glu16	Tyr421	Ar-HB	1.75
LL-37 (PDB ID: 2K6O)	**SARS-CoV-2 RBD (PDB ID: 6M0J)**		
Asp26	Tyr505	Ar-HB	2.96
Phe27	Gly496	Ar-HB	2.82
Arg23	Tyr505	HB	2.74
Arg23	Glu406	Salt Bridge	3.25
Glu16	Phe456	Ar-HB	3.38
Glu16	Lys417	Salt Bridge	3.09
Glu16	Tyr421	Ar-HB	2.25
Glu16	Tyr421	HB	1.94

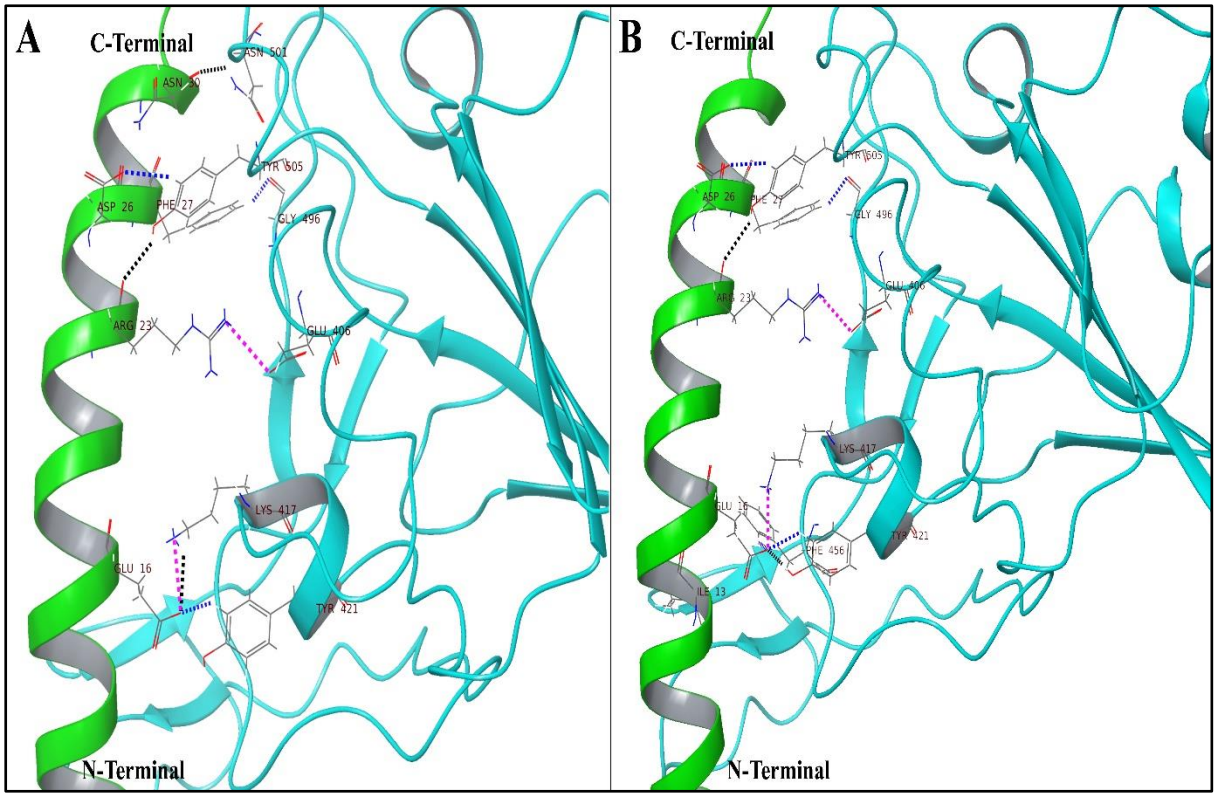

Figure 4. Intermolecular interactions and binding mode of LL-37 (green color) with **(A)** SARS-CoV-2 RBD from 6LZG PDB (cyan color**) (B)** SARS-CoV-2 RBD from 6M0J PDB (cyan color). Hydrogen bonds, aromatic hydrogen bonds and salt bridges are represented by dashed black, blue and pink colored lines respectively.

Figure 5. Legends on next page.

Figure 5. Diagrammatic representation of molecular docking of the three peptides to RBD.

(A) LL-37 bound to RBD OF 6MOJ (Test1); Sequence: RBD sequence (319 to 541 of UniProt ID PODTC2 common to 6MOJ and 6LZG); Red sequence: RBM sequence; Large font bold residues: Contacted by LL-37; Black/red residues: Common with LL-37_RBD of 6LZG docking; Green residue : Not common with LL-37_RBD of 6LZG docking; Blue bar: Contact region covered by LL-37; Decimals below residues: Mutation frequency.

(B) LL-37 bound to RBD OF 6LZG (Test2); Sequence: RBD sequence (319 to 541 of UniProt ID PODTC2 common to 6MOJ and 6LZG); Red sequence: RBM sequence; Large font bold residues: Contacted by LL-37; Black/red residues: Common with LL-37_RBD of 6MOJ docking; Green residue : Not common with LL-37_RBD of 6MOJ docking; Blue bar: Contact region covered by LL-37; Decimals below residues: Mutation frequency.

(C) TDS-2 bound to RBD OF 6LZG (Negative1); Sequence: RBD sequence (319 to 541 of UniProt ID PODTC2 common to 6MOJ and 6LZG); Red sequence: RBM sequence; Large font bold residues: Contacted by TDS-2; Blue bar: Contact region covered by TDS-2.

(D) GLP-2 bound to RBD OF 6LZG (Negative2); Sequence: RBD sequence (319 to 541 of UniProt ID PODTC2 common to 6MOJ and 6LZG); Red sequence: RBM sequence; Large font bold residues: Contacted by TDS-2; Blue bar: Contact region covered by GLP-2.

Table 4. Detailed intermolecular interaction analysis of the TDS-2 and GLP-2 with SARS-CoV-2 RBD, and LL-37-RBD (region adjacent to RBD) complex.

Interacting Residues		**Bond Type**	**Bond Distance (Å)**
TDS-2	**SARS-CoV-2 RBD**		
Val7	Tyr473	HB	2.47
Ala10	Phe456	2 Ar-HB	3.14, 3.26
Val11	Phe456	Ar-HB	3.35
Phe24	Tyr505	Pi-Pi	3.85
Arg32	Thr500	HB	2.61
GLP-2	**SARS-CoV-2 RBD (PDB ID: 6LZG)**		
Phe22	Asn450	Ar-HB	2.85
His1	Asn343	HB	2.23
His1	Glu340	Ar-HB	3.28
LL-37	**SARS-CoV-2 Region Adjacent to RBD (PDB ID: 6LZG)**		
Phe17	Phe238	Pi-Pi	4.80
Arg7	Asp138	Salt Bridge	3.13

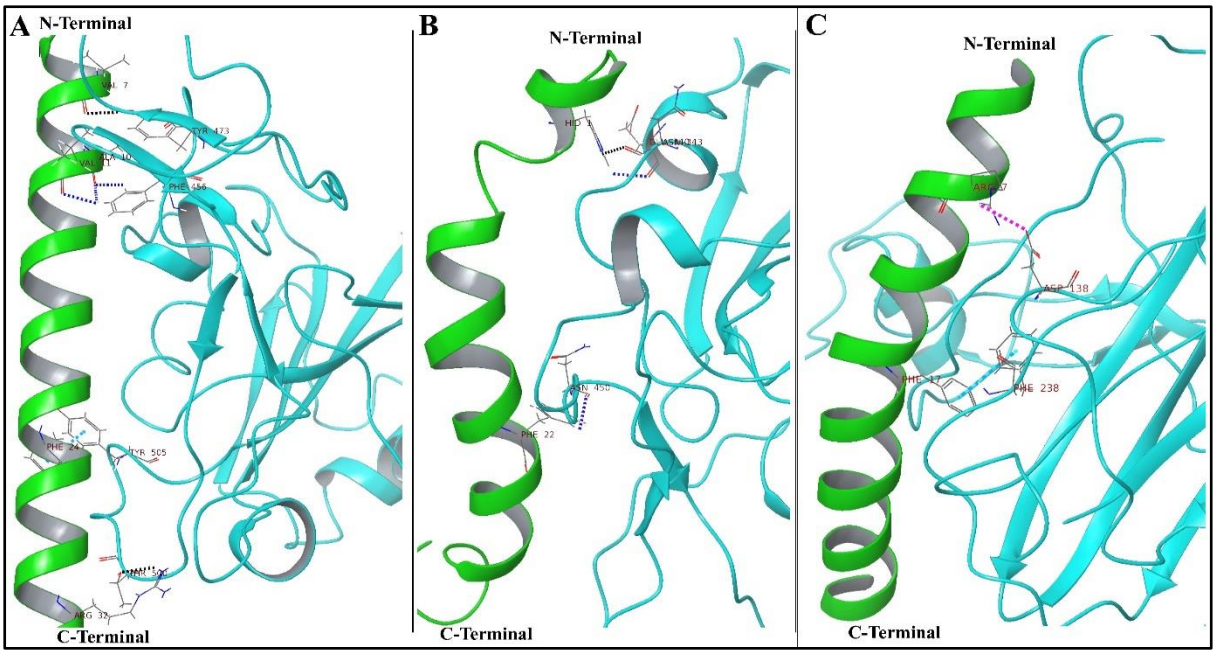

Figure 6. Intermolecular interactions and binding mode of **(A)** TDS-2 with SARS-CoV-2 RBD, **(B)** GLP-2 with SARS-CoV-2 RBD, and **(C)** LL-37 with SARS-CoV-2 region (96-318) which is adjacent to RBD. Hydrogen bonds, aromatic hydrogen bonds, Pi-Pi stacking and salt bridges are represented by dashed black, blue, sky blue and pink colored lines respectively.