Virtual screening of knottin and defensin peptides perceives hits against the SARS CoV-2 RBD domain and hACE2 interaction

Nitesh Mani Tripathi, and Anupam Bandyopadhyay*

Biomimetic Peptide Engineering Laboratory, Department of Chemistry, Indian Institute of Technology, Ropar, Punjab-140001, India.

*Correspondence: AB, anupamba@iitrpr.ac.in

Abstract

Targeting protein-protein interactions (PPIs) between the receptor binding domain (RBD) of SARS-CoV-2 and human angiotensin converting enzyme (ACE2) has been an attractive therapeutic target for peptide or protein drug discovery to inhibit the entry of SARS-CoV-2 into the host cells. Developing inhibitors for such extensive (RBD and human ACE2 interaction) contact surfaces involves multiple challenges. We sought to start *in silico* investigation with well-known knottins peptides from its databank for the first time and host defense peptides, defensins, due to their multifaceted antimicrobial activity. The molecular-level examination resulted in a handful of peptides binding selectively with the spike glycoprotein at low nanomolar potency. They exhibited a high thermodynamic binding stability in an MD simulation study. Noteworthy, Kalata B1 and human β-defensin 4 (HBD4) showed excellent interaction parameters calculated through an alanine scan of hot spot residues. These natural source peptide inhibitors could be a promising lead for developing SARS-CoV-2 prophylactic after an ongoing experimental assay test.

Keywords: COVID19, SARS-CoV-2, Human Angiotensin-Converting Enzyme 2 (hACE2), Cyclotides, peptide inhibitors, protein-protein interactions, Molecular Mechanics Poisson-Boltzmann Surface Area (MMPBSA).

Introduction

The emergence of a novel severe acute respiratory syndrome (SARS)-like coronavirus (SARS-CoV-2) caused ~ 6.7 million deaths around the globe due to the coronavirus disease 2019 (COVID-19) pandemic[1]. SARS CoV-2 is a β-coronavirus that consists of four structural proteins, including spike (S), envelope (E), membrane (M), and nucleocapsid (N). Among them, spike proteins on the virus envelope play the most critical roles in viral attachment, fusion, and entry to the host cell through the receptor angiotensin-converting enzyme 2 (ACE2)[2]. The spike protein's receptor-binding domain (RBD) selectively recognizes ACE2 as its receptor with an affinity in the low nanomolar range[3-4]. Therefore, immense interest in inhibiting RBD-ACE2, virus-host protein-protein interaction (PPI), has been a central focus in controlling SARS-CoV-2 entry during the outbreak (Figure 1A). Since inhibiting a large PPI interface is tedious with conventional small molecules, an immense interest has inclined in engineering peptide- and protein-based molecules to inhibit RBD-ACE2 interaction. There are numerous reports of SARS CoV-2 spike's RBD domain peptide inhibitors, including human native ACE2 α-helix and its decoys[5-6] , de-novo designed high affinity binders[7-8], D-peptide mimicking ACE2 helix[9], helical bundle[10,11], computationally modeled peptides[12,13,], *in silico* screening of peptides possessing antimicrobial properties such as antiviral[14,15], cyclic[16,17,] & stapled peptides[18,19] as well as covalently engineered minibinders[20]. Also, there are multiple earlier reports of *in silico* investigations and *in vitro* studies of human defensins[21–23], and an engineered defensin α-helix[24] provided a straightforward way to identify potential peptide-based therapeutics.

At the beginning of the COVID-19 outbreak, we hypothesized that disulphide-rich antimicrobial peptides might play a significant role in viral inhibition. Cyclic

knottin peptides (cyclotides) struck our mind and could show high binding affinity and selectivity to spike protein. Cyclotides are head-to-tail cyclized peptides, and their backbone are stabilized by disulphide cystine knots, which provide a unique structural scaffold and are exceptionally resistant to enzymes[25]. These macrocyclic knottin peptides were discovered in major families of *Rubiaceae, Violaceae* and *Cucurbitaceae,* which comprise ~30 amino acids and are considered the most prominent family of circular mini-proteins. We also piqued interest in human host defense disulphide-rich peptides and defensins to design RBD-ACE2 inhibitors. Because α- and β-defensins, two subfamilies, have been long recognized as natural antimicrobial peptides and are primarily involved in the host's innate immunity[26,27].

We started our investigation by considering deposited cyclotide and defensin PDB structures with a virtual screening of cyclotide databases and defensins. Its workflow and steps are shown in Figure 1B-C. Interestingly, a few knottin from the Kalata family (B1, B2 and B7) and β-defensin 1 and 4 exhibited potential binding affinity with the RBD domain of SARS CoV-2. In parallel, a few knottin and defensin enriched as selective human ACE2 (hACE2) binders. MD simulation has manifested thermodynamic parameters of RBD-peptide and hACE2-peptide complexes along with molecular interactions. In MD simulation, potential binder confidence was further expanded through an alanine scan of hot spot residues in peptide hits.

Results & Discussion

Virtual screening and MD simulation study of cyclotides and defensins

In order to find the potential peptide binders against RBD-hACE2 interaction, cyclotides and human defensins (from the online database, see material and methods ESI) were employed in HADDOCK 2.4 for docking studies, distinctly with RBD and hACE2 protein target of interest. Subsequently, the HADDOCK webserver generates 200 complexes thus larger clusters contains refined models within a cluster and can differ more in other sets of parameters such RMSD from the overall lowest structure (Angstrom, Å), HADDOCK score, Z-score, and various energy terms. The best cluster shows a more negative Z-score, which determines the number of standard deviations from the mean cluster for a specific model cluster. The most reliable cluster was chosen based on the lowest HADDOCK score and Z-score (Table S1-S4).

Additionally, we used a systematic approach towards peptide discovery based on MD simulations (MDS) study. The system stability was evaluated using the RMSD of the Cα- atoms with respect to the structures obtained from the 200 ns MDS trajectories. Based on the structural investigations and chosen parameters to shortlist the best peptide binders among the complexes, Kalata B1, B2, B7, β-defensin 1 (β-HD1) and β-defensin 4 (β-HD4) peptide found to be interacted the binding pocket of RBD domain. Therefore, we considered this could be potential peptide candidates blocking the RBD-hACE2 interaction (Table 1 and S1-S4). The binding affinity (K_d) and binding free energy (ΔG) were also calculated before MDS and after molecular docking in order to estimate the best peptide binder, to further perform the post MD simulation studies. Among five peptide candidates, we obtained β-defensin 4 (β-HD4) as "best peptide binder" in the terms of binding affinity (K_d) and binding free energy (ΔG) predicted by PRODIGY webserver[28,29].

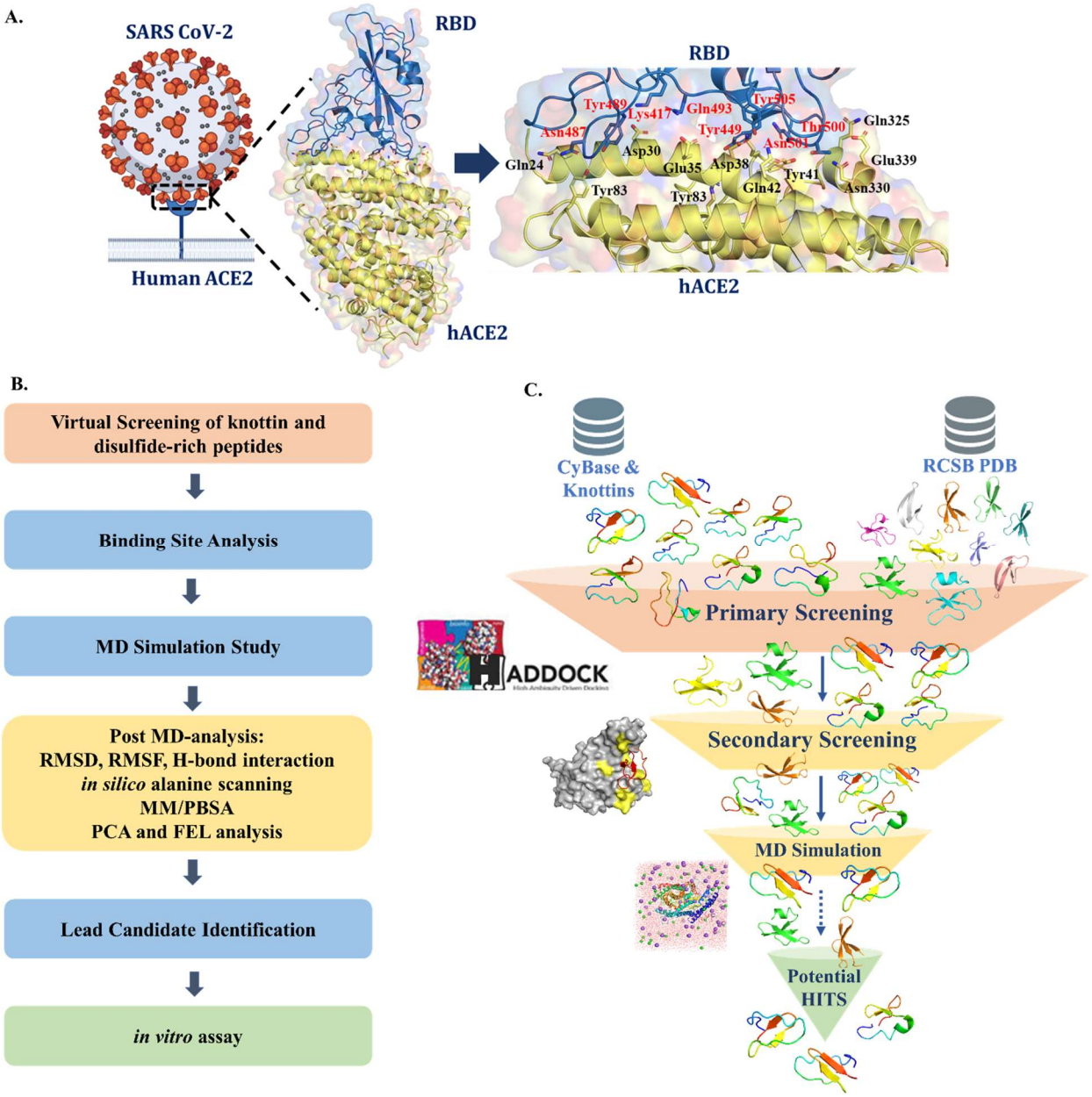


Figure 1: A) Molecular interaction of SARS CoV-2's Spike protein and human ACE2 receptor; Cartoon representation of SARS CoV-2 RBD domain (Pale blue) and human ACE2 (hACE2) receptor (Yellow); Key residues at the interface between the SARS-CoV-2 RBD domain and ACE2 receptor. Schematic representation of virtual screening of knottins and disulphide-rich peptides; B) Step-by-step flow chart demonstrating an *in silico* study performed here; C) Schematic workflow of peptide screening.

MD Simulation analysis

Root-mean-square deviation analysis (RMSD):

An investigation based on structural characteristics of the potential knottins and disulfide-rich peptides were subjected to 200 ns MDS to adequately sample the stable complex conformational space. The root-mean-square deviation (RMSD) of the atomic position, relative to the initial structure was evaluated by least-squares fitting for the C*α*-atoms to investigate the overall conformational stability of the protein−ligand complex in the MD simulation. Generally, when the system obtains a stable RMSD curve, which means that it reaches an equilibrated state; the other side, the sharp fluctuations of RMSD curve of the Cα atoms indicate low structural stability. Figure 1 and Figure S2-S5 depicts the time-dependent RMSD curves of the reference (ACE/RBD and ACE2/Cathelicidin) and potential knottins and disulfide-rich peptide inhibitors.

RBD-bound peptide complexes depict the smoothed fluctuation curves of the average RMSD in a stable state throughout the simulation process with an average RMSD value of < 0.4 nm. For the control ACE2/RBD complex, the RMSD value varies in between 168 to 187 ns from 0.4 to 1.5 nm but later falls to 1 nm after 185 ns (Figure S1). Whereas, the other ACE2/Cathelicidin showed slight fluctuation curves of the average RMSD value < 0.5 nm (Figure S2).

Table 1: HADDOCK Molecular Docking results of Cyclotides and defensins against RBD domain and hACE2 receptor.

Complex	Interface Binding	HADDOCK (Before Simulation)		Interface Binding	After MD Simulation (Upto 200ns)	
ACE2-Cyclotides complex		**ΔG (kcal/mol)**	**K_d (nM)**		**ΔG (kcal/mol)**	**K_d (nM)**
ACE2_RBD	+	-12.7	0.49	+	-8.6	480
ACE2_Cathelicidin	+	-10.1	36	+	-8.8	340
ACE2_1nb1	+	-10.2	34	-	NB	NB
ACE2_1pt4	+	-10.4	23	-	NB	NB
ACE2_1znu	+	-10.1	36	+	-8.9	270
ACE2_2lur	+	-10.6	17	-	NB	NB
ACE2_1kal	+	-8.8	340	-	NB	NB
ACE2_1vb8	+	-10.9	10	-	NB	NB
ACE2_2kuk	+	-10.5	20	-	NB	NB
ACE2_2f2i	+	-10.9	9.6	-	NB	NB
RBD-Cyclotides complex		**ΔG (Kcal/mol)**	**K_d (nM)**		**ΔG (Kcal/mol)**	**K_d (nM)**
RBD_ Cathelicidin	+	-10.3	13	-	NB	NB
RBD_1nb1	+	-10.8	11	-	NB	NB
RBD_1znu	+	-11.2	5.6	-	NB	NB
RBD_1pt4	+	-10.9	11	+	-10.1	39
RBD_2lur	+	-12.2	1.1	-	NB	NB
RBD_2m9o	+	-10.5	20	+	-9.8	70
RBD_2lam	+	-8.7	440	+	-7.4	4000
RBD_1df6	+	-7.3	4600	+	-8.5	550
RBD_1k48	+	-10.9	11	+	-10.1	42
ACE2-Defensins complex		**ΔG (Kcal/mol)**	**K_d (nM)**		**ΔG (Kcal/mol)**	**K_d (nM)**
ACE2_α-HD1	+	-10.2	34	-	NB	NB
ACE2_ α-HD5	+	-10.4	23	+	-8.5	630
ACE2_β-HD2	+	-10.5	19	-	NB	NB
ACE2_ β-HD3	+	-11.7	28	-	NB	NB
ACE2_ β-HD4	+	-12.2	11	-	NB	NB
RBD-Defensins complex		**ΔG (Kcal/mol)**	**K_d (nM)**		**ΔG (Kcal/mol)**	**K_d (nM)**
RBD_ α-HD1	+	-10.8	11	+	-9.4	120
RBD_ α-HD2	+	-11.2	5.6	+	-9.9	57
RBD_ α-HD3	+	-10.9	11	-	NB	NB
RBD_ α-HD4	+	-12.2	1.1	-	NB	NB
RBD_ β-HD1	+	-10.9	11	+	-10.4	28
RBD_ β-HD4	**+**	-11.3	5.3	**+**	**-10.9**	**11**

(*NB = No Binding was observed at the active site of human ACE2 & RBD domain of SARS CoV-2 spike protein)

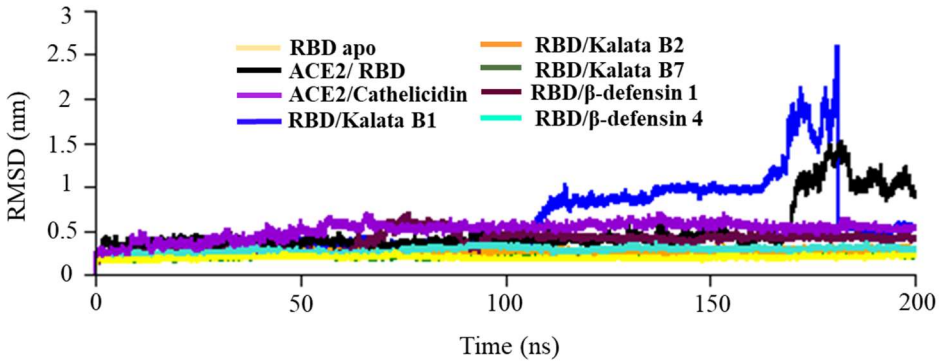


Figure 2: RMSD plot showing the average value of the RBD/peptide complexes. RMSD curve depicts the system stability, which means that it has reached an equilibrated state; the other side, the sharp fluctuations of RMSD curve indicate low structural stability.

Further for ACE2- and RBD-bound peptide complexes, we observed RBD/Kalata B1 show slight fluctuation in the RMSD value during the simulation process after 110 to 180 ns of 1 to 2.5 nm but later RMSD falls within range 0.3-0.5 nm which occurs due to slight shifting of the peptide from its original binding site of RBD domain that caused increase in RMSD value from 0.5 to 2.3 nm (Figure 2). Similarly, in the case RBD/β-defensin-1 peptide inhibitor was observed with variation in RMSD value of 0.25 to 0.5 nm for short period of time (i.e., in-between 60 to 90 ns). ACE2/KALATA B1 shows the average RMSD value <0.4 nm, ACE2/α-defensin 5 (α-HD5) shows a slight fluctuation for short period of 0.4-0.5 nm in 145-160 ns trajectory other than that average RMSD value < 0.4 nm (Figure 1, S2 and S3). RBD/α-defensin 1 (α-HD1) shows a fluctuation of 0.2-0.4 from 70-145 ns and at the end of the simulation value rises to 0.3 nm (Figure S4). For RBD/ α-defensin 2 (α-HD2) shows a slight fluctuation in the beginning of simulation 0.2-0.3 nm and after 15 ns it becomes stable and average RMSD value lies within 0.35 nm (Figure S5). Other than control ACE2/RBD and RBD/Kalata B1 (1K48) complex, other peptide inhibitors remain in equilibrium and maintains the RMSD value within <0.4 nm until the end of the MD simulation. In this case, it has been provided that the conformation of the system is in a dynamic equilibrium state under the simulation conditions. Therefore, with this MD simulation trajectories was found to be suitable for further energy analysis.

***Root-mean square fluctuation (RMSF) study**:*

To investigate the structural flexibility in the entire simulation, the root-mean square fluctuation (RMSF) of the main chain Cα atoms, reflecting the fluctuation of the atoms from their average position, was calculated on a stable trajectory. Generally, the higher value of the RMSF implies that the residue has larger flexibility. Similarly, the lower value of the RMSF means the restriction of residue movement, thereby reducing their flexibility. As shown in Figure 3B, for the different RBD-bound peptide complexes the major fluctuations can be observed within a residue 48-61, 71-79 in the helix region and residues 106-115, 123-157, 166-173, and 183-188 in the loop region of RBD show large fluctuations with RMSF values larger than 1.5 nm, indicating that these regions are flexible. In the case of peptides, there is no such fluctuations can be observed during the MDS except the loop region in the knottins and the terminal residue of defensin peptides which are termini atoms shows the perturbation in the RMSF curve as shown in Figure 3Aiv-v.

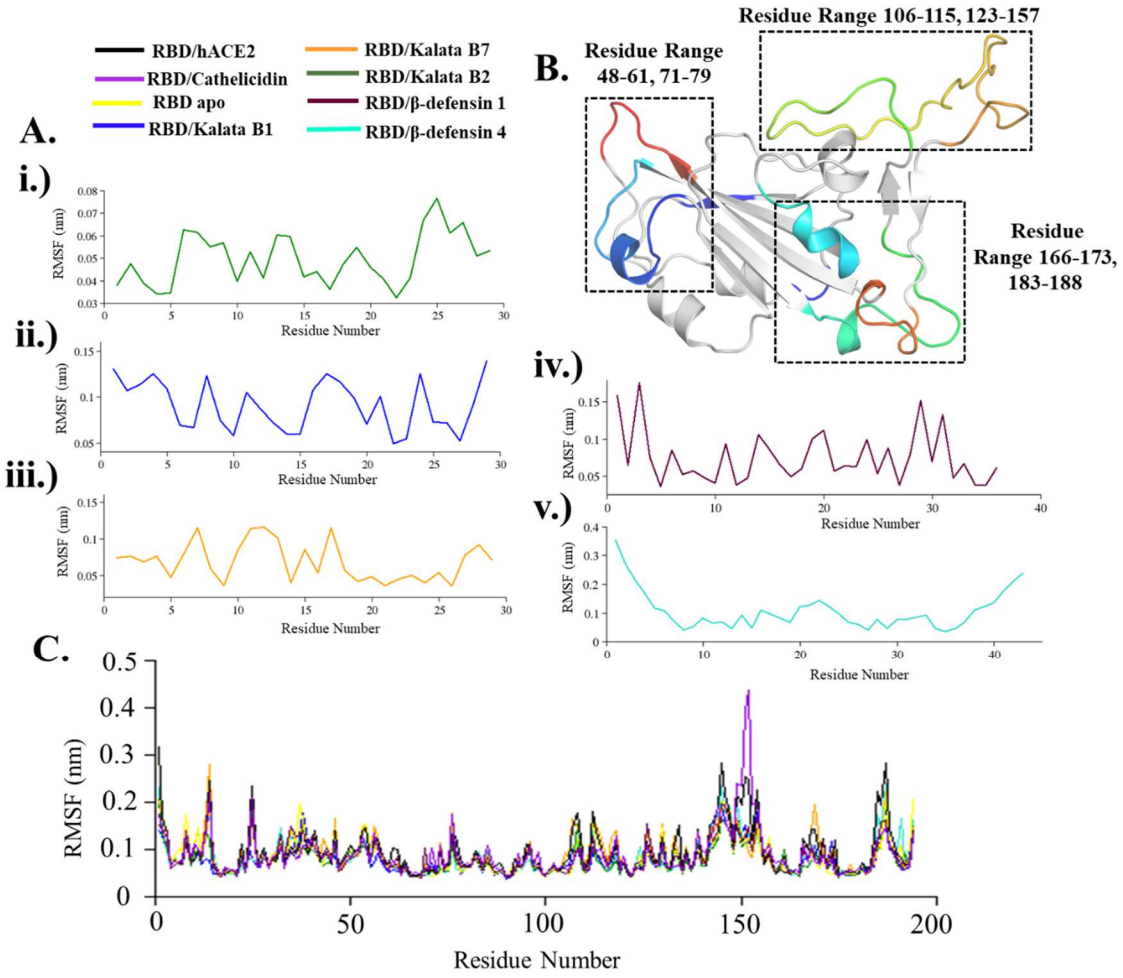


Figure 3: RMSF Plots of the Cα-atoms of the RBD/knottins and disulfide-rich peptide complexes. **A)** Ligand RMSF plot of the RBD-bound peptide complexes; **B)** Cartoon representation of loop region within RBD protein; **C)** Protein RMSF plot of the RBD-bound peptide complexes.

Due to rigid conformation and three disulfide linkage in the knottins and defensins peptide, the Ligand RMSF values of all residues of all inhibitors show small fluctuations, and their values are all <2.0 nm, which is consistent with the RMSD analysis, indicating that the peptide inhibitor are in a stable state and therefore exhibiting a lower RMSF values.

Radius of gyration (Rg) and solvent accessible surface Area (SASA) analysis

To understand the structural stability of the RBD-bound peptide complexes, we further determined the compactness of the protein structure by computing the radius of gyration (Rg). The Rg plots represented in Figure 4A as well as Figure S2-S5 in the Supporting Information show that the structural dynamics of the controls and RBD-bound peptide complexes remain quite stable throughout the 200 ns simulation time. The average Rg value for controls ACE2/RBD, ACE2/Cathelicidin and RBD apo protein was 3.35 ± 0.15 nm, 2.65 ± 0.10 nm, 1.82 ± 0.05 nm respectively. A slight deviation for controls ACE2/RBD can be seen during 163-177 ns with average Rg value of 3.60 ± 0.10 nm and ACE2/Cathelicidin during duration 131-163 ns with average Rg value of 2.65 ± 0.08 nm. The structural integrity of RBD/Kalata B2, B7 and RBD/β-defensin-1 peptide complex was observed to be intact with an average Rg value of 1.95 ± 0.05 nm, 2.00 ± 0.08 nm, 2.10 ± 0.06 nm, respectively. Whereas a slight deviation was observed for RBD/Kalata B1 can be seen during 165-183 ns with rise in Rg value, other than this deviation the average Rg value remains around 1.93 ± 0.08 nm. A steady equilibrium is noted till the end of simulation at 200 ns,

which signifies the stable structural dynamics for RBD/β-defensin-4 peptide complex with an average Rg value 2.05 ± 0.05 nm. The perturbation in between the simulation trajectories for RBD/Kalata B1 and controls (ACE2/RBD and ACE2/Cathelicidin) may imply the spatial adjustment of the peptide ligand in the binding site of RBD domain.

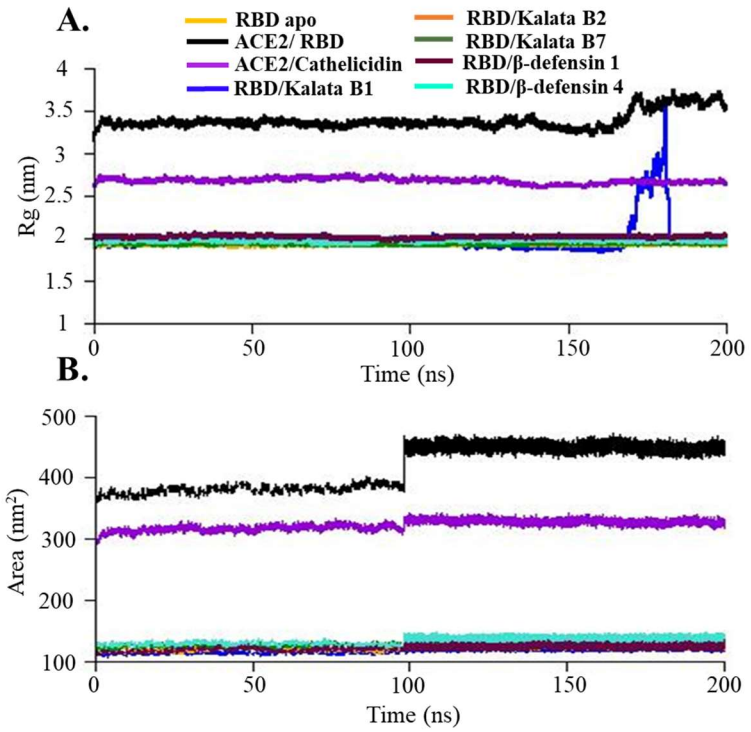


Figure 4: Rg and SASA Plots of the controls (ACE2/RBD, ACE2/Cathelicidin and RBD apo protein) and RBD-bound peptide complexes. **A.)** Radius of gyration (Rg) analysis plot of the controls and RBD-bound peptide complexes; **B.)** SASA analysis plot of the controls and RBD-bound peptide complexes.

Further investigation of another important parameter i.e., SASA (solvent accessible surface area) values was measured to understand the conformational stability of the RBD-bound peptide complexes. The stability, orientation, and processes governing protein-ligand interaction are mostly mediated by the solvent environment surrounding the protein. Figure 4B as well as Figure S2-S5 in the Supporting Information shows investigation of the peptide's solvation impact on the RBD-bound peptide complexes solvation behaviour. The SASA trajectories of the controls (ACE2/RBD and ACE2/Cathelicidin), RBD apo protein and RBD-bound peptide complexes were monitored, a lower average SASA values for the controls (ACE2/RBD and ACE2/Cathelicidin), RBD apo protein and RBD-bound peptide complexes was observed 382 ± 2.0 nm^2, 314 ± 1.5 nm^2, 119 ± 1.0 nm^2 and 120 ± 1.3 nm^2, particularly within the first 98 ns MDS. A peptide may be preferentially confined within the protein pocket, according to its dynamic behaviour. An elevated SASA trajectories might confer the migration of peptides towards the solvent side within the simulation time frames 100-200 ns where the RBD domain pocket became highly solvated and minimally compacted. With the equilibrated and stable conformational dynamics of RBD/β-defensin-4 peptide (136 ± 0.2 nm^2) and the other RBD-bound peptide complexes showed 124 ± 1.0 nm^2. Since the latter is a solvent-substitution process, it suggests stronger interaction peptides at the binding site of RBD domain.

H-bond interaction analysis

Hydrogen bond (H-bond) interaction analysis was performed to determine the RBD-bound peptide interaction types, specificity, and the binding strength of the peptide inhibitors. H-bonding is essential for determining stable contacts that support functional roles and prolong intermolecular contact between the complexes. On an

average five to eight H-bonds formed during the 200 ns simulations for each complex including the controls. These hydrogen bonding interactions were stable and the residues (Gly9, Tyr10, Ile26, Arg28, Pro30 and Thr32) involved in the bonding played an important role in β-defensin-4 peptide binding interface as shown in Figure 5B.

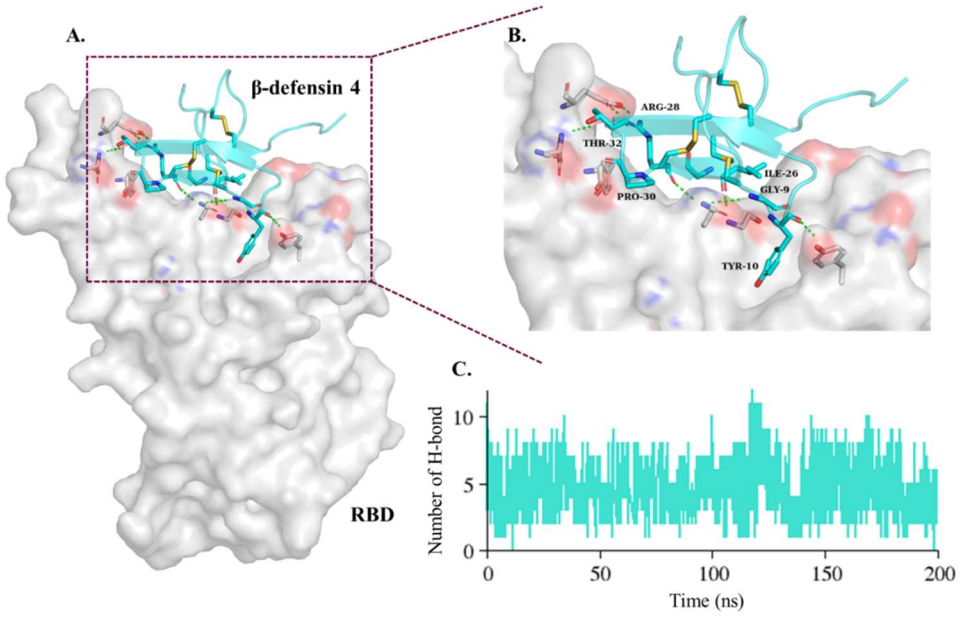


Figure 5: Detailed interactions between the RBD domain and the β-defensin-4 peptide inhibitor. **A)** Molecular interaction of RBD- β-defensin-4 peptide complex; **B)** A zoomed view of molecular interaction of RBD domain and the β-defensin-4 peptide complex; **C)** Average number of H-bonds formed during 200 ns MD simulation trajectory. (RBD domain showed in white surface, Binding site on RBD is highlighted by red (Oxygen atom) and blue (Nitrogen atom) color. β-defensin-4 peptide is shown in turquoise color, key residues blocking the binding site of RBD domain are shown in sticks).

The key interface residues from RBD domain such as Tyr117, Glu152, Asn155, Gln161, and Ser162 forms an important H-bond as well as salt bridges with Gly9, Tyr10, Arg28, Thr32 of β-defensin-4 peptide inhibitor are shown in Table 2, as well as Figure S1 in the Supporting Information. Further, using a Python script *readHBmap.py*[30] and gmx hbond program in the GROMACS 2020.4 analyzed and calculated the percentages of hydrogen bond occupancy between the RBD-bound complexes during the 200 ns simulation trajectory. Based on acceptor-donor atom lengths of less than 3.5 Å and acceptor H-donor angles of more than 120˚, the hydrogen bonds were determined. Table 2 as well as Figure S5-S8 in the Supporting Information shows the percentages of the occupancy of H-bonds between RBD-bound peptide complexes and key interface residues interacting in the all RBD-bound peptide complexes. The key residues noted in the H-bond formation in the β-defensin-4 peptide complex involved Ile26, Tyr10 with more than 50% occupancy. Whereas, hydrogen bond formation with less than 50% occupancy was observed with Arg28, Trp40, Thr32 and Glu42. Whereas several other interface residues assist in H-bonding during 200 ns simulation, by repositioning the flexible sidechains with the surrounding interface residues with the 1-15% occupancy range. Investigating the differential dynamic's behaviour has been depicted for each investigated peptides across the designated MD simulation trajectories and the time evolution of the binding interactions between peptides bound to RBD domain provided the necessary frameworks for understanding the stability of RBD-bound peptide complexes.

Table 2: H-bond Occupancy (%) of RBD/ β-defensin-4 during 200 ns MD simulation trajectory.

Donor	Acceptor	Occupancy (%)
162SER(HN)	26 ILE(O)	96.4
173TYR(HH)	10TYR(O)	60.1
28ARG(HE)	152GLU(OE2)	46.5
28ARG(H21)	152GLU(OE1)	44.5
28ARG(HE)	152GLU(OE1)	29.4
169ASN(D21)	40TRP(O)	26.8
155ASN(HN)	32THR(OG1)	25.1
28ARG(H21)	152GLU(OE2)	21.6
170GLY(HN)	42GLU(OE2)	15.3
170GLY(HN)	42GLU(OE1)	13.4
168THR(HG1)	40TRP(NE1)	8.2
157TYR(HH)	32THR(N)	7.7
169ASN(D21)	42GLU(OE1)	6.5
169ASN(D21)	42GLU(OE2)	6.2
10TYR(HH)	74GLU(OE2)	5.8
71ARG(H11)	10TYR(O)	5.8
173TYR(HH)	42GLU(OE2)	5
166GLN(E21)	40TRP(O)	4.9
157TYR(HH)	30PRO(O)	4.6
157TYR(HH)	32THR(OG1)	4.4
10TYR(HH)	74GLU(OE1)	3.6
173TYR(HH)	42GLU(OE1)	3.2
28ARG(H21)	152GLU(OE1)	1.7
28ARG(H11)	152GLU(OE2)	1.4
166GLN(E21)	40 TRP(NE1)	1.3

In silico alanine mutation

We implemented an *in silico* alanine scanning to discover the "HOTSPOT" residues that are crucial for interface binding and to analyze the contributions of each individual residue to the binding interface. We mainly focused on enhancing the binding affinity of these disulfide-rich peptides by the substitution with alanine residue. The alanine substitution replaces the side-chain without significantly changing the shape of the molecule to alter the mode of interaction. Table 3 shows the key interface binding residues (Cys8, Gly9, Tyr10, Thr12, Tyr24, Ile26, Gly27, Arg28, Cys29, Pro30, Asn31, Thr32, Tyr33, Cys35

and Leu37) of β-HD4 interacting with RBD domain. Likewise, the key interface residues for the Kalata B1 (Thr16, Val18, Thr21, Thr24, Thr28, Cys29, Ser30 and Trp31), Kalata B2 (Glu3, Thr4, Phe6, Ile21, Arg24 and Leu27), Kalata B7 (Gly1, Thr20, Ser22, Trp23, Lys27, Arg28 and Asn29) and β-HD1 includes Val6, Gln11, Leu13, Trp14, Thr21 and Lys36 has been tabulated in Table S9-S12, respectively.

Table 3: *in silico* alanine scanning of β-defensin 4 binding site for HOTSPOT residue prediction.

Wild-type	mutant	KFC2-A Class	KFC2-A Conf	mCSM-PPI2 prediction (affinity)	DrugScore-PPI
CYS8	ALA	-	-1.41	-	0.42
GLY9	ALA	-	-0.8	-	-
TYR10	ALA	HOTSPOT	0.66	-1.127 (decreasing)	2.92
THR12	ALA	-	-1.89	-0.115 (decreasing)	0.11
TYR24	ALA	-	-2.13	-0.83 (decreasing)	1.6
ILE26	ALA	HOTSPOT	0.34	-1.357 (decreasing)	0.8
GLY27	ALA	-	-1.13	-	-
ARG28	ALA	HOTSPOT	0.39	-1.409 (decreasing)	1.37
CYS29	ALA	-	-0.29	-	0.39
PRO30	ALA	-	-1.23	-	-
ASN31	ALA	-	-2.26	-	0.34
THR32	ALA	-	-0.12	-0.719 (decreasing)	0.69
TYR33	ALA	-	-2.37	-0.608 (decreasing)	0.14
CYS35	ALA	HOTSPOT	0.53	-	0.34
LEU37	ALA	-	-1.08	-0.688 (decreasing)	0.39

('-' means webserver could not predict the value for that residue)

KFC2 and DrugScore-PPI webserver predicted TYR10, ILE26, ARG28 and CYS35 residues for β-HD4 peptide complex as "HOTSPOT" residues with confidence value (KFC2-A Conf > 0) within the peptide inhibitor. mCSM-PPI2 webserver predict the effects of alanine substitution on protein-protein interaction relative binding either by increasing or decreasing the affinity, and the value lies within the range $-1.5 \leq \Delta\Delta G \leq 2.0$ kcal/mol for β-defensin 4 peptide inhibitor. In the case of HOTSPOT residues predicted the $\Delta\Delta G < -1.0$ kcal/mol which means interaction is lost due to mutation of that residue with alanine. From an *in silico* alanine scanning, it was established that these key residues are those that significantly lower the binding free energy by at least 2.0 kcal/mol. The strong binding residues should atleast lower the binding free energy to 4 kcal/mol or less[31,32].

MM-PBSA analysis

Binding Free-Energy Calculation

Further, in order to determine the binding free energy of ligand molecules to biological macromolecules, the Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) method is frequently employed[33]. In this study, gmx_MMPBSA tool was implemented to calculate the total binding free energy of RBD-bound peptide complexes by extracting the MD trajectories from the last 10 ns of the simulation systems. Primarily, the lower binding energy is the more stable complex, also a negative value signifies a favorable interaction, whereas a positive value signifies an unfavorable interaction. The gmx_MMPBSA calculation results show the binding free energies of Kalata B1, B2 and B7 and β-defensin 1, 4 are, −5.55, −40.44, −31.00, -29.28 and -34.87 kJ/mol, respectively. The results of predicted binding free energy and each energy term are shown in Table 4, and Table S13. Through MM-PBSA method, binding free energy analysis revealed that electrostatic interactions and van der Waals interactions provided the substantial driving force for the binding process along with the calculation the absolute binding free energies and per-residue decomposition analysis.

Table 4: Binding free energy (BFEs) calculation of RBD-bound peptide complexes using MM/PBSA method. All the energy components are reported in kcal/mol.

Peptide Complex	Energy Components				
	Van der Waals Energy ($\Delta E_{vdw,\ kcal/mol}$)	Electrostatic Energy ($\Delta E_{ele,\ kcal/mol}$)	Polar Solvation Energy ($\Delta E_{polar,\ kcal/mol}$)	Non-polar Solvation Energy ($\Delta E_{nonpolar,\ kcal/mol}$)	Binding free energy ($\Delta G_{bind,\ kcal/mol}$)
ACE2_RBD	-29.04	-400.88	412.62	-3.33	-20.63
ACE2_Cathelicidn	-63.81	114.21	-80.32	-8.14	-38.06
RBD_1K48	-13.58	-9.59	19.43	-1.81	-5.55
RBD_1PT4	-49.39	-74.83	89.84	-6.06	-40.44
RBD_2M9O	-55.27	-97.19	127.59	-6.13	-31.00
RBD_βHD1	-51.24	42.30	-13.40	-6.92	-29.28
RBD_βHD4	-74.46	-71.31	119.44	-8.55	-34.87

In order to clarify the various contribution of energies on the binding affinity of RBD, a total of 1000 structures from the MD trajectory with stable conformations were selected to calculate the contribution of individual components such as ΔE_{vdw}, ΔE_{ele}, ΔE_{polar}, and $\Delta E_{nonpolar}$, to obtain the total binding free energy ΔG_{bind}. The binding free energy of Kalata B2 and β-HD4 peptides is showing significantly lower than control ACE2/RBD complex, indicating the binding interaction between selected peptides with RBD is stronger than human ACE2 with RBD domain of SARS CoV-2's spike protein.

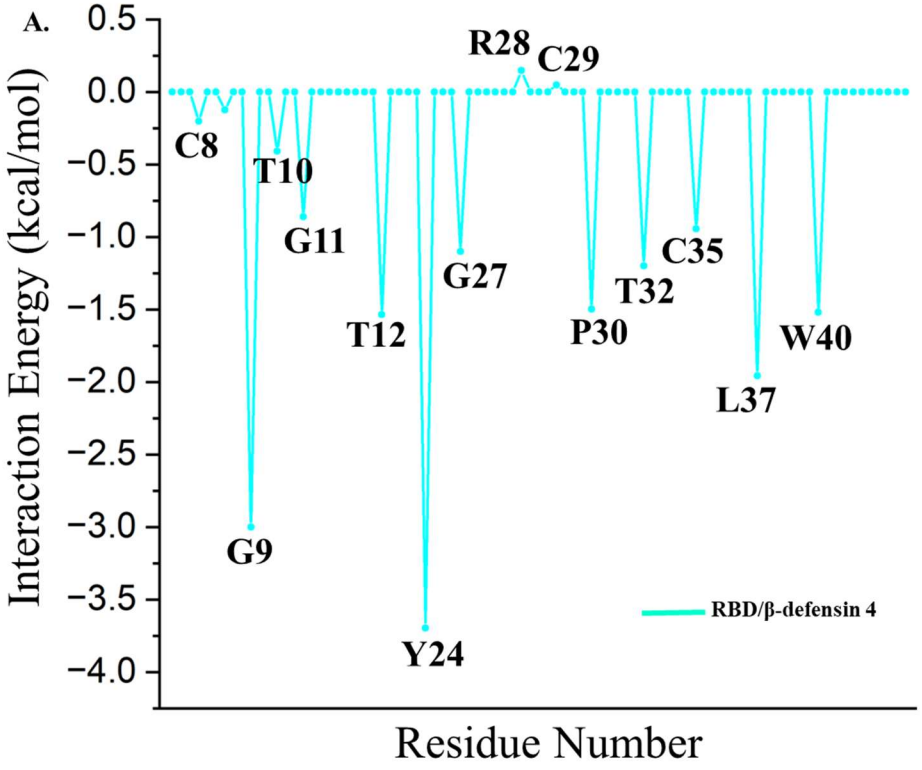
Per-residue Energy Decomposition Analysis

The residue-based energy decomposition of ΔG_{bind} is another effective approach to evaluate the energy contribution of key or active site residues. The energy contributions of these residues to the total binding energy are shown in Figure 6B. The key residues with favorable energy contribution mainly include the acidic amino acids of peptide inhibitors, such as aspartic acid and glutamic acid, and the basic amino acids of knottin and defensin peptides, such as arginine and lysine. These residues are generally involved in the electrostatic interaction during the MD simulation process, which means that the electrostatic interaction drives the binding. The line plot of free energy decomposition analysis shows the active site residues G9, T11, G10, T12, Y24, G27, P30, T32, L37 and W40 energetically favor the binding stability of β-defensin 4 to RBD domain of SARS CoV-2 spike protein as shown in Figure 6A.

The RBD/Kalata B2, B7 and β-defensin 1 complexes with binding affinities -40.44 Kcal/mol, -31.00 Kcal/mol and -29.28 Kcal/mol, respectively had higher binding affinity than RBD/Kalata B2 (-5.55 Kcal/mol) that agreed with the molecular docking and MD simulation results. Binding free energies and per-residue energy contribution of these peptide complexes have also been calculated in Figure S7-S9 in Supporting information. According to the per-residue energy contributions of RBD/Kalata B1 complex only E15 with interaction energy –0.12 kcal/mol interacted with RBD binding site. Whereas Kalata B2 (E3, T4, F6 and I21), Kalata B7 (L11, S22, W23, P24 and K27) and β-defensin 1 (V6, S7, Q11, L13 and P18) have strongly interacted and contributed more in peptide's binding site of RBD domain. Based on the present study, β-defensin 4 is the "best" peptide as well as potent inhibitor among the other complexes, with the highest inhibitory can be a potential peptide candidate against SARS CoV-2 infection.

PCA and FEL analysis

The overall motion of the RBD apo protein and RBD-bound peptide complexes was analyzed by principal component analysis (PCA) using construction of the eigenvectors. PCA was performed to the coordinates of the Cα-atoms on RBD-bound peptide inhibitors and RBD apo protein. Figure 7 as well as Figure S10 in Supporting information shows the comparison of PCA for RBD apo protein and RBD-bound knottins & defensins peptide complexes has shown that the motion properties described by the first two eigenvectors are different obtained by diagonalization of the atomic fluctuation's covariance matrix for RBD-bound peptide complexes. They have exhibited more structural motions than the RBD apo protein and is occupying a larger space. According to Figure 7A and B, the energy minima basins were shown from red to blue, and the blue zone denotes more stable conformation with minimal energy.



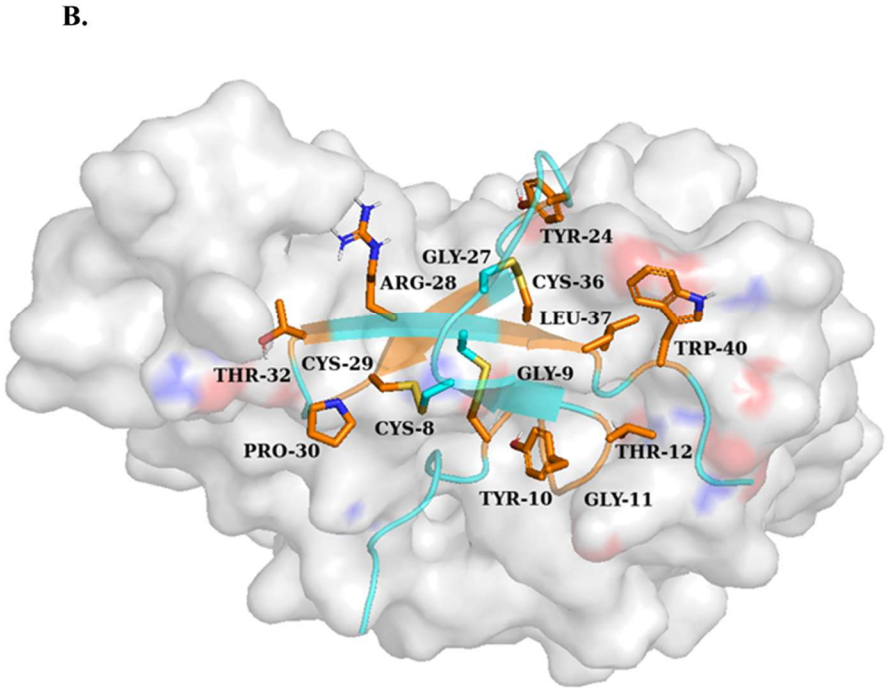


Figure 6: Per-residue decomposition projecting the minimum energy complexes structures; **A)** Binding free energy contribution of each residue in RBD- β-HD4 peptide complex; **B)** Per-residue binding of β-HD4 peptide into binding pocket of RBD domain drawn from MD snapshots (RBD protein shown in surface; Orange color denotes β-HD4 (peptide shown as turquoise color) interface residues at binding site).

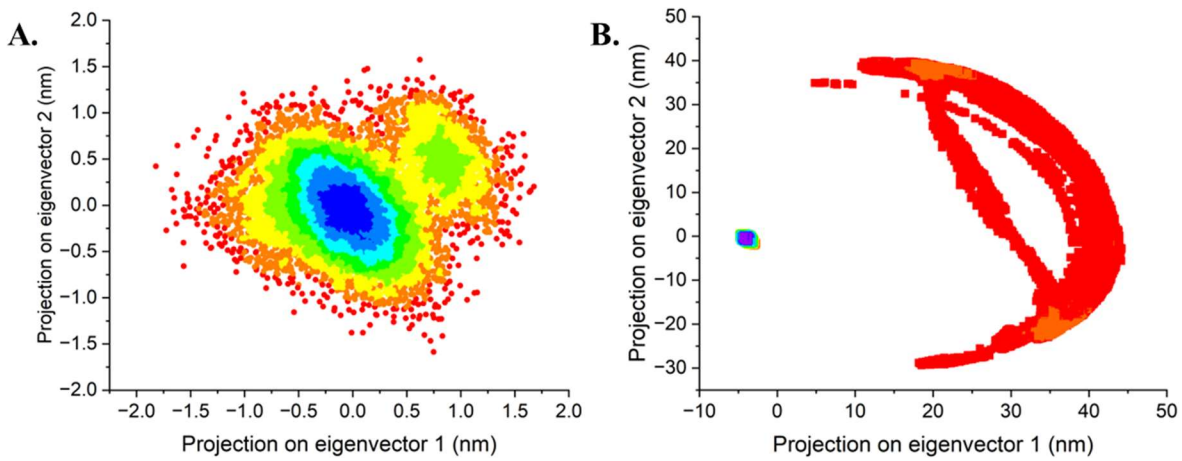


Figure 7: PCA scatter plot generated along the first two principal components from MD simulation trajectories.; **A)** RBD apo protein; **B)** RBD- β-HD4 peptide complex.

The 3D free energy landscapes (FELs) of the RBD apo and RBD-bound peptide inhibitor for the first two principal component (PCs) is represented in Figure 8. It is obvious from Figure 8 as well as Figure S11 in Supporting Information, suggests that the patterns of a basin on the FEL vary depending on the complexes. Figure S11 shows that while the RBD apo and other RBD-bound complexes explore broader conformational spaces, in case of RBD/betaHD4 examines a rather more constrained conformational space. These results indicate that the flexibility of RBD apo and other RBD-bound complexes is increasing, which is consistent with the RMS fluctuation data. The analysis of free energy landscape (FEL) has been used to identify lower-energy basins (minima) in RBD apo protein and RBD-bound knottins & defensins peptide complexes during the MD simulation trajectory. The purple-blue region in all the figures represents the lowest Gibbs free energy. The size and shape of FEL varies among the different RBD-bound peptide complexes, which indicates the stability differences of each complex. The more centralized, deeper, and smaller the purple-blue areas of the "valley," the more stable is the complex. As shown in Figure 9B, the RBD-β-HD4 peptide complex exhibits only one narrow and deep energy minimum, which indicates that this complex is stabilized in a minimum energy conformational region. Whereas the other complexes (knottins and β-HD1) and RBD apo protein principally exhibited two energy minima, which demonstrates that these two complexes cross two subspaces and transform between distinct conformational states.

For the knottins and β-HD1 peptide complexes, one energy minimum spread out over a large portion of the free energy space was detected, which indicates that these complexes have large conformational changes in the surrounding regions. Even though the depth of the "valley" of each system was like each other, the energy minimum value was positively correlated with its stability. The RBD/Kalata B1, B2 and β-HD1 complexes has a slightly deeper energy minimum compared to Kalata B7 peptide complex. Therefore, these analyses illustrated that the binding of the β-HD4 peptide complex results in a higher stability, which is in accord with our docking and molecular dynamics study.

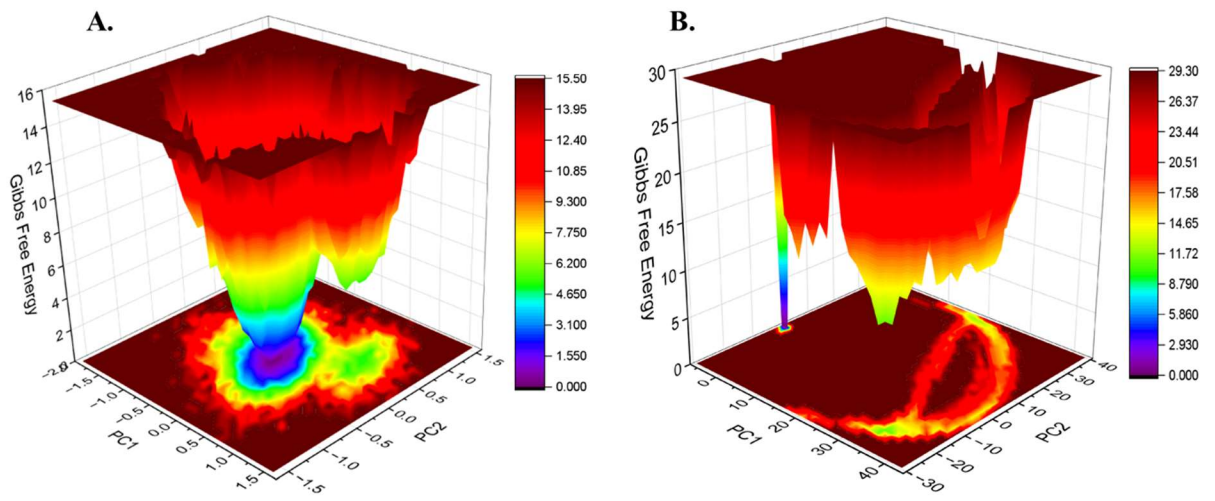


Figure 8: Free energy landscapes (FELs) projecting the minimum energy complexes structures; **A)** RBD apo protein; **B)** RBD-β-HD4 peptide complex.

Conclusion

To summarize, a virtual screening of a potential knottins and defensin peptides was performed, which resulted Kalata B1, B2 and B7 and β-defensin 1 and 4 as potential binders. These peptide binders are competitively binding to RBD domain and thus could be considered as potential peptide inhibitors of SARS-CoV-2 infections from our *in silico* study. We started preliminary verification of the molecular interaction and binding stability by molecular dynamics study, which was later subjected to *in silico* alanine screening to identify crucial residue at the peptide binding interface. Further, MM/PBSA binding free energy (BFEs) and per-residue energy contributions identified the binding mode and key residues as well as PCAs and FELs analysis from the MD simulation trajectories reveals the complex stability and difference in binding affinity between the five peptide inhibitors and RBD domain. β-defensin 4 (β-HD4) emerged out to be a better peptide binder compared to other defensins and control, based the on the *in silico* studies of thermodynamic parameters of bound complex. Subsequently, we are in process to validate the binding affinity of the acquired *in silico* screening of peptide candidates against the RBD domain of SARS-CoV-2 by an *in vitro* assay.

Acknowledgement

AB acknowledges financial support from IIT-Ropar, ISIRD grant, India. NMT acknowledges DST-INSPIRE for a doctoral fellowship. The authors thank IIT Ropar and Supercomputing facility for Bioinformatics and Computational biology (SCFBio), IIT Delhi for providing HPC access for all the computational study to carry out this work.

Conflict of interest

The author declares no conflict of interest.

References

1. WHO Coronavirus (COVID-19) Dashboard | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data. https://covid19.who.int/.

2. Jackson, C. B., Farzan, M., Chen, B. & Choe, H. Mechanisms of SARS-CoV-2 entry into cells. *Nature Reviews Molecular Cell Biology* vol. 23 3–20 Preprint at https://doi.org/10.1038/s41580-021-00418-x (2022).

3. Nguyen, H. L. *et al.* Does SARS-CoV-2 bind to human ACE2 more strongly than does SARS-CoV? *Journal of Physical Chemistry B* **124**, 7336–7347 (2020).

4. Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281-292.e6 (2020).

5. Linsky, T. W. *et al.* De novo design of potent and resilient hACE2 decoys to neutralize SARS-CoV-2. *Science (1979)* **370**, 1208–1214 (2020).

6. Zhang, L. *et al.* 1,5,* ,. doi:10.15252/emmm.202216109.

7. Pomplun, S. *et al.* De Novo Discovery of High-Affinity Peptide Binders for the SARS-CoV-2 Spike Protein. *ACS Cent Sci* **7**, 156–163 (2021).

8. Cao, L. *et al.* De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science (1979)* **370**, (2020).

9. Kim_D-Peptide Inhibitors of SARS-CoV-2.pdf.

10. Khatri, B. *et al.* A dimeric proteomimetic prevents SARS-CoV-2 infection by dimerizing the spike protein. *Nat Chem Biol* (2022) doi:10.1038/s41589-022-01060-0.

11. Wu, J., Zhang, J. & Zhang, H.-X. Computational Design of Miniprotein Inhibitors Targeting SARS-CoV-2 Spike Protein. *Langmuir* **38**, 10690–10703 (2022).

12. Sitthiyotha, T. & Chunsrivirot, S. Computational design of SARS-CoV-2 peptide binders with better predicted binding affinities than human ACE2 receptor. *Sci Rep* **11**, 1–14 (2021).

13. Polydorides, S. & Archontis, G. A computational protein design protocol for optimization of the SARS-CoV-2 receptor-binding-motif affinity for human ACE2. *STAR Protoc* **3**, 101254 (2022).

14. Maiti, B. K. Potential Role of Peptide-Based Antiviral Therapy against SARS-CoV-2 Infection. *ACS Pharmacology and Translational Science* vol. 3 783–785 Preprint at https://doi.org/10.1021/acsptsci.0c00081 (2020).

15. Gurung, A. B. *et al.* Potential of antiviral peptide-based SARS-CoV-2 inactivators to combat COVID-19. *PLoS One* **17**, 1–15 (2022).

16. Norman, A. *et al.* Discovery of Cyclic Peptide Ligands to the SARS-CoV-2 Spike Protein Using mRNA Display. *ACS Cent Sci* **7**, 1001–1008 (2021).

17. Johansen-Leete, J. *et al.* Antiviral cyclic peptides targeting the main protease of SARS-CoV-2. *Chem Sci* **13**, 3826–3836 (2022).

18. Choudhury, A. R., Maity, A., Chakraborty, S. & Chakrabarti, R. Computational design of stapled peptide inhibitor against SARS-CoV-2 receptor binding domain. *Peptide Science* (2022) doi:10.1002/pep2.24267.

19. Calugi, L., Sautariello, G., Lenci, E. & Trabocchi, A. Identification of a short ACE2-derived stapled peptide targeting the SARS-CoV-2 Spike protein. **2**,.

20. Han, Y. *et al.* Covalently Engineered Protein Minibinders with Enhanced Neutralization Efficacy against Escaping SARS-CoV-2 Variants. *J Am Chem Soc* **144**, 5702–5707 (2022).

21. Zhang, L. *et al.* HBD-2 binds SARS-CoV-2 RBD and blocks viral entry: Strategy to combat COVID-19. *iScience* **25**, (2022).

22. Kudryashova, E. *et al.* Inhibition of SARS-CoV-2 Infection by Human Defensin HNP1 and Retrocyclin RC-101. *J Mol Biol* **434**, (2022).

23. Whisenant, J. & Burgess, K. Blocking Coronavirus 19 Infection via the SARS-CoV-2 Spike Protein: Initial Steps. *ACS Medicinal Chemistry Letters* vol. 11 1076–1078 Preprint at https://doi.org/10.1021/acsmedchemlett.0c00233 (2020).

24. Fernandes, L. A. *et al.* Engineering Defensin α-helix to produce high-affinity SARS-CoV-2 Spike protein binding ligands. doi:10.1101/2022.02.09.479781.

25. Craik, D. J. & Conibear, A. C. The chemistry of cyclotides. *Journal of Organic Chemistry* vol. 76 4805–4817 Preprint at https://doi.org/10.1021/jo200520v (2011).

26. Wilson, S. S., Wiens, M. E. & Smith, J. G. Antiviral mechanisms of human defensins. *Journal of Molecular Biology* vol. 425 4965–4980 Preprint at https://doi.org/10.1016/j.jmb.2013.09.038 (2013).

27. Jarczak, J. *et al.* Defensins: Natural component of human innate immunity. *Human Immunology* vol. 74 1069–1079 Preprint at https://doi.org/10.1016/j.humimm.2013.05.008 (2013).

28. Darnell, S. J., Page, D. & Mitchell, J. C. An automated decision-tree approach to predicting protein interaction hot spots. *Proteins: Structure, Function, and Bioinformatics* **68**, 813–823 (2007).

29. Xue, L. C., Rodrigues, J. P., Kastritis, P. L., Bonvin, A. M. & Vangone, A. PRODIGY: A web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics* **32**, 3676–3678 (2016).

30. vilas/readHBmap.py at master · quytruong1808/vilas · GitHub. https://github.com/quytruong1808/vilas/blob/master/vilas/analyzer/readHBmap.py.

31. Morrison, K. L. & Weiss, G. A. Combinatorial alanine-scanning. *Curr Opin Chem Biol* **5**, 302–307 (2001).

32. Massova, I. & Kollman, P. A. Computational alanine scanning to probe protein-protein interactions: A novel approach to evaluate binding free energies. *J Am Chem Soc* **121**, 8133–8143 (1999).

33. Genheden, S. & Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discov* **10**, 449 (2015).