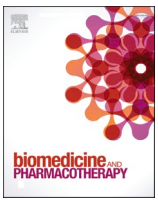
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Biomedicine & Pharmacotherapy 151 (2022) 113190

Contents lists available at ScienceDirect

Biomedicine & Pharmacotherapy

journal homepage: www.elsevier.com/locate/biopha



Self-derived peptides from the SARS-CoV-2 spike glycoprotein disrupting shaping and stability of the homotrimer unit

Monikaben Padariya [a,*], Alison Daniels [b,c], Christine Tait-Burkard [c], Ted Hupp [a,d], Umesh Kalathiya [a,*]

[a] International Centre for Cancer Vaccine Science, University of Gdansk, ul. Kładki 24, 80-822 Gdansk, Poland
[b] Department of Infectious Disease, Edinburgh, Scotland EH4 2XR, United Kingdom
[c] The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian, United Kingdom
[d] Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, Scotland EH4 2XR, United Kingdom

ARTICLE INFO

Keywords:
Self-derived peptides
Linear motif
SARS-CoV-2
Spike
COVID
Vaccine
Structural folding
RBD
Molecular dynamics

ABSTRACT

The structural spike (S) protein from the SARS-CoV-2 β-coronavirus is shown to make different pre- and post-fusion conformations within its homotrimer unit. To support the ongoing novel vaccine design and development strategies, we report the structure-based design approach to develop self-derived S peptides. A dataset of crucial regions from the S protein were transformed into linear motifs that could act as the blockers or stabilizers for the S protein homotrimer unit. Among these distinct S peptides, the pep02 (537-QQFGRDIAD-545) and pep07 (821-RDLICAQKFNGLTVLPPLLTDE-842) were found making stable folded binding with the S protein (550–750 and 950–1050 regions). Upon inserting SARS-CoV-2 S variants in the peptide destabilized the complexed S protein structure, resulting an allosteric effect in different functional regions of the protein. Particularly, the molecular dynamics revealed that A544D mutation in the pep02 peptide induced instability for the complexed S protein, whereas the N943K variant from pep09 exhibited an opposite behavior. An increased protein-peptide binding affinity and the stable structural folding were observed in mutated systems, compared to that of the wild type systems. The presence of mutation has induced an "up" active conformation of the spike (RBD) domain, responsible for interacting the host cell receptor. Among the lower affinity peptide datasets (e.g., pep01), the S1 and S2 subunit in the protein formed an "open" conformation, whereas with higher affinity peptides (e.g., pep07) these domains gained a "closed" conformation. These findings propose that our designed self-derived S peptides could replace a single S protein monomer, blocking the homotrimer formation or inducing stability.

1. Introduction

The severe acute respiratory syndrome coronavirus 1 (SARS-CoV-1) and SARS-CoV-2 (COVID) belonging to the beta coronavirus, are responsible for SARS epidemic in 2003 and 2019, respectively [1,2]. Among their different structural proteins, the spike (S) glycoprotein is the key component in these coronaviruses, which is responsible for host cell receptor binding. The S protein enhances the virus uptake and fusion with the target cell, and hence, makes it a main target of the immune system [3]. Particularly, SARS-CoV-2 spike gene encodes N-linked 22 glycan sequence per protomer and these glycans play crucial roles within S protein conformational changes responsible for immune evasion [4–6]. In the SARS coronaviruses family, the spike protein has been found in its homotrimer form, in which individual monomers play an important role in the functional structure folding [7,8]. Two conformations for S homotrimeric form have been identified, i.e., pre- and post-fusion [3,9,10]. The pre-fusion S protein is found comparatively more unstable [2,10], the receptor binding domains (RBD) alternate between an "up" (open) and "down" (closed) conformations [11]. Together the receptor binding domains and N-terminal domains (NTDs) make significant conformational changes inducing flexibility within the structure. These domains together form a "closed" or "open" conformation; the later one ("up") is responsible to bind with an angiotensin-converting enzyme 2 (ACE2) receptor from the host cell [11–13]. In addition, the S protein of the SARS-CoV family is cleaved by a host cell protease, the transmembrane protease / serine subfamily member 2 (TMPRSS2) [14] and this cleavage can be necessary for binding of the S protein to the ACE2 receptor.

* Corresponding authors.
E-mail addresses: monikaben.padariya@ug.edu.pl (M. Padariya), umesh.kalathiya@ug.edu.pl (U. Kalathiya).

https://doi.org/10.1016/j.biopha.2022.113190
Received 12 April 2022; Received in revised form 20 May 2022; Accepted 22 May 2022
Available online 24 May 2022

There have been several studies conducted revealing the importance of the spike homotrimeric form, and its different pre- and post-fusion conformations [15–18]. Xiong et al. [2], suggested that introducing point mutations in the spike protein allows the production of thermostable, disulfide-bonded S protein trimers which are trapped in the closed-conformation, i.e., in the pre-fusion state. In addition, it has been highlighted those murine antibodies (polyclonal) block the S protein fusion into the host cell, demonstrating that such cross-neutralizing

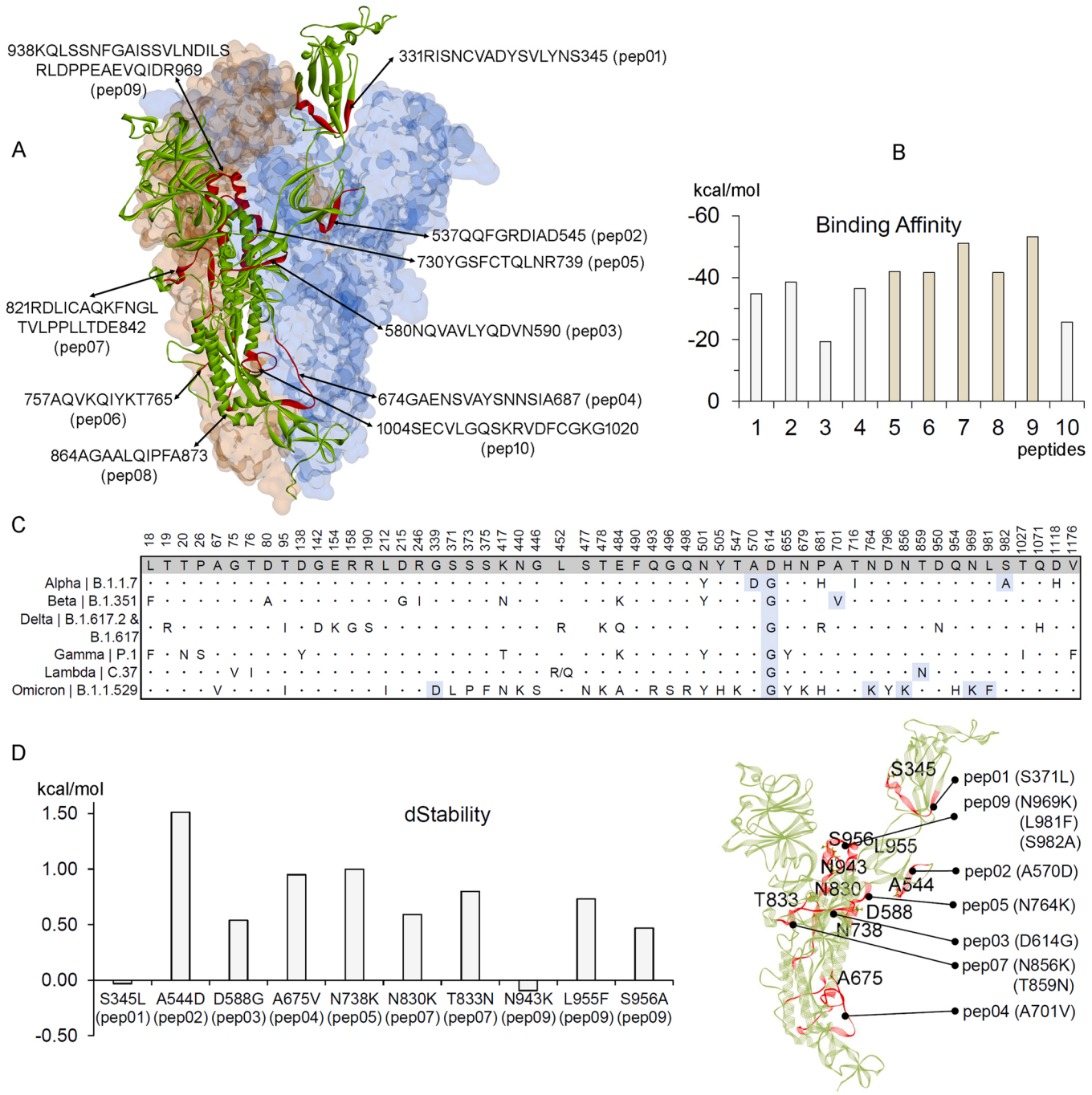


Fig. 1. Recognized self-derived SARS-CoV-2 spike (S) peptides linear motifs, and their binding affinity with the spike protein (monomers). (A) Average structure of the SARS-CoV-2 spike homotrimer retrieved from our previous study [10]. Considering different possible intermolecular interactions between spike monomers of the homotrimer unit, a set of 10 different self-derived peptides were constructed (marked in red color with labeled residues; chain A (monomer 1) in green, chain B (monomer 2) in blue, and chain C (monomer 3) in brown). (B) The binding affinities of 10 linear motif peptides (from monomer 1; chain A) with their respective S protein monomer (in light gray screened against monomer 2 (chain B), and light brown with monomer 3 (chain C). (C) A few infectious known SARS-CoV-2 virus variants, and emerging spike protein mutations from the Alpha (B.1.1.7), Beta (B.1.351), Delta (B.1.617.2 & B.1.617), Gamma (P.1), Lamba (C.37), and Omicron (B.1.1.529) variants [24,25]. The highlighted mutations in blue background belong to the linear S peptide motifs traced binding with spike protein monomers. (D) The left panel describes the change in binding affinity (dStability or ΔStability) upon inserting mutations in a particular peptide, identified using the residue scan pipeline implemented in the Molecular Operating Environment (MOE; Chemical Computing Group Inc., Montreal, QC, Canada)package. The right panel shows different mutations residing in a particular peptide (pep01; S345L, pep02; A544D, pep03; D588G, pep04; A675V, pep05; N738K, pep07; N830K / T833N, and pep09; N943K / L955F / S956A) with their location over the S protein (a 26 amino acid (aa) numbering difference shall be considered when comparing these data with the crystal structure of spike protein pdb id.: 6vsb [26,27]).

antibodies binding conserved epitopes can be generated upon vaccination [19]. Moreover, the pre-fusion S protein conformation in the presence of the infectious particle contains the epitopes for neutralizing antibodies [11,15,16,20,21]. The S protein structural transition from pre-fusion to post-fusion can be achieved by the simulators such as upon receptor binding, antibody binding, small molecule binding, and protease digestion [2,10,11,17,22,23].

Despite different structural approaches to target SARS-CoV-2 virus, there lacks a detailed understanding in the direction that would block the homotrimer formation of the SARS-CoV-2 S protein. Considering recent studies on the S protein fusion mechanisms, we sought to block the homotrimer formation or trapping stabilized pre-fusion conformation by self-derived S peptides or linear motifs. Since homotrimer is the functional unit of the spike protein, disrupting this homotrimer formation could block the viral entry into the host cell. In this line, we proposed linear motifs (*in silico* designed) that were found involved in the protein-protein intramolecular interaction between monomers of the S homotrimer. In our previous studies [10], we have generated an optimized SARS-CoV-2 S protein structure that was used in current work to identify interactions between monomers and designing self-derived linear peptide motifs (Fig. 1A). If successful, blocking the trimer functional unit could have several significant impacts on understanding the SARS-CoV-2; it can induce immune response, a critical component for vaccinology, and a blocked monomer or stabilized homotrimer can induce production of antibody repertoire different from the "open" conformation. Additionally, we investigated these linear motif self-derived peptides with respect to different mutations derived from the SARS-CoV-2 virus variants (Alpha | B.1.1.7, Beta | B.1.351, Delta | B.1.617.2 & B.1.617, Gamma | P.1, Lambda | C.37, and Omicron | B.1.1.529) [24,25].

2. Material and methods

2.1. System build-up and mutation landscape designing

The optimized SARS-CoV-2 spike protein homotrimer structure from our previous study [10], was used to trace stable intramolecular interactions between S monomers (Fig. 1A; the crystal structure of spike protein pdb id.: 6vsb [26,27]). Common high occupancy interactions between S monomers were considered to construct 10 self-derived linear motifs / peptides, that could mimic the interactions within the S monomers of the homotrimer unit. These intermolecular interactions were traced using the BIOVIA Discovery Studio visualizer (Dassault Systèmes, BIOVIA Corp., San Diego, CA, USA) package. Upon identifying self-derived peptides, the model structures were constructed and optimized using the Molecular Operating Environment (MOE; Chemical Computing Group Inc., Montreal, QC, Canada) pipelines. The following peptides (Fig. 1A) were energy minimized applying the CHARMM27 forcefield [28]: 331-RISNCVADYSVLYNS-345 (pep01), 537-QQFGRDIAD-545 (pep02), 580-NQVAVLYQDVN-590 (pep03), 674-GAENSVAYSNNSIA-687 (pep04), 730-YGSFCTQLNR-739 (pep05), 757-AQVKQIYKT-765 (pep06) 821-RDLICAQKFNGLTVLPPLLTDE-842 (pep07), 864-AGAALQIPFA-873 (pep08), 938-KQLSSNFGAISSVLNDILSRLDPPEAEVQIDR-969 (pep09), and 1004-SECVLGQSKRVDFCGKG-1020 (pep10). In our studied S protein structure [10], there is 26 amino acid (aa) numbering difference with respect to that of the crystal structure pdb id.: 6vsb [26,27].

The optimized peptide structures were transformed into a peptide library in MOE (Chemical Computing Group Inc., Montreal, QC, Canada) package, which were further screened (*in silico*) against individual S protein monomers tracing their binding affinities. The spike protein-peptide docking was implemented keeping the S protein as rigid and allowing S peptides high degree of freedom [29]. During protein-protein docking 500 conformations were produced by Generalized Born/Volume Integral (GB/VI, kcal/mol; Fig. 1B) [30] binding energies and applying CHARMM27 forcefield [28]. The best binding affinity conformation of the S peptide liner motifs with either of the S protein monomers were further investigated by molecular dynamics simulation (MDS) technique.

For the emerging mutations in the SARS-CoV-2 virus variants, we traced change in the binding affinity of S peptides (Fig. 1A) with the protein upon inserting point mutations. Mutations were retrieved from the following SARS-CoV-2 virus variants: Alpha (B.1.1.7), Beta (B.1.351), Delta (B.1.617.2 & B.1.617), Gamma (P.1), Lambda (C.37), and Omicron (B.1.1.529) [24,25] (Fig. 1C). Particularly, the effect of mutations in S peptides were investigated in the presence of wild type S protein and they were: pep01 (S345L), pep02 (A544D), pep03 (D588G), pep04 (A675V), pep05 (N738K), pep07 (N830K/T833N), and pep09 (N943K / L955F / S956A). The "Residue Scan" protocol from the protein design module in the MOE (Chemical Computing Group Inc., Montreal, QC, Canada) was implemented to insert mutations in the peptide, and to trace change in stability / binding affinity with the S protein (Fig. 1D). Mutations showing significant changes in the protein-peptide binding affinity were further investigated by the MDS.

To understand the effect of mutation over the SARS-CoV-2 S homotrimer formation (Fig. 1C and Table S1), each amino acids from the S protein monomer were screened against 20 other amino acids (A, R, N, D, C, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y, and V). Individual variant occurring in a particular monomer was screened in the presence of other two monomers from the S homotrimer unit. These analyses could allow interpretation of the effect of naturally occurring SARS-CoV-2 variants, as well as different possible mutations that could induce stability / flexibility within the S protein trimer (frequency with probability of amino acids at the residual mutation sites) [31]. Applying "Low Mode MD" parameters, the energy window for each mutation was set to 10 kcal/mol and RMSD limit was set to 0.25 Å. In addition, the residues farther than 4.5 Å of the point mutation in S protein were kept as fixed [32].

2.2. Molecular dynamic simulations over wild type and mutant S protein-peptide constructs

The MD simulations over the 20 constructed wild type and mutated systems were performed using the GROMACS 4.6.5 [33] package and applying the CHARMM27 forcefield. All modeled S protein-peptide complexes were placed in a periodic boundary condition (PBC) dodecahedron box (10 Å thick), that was solvated using the simple point charge (SPC) water models [34] and $Na^{+}Cl^{-}$ counter ions. Individual simulation boxes were energy minimized using the steepest descent algorithm for 50,000 steps or till the local minima was achieved. Particle Mesh Ewald (PME) method [35] was used to compute electrostatic interactions and bond lengths between the atoms were constrained using the LINCS algorithm [36]. The 10 Å cut-off distance was implemented for van der Waals and Coulomb interactions. Subsequently, the simulation systems were equilibrated for 1000 ps using the NPT (isobaric-isothermal) ensemble, maintaining the temperature (300 K; V-rescale thermostat [36]) and pressure (1 bar; Parrinello-Rahman barostat [37]). A 100 ns production run for each system was performed using the leapfrog integrator [38], saving coordinates every 10 ps. Parameters for tracing hydrogen bonds: donor-acceptor cutoff distance and angle were set to 3.5 Å and $\geq$ 160–180°, respectively. The retrieved MD trajectories after 100 ns were analyzed using the GROMACS and VMD tools [39], as well as by the MOE (Chemical Computing Group Inc., Montreal, QC, Canada) and BIOVIA Discovery Studio (Dassault Systèmes, BIOVIA Corp., San Diego, CA, USA) packages.

3. Results and discussion

The set of self-derived S peptides (Fig. 1A) were screened against the S protein monomer 2 (chain B; pep01, pep02, pep03, pep04, and pep10), and peptides pep05, pep06, pep07, pep08, and pep09 were screened against monomer 3 (chain C). Individual S peptide-protein docked complexes were ranked according to their binding affinities (Fig. 1A;

GB/VI, kcal/mol). Peptides belonging in the category of binding with monomer 3 were found showing higher affinity, compared to that of peptides studied with monomer 2 (Fig. 1A). The S protein mutations emerging from different SARS-CoV-2 virus variants [24,25] (Fig. 1C and Table S1); Alpha (B.1.1.7), Beta (B.1.351), Delta (B.1.617.2 & B.1.617), Gamma (P.1), Lambda (C.37), and Omicron (B.1.1.529) were inserted in the self-derived S peptides. Majority of these mutation reduce structural stability of the peptide. Exceptionally, S371L (pep01) and N943K (pep09) induced a moderate stability within the S peptides amino acid, whereas the A544D, D588G, A675V, N738K, N830K, T833N, L955F, and S956A mutations had a contrast behavior. In presence of the S protein, the A544D mutation gained highest instability within its pep02 peptide structure. Particularly, for the pep09 peptide distinct mutations demonstrated a diverse effect on the peptide folding (Fig. 1C).

3.1. The S protein-peptide best affinity conformation were investigated in the solvent environment (wild type systems)

The findings from *in silico* screening of peptides with protein and mutational effect (over protein-peptide constructs) were further investigated using the MD simulation approach. At first, simulated SARS-CoV-2 protein-peptide wild type and mutated complexes were explored tracing change in the stability over MDS time course. The RMSD (root mean square deviation) a time dependent change in the non-hydrogen

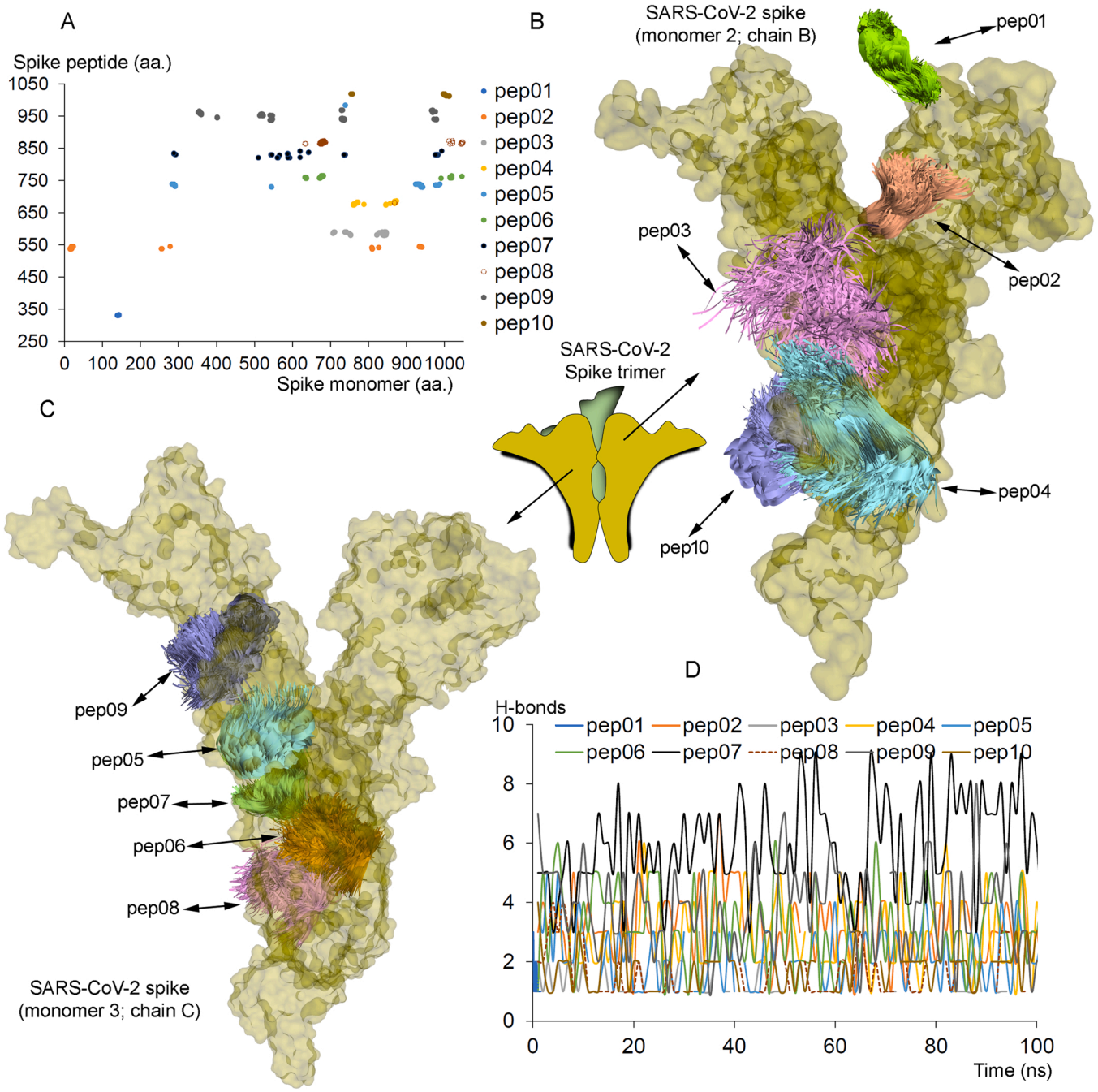


Fig. 2. Distinct self-derived S peptides screened (*in silico*) against the wild type S protein monomers. (A) The protein-peptide intermolecular hydrogen bonding residues identified during molecular dynamics simulations (MDS). (B) and (C) Different conformations of a particular peptide with its respective SARS-CoV-2 spike monomer (chain B or chain C). (D) The plot describing change in the protein-peptide interactions over the MD simulation time course. Parameters for tracing hydrogen bonds were donor-acceptor cutoff distance and angle was set to 3.5 Å and ≥ 160–180°, respectively.

atoms, and the root mean square fluctuations (RMSF; Cα atoms of each residue from the spike protein) were computed. The change in the binding free energy based on the hydrogen bond (H-bond) interactions of the S protein-peptide complex were analyzed, along with significant residues showing selectivity to sets of S peptides. Residues from the pep01 peptide were found interacting with the S protein residues ranging from 139 to 143 aa (Fig. 2A and Table S2). A set of few hot-spots were traced in the pep02-protein binding, and the residue D545 from S peptide demonstrated high occupancy binding with R18 and K938 from the S protein (Fig. 2A and Table S2). Common regions from S protein (700–900 aa) were found interacting with the pep03 and pep04 peptides, the N677 residue from peptides formed stable binding with I762 and K764 amino acids of the S protein (Fig. 2A and Table S2).

The R739 residue from the pep05 peptide was found to make high occupancy binding with D924 (84.43%) of the S protein. The S protein residues from 600 to 700 and 990–1050 aa range were found binding with the pep06 peptide, and the residues K764 and I762 were shown making high occupancy binding with N677 of S protein (42.42% and 39.72%, respectively). Additionally, these K764 and I762 residues from pep06 were also found binding with E676 and A675 from the S protein, respectively. Particularly, the residue pair R821(pep07)-E593(protein) was found showing high occupancy (91.02%), whereas the majority of

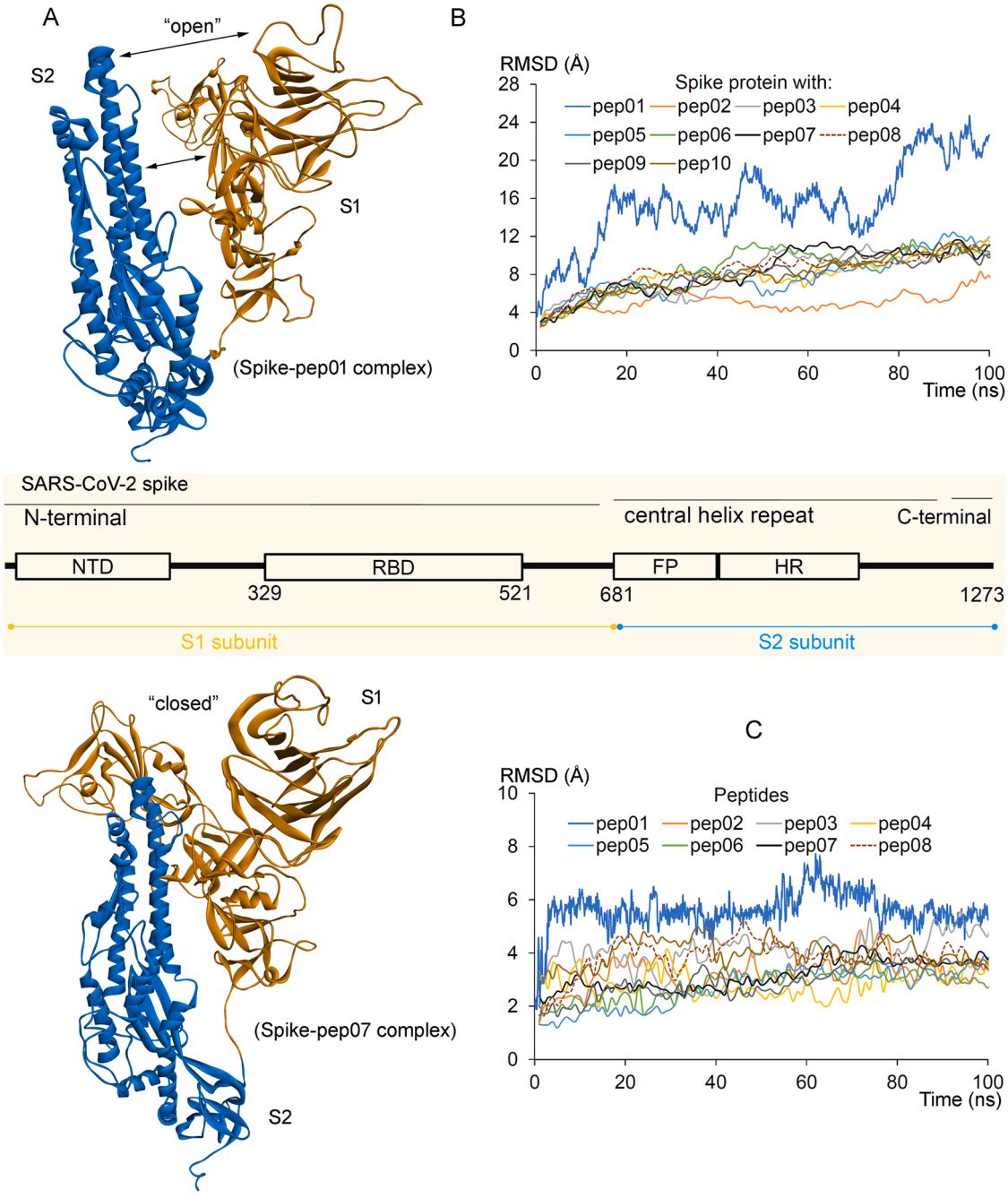


Fig. 3. The change in structural dynamics of the spike protein-peptide complexes over the MD simulation time course. (A) Different conformation of spike monomer observed when interacting with lower affinity (pep01) peptide, compared to a highly efficient binding peptide (pep07). In the presence of pep01, the S1 and S2 subunit from the S protein slightly shifted away from each other forming an "open" conformation. (B) The root-mean-square deviation (RMSDs) of each atomic position (excluding hydrogen atoms) for a spike monomer when complexed with different peptide. (C) Computed RMSDs of individual peptides, when complexed with the S protein monomer.

residue from the pep08 peptide had low occupancy interactions with the S protein. Peptide pep09 formed four binding sites with the S protein, and the peptide residues showing high occupancy binding the S protein were: K938, N952, S956, R957, and D959. Despite few residues involved in the peptide-protein binding, the R1013 residue from pep10 showed high frequency binding (throughout MD simulation time) with E1005 of the S protein (Fig. 2A and Table S2). Overall, a larger set of S protein residues ranging from 550 to 750 and 950–1050 aa, were found binding with different self-derived S peptides.

The structural folding of differently sized S peptide linear motifs revealed that except the pep02 and pep07 peptides, most of the studied peptides have obtained a high conformational movement with respect to their starting positions (Fig. 2B and C). In addition, these pep02 and pep07 peptides were found showing a higher number of interactions with the S protein (Fig. 2D). Furthermore, the pep04, pep06, and pep09 peptides showed a moderate sustained numbers of protein-peptide intermolecular interaction (Fig. 2D). The structural folding of the S protein in the presence of different peptides, revealed that peptides making higher binding affinity induced stable shaping or folding in the S protein structure (Fig. 3A). For example, a contrast conformation within the S protein is traced when put together with a peptide having less binding affinity (pep01), compared to that with the pep07 peptide. In the presence of pep01 molecule, the S1 and S2 subunit form the S protein slightly shifted away from each other forming an “open” conformation, whereas the S protein with pep07 no such confirmation was traced (“closed” conformation; Fig. 3A). These findings suggest that a particular S protein monomer require other two monomers for a stable homotrimer folding, or a self-derived S peptide may induce stability within the S protein structure folding.

The RMSDs of the S protein monomer in presence of different self-derived S peptides, suggest that protein with pep01 obtained higher flexibility compared to that with other peptides (Fig. 3B). The S protein monomers obtained a consistent stability after ~50 ns though with different peptides, and this stability ranged between ~8–10 Å (Fig. 3B). Particularly, the S protein with pep02 has obtained highest stability within its structure (Fig. 3B). Investigating different partners of the S protein, like the pep01 and pep03 has highest flexibility, whereas other peptides had RMSDs ranging between ~3–4 Å after 50 ns (Fig. 3C).

3.2. Mutant S protein-peptide compared with the wild type systems

The effect of different mutations occurring in the SARS-CoV-2 virus variants (Table S1) were investigated, with respect to our studied S protein-peptide components. Individual amino acid flexibility (RMSF; based on C-alpha (Cα) atoms) for each S protein monomers in presence of peptide, as well as from the wild type and the mutated systems are presented in Fig. 4. The change in the amino acids flexibility upon inserting S345L in the S peptide, suggests that the protein with pep01 has higher flexibility compared to other mutations (Fig. 4A). In particular, the mutated S345L amino acids in peptide have shown ~ 7 Å flexibility in the S protein compared to that of the wild type. Upon inserting A544D mutation in the peptide a stable S protein was observed, however, this positioned residue in the protein itself has shown instability (Fig. 4A).

Regions of the S protein in presence of different peptides and mutations have shown a similar trend in the change for the RMSF plots. For example, residue ranging 250–450, 680–760, and 915–1000 has induced stability for the S protein in presence of the pep02 and pep03 peptides, whereas contrast behavior was observed for the pep01, pep04, and pep05 peptides (Fig. 4A). For the N830K and T833N mutations in the pep07 peptide has shown a distinct behavior with the S protein, a higher flexibility is observed in the RMSF plots for the protein with T833N mutations. All mutations introduced in the pep09 (N943F, L955F, and S956A) peptide have induced stability in the complexed S protein structure. Furthermore, a significant increase in stability is observed in the protein structure upon N943K mutation in the pep09 peptide (Fig. 4A). Inserting mutations in the S peptide destabilized the complexed S protein structure, resulting an allosteric effect over different functional regions of the protein (Fig. 4A).

Measuring change in the amino acid stability from self-derived S peptides in the wild type and mutated forms, highlighted a significant difference among peptides (Fig. 4B). Particularly, for the pep01 peptide molecule the centered amino acids induced stability, whereas tailed residues from N-terminus or C-terminus has higher flexibility. The initial residues (N-terminus) from pep03 and pep05 has less stability compared to C-terminus, whereas the pep02 and pep04 peptides that formed high affinity binding with the S protein has induced stability (comparing the wild type and mutant systems; Fig. 4B). For the N830K and T833N mutant systems (Fig. 4A), the pep07 in N830K mutated system has conserved structure compared to that of the T833N system. Moreover, the N943F, L955F, and S956A point mutations have resulted in higher flexibility among the 944–955 residues for the pep09 peptide (Fig. 4B).

Tracing individual residue making interactions between the mutated S peptide and the wild type S protein, suggest that pep01 has shown a higher number of residues involved in the protein-peptide interactions in the mutated system (Fig. 5A). The mutated pep02 and pep03 peptides have slightly reduced the number of amino acids involved in the protein-peptide binding. No significant change for pep04 was observed in both forms, whereas the pep05 peptide has induced interactions in the mutated system. Particularly, the pep07 (T833N) mutant system is similar as in the wild type whereas pep07 (N830K) has fewer binding residues, and the pep09 peptide has formed an almost similar number of interactions in different mutated and wild type systems (Fig. 5A). Overall, for the mutant peptide system a higher number of residues were found binding the S protein, despite increased flexibility in the protein structure (Fig. 4A).

A change in the H-bond interactions over time highlighted that the pep01 (S345L) peptide has higher number of interactions with the protein in the mutated system, compared to that of the wild type system (Figs. 2D and 5B). The mutated L345 amino acid from the peptide was found binding with the R969 and D172 residues of the S protein (with occupancy 75.85%; Table S3). Among studied mutations, the pep07 (N830K or T833N) have shown a higher number of interactions between S peptide and protein (Fig. 5B). In the wild type system the N830 residue from the pep07 peptide was found interacting with T983, Q979, and Q736 residues of the S protein, whereas these interactions were disrupted upon N830K mutation (Fig. 5C). Moreover, inserting T833N in the pep07 peptide has induced N833(peptide)-D588 / G568 / Q288 (protein) interactions (Fig. 5D).

Inserting A544D mutation in pep02 have replaced A544(peptide)-N934(protein) with D544-R18 / K938, whereas A675V mutation in pep04 has almost similar binding frequency with the S protein (Fig. 5C and D). Conformational dynamics of the wild type and mutant residues at position 544 from the pep02 (A544D) peptide are presented in Fig. 5E. In the pep05 peptide, the residue positioned at 738 was found binding only in the mutated N738K form with N291 and F566 of the S protein (Fig. 5D). Residues L955 and S956 from the pep09 peptide binding with S protein in the wild type were disrupted upon inserting mutations, whereas in the pep09 (N943K) system the mutated residue itself interacted with the S protein (Fig. 5D). Visualizing dynamics of different protein-peptide complexes, it was observed that presence of mutation in the system has induced an “up” active conformation of the spike (RBD) domain (Fig. 5E), that could enhance the interaction with the ACE2 receptor.

A change in RMSD for the S protein in different mutated systems, suggest that the monomers with the pep05 (N738K) and pep04 (A675V) peptides have higher flexibility (Fig. 6A). Comparing protein dynamics when with two different pep07 mutations (N830K and T833N), the protein with pep07 (N830K) has stable structure (Fig. 6A). Conformational dynamics of individual mutated peptides (Fig. 6B) over the 100 ns of MDS time course, demonstrated that mutated peptides have stable structural folding compared to that of the wild type (Fig. 2B and C). The

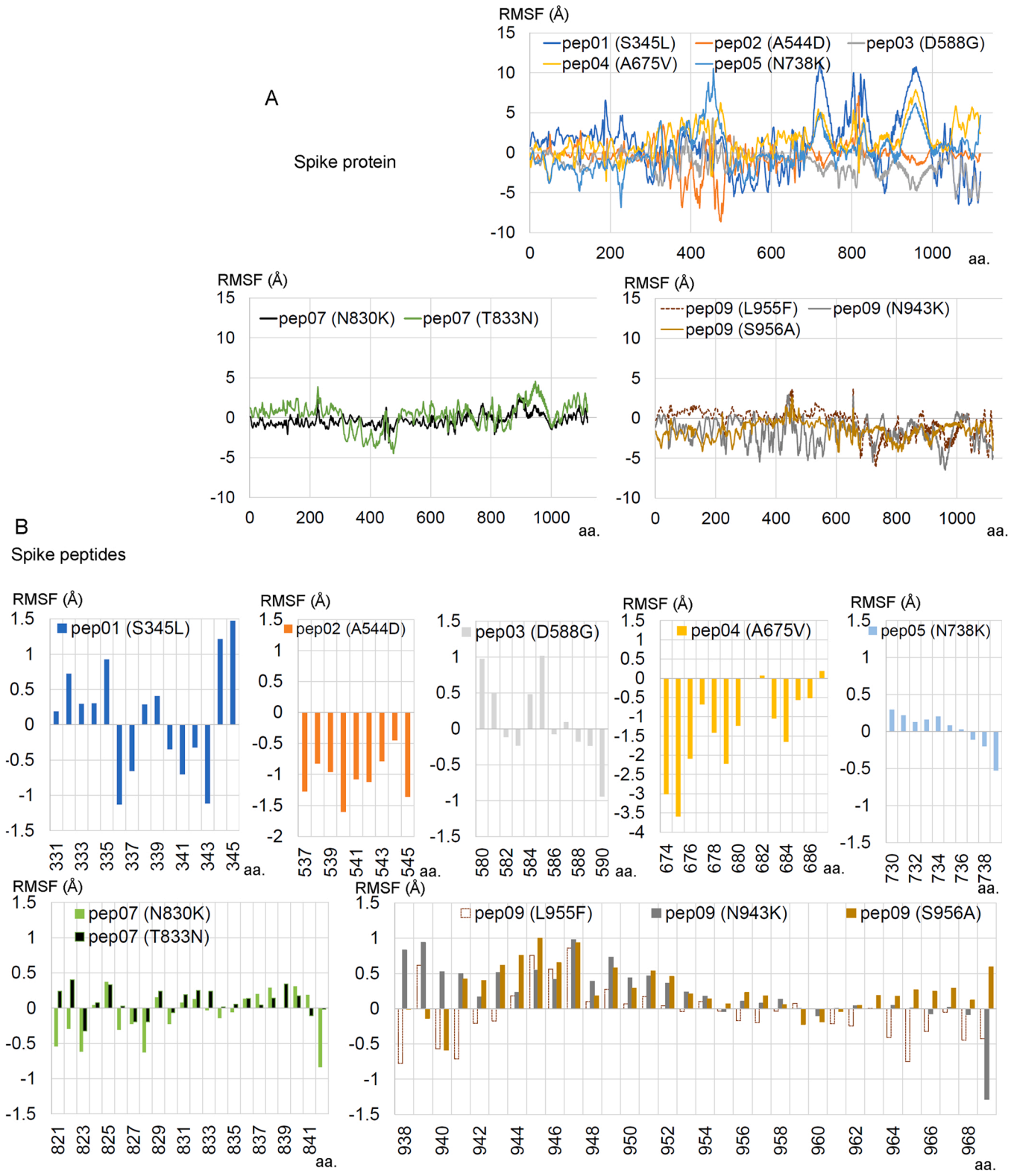


Fig. 4. Conformation dynamics of individual amino acids from the wild type and mutated systems (point mutations were inserted in the self-derived peptides). (A) The change / difference in the RMSF for the S protein, when complexed with different mutated S peptides. (B) Change in the RMSF for self-derived S peptides (compared with wild type and mutated systems). The following point mutations were inserted in the S peptides: S345L (pep01), A544D (pep02), D588G (pep03), A675V (pep04), N738K (pep05), N830K / T833N (pep07), and N943K / L955F / S956A (pep09). A 26 amino acids numbering difference shall be considered with respect to the crystal structure of the S protein pdb id.: 6vsb [26,27].

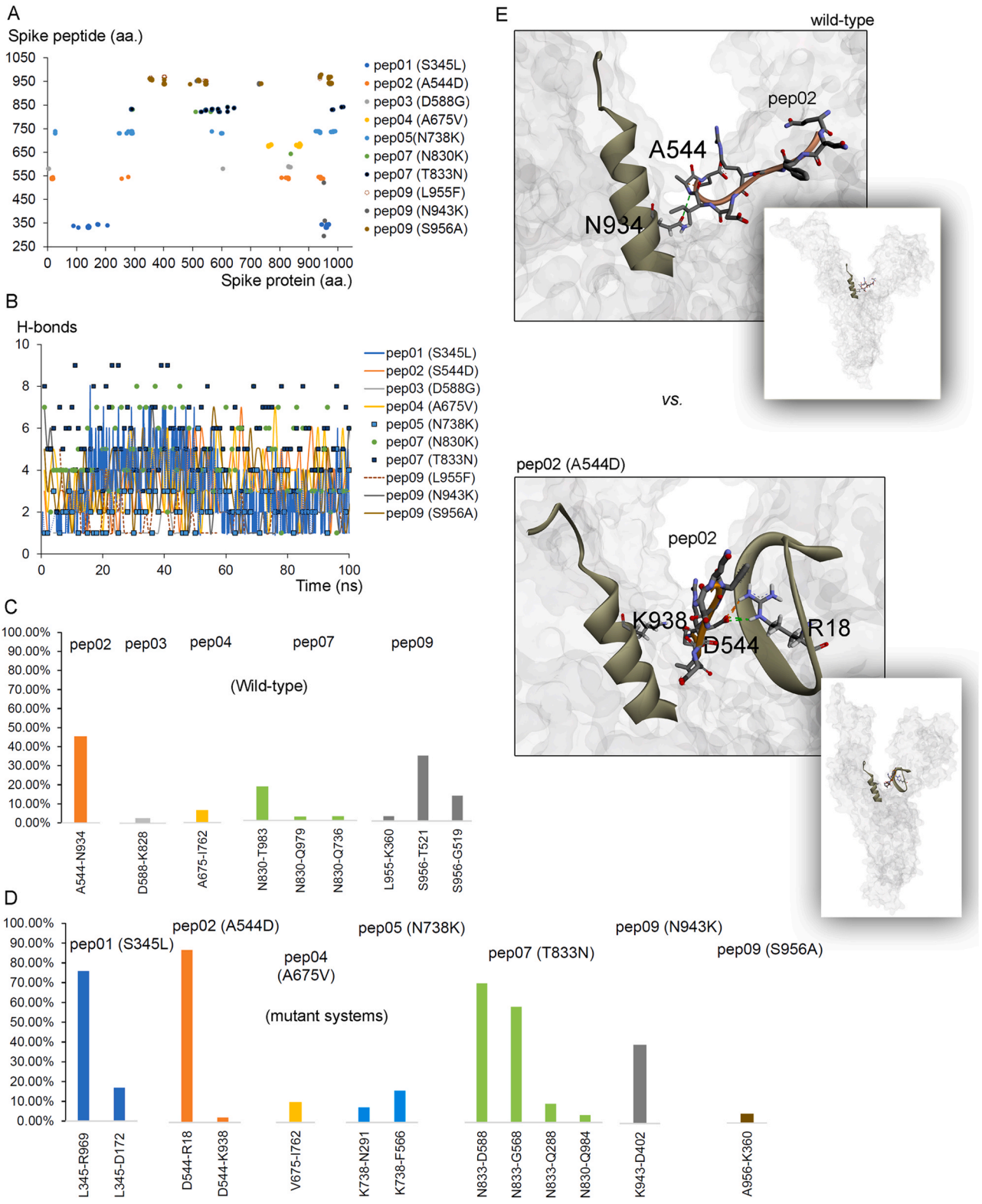


Fig. 5. Effect of mutations over the binding affinity between the SARS-CoV-2 spike peptides and protein. (A) The protein-peptide intermolecular hydrogen bonding residues identified from the systems containing the mutant peptide. (B) The hydrogen bond interaction plot describing change in the protein-peptide interactions over the MD simulation time course. Parameters for tracing hydrogen bonds were donor-acceptor cutoff distance and angle was set to 3.5 Å and ≥ 160–180°, respectively. (C) and (D) Particular amino acids from the peptide found mutated in different SARS-CoV-2 virus variants, and making interactions with the spike protein (interactions with occupancy ≥ 2% or 2 ns are shown). These residues are labeled in the peptide-protein order. (E) Conformational dynamics of the wild type and mutant residue at position 544 from the pep02 (A544D) peptide. Intermolecular interactions between protein and peptide are represented in the dotted line. Presence of mutation in the protein-peptide system has induced an "up" active conformation of the spike (RBD) domain that could be responsible for interacting with the ACE2 (angiotensin-converting enzyme 2) receptor.

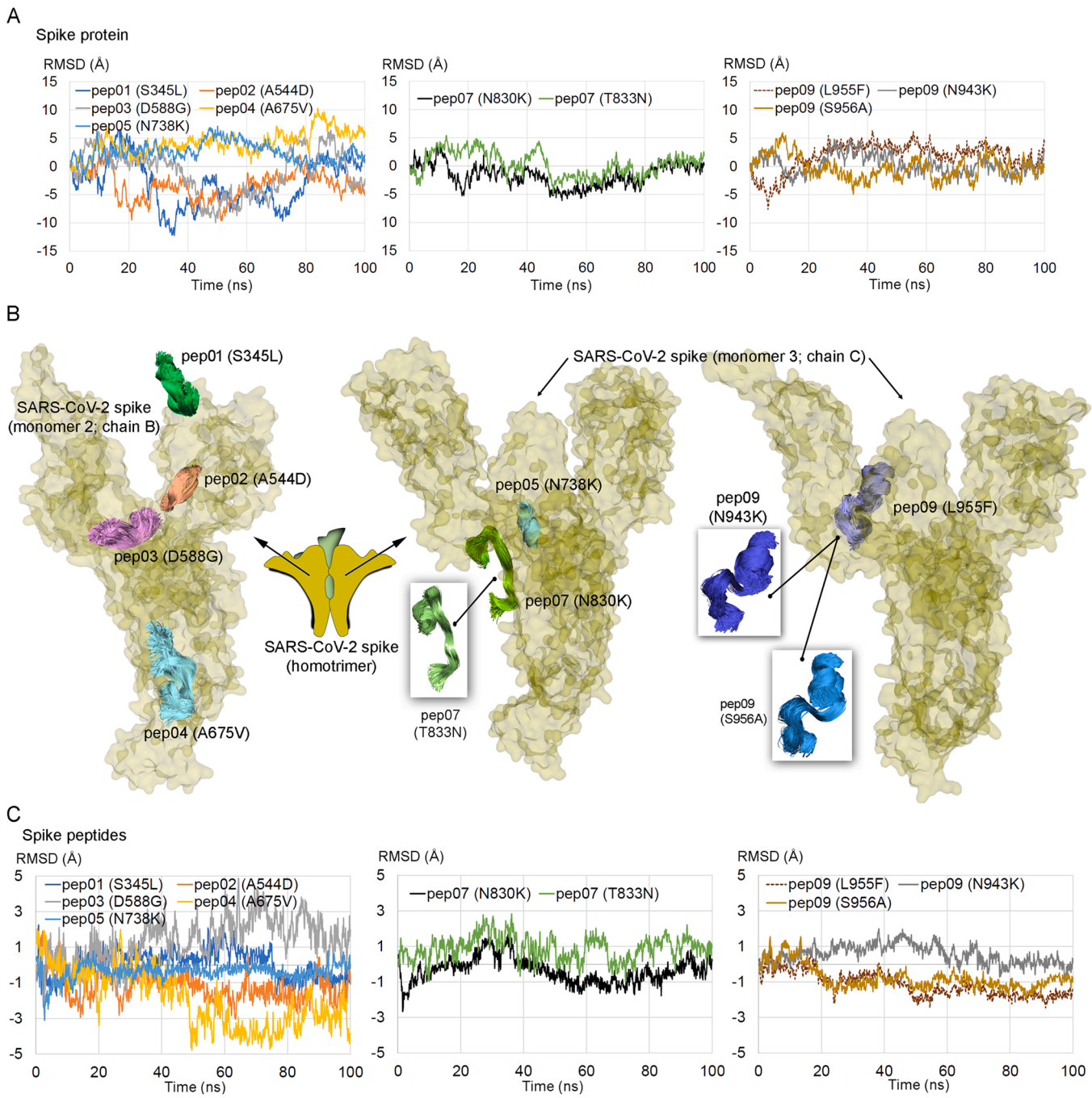


Fig. 6. Investigating the influence of different variants over the SARS-CoV-2 spike protein and peptide structures. (A) Difference in RMSDs for the spike protein (monomers) compared with the wild type and mutant systems, in the presence of different self-derived derived linear motif S peptides. (B) The conformation dynamics upon interesting point mutation in the SARS-CoV-2 S peptides. (C) The change in the RMSD values for linear motif S peptides upon inserting point mutations.

mutated pep02 (A544D) and pep04 (A675V) peptides, were found comparatively more stable after 50 ns (Fig. 6C). In addition, the pep07 (N830K) peptide was found less flexible compared to pep07 (T833N), whereas all three mutations (L955F, N943K, and S956A) in the pep09 peptide has obtained stability similar as that of the wild type (Fig. 6C).

3.3. Mutational landscape over the SARS-CoV-2 spike homotrimer

To understand the effect of different SARS-CoV-2 variants over the S protein homotrimer formation (Fig. 1C and Table S1), each amino acids from the S protein were screened against 20 other amino acids (A, R, N, D, C, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y, and V). These findings can help to interpret the frequency with probability of amino acids at the residual mutation sites of naturally occurring SARS-CoV-2 variants, as well as different possible mutations. Overlaying S protein mutation from SARS-CoV-2 variants with the protein-peptide interactions from the wild type system, highlighted that most mutations reside in regions where self-derived peptides interact with the S protein (Fig. 7A and Table S4). The frequency of mutations demonstrating protein-protein binding affinity change within the trimeric S protein, suggest significant changes in the residue ranging from 500 to 1000 aa, where a higher number of the protein-peptide interactions (H-bonds) were traced (Fig. 7B and Table S4). Particularly, the A544R (residing in the pep02) mutation induced binding affinity within the spike homotrimer monomers

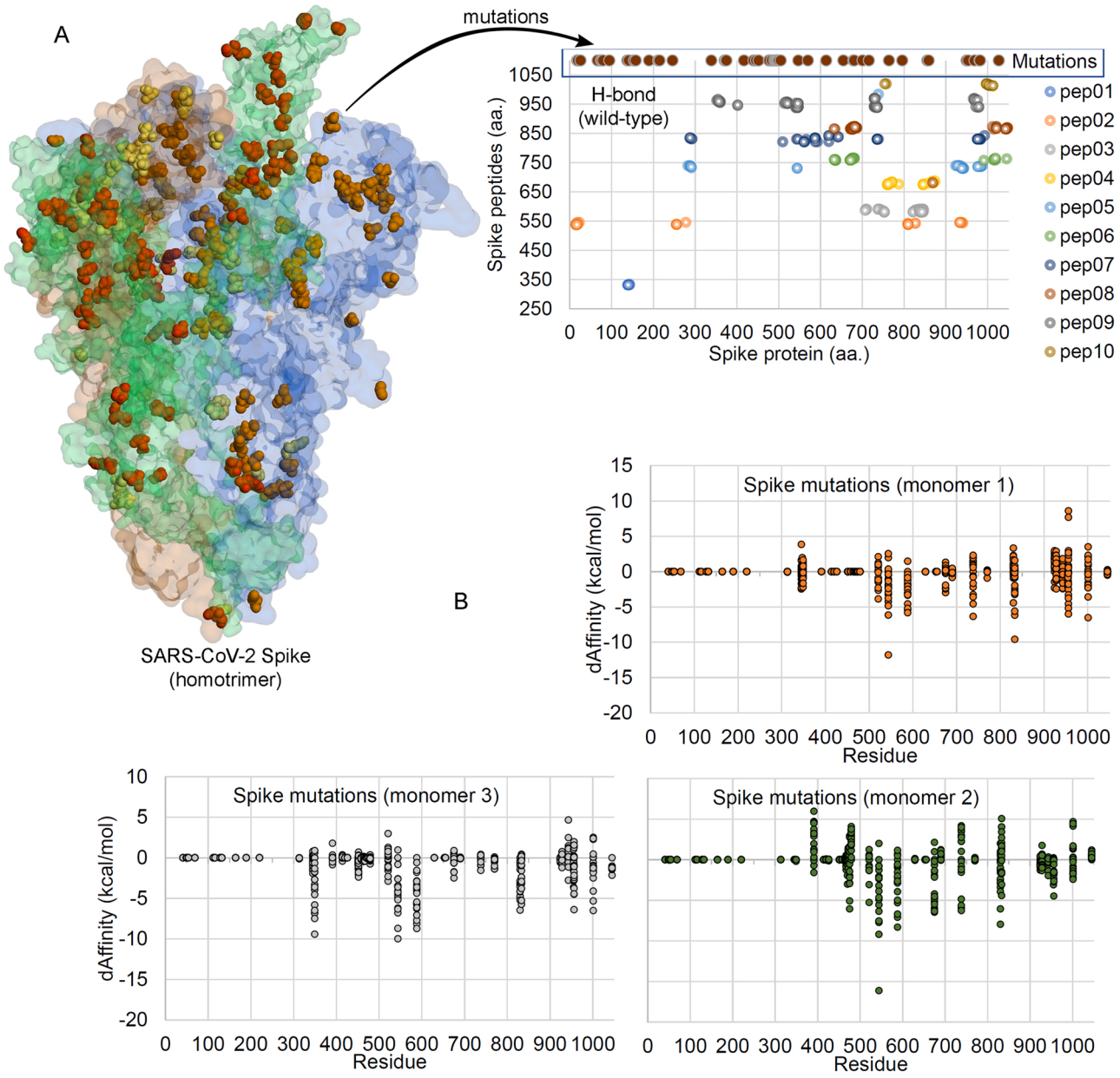


Fig. 7. A distinctive behavior of variants over the SARS-CoV-2 spike homotrimer unit. (A) Some of the infectious known variants of the SARS-CoV-2 virus: Alpha (B.1.1.7), Beta (B.1.351), Delta (B.1.617.2 & B.1.617), Gamma (P.1), Lamba (C.37), and Omicron (B.1.1.529) [2,3] represented over the S protein structure (mutated residues are represented as spheres in red / orange / yellow color). The right panel describes protein-peptide intermolecular hydrogen bonding residues identified for the wild type systems and compared with the mutations from SARS-CoV-2 variants. (B) Change in the binding affinity within the monomers of spike trimer, upon inserting point mutation in a respective S protein monomer. In addition, to the known variants of S protein (as shown in Fig. 1 and Table S1), these mutated amino acids were screened or mutated by 20 other amino acids (A, R, N, D, C, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y, and V).

(Fig. 7B), from which it can be postulated that by inserting such mutations (Table S4) could stabilize an "up" or "down" conformation of the functional spike protein homotrimer unit. A blocked monomer or stabilized homotrimer can induce production of antibody repertoire different from the "open" conformation.

4. Conclusions

Several structural based approaches have already been applied to develop vaccine targeting the SARS-CoV-2 virus, however, there lacks a detailed understanding in the direction to block the homotrimer formation of the S protein. Herein, we propose that disruption in homotrimer formation may block the viral entry into the host cell or reduce the kinetics of the S protein. We designed self-derived peptides from the S protein (monomer 1 of the homotrimer) and screened them against the second or third monomer of the homotrimer, with a goal that they could stabilize or block the trimer unit. Alongside, we investigated point mutations emerging from different SARS-CoV-2 virus variants with respect to change in the protein-peptide binding affinity. Molecular dynamics data revealed that among the hydrogen bond interacting residues, the R821(pep07)-E593(protein) residue pair formed high occupancy binding (91.02% in the wild type system). In addition, a larger set of S protein residues ranging from 550 to 750 and 950–1050 in the wild type systems were found interacting with different self-derived S peptides. Peptides pep02 and pep07 in the wild type systems were among high affinity binders with the S protein, and in particular, the point mutation in the pep02 peptide lowered the number of protein-peptide interactions. Considering different studied mutations in the

peptide dataset, the pep07 (N830K or T833N system) had a higher number of interactions between S peptide-protein, and the mutant A544D in the pep02 peptide have replaced A544(peptide)-N934 (protein) with D544-R18 / K938 residue pairs. Moreover, the A544D mutation from the pep02 peptide induced instability for the complexed S protein, whereas the point mutation N943K from pep09 had exhibited a contract behavior. A significant increase in the residue binding with S protein was observed in the mutant peptide systems, compared to that of the wild type system.

The dynamics of different protein-peptide complexes were investigated, suggesting the presence of mutation in the system has induced an "up" active conformation of the spike (RBD) domains that could enhance the interaction with the host cell receptor. Inserting mutation in the peptides destabilized the complexed S protein structure, which resulted in the allosteric effect over other different functional regions of the protein. Comparatively, two point mutations N830K and T833N in the pep07 peptide have distinctive effects over the structure, the N830K mutated system has conserved structures compared to the T833N system. The frequency of mutations in the protein-protein binding affinity changes within the trimeric S protein, suggest a significant changes in the region where protein-peptide forms a high number of H-bonds interactions. The A544R (pep02) mutation can have a significant impact on the binding affinity within the spike monomers of the homotrimer unit. We believe that our findings can allow better understanding of the SARS-CoV-2 spike homotrimer, and our data could help to design further experiments to develop self-derived peptides replacing a single S protein monomer from the trimer, blocking the homotrimer functional unit or inducing stability.

CRediT authorship contribution statement

Monikaben Padariya: Data curation, Formal analysis, Methodology, Visualization, Investigation, Project administration, Writing - original draft, Writing - review & editing. **Alison Daniels**: Writing - review & editing. **Christine Tait-Burkard**: Writing - review & editing. **Ted Hupp**: Conceptualization, Writing - review & editing. **Umesh Kalathiya**: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Visualization, Writing - original draft, Writing - review & editing.

Conflicts of interest statement

The authors declare no conflict of interest.

Acknowledgments

The International Centre for Cancer Vaccine Science (Fundacja na rzecz Nauki Polskiej: MAB/3/2017) project is carried out within the International Research Agendas programme of the Foundation for Polish Science co-financed by the European Union under the European Regional Development Fund. Authors would also like to thank the PL-Grid Infrastructure, Poland for providing their hardware and software resources.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.biopha.2022.113190.

References

[1] Coronaviridae Study Group of the International Committee on Taxonomy of Viruses, The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2, Nat. Microbiol. 5 (2020) 536–544, https://doi.org/10.1038/s41564-020-0695-z.

[2] X. Xiong, K. Qu, K.A. Ciazynska, M. Hosmillo, A.P. Carter, S. Ebrahimi, Z. Ke, S.H. W. Scheres, L. Bergamaschi, G.L. Grice, Y. Zhang, , CITIID-NIHR COVID-19 BioResource Collaboration, J.A. Nathan, S. Baker, L.C. James, H.E. Baxendale, I. Goodfellow, R. Doffinger, J.A.G. Briggs, A thermostable, closed SARS-CoV-2 spike protein trimer, Nat. Struct. Mol. Biol. 27 (2020) 934–941, https://doi.org/10.1038/s41594-020-0478-5.

[3] F. Li, Structure, function, and evolution of Coronavirus spike proteins, Annu. Rev. Virol. 3 (2016) 237–261, https://doi.org/10.1146/annurev-virology-110615-042301.

[4] Y. Watanabe, J.D. Allen, D. Wrapp, J.S. McLellan, M. Crispin, Site-specific glycan analysis of the SARS-CoV-2 spike, Science 369 (2020) 330–333, https://doi.org/10.1126/science.abb9983.

[5] Y. Cai, J. Zhang, T. Xiao, H. Peng, S.M. Sterling, R.M. Walsh Jr, S. Rawson, S. Rits-Volloch, B. Chen, Distinct conformational states of SARS-CoV-2 spike protein, Science 369 (2020) 1586–1592, https://doi.org/10.1126/science.abd4251.

[6] W. Yin, Y. Xu, P. Xu, X. Cao, C. Wu, C. Gu, X. He, X. Wang, S. Huang, Q. Yuan, K. Wu, W. Hu, Z. Huang, J. Liu, Z. Wang, F. Jia, K. Xia, P. Liu, X. Wang, B. Song, J. Zheng, H. Jiang, X. Cheng, Y. Jiang, S.-J. Deng, H.E. Xu, Structures of the Omicron spike trimer with ACE2 and an anti-Omicron antibody, Science 375 (2022) 1048–1053, https://doi.org/10.1126/science.abn8863.

[7] Y. Yuan, D. Cao, Y. Zhang, J. Ma, J. Qi, Q. Wang, G. Lu, Y. Wu, J. Yan, Y. Shi, X. Zhang, G.F. Gao, Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains, Nat. Commun. 8 (2017) 15092, https://doi.org/10.1038/ncomms15092.

[8] M. Gui, W. Song, H. Zhou, J. Xu, S. Chen, Y. Xiang, X. Wang, Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding, Cell Res. 27 (2017) 119–129, https://doi.org/10.1038/cr.2016.152.

[9] U. Kalathiya, M. Padariya, R. Fahraeus, S. Chakraborti, T.R. Hupp, Multivalent display of SARS-CoV-2 spike (RBD domain) of COVID-19 to nanomaterial, protein ferritin nanocages, Biomolecules 11 (2021) 297, https://doi.org/10.3390/biom11020297.

[10] U. Kalathiya, M. Padariya, M. Mayordomo, M. Lisowska, J. Nicholson, A. Singh, M. Baginski, R. Fahraeus, N. Carragher, K. Ball, J. Haas, A. Daniels, T.R. Hupp, J. A. Alfaro, Highly conserved homotrimer cavity formed by the SARS-CoV-2 spike glycoprotein: a novel binding site, J. Clin. Med. 9 (2020) 1473, https://doi.org/10.3390/jcm9051473.

[11] J. Juraszek, L. Rutten, S. Blokland, P. Bouchier, R. Voorzaat, T. Ritschel, M.J. G. Bakkers, L.L.R. Renault, J.P.M. Langedijk, Stabilizing the closed SARS-CoV-2 spike trimer, Nat. Commun. 12 (2021) 244, https://doi.org/10.1038/s41467-020-20321-x.

[12] A.C. Walls, M.A. Tortorici, B.-J. Bosch, B. Frenz, P.J.M. Rottier, F. DiMaio, F. A. Rey, D. Veesler, Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer, Nature 531 (2016) 114–117, https://doi.org/10.1038/nature16988.

[13] R.N. Kirchdoerfer, C.A. Cottrell, N. Wang, J. Pallesen, H.M. Yassine, H.L. Turner, K. S. Corbett, B.S. Graham, J.S. McLellan, A.B. Ward, Pre-fusion structure of a human coronavirus spike protein, Nature 531 (2016) 118–121, https://doi.org/10.1038/nature17200.

[14] A. Sternberg, C. Naujokat, Structural features of coronavirus SARS-CoV-2 spike protein: targets for vaccination, Life Sci. 257 (2020), 118056, https://doi.org/10.1016/j.lfs.2020.118056.

[15] M. Yuan, N.C. Wu, X. Zhu, C.-C.D. Lee, R.T.Y. So, H. Lv, C.K.P. Mok, I.A. Wilson, A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV, Science 368 (2020) 630–633, https://doi.org/10.1126/science.abb7269.

[16] P.J.M. Brouwer, T.G. Caniels, K. van der Straten, J.L. Snitselaar, Y. Aldon, S. Bangaru, J.L. Torres, N.M.A. Okba, M. Claireaux, G. Kerster, A.E.H. Bentlage, M. M. van Haaren, D. Guerra, J.A. Burger, E.E. Schermer, K.D. Verheul, N. van der Velde, A. van der Kooi, J. van Schooten, M.J. van Breemen, T.P.L. Bijl, K. Sliepen, A. Aartse, R. Derking, I. Bontjer, N.A. Kootstra, W.J. Wiersinga, G. Vidarsson, B. L. Haagmans, A.B. Ward, G.J. de Bree, R.W. Sanders, M.J. van Gils, Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability, Science 369 (2020) 643–650, https://doi.org/10.1126/science.abc5902.

[17] B.S. Graham, M.S.A. Gilman, J.S. McLellan, Structure-based vaccine antigen design, Annu. Rev. Med. 70 (2019) 91–104, https://doi.org/10.1146/annurev-med-121217-094234.

[18] Z. Ke, J. Oton, K. Qu, M. Cortese, V. Zila, L. McKeane, T. Nakane, J. Zivanov, C. J. Neufeldt, B. Cerikan, J.M. Lu, J. Peukes, X. Xiong, H.-G. Kräusslich, S.H. W. Scheres, R. Bartenschlager, J.A.G. Briggs, Structures and distributions of SARS-CoV-2 spike proteins on intact virions, Nature 588 (2020) 498–502, https://doi.org/10.1038/s41586-020-2665-2.

[19] A.C. Walls, Y.-J. Park, M.A. Tortorici, A. Wall, A.T. McGuire, D. Veesler, Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein, Cell 181 (2020) 281–292.e6, https://doi.org/10.1016/j.cell.2020.02.058.

[20] W.-H. Chen, P.J. Hotez, M.E. Bottazzi, Potential for developing a SARS-CoV receptor-binding domain (RBD) recombinant protein as a heterologous human vaccine against coronavirus infectious disease (COVID)-19, Hum. Vaccin. Immunother. 16 (2020) 1239–1242, https://doi.org/10.1080/21645515.2020.1740560.

[21] L. Liu, P. Wang, M.S. Nair, J. Yu, M. Rapp, Q. Wang, Y. Luo, J.F.-W. Chan, V. Sahi, A. Figueroa, X.V. Guo, G. Cerutti, J. Bimela, J. Gorman, T. Zhou, Z. Chen, K.-Y. Yuen, P.D. Kwong, J.G. Sodroski, M.T. Yin, Z. Sheng, Y. Huang, L. Shapiro, D. D. Ho, Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike, Nature 584 (2020) 450–456, https://doi.org/10.1038/s41586-020-2571-7.

[22] A.C. Walls, X. Xiong, Y.-J. Park, M.A. Tortorici, J. Snijder, J. Quispe, E. Cameroni, R. Gopal, M. Dai, A. Lanzavecchia, M. Zambon, F.A. Rey, D. Corti, D. Veesler, Unexpected receptor functional mimicry elucidates activation of Coronavirus fusion, Cell 176 (2019) 1026–1039.e15, https://doi.org/10.1016/j.cell.2018.12.028.

[23] A.C. Walls, M.A. Tortorici, J. Snijder, X. Xiong, B.-J. Bosch, F.A. Rey, D. Veesler, Tectonic conformational changes of a coronavirus spike glycoprotein promote membrane fusion, Proc. Natl. Acad. Sci. USA 114 (2017) 11157–11162, https://doi.org/10.1073/pnas.1708727114.

[24] Dataset retrieved from Acro Biosystems company, https://acrobiosystems.com.cn/ (accessed on 21/01/2022).

[25] L. Zhang, Z. Cui, Q. Li, B. Wang, Y. Yu, J. Wu, J. Nie, R. Ding, H. Wang, Y. Zhang, S. Liu, Z. Chen, Y. He, X. Su, W. Xu, W. Huang, Y. Wang, Ten emerging SARS-CoV-2 spike variants exhibit variable infectivity, animal tropism, and antibody neutralization, Commun. Biol. 4 (2021) 1196, https://doi.org/10.1038/s42003-021-02728-4.

[26] D. Wrapp, N. Wang, K.S. Corbett, J.A. Goldsmith, C.-L. Hsieh, O. Abiona, B. S. Graham, J.S. McLellan, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation, Science 367 (2020) 1260–1263, https://doi.org/10.1126/science.abb2507.

[27] P.W. Rose, B. Beran, C. Bi, W.F. Bluhm, D. Dimitropoulos, D.S. Goodsell, A. Prlic, M. Quesada, G.B. Quinn, J.D. Westbrook, J. Young, B. Yukich, C. Zardecki, H. M. Berman, P.E. Bourne, The RCSB Protein Data Bank: redesigned website and web services, Nucleic Acids Res. 39 (2011) D392–D401, https://doi.org/10.1093/nar/gkq1021.

[28] B.R. Brooks, C.L. Brooks III, A.D. Mackerell Jr, L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, M. Karplus, CHARMM: The biomolecular simulation program, J. Comput. Chem. 30 (2009) 1545–1614, https://doi.org/10.1002/jcc.21287.

[29] Molecular Operating Environment (MOE) 2011.10. Chemical Computing Group (2011) Montrea l, Quebec, Canada.

[30] P. Labute, The generalized Born/volume integral implicit solvent model: estimation of the free energy of hydration using London dispersion instead of atomic surface area, J. Comput. Chem. 29 (2008) 1693–1698, https://doi.org/10.1002/jcc.20933.

[31] D.B. Kitchen, H. Decornez, J.R. Furr, J. Bajorath, Docking and scoring in virtual screening for drug discovery: methods and applications, Nat. Rev. Drug Discov. 3 (2004) 935–949, https://doi.org/10.1038/nrd1549.

[32] P. Labute, LowModeMD–implicit low-mode velocity filtering applied to conformational search of macrocycles and protein loops, J. Chem. Inf. Model. 50 (2010) 792–800, https://doi.org/10.1021/ci900508k.

[33] A.D. MacKerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, J. Phys. Chem. B 102 (1998) 3586–3616, https://doi.org/10.1021/jp973084f.

[34] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, J. Hermans, Interaction models for water in relation to protein hydration, in: The Jerusalem Symposia on Quantum Chemistry and Biochemistry, Springer Netherlands, Dordrecht, 1981, pp. 331–342.

[35] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092, https://doi.org/10.1063/1.464397.

[36] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. 126 (2007), 014101, https://doi.org/10.1063/1.2408420.

[37] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: a new molecular dynamics method, J. Appl. Phys. 52 (1981) 7182–7190, https://doi.org/10.1063/1.328693.

[38] W.F. Van Gunsteren, H.J.C. Berendsen, A leap-frog algorithm for stochastic dynamics, Mol. Simul. 1 (1988) 173–185, https://doi.org/10.1080/08927028808080941.

[39] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38, https://doi.org/10.1016/0263-7855(96)00018-5, 27–8.