RESEARCH ARTICLE

Probing structural basis for enhanced binding of SARS-CoV-2 P.1 variant spike protein with the human ACE2 receptor

Surabhi Lata | Mohd. Akif

Laboratory of Structural Biology, Department of Biochemistry, School of Life Sciences, University of Hyderabad, Hyderabad, Telangana, India

Correspondence
Mohd. Akif, Laboratory of Structural Biology, Department of Biochemistry, School of Life Sciences, University of Hyderabad, Hyderabad, Telangana, India.
Email: akif@uohyd.ac.in

Abstract

The initial step of infection by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) involves the binding of receptor binding domain (RBD) of the spike protein to the angiotensin converting enzyme 2 (ACE2) receptor. Each successive wave of SARS-CoV-2 reports emergence of many new variants, which is associated with mutations in the RBD as well as other parts of the spike protein. These mutations are reported to have enhanced affinity towards the ACE2 receptor as well as are also crucial for the virus transmission. Many computational and experimental studies have demonstrated the effect of individual mutation on the RBD-ACE2 binding. However, the cumulative effect of mutations on the RBD and away from the RBD was not investigated in detail. We report here a comparative analysis on the structural communication and dynamics of the RBD and truncated S1 domain of spike protein in complex with the ACE2 receptor from SARS-CoV-2 wild type and its P.1 variant. Our integrative network and dynamics approaches highlighted a subtle conformational changes in the RBD as well as truncated S1 domain of spike protein at the protein contact level, responsible for the increased affinity with the ACE2 receptor. Moreover, our study also identified the commonalities and differences in the dynamics of the interactions between spike protein of SARS-CoV-2 wild type and its P.1 variant with the ACE2 receptor. Further, our investigation yielded an understanding towards identification of the unique RBD residues crucial for the interaction with the ACE2 host receptor. Overall, the study provides an insight for designing better therapeutics against the circulating P.1 variants as well as other future variants.

KEYWORDS
ACE2 receptor, COVID-19, molecular dynamic simulation, protein network analysis, SARS-CoV-2, spike protein, variant

1 | INTRODUCTION

The COVID-19 disease is caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) which emerged in late December 2019, but is still a global pandemic (World Health Organization [WHO] report, 2021). The SARS-CoV-2 belongs to the beta coronavirus genus which also includes SARS-CoV and Middle East respiratory syndrome coronavirus (MERS-CoV) which have caused a similar outbreak, but at a small transmission rate.[1-3] Coronaviruses get entry inside the host cell with the help of spike glycoprotein present on the outer envelope through a host cellular receptor, angiotensin converting enzyme 2 (ACE2).[4,5] Spike protein has been one of the important therapeutic targets for drug design and vaccine development. It is a homo-trimeric protein consisting of S1 and S2 subunits, the S1 subunit aids in the binding to the ACE2 receptor and the S2 subunit is involved in the fusion of virus to the host cellular membrane.[6] The structural organization of the S1 domain includes N-terminal domain (NTD), C-terminal domain (CTD), sub-domain 1 (SD1) and sub-domain 2 (SD2). The CTD consists of receptor binding domain (RBD) which is a main component that interacts with the ACE2 receptor. Previously, several studies have demonstrated that virus replication, infectivity, and transmission ability depend on the binding affinities of the RBD to the ACE2 receptor.[6,7] COVID-19 pandemic has been more challenging due to SARS-CoV-2 emerging variants. Since December 2020, many variants have been detected with a high number of mutations mainly in the spike protein.[8] The WHO has classified these variants either as variants of concern (VOC) or variants of interest, where three variants have gained more attention: B.1.1.7 (VOC 202012), B.1.351 (501Y.V2) and P.1 (B.1.1.28.1).[9-11] The B.1.1.7 was first detected in the United Kingdom in September 2020 and has been reported with 7 mutations in the Spike protein.[8,12] Within a month a South African variant was detected with 9 mutations.[8,11] Later on 2 January 2021 Japan/Brazil variant named P.1 (B.1.1.28.1) was in circulation with 12 mutations in the Spike protein.[8,13] The two mutations N501Y and D614G in spike protein were common in all the three VOC. Very recently, Omicron has also been identified as a VOC and reported to have almost 30 mutations in the spike protein.[14] A recent report on phylogenetic analysis of the RBD region of Omicron suggests that Omicron is closely related to the P.1 variant. Moreover, crystallographic structure and biophysical binding kinetics experiments suggest that the RBD of P.1 variant displays a greater affinity as compared to wild type and other circulating VOCs such as Delta and Omicron.[15] The functional role of many mutations associated with the VOCs have been

demonstrated using deep mutational scanning and other biochemical methods. Most of the mutations associated with the variants have been reported to have enhanced affinity towards the ACE2 receptor as well as are also crucial for the virus transmission.[16,17] These mutations are also demonstrated to be ineffective towards neutralizing antibodies and evading from neutralization.[18-20] The underlying cause of the enhanced binding affinity in the variants has not been associated with significant change in three dimensional structures of spike as well as ACE2 receptor. However, local and global rearrangement of the interaction sites and allosteric sites cannot be overlooked. Mutations away from the RBD facilitate spike protein in the open state that binds efficiently with the ACE2 receptor.[21] Most of the previous studies focus on the effect of individual mutation. However, the cumulative effect of the mutations present on the RBD and away from the RBD such as NTD, SD1 and SD2 of the S1 domain was not investigated in detail. It is assumed that the cumulative effect of the mutations may provide a significant rearrangement at protein contacts level that may have an effect on increased binding affinity with the ACE2 receptor.

Protein contact network (PCN) and Molecular dynamic simulation (MDS) have been very useful approaches for analyzing nominal conformational changes crucial for large conformational effects on the function and inhibitor binding of proteins.[22] These methods have already been used in various studies such as understanding the dynamics of mutant proteins,[23,24] structural flexibility[25] structure-function relationship[26] identifying crucial functional residues in proteins.[27] Recently, the PCN approach has also been applied to study a structural comparison of SARS-CoV-1 and SARS-CoV-2 M^{pro} in apo and inhibitor bound states.[28] These approaches have been important for analyzing evolution as well conformational dynamics of SARS-CoV-2 spike protein[29] and prediction of its binding affinity with ACE2 receptor.[30,31] Recently, the PCN along with perturbation scanning based elastic network methods provided evidence of the allosteric sites and their involvement in the regulation of functional dynamics of the spike protein.[21,32-34] There have been many integrative computational efforts towards better understanding of SARS-CoV-2 and its variants interaction with the ACE2.[35-39] Moreover, efforts are also being put into designing inhibitors focusing on the RBD domain of spike protein.[36,37,40] However, the mutations associated in the spike protein render the current therapeutics a challenge to combat SARS-CoV-2 and possibly be ineffective to other similar future pandemics. It is therefore really warranted to have an insight into the atomic details and effect of these mutations on the structural

communication and dynamics of the RBD and truncated S1 domain of the spike protein in context to its interaction with the ACE2 receptor.

To probe the effect of all mutations present on the RBD as well as mutations from other part of the truncated S1 domain, we report, here, a comparative analysis on the structural communication and dynamics of the RBD and truncated S1 domain of spike proteins from the SARS-CoV-2 wild type and P.1 variant. However, we emphasize a limitation of our study. We have not considered the full length of spike protein. Since our aim was to understand the cumulative effect of mutations on the RBD as well as truncated S1 domain, assuming that this will not affect our comparison with the available experimental data for the P.1 variants, as the experimental studies were also performed mostly with the same domains.[18] We applied an integrative network and dynamics approaches to investigate the subtle conformational changes in the RBD as well as the truncated S1 domain of the P.1 variant spike protein responsible for the increased affinity with the ACE2 receptor. Our study on the protein contact network identifies changes in the protein structures at the residue contact level which are otherwise not easily detectable. Moreover, our study also highlights the identification of commonalities and differences in the dynamics of the interactions between spike proteins of SARS-CoV-2 and its P.1 variant with the ACE2 receptor. Together with our structural communication analysis and dynamics study provide an understanding towards interaction with the ACE2 receptor. Overall, this may be utilized for designing better therapeutics against the circulating P.1 variants as well as other future variants.

2 | METHODS

2.1 | Construction of protein structure graph (PSG)

The PCN identifies the change in protein structures at contact level which are otherwise not easily detectable.[23,28] The 3D coordinates of the RBD-ACE2 complexes spike protein from SARS-CoV-2 wild type [6MOJ[7]] and its P.1 variants [7NXC[18]] were extracted from the PDB. The PSG was constructed on the complex structures using the PSN and ENM-NMA approaches implemented in the WebPSN.[41] Each amino acid is depicted as a node and each node is connected to other nodes in a protein structure through an edge. The interaction strength between two connecting nodes is defined as

$$Iij = \frac{nij}{\sqrt{NiNj}}100$$

Where (Iij) interaction percentage of nodes i and j. It follows the number of side chain atoms pairs within (4.5 Å) cutoff, Ni and Nj are normalization factors.[42,43] It constructs a PSG on the basis of the atomic cross correlation motions using the ENM-NMA.

2.2 | Network parameters

Three important network parameters such as Hubs, Modularity and Structural communication pathways were calculated for the RBD-ACE2 and S1-ACE2 complexes from the wild type as well as P.1 variant. PYMOL was used to visualize these parameters. Briefly, **hubs** are nodes with highest degree. **Modularity** is represented as **c**ommunities with more interconnected nodes and the nodes of same community are highly connected to each other than the poorly connected nodes of outside the community. **Shortest Path** is the smallest number of links required to travel from one node to other. It is calculated on the basis of the Dijkstra's algorithm.

2.3 | Protein stability and flexibility analysis

To analyze the effect of all mutations present on the RBD domain and outside the RBD (truncated S1 domain) of the spike protein, free energy changes (ΔΔG) and vibrational entropy difference (ΔΔS) were calculated using the DYNAMUT tool.[44] Based on these calculations, protein stability and flexibility were analyzed on the complexes by utilizing the NMA based elastic network contact model (ENCoM).

2.4 | Perturbation residue scanning (PRS)

PRS was performed using the pPerturb server.[45] It allows the mutation of one or more residues to alanine and generates a perturbation profile (ΔQ vs. Calpha-Calpha) distance from the perturb site. The perturbation effect can be analyzed as a distance connecting the perturb residue to its nearby residues or on the interaction network strength.

2.5 | Molecular dynamics simulation

To study the dynamic behavior of the complexes (RBD-ACE2 and truncated S1-ACE2) from the wild type and P.1 variant with respect to time, MDS was performed for the time scale of 100 ns. To ensure the convergence of simulations, we performed three independent 100 ns production runs for the wild type and P.1 variant RBD-ACE2 complex. The co-ordinates of the complexes were obtained from the PDB and used as starting models for the MDS using the GROMACS package version 5.1.4.[46] The crystal waters were removed from the co-ordinates. For setting the periodic boundary conditions the system was soaked in a cubic box with a dimension of 1.5 nm, further the box was filled with SPC water molecules and OPLSAA force field was applied.[47] For neutralization the appropriate water molecules were replaced by counter ions and the system details are described in the Table S5. To remove the steric clash the systems were subjected to energy minimization with the steepest descent method for 50 000 steps until the largest force was smaller than 1000 kj/mol/nm. Then the minimized system was equilibrated in the NVT and NPT ensemble for 250 ps each. Last the production run was performed for 100 ns at 300 K where the leapfrog integrator was used for time evolution trajectories. A constant temperature on the system was maintained at 300 K using modified Berendson Thermostat[48] and Parinello-Rahman barostat[49] pressure was maintained at 1 bar during the simulations. The analysis of the trajectory files were done with gmx rmsd, gmx rmsf, gmx gyrate, gmx sasa, gmx hbond, gmx covar, gmx anaeig, gmx do_dssp for the root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), solvent accessible surface area (SASA), hydrogen bond (H-bond), principal component and secondary structures, respectively. Graphs were generated using pymol, XMGRACE and GNUPLOT.

2.6 | Principal component analysis (PCA)

To analyze dominant and collective motions in the RBD-ACE2 and S1-ACE complexes from the wild type and P.1 variant, PCA was performed. Mass weighted covariance matrix on the coordinates of the MD trajectories was calculated. Diagonalization of the covariance matrix was used to get a set of eigenvectors and eigenvalues from the G_covar function implemented in the GROMACS. The g_anaeig function was utilized to obtain trajectories on the eigenvectors.[50] Backbone atoms were considered for the PCA analysis. The coordinates of the residues were projected on the first two Principal Components that represent the highest eigenvalues and captured the overall motion of the protein.

2.7 | Binding free energy calculations

Molecular mechanism-Poisson-Boltzmann surface analysis (MMPBSA) method was used to estimate the binding free energy between the RBD-ACE2 complexes from the wild type and P.1 variant. Last 10 ns time frames of MD simulations were integrated to calculate the free energy using the g_mmpbsa tool of GROMACS.[51] The molecular mechanics energy includes electrostatic, Van-der-Waals interactions and polar and nonpolar solvation energy.

2.8 | Normal mode analysis (NMA)

NMA was performed on the stable simulated trajectories coordinates by using the I-MODS tool.[52] The NMA calculated the lowest frequency modes and identified atoms moving together as a rigid body. The internal modes were estimated on the basis of the Lagrangian equation of motions.

3 | RESULTS AND DISCUSSION

Due to emerging variants of the COVID-19, the current pandemic seems to be quite alarming with continuously rising waves. The emerging variants are associated with many mutations mostly accumulated in the spike protein of SARS-CoV-2. These mutations have been functionally mapped for the increased affinity toward the ACE2 receptor. The increased affinity attributes variants to be more infectious and have a high transmission. This renders the current therapeutics a challenge to combat SARS-CoV-2. Therefore, it is warranted to have a better understanding of the structural basis of interaction in variant of spike protein with the ACE2 receptor. In our study, we have chosen the P.1 variant of SARS-CoV-2 that has almost ten mutations in the spike protein. In spite of having three mutations on the binding interface of the RBD, the calculated cross structure RMSD between the complex of the RBD-ACE2 from the wild type and P.1 variant was 0.29 Å which is almost negligible (Figure S1), suggesting no overall structural changes among the two. Still it is a matter of concern. Hence, to investigate further, we considered the PCN and dynamics of the complexes, assuming that the changes at contact level may cause the significant variations noted for the increased affinity with the ACE2 receptor.[29,34] Since, the overall structural changes are negligible hence, we

performed a protein structure network based approach to examine the small variations spanned throughout the protein.

3.1 | Protein structure network analysis of wild type and P.1 variant protein structures

The PSN construction calculated three important network parameters such as hubs, modularity and shortest communication pathway and yielded a global average network parameter values for the wild type and P.1 variant (Table S1). These values were quite comparable suggesting diminutive changes among the two structures. This also supports the observation of negligible cross structure RMSD deviation between the wild type and P.1 variant complexes. Further, for the better understanding of the effect of mutations on the subtle conformational variations, residue specific local variations among the network parameters were calculated for the wild type and P.1 variant.

3.1.1 | Hubs

The calculated hubs were mapped on the RBD-ACE2 complex from the wild type and P.1 variant. A total number of 172 and 86 hubs were observed in wild type and P.1 variant complex, respectively. This indicates a significant change in the hub formation among the two. Moreover, a reduction of hub residues in the P.1 variant complex observed to be localized in the RBD region. A significant loss of hub residues was also observed at the interface region of RBD-ACE2 complex from the P.1 variant and these regions are highlighted in the circle as shown in the Figure 1. The interacting residues from this region, Y449, Y453, L455, Q493, Y495, F497, Q498, N501, Y505, were found to be involved in hubs formation in the wild type, but were absent in the P.1 variant. Additionally, a difference in the hub formation was also observed on the centrally located antiparallel β-sheets (Figure 1). This difference suggests that three interface mutations (K417T, E484K, and N501Y) associated in this region of the P.1 variant may have significantly perturbed this hub zones. All the residues involved in the hub formation among the RBD from wild type and P.1 variant are listed in (Table S2). Few unique hubs (Y41, K353, and R357) were also observed at the interface of the ACE2 from the P.1 variant complex (Table S2). The observation of the variation among the hub residues from wild type and P.1 variant suggests differences in the structural connectivity. This may be the consequence of loss of interactions among the domain which may account for the flexibility and better conformational adaptability in case of the P.1 variant. Structural dynamics/variability in the spike protein in context to open and close state has been investigated by analyzing unique hubs among the two state.[53]

3.1.2 | Modularity

The community structure analysis results showed a total 22 communities in the wild type and 20 communities in the P.1 variant. The residue-wise communities in the wild type and P.1 variant are shown in Figure 2. The largest community (C1) was observed in the ACE2 from

(A) Wild type   (B) P.1 variant

ACE2

RBD
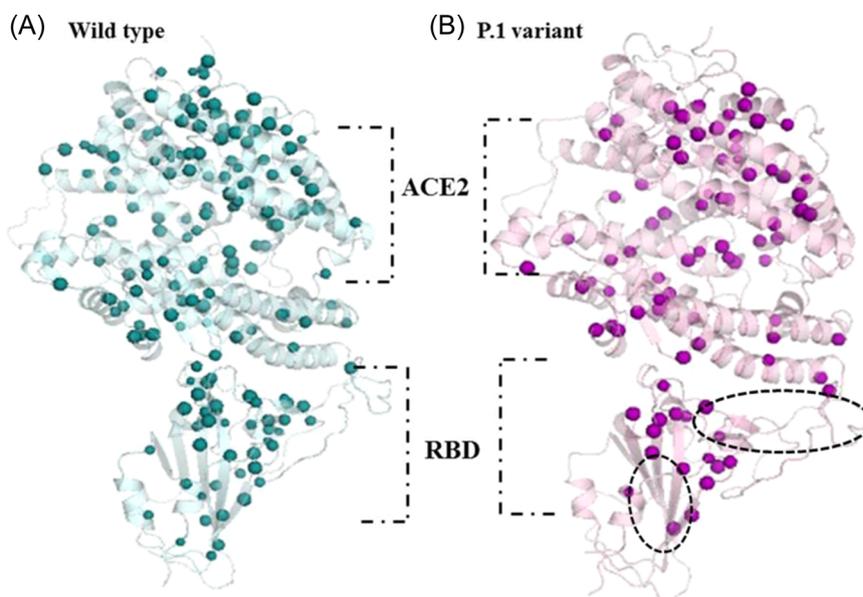
FIGURE 1 Representation of hubs arrangement on RBD-ACE2 complexes (A) wild type and (B) P.1 variant. Hubs are represented as cyan and purple spheres in wild type and P.1 variant, respectively. Loss of hubs from the interface RBM and beta-sheet core of the RBD of P.1 variant are encircled. ACE2, angiotensin converting enzyme 2; RBD, receptor binding domain
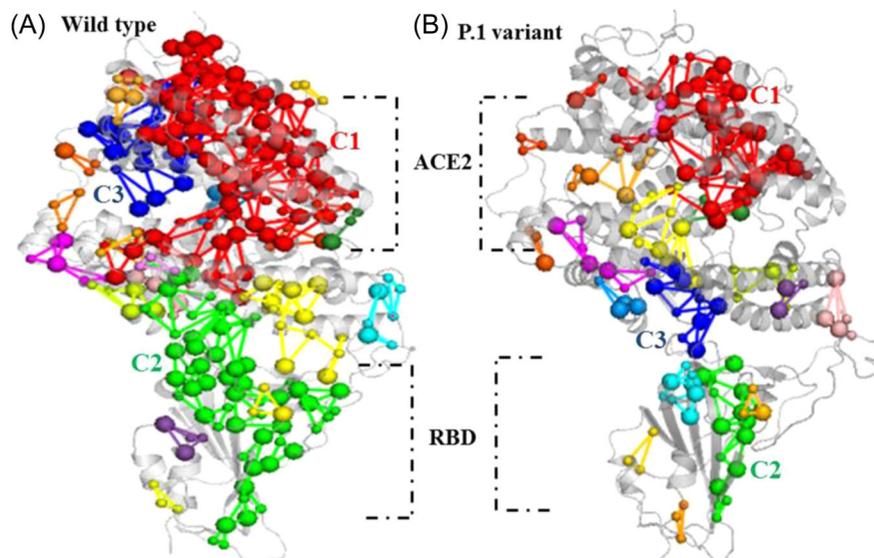
FIGURE 2 Comparative community structures in RBD-ACE2 complexes. (A) Wild type (B) P.1 variant. Communities are represented in different colors where first three largest communities (C1, C2, C3 are in red, green, blue colors), respectively. Nodes and links involved in these communities were reduced in the P.1 variant. ACE2, angiotensin converting enzyme 2; RBD, receptor binding domain

the both complexes. Interestingly, the total number of nodes and links involved in this community formation were observed to be reduced in the ACE2 of the P.1 variant complex. This community consisted of 114 nodes, 192 links and 74 hubs in the wild type complex, while the same was reduced to 54 nodes, 82 links and 30 hubs in the P.1 variant complex. The second largest community (C2) was observed in the RBD domain and the same was extended throughout the interface of the ACE2 in the wild type complex. In contrast, this region in the P.1 variant was divided into two different communities (Figure 2). Overall this change in community arrangement reflects subtle conformational changes due to mutations which were not evident at the structural level. Moreover, the rearrangements in modules were spanned throughout the structure which reflects the perturbation at global level.

3.1.3 | Shortest communication pathways

The total length of the shortest path in wild type was 1 210 620 while a total of 523 853 paths were observed in the communication channel of the P.1 variant. This observation indicated a decrease in the pathways of the RBD-ACE2 complex in the P.1 variant. Moreover, the average path hub% was also significantly decreased in the P.1 variant (Table 1). Major difference was observed in the interface residues nearby mutation N501Y and the N-terminal region of the RBD. The interface residues such as E497, Q498, P499, and T500 from the wild type were observed as nodes in the shortest path. In contrary, these residues were not observed to be involved in the shortest pathway in the P.1 variant (Figure S2). Possibly the N501Y mutation from the P.1 variant may have

TABLE 1 Network components and parameters involved in the average shortest paths in the RBD-ACE2 complexes from the wild type and P.1 variant

Path summary	Wild type	P.1 Variant
No. of nodes in path	69	75
No. of links in path	68	74
No. of shortest path	1 210 620	523 853
Average path length	20.7	24.19
Average path hub %	61.03	41.12

Abbreviations: ACE2, angiotensin converting enzyme 2; RBD, receptor binding domain.

perturbed this channel. Another prominent change was observed in the N-terminal region (Res. C336, F338, F374) of the wild type RBD were involved on paths but were absent in the P.1 variant (Figure S2). These changes clearly suggests a significant perturbation in inter and intra communication within the complex due to the mutations.

3.2 | Perturbation in PSG due to mutation outside the RBD domain

Mutations present outside the RBD of spike protein may also confer an effect on the binding affinity with the ACE2 receptor. Hence, to analyze the effect of these mutations on the structural flexibility, the PSG for the truncated S1 domain-ACE2 complex from the wild type and P.1 variant was generated. Except few differences, the hub arrangement was almost similar in both the complexes (Figure S3). Five out of seven mutations in the NTD of the S1 domain found to have altered hubs

arrangement compared to wild type. Interestingly, few interface hubs were observed in the same complex of the P.1 variant, which were absent in the PCN analysis of the RBD-ACE2 complex. Moreover, the N501Y mutation in P.1 variant was also observed to engage in hub formation. The altered hub in the P.1 variant may have some role in providing a better efficiency towards the ACE2 receptor. In fact, the functional role of N501Y towards the binding affinity with the ACE2 receptor has already been demonstrated.[18]

A total of 39 communities were observed in the truncated S1 domain of wild type while the corresponding domain in the P.1 variant showed up only 41 communities. In both the complexes, the ACE2 region was observed to form the largest community and RBD region contained a second largest community (Figure S4). A significant perturbation was observed in the community arrangement of the NTD domain of P.1 variant complex suggesting an extensive reorganization in node's membership as compared to the wild type. Alteration of nodes and links was observed in the NTD of the P.1 variant compared to the wild type complex. Moreover, residues participating in the shortest path of the NTD from both the cases were also different. The pathways from this region were observed to be slightly shifted in the P.1 variant (Figure S5). A total number of nodes and links in the communication pathway from the P.1 variant were more compared to the wild type. The total numbers of the shortest pathways along with the average path length were also slightly high in the P.1 variant (Table S3). These observations form this region clearly indicate that the mutations may have perturbed the core of the NTD communication in the P.1 variant.

3.3 | Effect of mutations on the thermodynamic of protein

To analyze the effect of mutations on protein structures, change in vibrational entropy ($\Delta\Delta S$) and change in free energy ($\Delta\Delta G$) were calculated. All three interface mutations (E484K, K417T and N501Y) of the RBD provided a destabilizing effect on the protein and found to have negative free energy values (Table S4). Moreover, change of vibrational entropy value for the K417T mutant was observed as $0.514\,\mathrm{kcal\,mol^{-1}\,K^{-1}}$ which was comparatively higher than the E484K and N501Y mutations (Table 2). The K417T mutation attributed towards more flexibility as compared to the other two mutations and its effect was extended to the other part of the protein as well. Though, the K417T is present near 3/10 α-helix, but observed to impart flexibility to the loop region of RBM as well as N-terminal region residues

(Figure 3A). In contrast to the K417T mutant, the effect of the E484K and N501Y mutations were localized at the nearby residues only (Figure 3B,C). These vibrational entropy ($\Delta\Delta S$) observation for each RBD interface mutation showed destabilization effect on the RBD, which perfectly corroborates with an in vitro study.[54] Further, the vibrational entropy change was also observed for the mutations localized outside the RBD domain (Figure S6). The effect of the R190S mutation from the NTD was significant and observed to make the complete domain flexible. Slightly less flexibility was observed for the T20N mutation. Interestingly, a major effect was noticed for the D614G mutation that provided a flexibility to the RBD as well as the NTD region of S1 domain, suggesting that the D614G mutation is associated with a global flexibility in the protein. These changes may provide a conformational plasticity to the RBD domain and other parts of the protein enabling efficient binding with the ACE2 receptor. Similar observation was reported in many previous studies signifying that D614G mutation has increased the binding affinity for ACE2 receptor.[55,56]

Additionally, to probe the effect of the interface mutations on the allosteric residue, PRS was analyzed. Similar to the entropy change analysis, the PRS analysis of the K417T mutation showed maximum effect on the allosteric residues (Figure S7). This observation suggests that mutation at one part of the protein may provide an effect on the other part to make it more adaptable. The flexibility associated with the mutations in the ACE2 complexes was further investigated by MDS at 100 ns time-frames. The cumulative local as well as allosteric effects of the mutations on the conformational dynamics behavior of the interfacial binding residues of RBD-ACE2 complexes were mapped.

3.4 | Probing of conformational dynamics of RBD-ACE2 complex

The calculated average backbone RMSD values for each three replicates of the RBD-ACE2 complexes from the wild type and P.1 variant were observed as 0.3 and

TABLE 2 Predicted change in entropy values associated with the interface mutations on the RBD

Mutations	$\Delta\Delta S$ (kcal mol^{-1} K^{-1})
K417T	0.514
E484K	0.179
N501Y	0.145

Abbreviation: RBD, receptor binding domain.
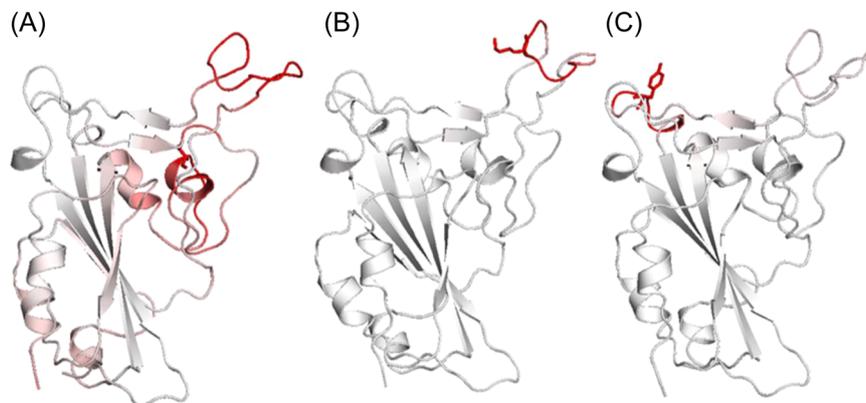
FIGURE 3 Visual representation of the interface mutations on dynamicity and plasticity of the RBD. (A) RBD structure with the K417T mutation (B) RBD structure with the E484K mutation (C) RBD structure with the N501Y mutation. The RBD is represented as cartoon structure and mutations are shown as stick model. Red color in the cartoon structure indicates the flexibility in the protein. RBD, receptor binding domain
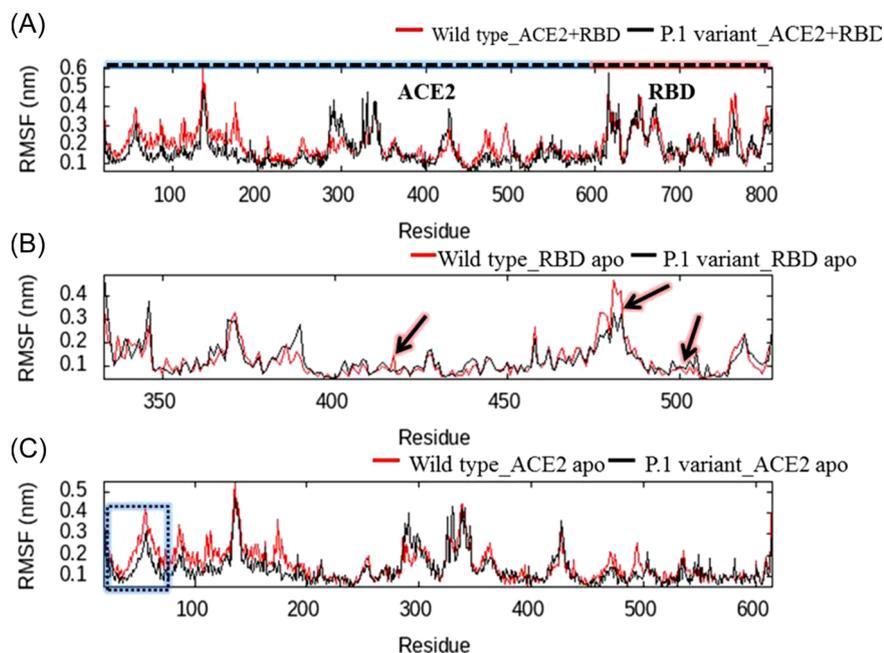
FIGURE 4 RMSF trajectories generated for the wild type and P.1 variant. (A) RBD-ACE2 complex. Region of ACE2 and RBD from the complex is highlighted and labeled at the top of the graph. (B) Apo RBD. Arrows indicate three mutations of the RBD interface (C) Apo ACE2. Interface interacting residues of the ACE2 are depicted in the dotted rectangle. ACE2, angiotensin converting enzyme 2; RBD, receptor binding domain; RMSF, root mean square fluctuation

0.35 nm, respectively. The RMSD values clearly indicate that there is no significant change in the trajectory of both the complexes and convergence point was reached after 50 ns both for wild type and P.1 variant complexes (Figure S8). In addition, we also estimated the cosine content of the first two eigenvectors derived from PCA of the individual trajectory. The cosine similarity of the first few eigenvectors is a good indicator to assess whether the conformational sampling is converged or not.[57] For each replicas of wild type and P.1 variant simulation systems, the eigenvector cosine contents were found to be less than 1. In fact in all cases the cosine contents fall between 0.2 and 0.6 values (Table S6). This value indicates that the simulation are converge for both the

systems. Similarly, the RMSF profile of the RBD-ACE2 complexes from the wild type and P.1 variant did not show any significant change throughout the simulation time (Figure 4A). Except few differences, the trajectory of both the complexes was observed quite overlapping to each other with a maximum fluctuation of less than 0.45 nm. The differences in the RMSF were observed when only RBD (apo-RBD) or only ACE2 (apo-ACE2) was considered for the analysis. For apo-RBD, the RMSF values of the regions—Res380-388 from the P.1 variant and Res470–480 from the wild type, were slightly high (Figure 4B). Additionally, the fluctuations of the mutated residues at the interface of the apo-RBD from the P.1 variant were observed to be reduced. The region,

Res19–200, of the ACE2 from the wild type complex displayed comparatively higher fluctuation (Figure 4C). Expectedly the interface interacting residues from apo-ACE2 were less fluctuating in the P.1 variant complex suggesting a rigidity in the residues upon binding with the RBD of the P.1 variant and results in better binding.

The SASA analysis of both the complexes indicated a slight difference in the overall values (Figure 5A). In fact, the SASA value for the RBD-ACE2 complex of the P.1 variant was observed to be higher than the same complex from the wild type. The probability distribution function of the SASA also showed a slightly different distribution pattern for the two (Figure 5B). The residues which are buried in the hydrophobic core of the protein are the driving force in the protein folding and the slight change of SASA suggests a conformational change in the protein during the course of 100 ns of simulation.

Inter/intramolecular hydrogen bond interactions play an important role for the stability and integrity of protein. Compared to the RBD-ACE2 complex from the P.1 variant, the same complex from the wild type possessed a significantly higher number of hydrogen bonds (H-bonds) (Figure 6A,B). The average number of inter H-bonds for the wild type and P.1 variant were 13 and 9, respectively. Similarly, the average number of intra H-bonds in the RBD-ACE2 complex of the P.1 variant was 113 while the same for the wild type was 118.4. The decrease in total H-bonds in the RBD-ACE2 complex of the P.1 variant indicates comparatively lesser stability which suggests that this complex may have slightly higher overall structural flexibility that may be required for the better adaptability. Further, a change in the Rg was observed among the two indicating a variation in the compactness also. The compactness

during 0–20 ns for the wild type and P.1 variant were almost same and the average value of complexes was observed approx. 3.15 Å (Figure 6C). Later, the structural compactness of the RBD-ACE2 complex from the P.1 was lost. Its average value was increased to 3.25 Å and maintained till the end of simulations. In contrast, the RBD-ACE2 complex from the wild type was observed to have higher average value during the initial frame of the simulation (0–40 ns) and later it decreased. This was maintained till the end of the 100 ns simulation. This suggests that the wild type complex is comparatively stable and the loss of compactness of the P.1 variant complex may be associated with more structural flexibility and plasticity.

To probe dominant motions and conformational changes among the RBD-ACE2 complexes from wild type and P.1 variant, the principal components were extracted from the MD simulation trajectories. Covariance analysis of the atomic correlation plot indicates that the RBD-ACE2 complex of the wild type mostly possessed dominant anticorrelation motion. Whereas the same complex from the P.1 variant was associated with uncorrelated motion (Figure S9A,B). This difference may be accounted for due to the loss of few contacts and slight conformational variations associated in the P.1 variant. This observation is in correlation with our PCN and MDS analysis. Further, nature of motions of the RBD from the wild type and P.1 variant was analyzed through the NMA. The constructed porcupine plots extracted from the stable trajectories indicate an opposite motion for the RBD from the complex from wild type and P.1 variant. Interestingly, the RBM domain and N-terminal region of the RBD from the wild type appeared to move in the same direction but the same
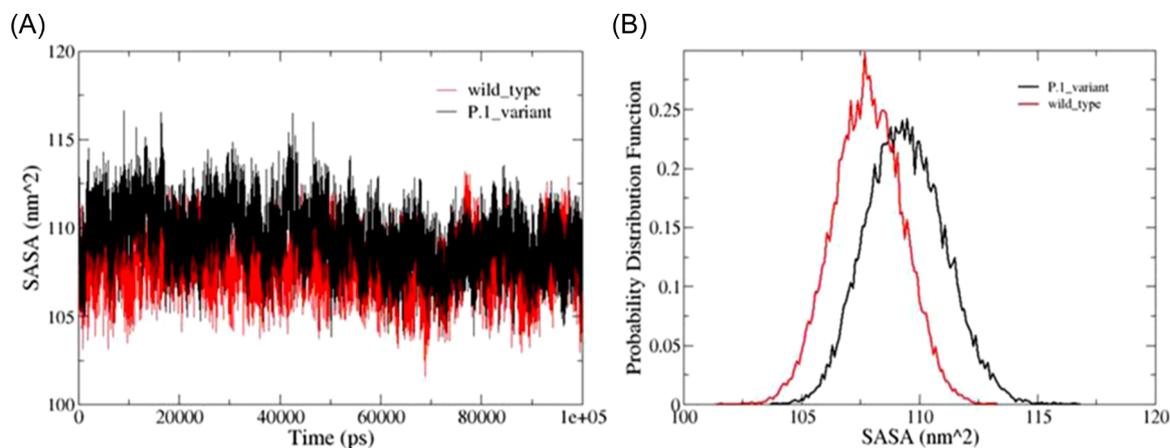
FIGURE 5 Graphical representation of the conformational dynamics. (A) SASA values against time for wild type and P.1 variant RBD-ACE2 complex. (B) Probability distribution function of SASA values. ACE2, angiotensin converting enzyme 2; RBD, receptor binding domain; SASA, solvent accessible surface area
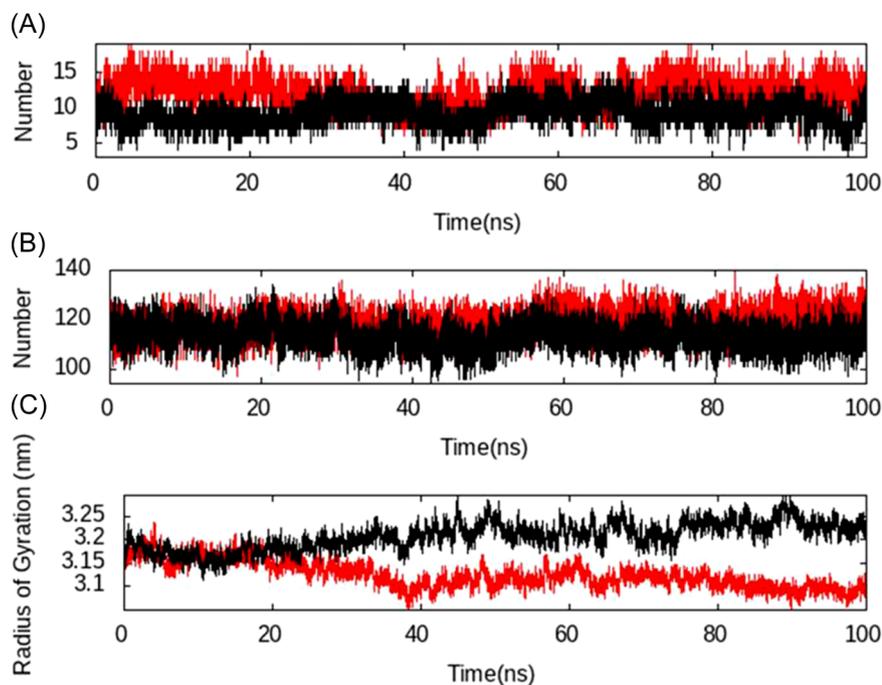
FIGURE 6 Plots representing strength and compactness in the RBD-ACE2 complexes. (A) Number of Inter hydrogen bonds vs time between RBD and ACE2. (B) Number of Intra hydrogen bonds vs time within the RBD. (C). Radius of gyration of RBD-ACE2 complex. Wild type (red plot) and P.1 variant (black plot). ACE2, angiotensin converting enzyme 2; RBD, receptor binding domain

regions in P.1 variant was observed to have motion opposite to each other (Figure S10A,B).

To estimate the binding affinities between RBD and ACE2 complex from the wild type and P.1 variant, free energy of association for the complex was estimated using the MMPBSA method. The MMPBSA calculation clearly indicates that the complex from P.1 variant possessed a higher negative free energy value of −1804 kJ/mol compared to the complex from wild type which was estimated as −1325 kJ/mol (Table S7). Comparatively better binding energy for the P.1 variant suggests a better affinity towards ACE2. This observation is in accordance with the experimental data reported about the P.1 variant.[15,18] The better binding is mostly contributed by the Van-der-Waals, the SASA and the electrostatic interactions. All these interactions provide a negative binding energy towards the total free energy. However, the polar solvation energy was positively contributing toward the total binding free energy (Table S5). Additionally, the number and type of interactions from the interface of the RBD-ACE2 complexes of the wild type and P.1 variant during 70–100 ns time frames were plotted. The plot clearly indicate that the hydrophobic interactions were significantly high in P.1 variant as compare to wild type. In contrast, the interface H-bonds in the complex from P.1 variant were less (Figure S11).

A clustering analysis (GROMOS method with cut-off 0.2 nm) was performed to extract a representative structure from 100 ns time frame of MDS. The extracted structures from wild type and P.1 variant were compared for the interface interactions. The rearrangement of the hydrogen bonds were observed in interface of the complexes. Interestingly, a part of common hydrogen bonds, few unique hydrogen bonds were appeared in the P.1 variant. Few residues from RBD of P.1 variant such as Phe490, Gln493 and Leu492 were observed to form hydrogen bonds with Lys31 of ACE2. Moreover, Asn487 from P.1 variant displayed a hydrogen bond interaction with Gln24 of ACE2. In addition, another unique interactions were observed between Asn487, Tyr495 of RBD and Tyr83 and Lys353 of ACE2, respectively (Figure 7). Involvement of these unique interactions in the P.1 variant may provide an energetically favorable binding with the ACE2. Hence, these interactions may be utilized to target with the specific therapeutics to disturb the RBD-ACE2 binding in the variant.

3.5 | Probing the conformational variations in the truncated S1 domain

To study the conformational variations attributed with the mutations present outside the RBD, 100 ns simulation was performed for the truncated S1 domain-ACE2 complexes from the wild type and P.1 variant. Residue-wise RMSD calculation quantified the flexibility due to mutations and its cumulative effect was analyzed. For both wild type and P.1 variant, the NTD of truncated S1 domain represented a high RMSD but the extent of deviation was relatively slightly more in the P.1 variant (Figure 8A). Interestingly we observed that the anchoring
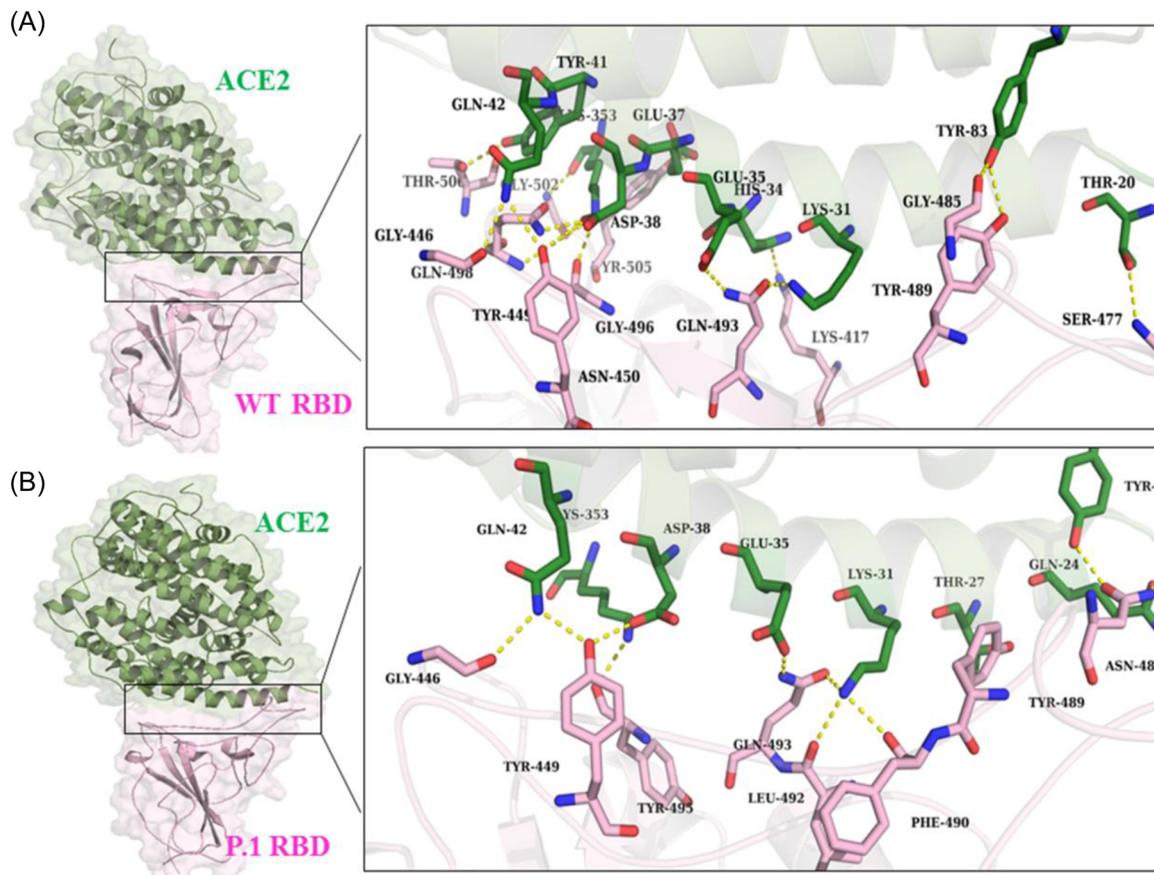
FIGURE 7 Representation of interface interacting residues involved in hydrogen bond interaction in the representative structure of wild type and P.1 variant complex obtained after clustering analysis. RBD and ACE2 are represented as cartoon structure and surface diagram with color hot-pink and smudge green, respectively. The interacting region is denoted in a rectangular box. Interacting residues are represented in stick model with the respective color. Hydrogen bonds are shown in yellow dashed lines. ACE2, angiotensin converting enzyme 2; RBD, receptor binding domain

RBM loop (Res 437–508) region showed comparatively high flexibility in the P.1 variant, possibly because of the interface mutations (K417T and E484K, N501Y) present on the loop. In contrast, this loop was observed to be slightly rigid in the wild type. Moreover, D614G mutation on the SD2 region of the P.1 variant also incorporates a significant high RMSD (Figure 8A). These observations suggest that mutations outside the RBD have a role in the conformational heterogeneity and may provide a conformational plasticity for the efficient binding with the ACE2 receptor. Similar studies have been investigated to analyse conformational variation in RBD as well as truncated S1 domain in complex with ACE2. Ray et al.[21] stated the effect of D614G on the conformational variation facilitating the RBD opening process from down to up state. Conformational flexibility has also been analyzed within three important loops of the RBD (Res 474–485, 488–490, 494–505) and speculated to have a conformational variation for better binding with the ACE2.[36] MDS of similar complexes has been investigated

by another group.[33] The observed flexibility in the RBM of the truncated S1 domain complex was similar to our observation. Overall, the effect of the mutations on the conformation flexibility in the spike protein supports the earlier studies. This feature in the variants especially P.1 variant may provide an improved adaptability for binding with the ACE2.

Further, structural flexibility of the truncated S1 domain was also probed in terms of B-factor or thermal displacement. This has been an important parameter to analyze displacement of amino acid residues due to the thermal vibrations. The B-factor of each amino acid of the wild type and P.1 variant clearly indicates few variation among the two (Figure 8B,C). The NTD of the both wild type and P.1 variant showed comparable B-factor values. However, significant thermal fluctuations were noticed in the connecting loop of the NTD and RBD. Similar fluctuation was also observed in the RBM region as well as the SD1-2 region of the P.1 variant (Figure 8C). These regions showed reduced B-Factor in
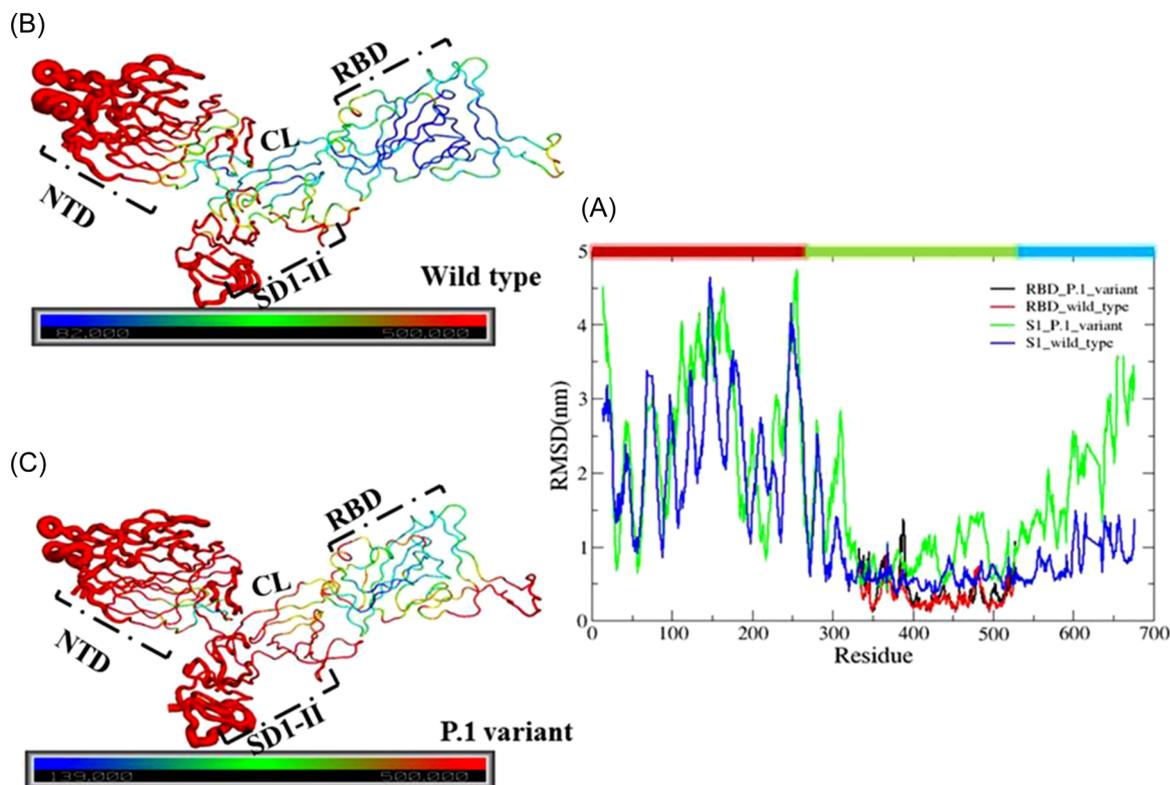
FIGURE 8 Representation of the conformational variation. (A) Residue-wise RMSD for the RBD as well as the truncated S1 domain. Shaded part at the top of the plot shows different regions of the S1 domain. Maroon, green, and cyan represents NTD, RBD, and CTD, respectively. (B) Residue-wise B-factor for the wild type (C) Residue-wise B-factor for the P.1 variant of the S1 domain. B factor is represented according to the color thickness scheme where blue region represent the rigid residues and red corresponds to the most flexible residues. Green and blue line is for complete S1 domain of Spike protein. Black and red is only for RBD of Spike protein. CL, connecting loop; NTD, N-terminal domain; RBD, receptor binding domain; SD1-II, Sub domain 1 and 2

the wild type (Figure 8B). In brief, our B-factor analysis for the truncated S1 domain structures also supports the similar flexibility changes with respect to our residue-wise RMSD deviation (Figure 8).

PCA of S1 domain complex with ACE2 from wild type and P.1 variant was performed to probe the dominant motions and conformational changes. Through PCA we can investigate the global motions of protein characterized into a few principal motions in terms of eigenvectors and eigenvalues. Usually the first two eigenvectors having most of the varying eigenvalues were considered for the analysis as they represent the most dominant motions of the protein. The eigenvalues of the first 10 corresponding eigenvector for wild type and P.1 variant are shown in Figure S12. This clearly depicted that the first two eigenvalue contributed the maximum variance of the projected points. Then further to know the global motions of backbone atoms of wild type and P.1 variant complexes were projected by first two principal components into the essential subspace for the conformational sampling. It was observed that the S1 domain of wild type covers comparatively smaller space

along the PC1 and PC2 than the same domain of P.1 variant structures (Figure 9). The mutant structure has occupied more conformational space with trace value of 171.07 nm^2 as compare to wild type (trace value of 121.40 nm^2). This high trace value in mutant complex showed an overall increased flexibility compared to wild type. Analysis clearly depicted that the P.1 variant covered a wide range of phase space along PC1 and PC2 which suggests that the mutations outside the RBD are also involved in providing the conformational heterogeneity or flexibility to the P.1 variant complex.

The change of secondary structure content analysis in the truncated S1 domain complexes from the both did not provide any noticeable difference in 100 ns time scale simulation. This suggests that the mutations do not incorporate any noticeable deviation in the secondary structure content. The mutations has sustained the stability at secondary structure content level throughout the 100 ns of MDS (Figure S13). However, calculated electrostatic charge distribution in S1-ACE2 complex in wild type and P.1 variant display a change in electrostatic charge distribution. Complex from P.1 variant observed
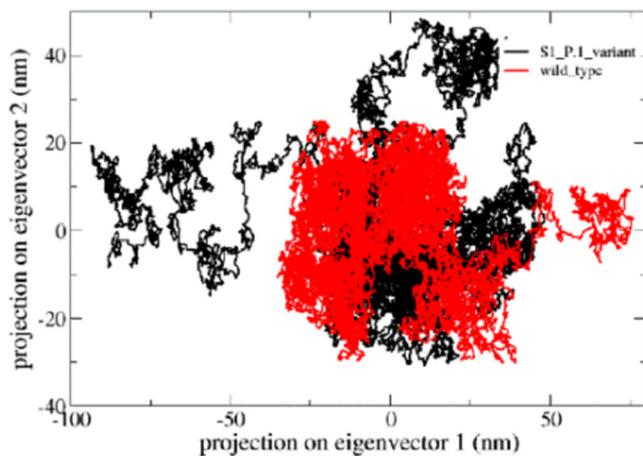
FIGURE 9 Representation of the conformational variation. Projection of dominant motions of backbone atoms in essential subspace along the first two principal eigenvectors of wild type and P.1 variant

to display more positive charge patches at the interface, compared to wild type (Figure S14). Positive charge in the P.1 variant may contribute for the better binding affinity. Earlier reports also suggested that positive charge patch on RBD of SARS-CoV-2 provide better binding with ACE-2 compared to SARS-CoV.[7] Other study have also demonstrated that mutations at RBM affect the charge distribution at the interface.[35]

4 | CONCLUSIONS

Emerging SARS-CoV-2 variants in each successive wave of infection have made this pandemic more threatening; where most of the mutations accumulated mainly in the Spike (S) glycoprotein. Since the Spike protein is the primary target of infection, concentrating on mutations of S protein leads to increase in transmission efficiency and escape from neutralizing antibodies. P.1 variant contains ten mutations on the S1 domain of the spike protein. Mutations on the RBD were reported to have an impact on increased affinity for ACE2 receptor. So in our study we have used SARS-CoV-2 S protein wild type and its P.1 variant RBD complex with ACE2 to uncover the structural basis responsible for the increased affinity towards the ACE2 receptor.

We have applied integrative networking and dynamics approaches to study the small conformational variations. Our networking study highlighted a node wise rearrangement in the network parameters. Significant reduction in nodes and links were observed related to each parameter of the PCN. The reduction of nodes and links in P.1 variant signifies a loss of interactions. Additionally, the mutations outside the RBD were observed to perturb the network properties of the NTD of the truncated S1 domain

of P.1 variant. Further, the effect of mutations were also predicted on the protein stability and flexibility through the vibrational entropy change ($\Delta\Delta S$) and free energy change ($\Delta\Delta G$) calculations. Both the analyses concluded that the interface RBD mutations are having a destabilizing effect. Comparative conformational variations among the RBD-ACE2 and truncated S1 domain-ACE2 were highlighted by MDS studies. Our MMPBSA analysis suggested the energetically favorable binding of the RBD with the ACE2 in the P.1 variant. Moreover, the conformational variations have also rearranged the type of interaction and interface interacting residues. The PCA of the S1 domain-ACE2 complex indicates an increased motion in the P.1 variant, which may be associated with the conformational heterogeneity in the P.1 variant complex. Overall, our study provides the structural basis of better interaction of the P.1 variant and identifies unique RBD residues crucial for the interaction with the ACE2 receptor. This study may be useful for designing better therapeutics against the circulating P.1 variant and other future variants as well.

AUTHOR CONTRIBUTIONS
Surabhi Lata and Mohd. Akif: conceptualize the study. **Surabhi Lata**: performed the study and analyzed the data. **Mohd. Akif**: supervised the study. **Surabhi Lata**: wrote the manuscript. **Mohd. Akif**: reviewed and edited the manuscript.

CONFLICTS OF INTEREST
The authors declare no conflicts of interest.

REFERENCES
1. Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol.* 2019;17(3):181-192.
2. Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease in China. *Nature.* 2020;579(7798): 265-269.
3. Zhou P, Yang XLou, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature.* 2020;579(7798):270-273.
4. Gui M, Song W, Zhou H, et al. Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a

prerequisite conformational state for receptor binding. *Cell Res*. 2017;27(1):119-129.

5. Song W, Gui M, Wang X, Xiang Y. Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. *PLoS Pathog*. 2018;14(8):1-19.

6. Wang Q, Zhang Y, Wu L, et al. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell*. 2020;181(4): 894-904.e9.

7. Lan J, Ge J, Yu J, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*. 2020;581(7807):215-220.

8. Gómez CE, Perdiguero B, Esteban M. Emerging sars-cov-2 variants and impact in global vaccination programs against sars-cov-2/covid-19. *Vaccines*. 2021;9(3):1-13.

9. Fiorentini S, Messali S, Zani A, et al. First detection of SARS-CoV-2 spike protein N501 mutation in Italy in August, 2020. *Lancet Infect Dis*. 2021;21(6):e147.

10. Buss LF, Prete CA, Abrahim CMM, et al. Three-quarters attack rate of SARS-CoV-2 in the Brazilian Amazon during a largely unmitigated epidemic. *Science*. 2021;371(6526): 288-292.

11. Tegally H, Wilkinson E, Lessells RJ, et al. Sixteen novel lineages of SARS-CoV-2 in South Africa. *Nat Med*. 2021;27(3): 440-446.

12. Davies NG, Abbott S, Barnard RC, et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science*. 2021;(80-) 372(6538).

13. Sabino EC, Buss LF, Carvalho MPS, et al. Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *Lancet*. 2021;397(10273):452-455.

14. Kannan SR, Spratt AN, Sharma K, Chand HS, Byrareddy SN, Singh K. Omicron SARS-CoV-2 variant: Unique features and their impact on pre-existing antibodies. *J Autoimmun*. 2022;126:102779.

15. Han P, Li L, Liu S, et al. Receptor binding and complex structures of human ACE2 to spike RBD from omicron and delta SARS-CoV-2. *Cell*. 2022;185(4):630-640.e10.

16. Chan KK, Dorosky D, Sharma P, et al. Engineering human ACE2 to optimize binding to the spike protein of SARS coronavirus 2. *Science*. 2020;369(6508):1261-1265.

17. Starr TN, Greaney AJ, Hilton SK, et al. Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. *Cell*. 2020;182(5): 1295-1310.e20.

18. Dejnirattisai W, Zhou D, Supasa P, et al. Antibody evasion by the P.1 strain of SARS-CoV-2. *Cell*. 2021;184(11):2939-2954.e9.

19. Wibmer CK, Ayres F, Hermanus T, et al. SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. *Nat Med*. 2021;27(4):622-625.

20. Xie X, Liu Y, Liu J, et al. Neutralization of SARS-CoV-2 spike 69/70 deletion, E484K and N501Y variants by BNT162b2 vaccine-elicited sera. *Nat Med*. 2021;27(4):620-621.

21. Ray D, Le L, Andricioaei I. Distant residues modulate conformational opening in SARS-CoV-2 spike protein. *Proc Natl Acad Sci USA*. 2021;118:43.

22. Srivastava A, Birari V, Sinha S. Small Conformational Changes Underlie Evolution of Resistance to NNRTI in HIV Reverse Transcriptase. *Biophys J*. 2020;118(10):2489-2501.

23. Kandhari N, Sinha S. Complex network analysis of thermostable mutants of Bacillus subtilis Lipase A. *Appl Netw Sci*. 2017;2:1.

24. Srivastava A, Sinha S. Thermostability of in vitro evolved Bacillus subtilis Lipase a: a network and dynamics perspective. *PLoS One*. 2014;9:8.

25. Jacobs DJ, Rader AJ, Kuhn LA, Thorpe MF. Protein flexibility predictions using graph theory. *Proteins: Struct, Funct, Genet*. 2001;44(2):150-165.

26. Srivastava A, Sinha S. Uncoupling of an ammonia channel as a mechanism of allosteric inhibition in anthranilate synthase of Serratia marcescens: dynamic and graph theoretical analysis. *Mol BioSyst*. 2017;13(1):142-155.

27. Amitai G, Shemesh A, Sitbon E, et al. Network analysis of protein structures identifies functional residues. *J Mol Biol*. 2004;344(4):1135-1146.

28. Lata S, Akif M. Comparative protein structure network analysis on 3CLpro from SARS-CoV-1 and SARS-CoV-2. *Proteins Struct Funct Bioinforma*. 2021;89(9):1216-1225.

29. Verkhivker G. Coevolution, dynamics and allostery conspire in shaping cooperative binding and signal transmission of the SARS-CoV-2 spike protein with human angiotensin-converting. *Enzyme*. 2020a;2:1-31p.

30. Laurini E, Marson D, Aulic S, Fermeglia M, Pricl S. Computational alanine scanning and structural analysis of the SARS-CoV-2 spike protein/angiotensin-converting enzyme 2 complex. *ACS Nano*. 2020;14(9):11821-11830.

31. Wang Y, Liu M, Gao J. Enhanced receptor binding of SARS-CoV-2 through networks of hydrogen-bonding and hydrophobic interactions. *Proc Natl Acad Sci U S A*. 2020b;117(25): 13967-13974.

32. Paola LDi, Hadi-Alijanvand H, Song X, Hu G, Giuliani A. The discovery of a putative allosteric site in the SARS-CoV-2 spike protein using an integrated structural/dynamic approach. *J Proteome Res*. 2020;19(11):4576-4586.

33. Verkhivker GM, Agajanian S, Oztas DY, Gupta G. Comparative perturbation-based modeling of the SARS-CoV-2 spike protein binding with host receptor and neutralizing antibodies: structurally adaptable allosteric communication hotspots define spike sites targeted by global circulating mutations. *Biochemistry*. 2021;60(19):1459-1484.

34. Verkhivker GM, Paola LDi. Dynamic network modeling of allosteric interactions and communication pathways in the SARS-CoV-2 spike trimer mutants: differential modulation of conformational landscapes and signal transmission via cascades of regulatory switches. *J Phys Chem B*. 2021;125(3): 850-873.

35. Dehury B, Raina V, Misra N, Suar M. Effect of mutation on structure, function and dynamics of receptor binding domain of human SARS-CoV-2 with host cell receptor ACE2: a molecular dynamics simulations study. *J Biomol Struct Dyn*. 2021;39(18):7231-7245.

36. Khan A, Gui J, Ahmad W, et al. The SARS-CoV-2 B.1.618 variant slightly alters the spike RBD-ACE2 binding affinity and is an antibody escaping variant: a computational structural perspective. *RSC Adv*. 2021;11(48):30132-30147.

37. Periwal N, Rathod SB, Pal R, et al. In silico characterization of mutations circulating in SARS-CoV-2 structural proteins. *J Biomol Struct Dyn*. 2021:1-16. doi:10.1080/07391102.2021.1908170

38. Teruel N, Mailhot O, Najmanovich RJ. Modelling conformational state dynamics and its role on infection for SARS-CoV-2 Spike protein variants. *PLoS Comput Biol*. 2021;17(8):1-25.

39. Verkhivker GM. Molecular simulations and network modeling reveal an allosteric signaling in the SARS-CoV-2 spike proteins. *J Proteome Res*. 2020b;19(11):4587-4608.

40. Razizadeh M, Nikfar M, Liu Y. Small molecule therapeutics to destabilize the ACE2-RBD complex: a molecular dynamics study. *Biophys J*. 2021;120(14):2793-2804.

41. Felline A, Seeber M, Fanelli F. webPSN v2.0: a webserver to infer fingerprints of structural communication in biomacro-molecules. *Nucleic Acids Res*. 2020;48(W1):W94-W103.

42. Brinda KV, Vishveshwara S. A network representation of protein structures: Implications for protein stability. *Biophys J*. 2005;89(6):4159-4170.

43. Kannan N, Vishveshwara S. Identification of side-chain clusters in protein structures by a graph spectral method. *J Mol Biol*. 1999;292(2):441-464.

44. Rodrigues CHM, Pires DEV, Ascher DB. DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Res*. 2018;46(W1):W350-W355.

45. Gopi S, Devanshu D, Rajasekaran N, Anantakrishnan S, Naganathan AN. PPerturb: a server for predicting long-distance energetic couplings and mutation-induced stability changes in proteins via perturbations. *ACS Omega*. 2020;5(2):1142-1146.

46. Pronk S, Páll S, Schulz R, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*. 2013;29(7):845-854.

47. Berendsen HJC, Grigera JR, Straatsma TP. The missing term in effective pair potentials. *J Phys Chem*. 1987;91(24):6269-6271.

48. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem Phys*. 2007;126:1.

49. Parrinello M, Rahman A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J Appl Phys*. 1981;52(12):7182-7190.

50. J. CCD and D. Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins (2014):193-226 p.

51. Kumari R, Kumar R, Lynn A. G-mmpbsa -A GROMACS tool for high-throughput MM-PBSA calculations. *J Chem Inf Model*. 2014;54(7):1951-1962.

52. López-Blanco JR, Aliaga JI, Quintana-Ortí ES, Chacón P. IMODS: internal coordinates normal mode analysis server. *Nucleic Acids Res*. 2014;42(W1):271-276.

53. Halder A, Anto A, Subramanyan V, Bhattacharyya M, Vishveshwara S, Vishveshwara S. Surveying the Side-Chain Network approach to protein structure and dynamics: the SARS-CoV-2 spike protein as an illustrative case. *Front Mol Biosci*. 2020;7(December):1-15.

54. Bayarri-Olmos R, Jarlhelt I, Johnsen LB, et al. Functional effects of receptor-binding domain mutations of SARS-CoV-2 B.1.351 and P.1 variants. *Front Immunol*. 2021;12(October):1-10.

55. Gobeil SMC, Janowska K, McDowell S, et al. D614G mutation alters SARS-CoV-2 spike conformation and enhances protease cleavage at the S1/S2 junction. *Cell Rep*. 2021;34(2):108630.

56. Ozono S, Zhang Y, Ode H, et al. SARS-CoV-2 D614G spike mutation increases entry efficiency with enhanced ACE2-binding affinity. *Nat Commun*. 2021;12:1.

57. Papaleo E, Mereghetti P, Fantucci P, Grandori R, Gioia LDe. Free-energy landscape, principal component analysis, and structural clustering to identify representative conformations from molecular dynamics simulations: The myoglobin case. *J Mol Graph Model*. 2009;27(8):889-899.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.