Structural communication fingerprinting and dynamic investigation of RBD-hACE2 complex from BA.1 x AY.4 recombinant variant (Deltacron) of SARS-CoV-2 to decipher the structural basis for enhanced transmission

Abstract

The BA.1 x AY.4 recombinant variant (Deltacron) continues to inflict chaos globally due to its rapid transmission and infectivity. To decipher the mechanism of pathogenesis by the BA.1 x AY.4 recombinant variant (Deltacron), a protein coupling, protein structural graphs (PSG), residue communication and all atoms simulation protocols were used. We observed that the bonding network is altered by this variant; engaging new residues that helps to robustly bind. The protein structural graphs revealed variations in the hub residues, number of nodes, inter and intra residues communities, and path communication perturbation caused by the acquired mutations in the Deltacron-RBD thus alter the binding approach and infectivity. Moreover, the dynamic behaviour reported a highly flexible structure with enhanced residues flexibility particularly by the loops required for interaction with ACE2. It was observed that these mutations have altered the secondary structure of the RBD mostly transited to the loops thus acquired higher flexible dynamics than the native structure during the simulation. The total binding free energy for each of these complexes i.e. WT-RBD and Deltacron-RBD were reported to be -61.38 kcal/mol and -70.47 kcal/mol. Protein's motion revealed a high trace value in the Deltacron variant that clearly depict more structural flexibility. The broad range of phase space covered by the Deltacron variant along PC1 and PC2 suggests that these mutations are important in contributing conformational heterogeneity or flexibility that consequently help the variant to bind more efficiently than the wild type. The current study provide a basis for structure-based drug designing against SARS-CoV-2.

Keywords: Deltacron; Variant; Binding; Simulation; MM/GBSA

Introduction

With the advent of the COVID-19 pandemic in 2019, by the SARS-CoV-2 virus, which obtained the ability to infect the human race, marked a new era in history [1]. The transmission, of the disease, principally occurred from an infected individual through droplets, both aerosol and surface-to-surface [2, 3]. Despite the proofreading capability, the mutation frequency during viral replications is exceptionally high, owing to insertions, deletions in the viral genome and co-infection of the same cell by the two different viral strains spontaneously developed new SARS-CoV-2 variants. Some of the new mutations added to the increased ability of the virus to infect host cells effectively, and evade the immunity[3][4][5]. The COVID-19 pandemic has ravaged globally, jobs, academics, social interactions, healthcare systems and smashed the economic growth of the developed countries. The economic cadaverous has headed the world towards the greater inflation.

Regardless of the innovative advances, in the fields of biomedical and clinical sciences, vaccines, combination of drugs, the research in both sectors remained futile in precluding the emergence of new variants. With the constant emergence of novel variants, some of the countries have experienced the fourth and fifth wave of COVID-19 pandemic; the situation is further pushing the end to this pandemic, creating an intimidating concern for the world. Thus far, among the reported COVID-19 variants are the VOC Delta (δ)+ (AY.1 or lineage B.1.617.2.1) variant, descendent of the delta variant has acquired L452R, T478K mutations in the receptor-binding domain (RBD) and an additional δ+ variant's mutation is K417N. [4]. Similarly, the δ variant discovered in Colombia, has acquired E484K, N501Y, and P681H mutations in the spike protein. The δ has an enhanced ability of infectivity, which further increased the COVID cases (reported January 2021). This particular δ variant, has also picked up other new mutations that includes R346K, Y144T,Y145S and 146N insertion [5]. Likewise, the Lambda (λ) or C.37 variant discovered in Pero, regarded as the "variant of interest" had mutations (L452Q and F490S in the RBD) assumed to be related with decreased antibody neutralizing susceptibility, predominantly due to the F490S mutation in the RBD [6, 7]. In India, the B.1.617.1 or the Kappa (κ) discovered is a VOI possessing the L452R variant also assumed to be involved in reduced antibody neutralizing by disrupting the respective conformational epitopes [8]. In early 2021, in New York City the discovered VOI Iota (ι) of the linage B.1.526, has the mutation E484K observed in P.1 variant. Experimental investigations has demonstrated the entirely or partial escape form the used therapeutic monoclonal antibodies (mAbs) and has decreased susceptibility to neutralization. [9]. The reported E484K mutation in the P.1, has enabled the direct interaction the host's hACE2 receptor [10]. Similarly, another C.12 variant discovered in South Africa, nominated as a "variant under monitoring" has no confirmed associated risk factors [11].

Recently, a recombinant variant of known as "Deltacron" has been reported to combine mutations from Omicron and Delta variants. Genomic sequencing of the isolated samples and biochemical analysis revealed the optimized binding of the hybrid variant with the host receptor[12]. As of March 10, an international database of viral sequences reported 33 samples of the new variant in France, eight in Denmark, one in Germany and one in the Netherlands. This variant has been reported to spread faster than any other reported variant until now. The variant is still under investigation and no information on the binding and infectivity are yet disclosed. Hence, deep analysis to understand the binding pattern and to disclose the other features are required. It is thus crucial to investigate whether the mutation has made significant changes in the structural integrity, the functional outcome and the binding deviations of the RBD-hACE2.

In the current study, to decipher the pathogenesis of the Deltacron variant, the protein-protein docking and all atoms simulation protocols were deployed by sequentially analyzing it with wild type. Detail investigation of the dynamics features such as a protein coupling, protein structural graphs (PSG), residue communication and all atoms simulation protocols were used to provide atomic level insights into the dynamic variation. In addition, we employed the MM/GBSA approach to demonstrate binding free energy to further validate the docking results. The current study is first of its kind to decipher the binding mechanism of Deltacron variant and provide basis for structure-based drug designing.

Material and Methods

Structural Modeling and Interaction Prediction

Structure of the wild type RBD in complex with ACE2 was retrieved using 6M0J from the Protein Databank (RCSB). The sequence of wild type RBD was manipulated and the reported mutations in the RBD of Deltacron were modelled using Chimera embedded Modeller software. For the template the reported co-crystal PDB ID: 6M17 was considered. HADDOCK enabled restraint docking of the wild type and mutant was considering the previous parameter [13-17]. A special interface, Guru Interface, was exploited to exercise all the available options for best docking [10, 17-20]. The protein complexes were generated by recruiting the Guru interface and visualized to check the electrostatic contacts, hydrogen bonds, and salt bridges using PDBsum web server [21].

Structural Fingerprints for hACE2-RBD Communication

To detect the inter-connectivity variations at atomic level a protein communication network (PCN) was constructed. For this purpose, webPSN v2.0 (http://webpsn.hpc.unimo.it/wpsn3.php) webserver was used which combine Protein Structure Network (PSN) and Elastic Network Model-Normal Mode Analysis (ENM-NMA)-based strategy (PSN-ENM) to demonstrate the structural communication information. Hubs, concisely, are nodes with the greatest degree. Modularity is expressed by communities with more linked nodes, and nodes within the same community are highly connected to each other compared to nodes outside the community that are poorly connected. The shortest path is the one that requires the lowest number of links to get from one node to the next. It is calculated by using Dijkstra's method. The wild type (hACE2-RBD) and Deltacron variant complexes were uploaded as PDB files to investigate the total number of nodes, edges, modularity and shortest communication paths. The server uses the following expression to construct PSN and its important parameters.

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

Where (*Ii*j) 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. It constructs a PSG based on the atomic cross correlation motions using the ENM-NMA.

Dynamics of the wild type and B.1.640.2-RBD Complexes

We performed 500ns simulation of each complex using AMBER20 employing FF19SB [22, 23]. Abbas et al., 2021, previously reported complete details on the system preparation and MD analysis [24]. Shortly, an OPC water box and the addition of Na+ ions for neutralization and solvation followed by 6000 and 3000 steps of minimization employing steepest descent and conjugate gradient algorithms. In the further process, heating at 300 K and equilibration for 50ns was achieved. Finally, a total of 1microsecond simulation was executed each complex of 500ns. Simulation trajectories were analyzed through the CPPTRAJ and PTRAJ modules of AMBER[25]. For structural stability root mean square deviation (RMSD) analysis as a function of time was performed using the following equation.

$$RMSD = \sqrt{\frac{\sum d^2 i=1}{N_{atoms}}} \quad \text{(i)}$$

Where:

d_i is the difference of position between atoms and i refers to the original and superimposed structure.

Whereas the root mean square fluctuation (RMSF) can be computed by employing B-factor [33], which is the most imperative constraint to compute the flexibility of all the residues in a protein. Mathematically the RMSF can be calculated by using the following equation.

$$Thermal\ factor\ or\ B-factor\ =\ [(8\pi\ **\ 2)/3]\ (msf) \quad \text{(ii)}$$

Estimation of Binding Free Energy

We calculated the binding free energy as MM/GBSA for each complex such as wild type and Deltacron variant using the MMPBSA.py script [26]. This widely employed approach gives estimation of vdW, electrostatic, GB and SA also used by other studies to calculate the total free energy of the RBD and ACE2 complexes [20, 27-32]. Mathematically the following equation was used to estimate the binding energy:

$$"\Delta \mathbf{G_{net\ binding\ energy}} = \Delta \mathbf{G_{complex\ binding\ energy}} - [\Delta \mathbf{G_{receptor\ binding\ energy}} + \Delta \mathbf{G_{ligand\ binding\ energy}}]"$$

Each of the above components of net binding energy can be split as follows:

$$"\mathbf{G} = \mathbf{G_{bonded}} + \mathbf{G_{van\ der\ waals}} + \mathbf{G_{polar\ solvation\ energy}} + \mathbf{G_{non-polar\ solvation\ energy}}"$$

Capturing the Protein Collective Motions During Simulation

The internal and localized motions of each trajectory were clustered by using Principal Component Analysis (PCA) approach [33, 34]. For the clustering of each trajectory, a CPPTRAJ module was used to compute the positional covariance matrix for eigenvectors and their atomic coordinates. Orthogonal coordinate's transformation was used to diagonalize the matrix of eigenvalues. Finally, the PCs were acquired based on eigenvalues and eigenvectors, which clustered the motions of each trajectory during the 500ns of simulation[35, 36].

Results and Discussion

Structural Modelling and Analysis

Since the inception of COVID-19 pandemic, the world is still struggling to cope with this prolonged aggravated condition. While progress in clinical research has led to an increased understanding of SARS-CoV-2 and its treatments, newly emerged variants remain an important concern and have caused multiple waves of the pandemic in several countries. Recently, a recombinant variant of known as "Deltacron" has been reported to combine mutations from Omicron and Delta variants. Genomic sequencing of the isolated samples and biochemical analysis

revealed the optimized binding of the hybrid variant with the host receptor[12]. As of March 10, an international database of viral sequences reported 33 samples of the new variant in France, eight in Denmark, one in Germany and one in the Netherlands. This variant has been reported to spread faster than any other reported variant until now. The variant is still under investigation and no information on the binding and infectivity are yet disclosed. Hence, deep analysis to understand the binding pattern and to disclose the other features are required. It is thus crucial to investigate whether the mutation has made significant changes in the structural integrity, the functional outcome and the binding deviations of the RBD-hACE2. The Deltacron variant continues to inflict chaos globally due to its rapid transmission and infectivity. The variant is still under investigation and no information on the binding and infectivity are yet disclosed. The Spike glycoprotein, which comprise of multiple domain is the prime virulent factor and is mostly targeted by the virus for the mutations **(Figure 1A)**. Therefore, in the current study, to decipher the pathogenesis of the Deltacron variant, the protein-protein docking of the RBD- hACE2 and all atoms simulation protocols were deployed by sequentially analyzing it with wild type. The reported mutations in the Deltacron RBD were identified and shown in **Figure 1B**. For the docking, the interface site was identified from the crystallographic structure and previous literature which was targeted for the interaction. The interface of RBD-ACE2 is shown in **Figure 1C**. The modeled structure of the Deltacron RBD was compared with the wild type RBD. Superimposition of the wild type and Deltacron RBD revealed an RMSD difference of 0.171Å, which demonstrate deviation in the structure. The superimposed structure of the wild type and Deltacron RBD is given in **Figure 1D**.

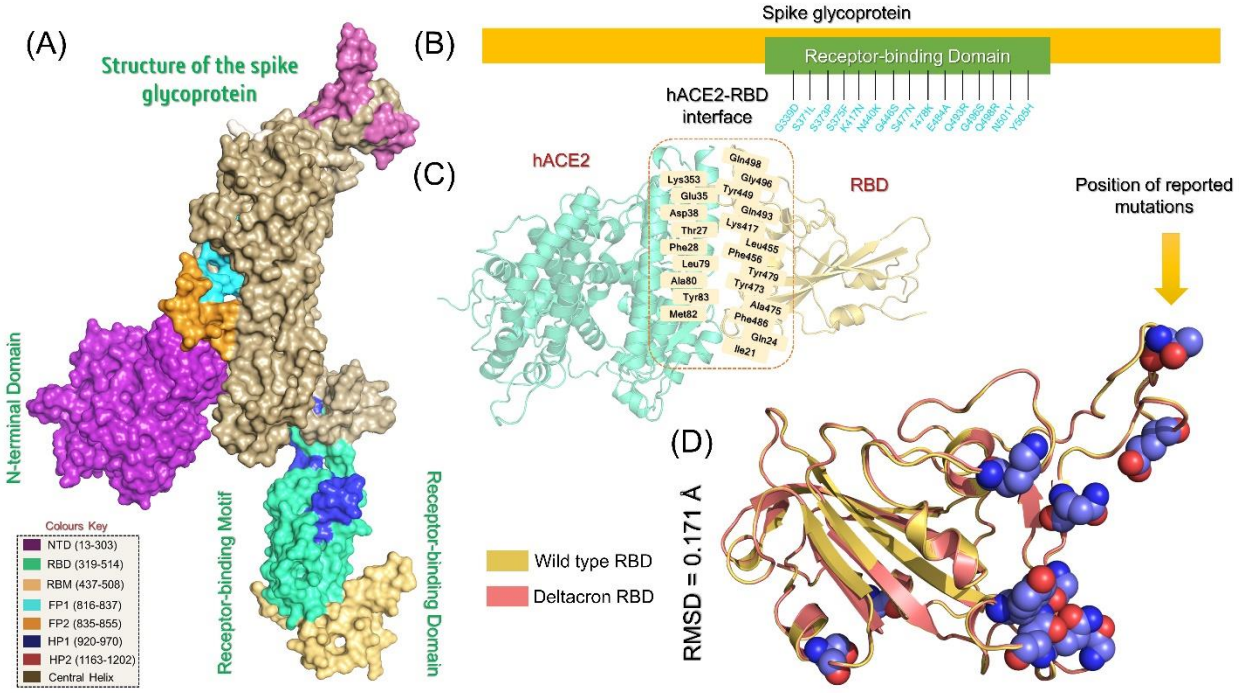

Figure 1: (A) domain mapping of the spike glycoprotein. **(B)** Mutations mapping on spike protein. **(C)** Interface residues between RBD and ACE2. **(D)** Superimposed structures of the wild type and Deltacron RBD whereas the spheres represent the location of mutations.

Docking of the wild type and Deltacron RBD with hACE2

Analysis of the binding variations between the wild type and the Deltacron variant were explored to provide deep insights in the mechanism of higher infectivity by the Deltacron variant. For instance, the HADDOCK docking score for the wild type has been previously reported to be -111.8 +/-1.5 kcal/mol. In contrast the docking score for the Deltacron variant was calculated to be -128.3 +/-2.5 kcal/mol. The docking score for the Deltacron variant is higher than the omicron (-118.3 +/-4.9 kcal/mol) and other variants previously reported by other studies [10, 16]. In the case of the wild type the Van der Waals energy has been reported to be -48.1 +/-1.3 kcal/mol while the electrostatic energy has been reported to -169.7 +/-13.2 kcal/mol[16]. Herein for the Deltacron variant the vdW was calculated to be -62.9 +/-4.4 kcal/mol while the electrostatic energy was calculated to be -175.0 +/-28.1 kcal/mol respectively. Hence this shows the stronger interaction of the Deltacron-RBD with the host receptor ACE2 than the wild type. The interaction analysis for the wild type and Deltacron variant was performed to see the binding differences. For the wild type a total of 10 hydrogen bonds with one salt bridge has been reported previously. For the Deltacron variant nine hydrogen bonds and one salt bridge was observed. The specific interactions involve Glu38-Asn487, Glu75-Asn417, Thr78-Arg403, Gln81-His505, Gln81-Arg403, Thr82-Tyr501, Thr82-Arg403, Glu87-Arg498 and Glu87-Thr500. The only salt bridge was reported between Glu87 and Arg498. The binding pattern of the wild type and Deltacron variant is shown in **Figure 2A** and **2B**. It can be seen that the Deltacron variant demonstrated a highly varied bonding network in contrast to the wild type. In the case of Deltacron variant the Thr21 is involved in interaction with Thr500 which is involved in interaction with Lys353 in the wild type. The Asn487 in interaction with Tyr83 in the wild type complex is also altered in the Deltacron complex. Herein Glu38 instead of Tyr83 is involve in interaction with Asn487. The interactions Glu75-Asn417, Thr78-Arg403, Gln81-His505, Gln81-Arg403, Thr82-Arg403, Glu87-Arg498 and Glu87-Thr500 are the newly reported interactions and only in the Deltacron variant but not reported in any previous variants [10, 16, 17, 20]. This consequently show that this particular variant uses different strategy to interact with the hACE2 and enter into the host cell. The current findings corroborate with the recent experimental report which claim an optimized binding of the Deltacron RBD with the host[12].

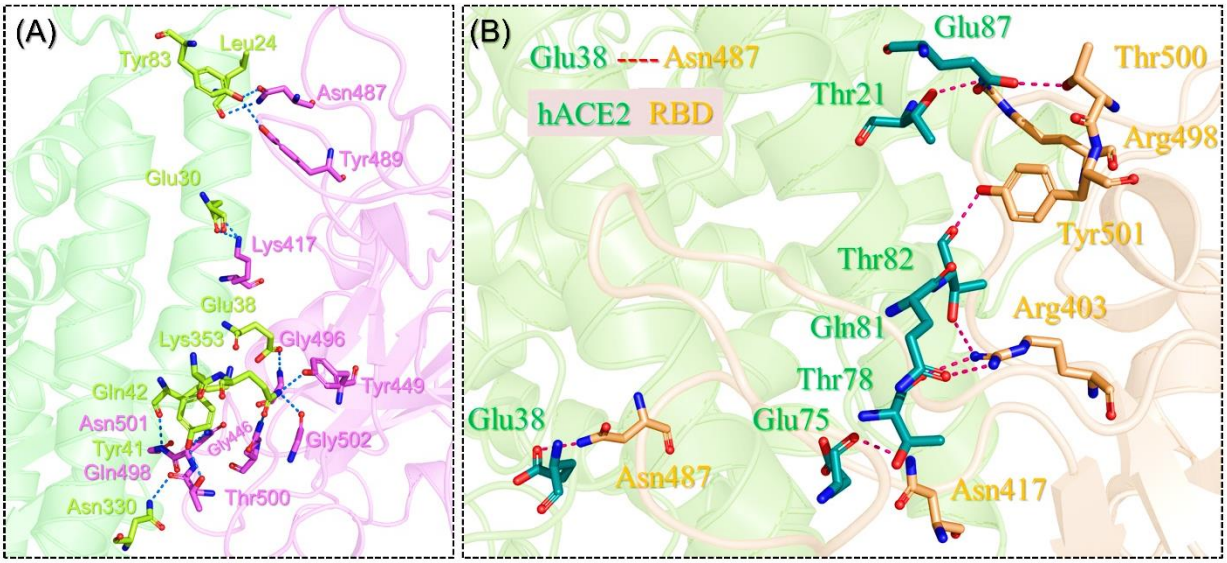

Figure 2: Structural analysis of the binding of the wild type and Deltacron-RBD with the hACE2. (A) Shows the binding pattern of the wild type RBD in complex with hACE2 while **(B)** Shows the binding pattern of the Deltacron RBD in complex with hACE2.

hACE2-RBD Structural Network Analysis

To derive knowledge regarding the regarding residues network specific variations caused by mutations the protein structure network analyses were performed. Assessment of the total number of hubs in each complex revealed 236 hub residues in the wild type and 136 hub residues in the Deltacron RBD-hACE2 complex. It indicate that due to the significant number of mutation in the Deltacron variant the hub residues variations are also significant. For instance, variations in the hubs are also reported in the P.1 variant where decrease in the hub residues in the variant complex was also observed [37]. Hence, it show the structural perturbation caused by these mutations, which consequently used alternate interaction pattern with the host receptor. The surface mapping of hub residues on the structure of the wild type hACE2-RBD complex and Deltacron RBD in complex with hACE2 are shown in **Figure 3A** and **3B**. Consistent with the previous results on P.1 variant the mutated residues perturbed the hub residues network and particularly in the RBD of Deltacron complex [37]. The stabilizing anti-parallel beta-sheets in the structure of RBD also demonstrated notable variation in the hub residues. This particular region in the wild type complex is enriched with the hub residues while in case of Deltacron complex significant decline in the hub residues was observed. Several novel hub residues i.e. L351 with an average force 7.24, R357 with an average force 11.71, Y454 with an average force 10.22 while H505 with an average force 7.29 were newly observed in the Deltacron hACE2-RBD complex only. Hence, this show that the acquired mutations does not only increase the binding but also affect the structural residues connectivity network which consequently opt the BA.1 x AY.4 recombinant variant to adapt the best conformational coordinated for enhanced binding and transmission.

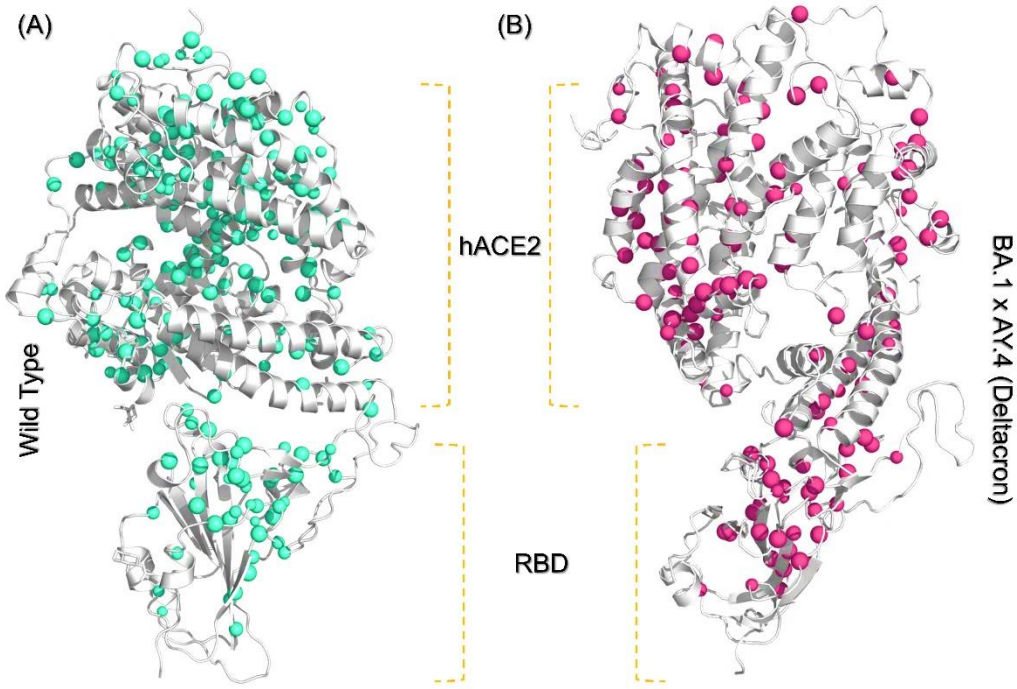

Figure 3: Structural Network analysis of the wild type and Deltacron-RBD with the hACE2. **(A)** Shows the hub residues occurrence and distribution in the wild type RBD in complex with hACE2. The green sphere represent the hub residues in the wild type, while **(B)** shows the hub residues occurrence and distribution of the Deltacron RBD in complex with hACE2. The dark purple sphere represent the hub residues in the Deltacron-RBD complex.

Residues Communities

The residues community or modularity in the protein structure show the sub-residue networks, which show the communication among different functional residues. It was observed that 19 communities were formed in the wild type complex while 25 communities in the Deltacron-RBD complex. In the wild type, the largest community was reported in the ACE2 structure where 196 nodes, 352 edges and 132 hub residues were involved. While in the case of Deltacron-RBD complex the largest community reside in ACE2 involve 49 nodes, 73 edges while 29 hub residues. The findings corroborate with the previous report where largest community of residues was reported in the ACE2 structure. The second largest community was reported in RBD in each complex where the wild type reported 47 nodes, 72 edges and 25 hub residues while the Deltacron complex reported 27 nodes, 41 edges and 18 hub residues. This show that the altered hub residues and community clusters in the variant helps in implying better efficiency for binding than the wild type. The detected communities in each complex are shown in **Figure 4A** and **4B**.

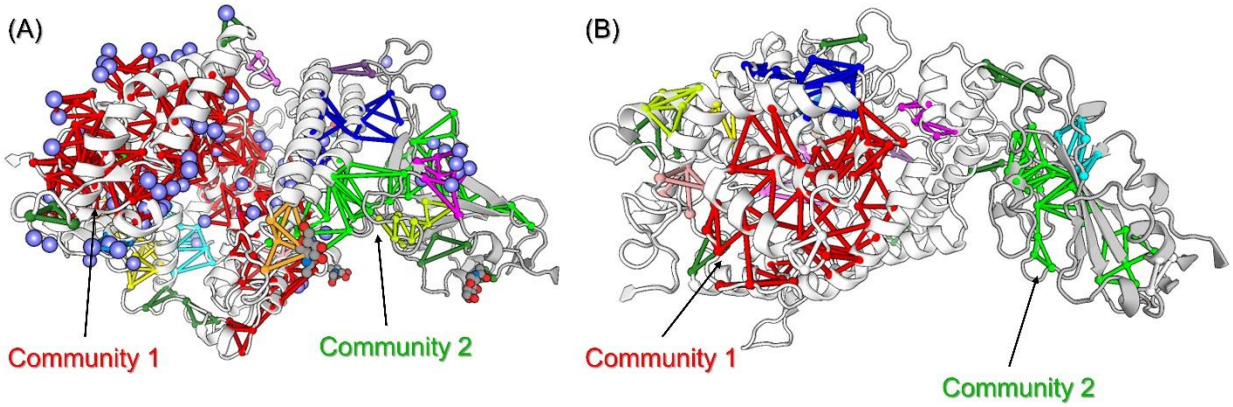

Figure 4: Residues sub-networks analysis in the wild type and Deltacron-RBD with the hACE2. (A) Shows the residues sub-networks analysis in the wild type RBD in complex with hACE2, while **(B)** shows residues sub-networks analysis the Deltacron-RBD in complex with hACE2. The largest two communities in each complex are shown with arrow.

Communication pathway analysis

Furthermore, we also calculated the shortest communication pathway to see how these complexes vary in the communication channel. It was observed the shorted path in the wild type RBD-ACE2 complex was 1878914 while in the Deltacron-RBD 984220 shortest path was detected. Moreover, the average path hub percentage was also observed to have decreased in the variant complex. For the wild type the communication path involve Y495→ R403→ Y505→ E37→ R393→ Y385→ Y381→ F400→ F397→ Y207→ I513→ E457→ and F512 residues while the Deltacron variant complex the communication path involve I402 → Y495 → Y453 → L79 → L455 → F456 → F72 → W69 → S40 → F390 → Q39 → R393 → Y385 → Y381 → M557 → F400 → F397 → Y207 → E398 → S511 → W203 → R460 and V506 residues. it can be seen that the variant complex involve mostly the RBD and mutated residues particularly for the inter and intra residues communication. This show the path communication perturbation caused by the acquired mutations in the Deltacron-RBD thus alter the binding approach and infectivity. The communication pathways for the wild type and Deltacron-RBD complexes are shown in **Figure 5A** and **5B**. The observed key parameters in the networks of each complex are given in **Table 1**.

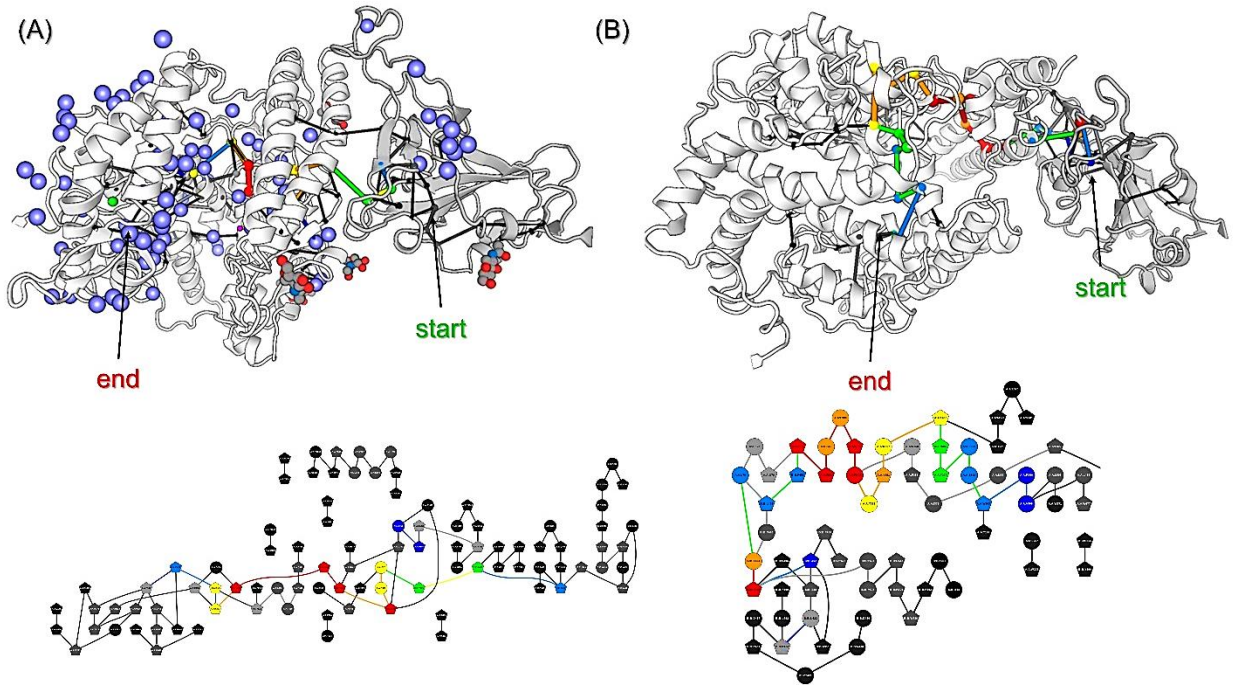

Figure 5: Shortest communication pathways analysis in the wild type and Deltacron-RBD with the hACE2. (A) Shows the shortest communication pathway in the wild type RBD in complex with hACE2, while **(B)** shows the shortest communication pathway in the Deltacron-RBD in complex with hACE2. The lower panels show the topographical representation of the communication pathway.

Table 1: Protein Network components and parameters.

Path Summary	Wild Type	Deltacron Variant
Number of nodes in path		
Number of links in path		
Number of shortest paths		
Average path length		
Average path hub %		

Assessment of Dynamic Stability (RMSD)

To compute the variations in the dynamic behaviour between the wild type and Deltacron variant we performed molecular dynamics simulation of each complex. RMSD is an important estimation to determine the complex stability in a dynamic environment, which can be tally with the binding strength, and stability. The role of dynamic stability in the enhanced infectivity has been previously deciphered for other variants i.e. B.1.1.7, P.1, B.1.351, B.1.617, B.1.1618 and B.1.1.529 [10, 16, 17, 20]. Hence employing the similar method, we also calculated the stability as RMSD. As given in **Figure 6A**, the wild type complex consistently reported a stable dynamic behaviour over the simulation time. The complex demonstrated an initial RMSD of 3.0 Å and reached the equilibrium at 10ns. The RMSD stabilized at 2.8 Å continues to follow the similar pattern until the end of simulation. No significant structural perturbation can be seen particularly after 250ns. During the

first part simulation (1-250) smaller deviations can be observed i.e. at 95ns, 105ns and 210ns. The complex then reported a uniform pattern in the later part of the simulation. This shows a stable dynamic behaviour of the wild type-RBD and hACE2 complexes validating the previous reports where a stable dynamic behaviour has been reported [10, 16, 17, 20]. An average RMSD for the wild type was estimated to be 2.92 Å. On the other hand, the RMSD of the Deltacron-RBD hACE2 complex reported a very different dynamic behaviour than the wild type and the previously reported variants until now (**Figure 6B)**. The RMSD of the Deltacron variant continues to gradually increase from the start of the simulation and during the first 200ns the RMSD reached 8.0 Å. The RMSD then abruptly increased for a shorter period (201-215ns) and reached 11.0 Å. The RMSD then decreased and stabilized at 9.0 Å. Afterwards the complex reported a very stable pattern with no deviation until the end of simulation. Despite the higher RMSD value the structure reached the stability and an average RMSD was calculated to be 8.24 Å. This is the first variant, which reported a higher RMSD than the wild type although reached the stability at the later time. For instance a strong correlation between the mutations induced stability in the RBD and infectivity has been explored by a study in the *Cell* journal[38]. They reported that mutations, which increases the RBD stability, also increases the infectivity and this relationship was observed in other variants too characterized by using biophysical approaches [10, 17, 20, 39]. The C432D has been reported to decrease the stability and thus reduces viral entry[38]. Since the behaviour of the Deltacron variant is different, it is not surprising that the trend may have been altered because of the complex game between the environment and organism survival. This unstable behaviour may be a cause of this accelerated transmission and optimized binding. As the destabilizing mutation cannot be benign but could produce radical functions claimed by an evolutionary and structural study on the immune evasion protein of the SARS-CoV-2[32]. Conclusively this fixed amino acid substitutions give a different ability to the Deltacron variant to interact with the host and render more rapidly than the other variants.

Assessment of Protein Packing through Rg

We examined the structural compactness in a dynamic environment by calculating the radius of gyration (Rg) as a function of time. As given in **Figure 7A**, the wild type complex initially demonstrated a higher Rg value was recorded. The Rg initially increased until 15ns, then followed a uniform pattern until 100ns and then continues to gradually decrease until 350ns. The Rg then continue to increase again until the end of simulation. An average Rg value for the wild type was reported to be 31.5 Å. A similar pattern of Rg has been previously reported for the wild type which demonstrated the higher number of binding and unbinding events happened during the simulation [16]. On the other hand, the Rg of the Deltacron complex reported a similar behaviour as the RMSD. The Rg increased gradually during the first 200ns and then continue decrease until 300ns. The Rg then completely stabilized and no significant deviation was observed until the end of simulation. This shows the minimal unbinding events experienced by the Deltacron complex during the simulation thus reveals a binding stability of the RBD in the later part of the simulation. An average Rg for the Deltacron complex was calculated to be 33.28 Å. The Rg for the Deltacron complex is given in **Figure 7B**.

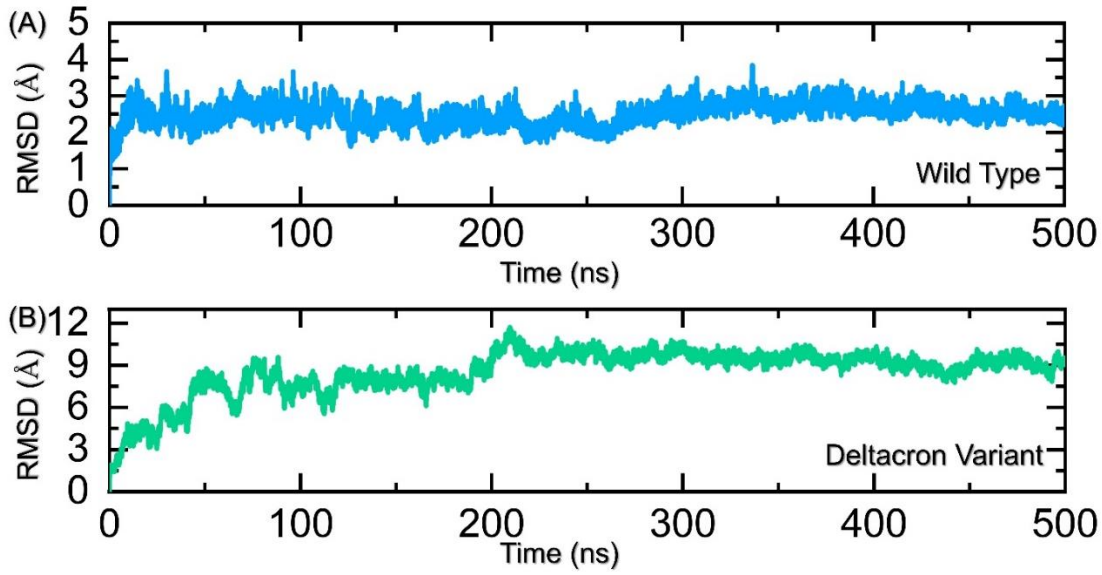

Figure 6: Structural stability of the binding of the wild type and Deltacron-RBD with the hACE2. (A) Shows the RMSD of the wild type RBD in complex with hACE2 while **(B)** Shows the RMSD of the Deltacron RBD in complex with hACE2.

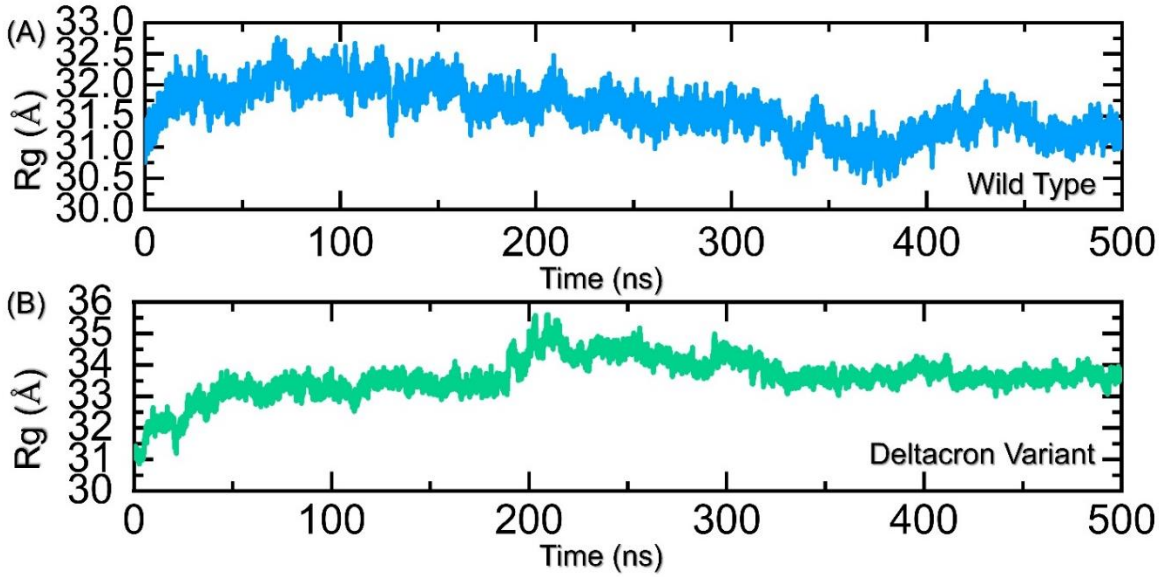

Figure 7: Structural stability of the binding of the wild type and Deltacron-RBD with the hACE2. (A) Shows the RMSD of the wild type RBD in complex with hACE2 while **(B)** Shows the RMSD of the Deltacron RBD in complex with hACE2.

Residues flexibility indexing

Knowledge regarding the protein's residues flexibility is key to deciphering the function of a protein. It helps to elucidate the role of essential residues required for molecular interactions, catalysis, protein design and engineering, protein-protein interaction and molecular recognition. Conformational alterations that span a wide variety of amplitude scales are typically linked to protein function. Protein dynamics has been shown to be crucial to molecular processes, since it is engaged in turnover rate modulation, ligand/target validation, binding, and product release. As a

result, knowing about protein flexibility is just as important as knowing about protein structure when it comes to understanding protein's function and improving drug development[40]. Considering the important role of residues flexibility herein we calculated the RMSF for each complex. As given in **Figure 8A**, the flexibility for the wild type is very minimal for the region particularly (1-200) which is the RBD. While the rest of the residues i.e. from 201-791 a more comparable flexibility can be observed. The flexibility of Deltacron is higher as compared to the wild type. The flexibility behaviour of the Deltacron variant is completely different than the wild type and other variants previously characterized using structural modelling approaches. The RMSF for RBD domain only is shown in **Figure 8B** while the ACE2 is shown in **Figure 8C**. In the case of RBD only the flexibility is completely altered while the region 200-300 in ACE2 of the Deltacron complex reported higher flexibility which is the binding site for RBD. This shows a better conformational optimization of the Deltacron variant for recognition and binding of the RBD to the ACE2, which consequently increases the infectivity.

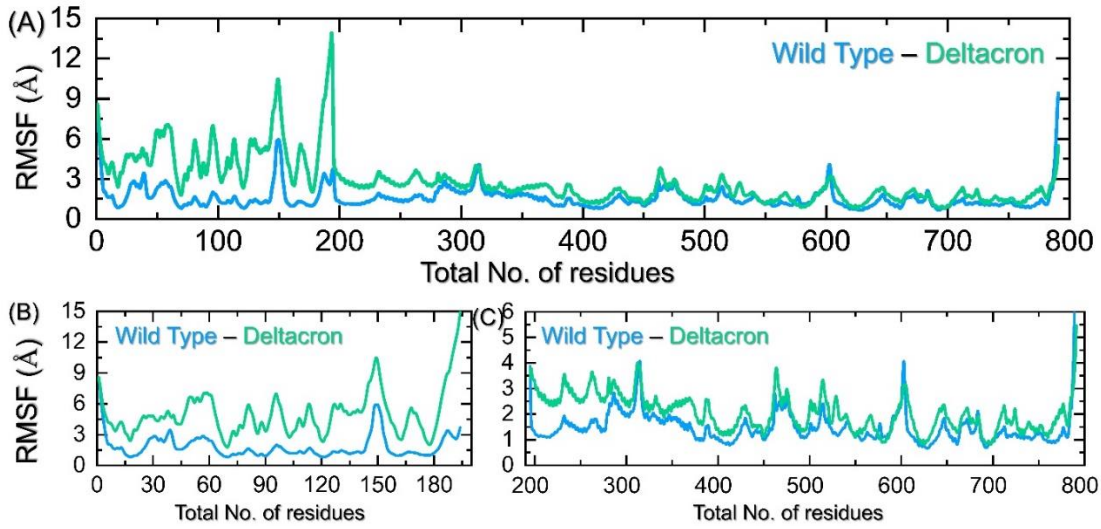

Figure 8: Residues flexibility analysis of the wild type and Deltacron-RBD with the hACE2. (A) Shows the RMSF of the wild type and Deltacron-RBD in complex with hACE2, **(B)** Shows the RMSF of the wild type and Deltacron-RBD only while **(C)** show the RMSF for wild type and Deltacron ACE2 only.

We examined the structures at different time intervals to check the highly flexible regions in the RBD domain of the Deltacron variant. At different time scale i.e. 50ns, 100ns, 150ns, 200ns, 250ns, 300ns, 350ns, 400ns, 450ns and 500ns the structures were retrieved from the trajectory and analyzed for highly moveable parts. As given in **Figure 9A-6D**, the highly dynamic regions are encircled and superimposed on the native structure. The region 471-490 (correspond to 139-158) are the binding loops previously reported to be required for the direct interaction with the hACE2 demonstrated higher flexibility. This region has also been previously reported to have higher flexibility, which consequently increases the binding affinity [10, 16]. Moreover, the two terminal tails i.e. 333-372 (correspond to 1-42) and 516-526 (correspond to 184-194) also demonstrated higher flexibility than the native structure. It can be also seen that these mutations have altered the secondary structure of the RBD mostly transited to the loops thus acquired higher flexible

dynamics than the native structure during the simulation. Consequently, the mutations has induced higher flexibility in the spike glycoprotein that in turn results in altered binding and dynamics to increase the infectivity.

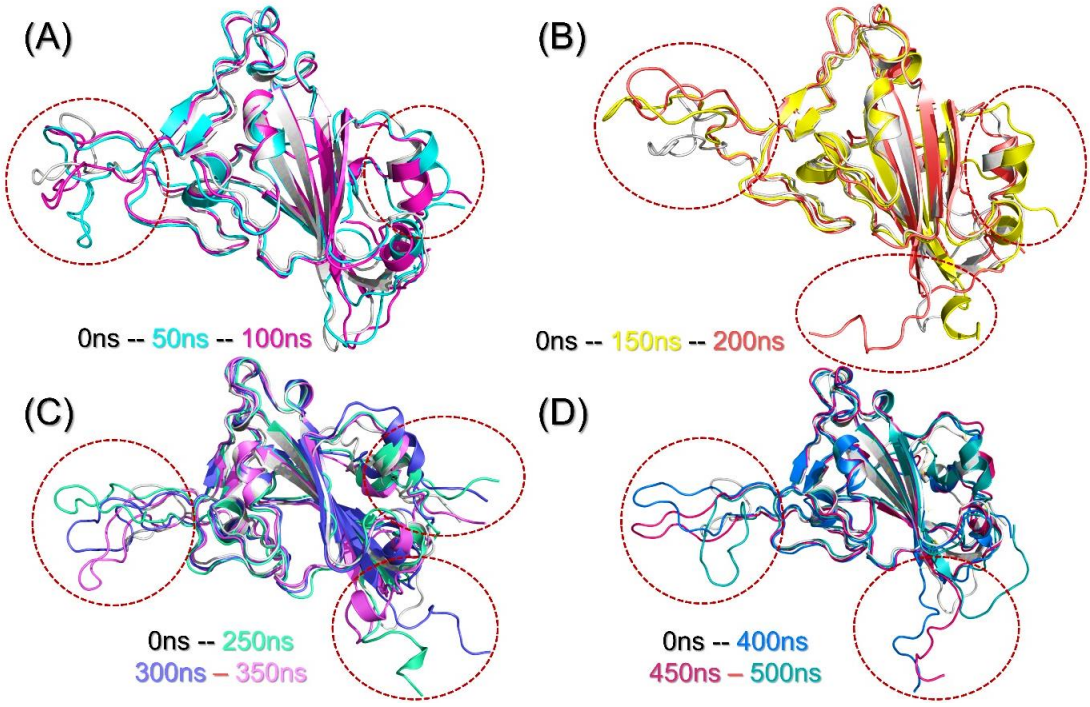

Figure 9: Residues flexibility analysis of the native and structures retrieved at different time intervals. (A) Show the superimposed structures including the native structure, 50ns, and 100ns. **(B)** Show the superimposed structures including the native structure, 150ns and 200ns. **(C)** Show the superimposed structures including the native structure 250ns, 300ns and 350ns while **(D)** Show the superimposed structures including the native structure, 400ns, 450ns and 500ns.

Hydrogen Bonding Analysis

Macromolecular complexes, particularly protein-protein coupling, are primarily driven by numerous factors, among which hydrogen bonding and hydrophobic contacts are essential. The environment of protein interfaces is enriched with water molecules that work with the residues to form hydrogen bonds [41]. The mechanisms underlying protein-protein interaction, as well as the ramifications for hydrogen bonding, are unclear [42]. Whether hydrogen bonds govern protein-protein docking in particular is a long-standing concern, and the mechanism is poorly understood [43, 44]. Thus, it is important to understand the hydrogen bonding landscape in the protein-protein association. For instance, previously, hydrogen bonding was predicted to estimate the strength of the association between two macromolecules, which shed light on the mechanism of pathogenesis induced by different mutations in SARS-CoV-2 variants, including B.1.1.7 B.1.351, P.1, B.1.617, and B.1.618. Here, we have employed a similar approach to understand the differences in hydrogen bonding between the wild type and Deltacron variant complexes. The hydrogen bonding over the simulation time (500ns) is shown in **Figure 10** where the average number of hydrogen bonds in the wild type were calculated to be 375 while the Deltacron variant reported average hydrogen

bonds of 386. This show that the binding of the Deltacron variant is increased during the simulation steered by hydrogen bonds.

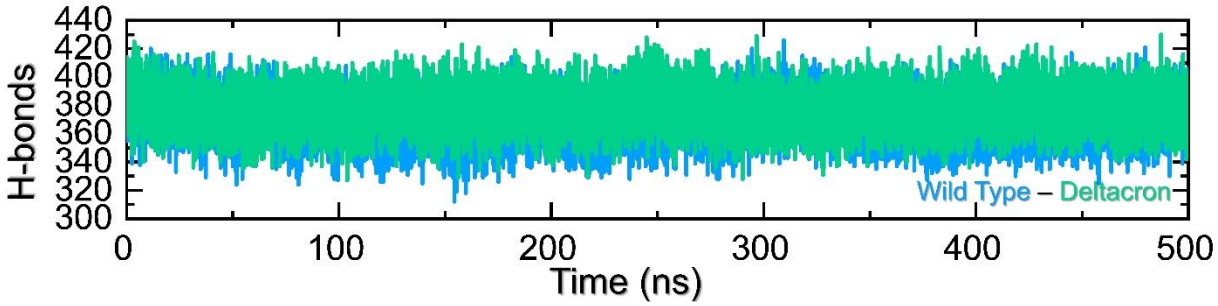

Figure 10: Hydrogen bonding analysis of the wild type and Deltacron variant during the 500ns simulation.

Binding Free Energy Estimation

Determination of the accurate binding energy and validation of docking conformation can be achieved by estimating the binding free energy of the molecular complex. It is a simulation based method which has been reported to be more accurate, cheaper and faster than the conventional approaches such as the alchemical method. The binding estimation for the other variants including the alpha variant, beta, gamma, delta, omicron and others are previosly reported [10, 13, 16, 17, 20]. Considering the accuracy of the MM/GBSA approach we also estimated the binding energy for the wild type and Deltacron variant-RBD with the hACE2. As given in **Table 1**, the binding free enrgy of the wild type RBD is less than the Deltacron varaint. The vdw for the WT-RBD and Deltacron-RBD were reported to be -80.20 kcal/mol and -120.26 kcal/mol respectively. The electrostatic energy for each complex was reported to be -610.36 kcal/mol and -897.15 kcal/mol. This show that the binding of the Deltacron-RBD has been increased due to both the vdW and electrostatic contacts. The total binding free enrgy for each of these complexes i.e. WT-RBD and Deltacron-RBD were reported to be -61.38 kcal/mol and -70.47 kcal/mol which consquently show the higher affinity of Deltacron-RBD for the hACE2 receptor and infectivity. These findings strongly corroborate with the previous published researches where the higher binding by the SARS-CoV-2 variants has been reported due to the acquired mutation in the RBD[10, 13, 16, 17, 20].

Table 2: Binding Free energy results obtained from MM/GBSA analysis.

Complexes Names	vdW	ELE	EGB	SASA	ΔG
Wild Type-RBD-hACE2	-80.20	-610.36	640.96	-11.78	**-61.38**
DELTACRON-RBD-hACE2	-120.26	-897.15	962.85	-15.91	**-70.47**

Trajectories Motion Mapping through PCA

The two PCs were used to construct scatter map of the protein trajectories to understand dominant motions and conformational changes. Due to significant contribution to the total global and dominant motions, only the first two eigenvectors were considered. The first ten eigenvectors for each complex are shown in **Figure 11A**. The first eigenvector contributed 27% (wild type) and 85% (Deltacron) of the total motion. Align with the previous research the first eigenvectors dominated the total motion of the proteins complexes. The eigenvectors were mapped onto scatter

plot where the conformational transition (blue to orange) are shown in **Figure 11B** and **11C**. In the case of wild type in contrast to the Deltacron variant, the structure has occupied less conformational trace space than the wild type. The trace value for the wild type was reported to be 210nm^2 while the Deltacron variant occupied more conformation trace space (300nm^2). The high trace value in the Deltacron variant clearly depict more structural flexibility. The broad range of phase space covered by the Deltacron variant along PC1 and PC2 suggests that these mutations are important in contributing conformational heterogeneity or flexibility that consequently help the variant to bind more efficiently than the wild type.

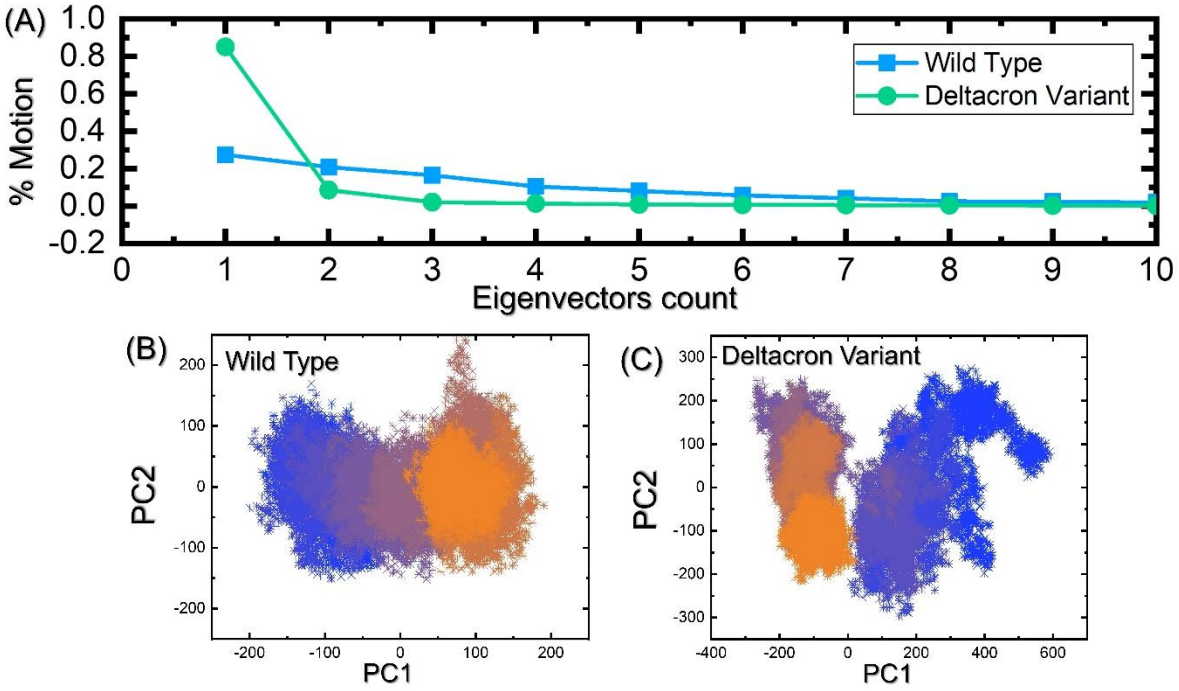

Figure 11: Clustering of the protein's motion in the simulation trajectories. (A) Show the motion contributed by each eigenvector to the total motion. **(B)** Scatter plot for the distribution of trajectories in PC1 and PC2 phase space for the wild type complex, while **(C)** show the scatter plot for the distribution of trajectories in PC1 and PC2 phase space for the Deltacron variant complex.

Conclusions

The Deltacron variant continues to inflict chaos globally due to its rapid transmission and infectivity. The variant is still under investigation and no information on the binding and infectivity are yet disclosed. Hence, deep analysis to understand the binding pattern and to disclose the other features are required. Our analysis revealed that despite the structural resemblance the Deltacron variant established a different bonding network by engaging new residues, which helps to robustly bind. The protein structural graphs revealed variations in the hub residues, number of nodes, inter and intra residues communities, and path communication perturbation caused by the acquired mutations in the Deltacron-RBD thus alter the binding approach and infectivity. Moreover, the dynamic behaviour reported a highly flexibility structure with enhanced residues flexibility particular by the loops required for interaction with ACE2. The binding free energy further

validated the stronger binding of Deltacron by sharing higher binding free energy. The current study is first of its kind to decipher the binding mechanism of Deltacron variant and provide basis for structure-based drug designing.

Funding

Dong-Qing Wei is supported by grants from the Key Research Area Grant 2016YFA0501703 of the Ministry of Science and Technology of China, the National Science Foundation of China (Grant No. 32070662, 61832019, 32030063), the Science and Technology Commission of Shanghai Municipality (Grant No.: 19430750600), as well as SJTU JiRLMDS Joint Research Fund and Joint Research Funds for Medical and Engineering and Scientific Research at Shanghai Jiao Tong University (YG2021ZD02).

Acknowledgements

The computations were partially performed at the PengCheng Lab. and the Center for High-Performance Computing, Shanghai Jiao Tong University.

Availability of data and material

All the data is available on RCSB, UniProt and any simulation data would be provided on reasonable demand. The accession numbers to access this data are given in the manuscript.

Ethics approval and consent to participate

N/A

Consent for publication

N/A

Competing interests

Declared none.

Authors' contributions

References

[1] C. Wang, P.W. Horby, F.G. Hayden, G.F. Gao, A novel coronavirus outbreak of global health concern, Lancet 395(10223) (2020) 470-473.

[2] K. Moelling, Within-Host and Between-Host Evolution in SARS-CoV-2-New Variant's Source, Viruses 13(5) (2021).

[3] J.A. Plante, B.M. Mitchell, K.S. Plante, K. Debbink, S.C. Weaver, V.D. Menachery, The variant gambit: COVID-19's next move, Cell Host Microbe 29(4) (2021) 508-515.

[4] S.R. Kannan, A.N. Spratt, A.R. Cohen, S.H. Naqvi, H.S. Chand, T.P. Quinn, C.L. Lorson, S.N. Byrareddy, K. Singh, Evolutionary analysis of the Delta and Delta Plus variants of the SARS-CoV-2 viruses, Journal of autoimmunity 124 (2021) 102715.

[5] S. Messali, A. Bertelli, G. Campisi, A. Zani, M. Ciccozzi, A. Caruso, F. Caccuri, A cluster of the new SARS‐CoV‐2 B. 1.621 lineage in Italy and sensitivity of the viral isolate to the BNT162b2 vaccine, Journal of Medical Virology (2021).

[6] P.L. Wink, F.C.Z. Volpato, F.L. Monteiro, J.B. Willig, A.P. Zavascki, A.L. Barth, A.F. Martins, First identification of SARS-CoV-2 Lambda (C. 37) variant in southern Brazil, Infection Control & Hospital Epidemiology (2021) 1-2.

[7] Z. Liu, L.A. VanBlargan, L.-M. Bloyet, P.W. Rothlauf, R.E. Chen, S. Stumpf, H. Zhao, J.M. Errico, E.S. Theel, M.J. Liebeskind, Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization, Cell host & microbe 29(3) (2021) 477-488. e4.

[8] T.-J. Yang, P.-Y. Yu, Y.-C. Chang, N.-E. Chang, Y.-X. Tsai, K.-H. Liang, P. Draczkowski, B. Lin, Y.-S. Wang, Y.-C. Chien, Structure-activity relationships of B. 1.617 and other SARS-CoV-2 spike variants, bioRxiv (2021).
[9] M.K. Annavajhala, H. Mohri, P. Wang, M. Nair, J.E. Zucker, Z. Sheng, A. Gomez-Simmonds, A.L. Kelley, M. Tagliavia, Y. Huang, Emergence and expansion of SARS-CoV-2 B. 1.526 after identification in New York, Nature 597(7878) (2021) 703-708.
[10] A. Khan, T. Zia, M. Suleman, T. Khan, S.S. Ali, A.A. Abbasi, A. Mohammad, D.Q. Wei, Higher infectivity of the SARS-CoV-2 new variants is associated with K417N/T, E484K, and N501Y mutants: An insight from structural data, J Cell Physiol 236(10) (2021) 7045-7057.
[11] C. Scheepers, J. Everatt, D.G. Amoako, H. Tegally, C.K. Wibmer, A. Mnguni, A. Ismail, B. Mahlangu, B.E. Lambson, S.I. Richardson, Emergence and phenotypic characterization of C. 1.2, a globally detected lineage that rapidly accumulated mutations of concern, medRxiv (2021) 2021.08. 20.21262342.
[12] P. Colson, P.-E. Fournier, J. Delerce, M. Million, M. Bedotto, L. Houhamdi, N. Yahi, J. Bayette, A. Levasseur, J. Fantini, Culture and identification of a Deltamicron SARS-CoV-2 in a three cases cluster in southern France, medRxiv (2022).
[13] A. Khan, A. Mohammad, I. Haq, M. Nasar, W. Ahmad, Q. Yousafi, M. Suleman, S. Ahmad, A. Albutti, T. Khan, S.K. Marafie, E. Alshawaf, S.S. Ali, J. Abubaker, D.Q. Wei, Structural-Dynamics and Binding Analysis of RBD from SARS-CoV-2 Variants of Concern (VOCs) and GRP78 Receptor Revealed Basis for Higher Infectivity, Microorganisms 9(11) (2021).
[14] I.M. Ibrahim, A.A. Elfiky, A.M. Elgohary, Recognition through GRP78 is enhanced in the UK, South African, and Brazilian variants of SARS-CoV-2; An in silico perspective, Biochemical and Biophysical Research Communications 562 (2021) 89-93.
[15] A.A. Elfiky, SARS-CoV-2 spike-heat shock protein A5 (GRP78) recognition may be related to the immersed human coronaviruses, Frontiers in pharmacology 11 (2020) 1997.
[16] A. Khan, H. Waris, M. Rafique, M. Suleman, A. Mohammad, S.S. Ali, T. Khan, Y. Waheed, C. Liao, D.-Q. Wei, The Omicron (B. 1.1. 529) variant of SARS-CoV-2 binds to the hACE2 receptor more strongly and escapes the antibody response: Insights from structural and simulation data, International Journal of Biological Macromolecules (2022).
[17] A. Khan, J. Gui, W. Ahmad, I. Haq, M. Shahid, A.A. Khan, A. Shah, A. Khan, L. Ali, Z. Anwar, 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 Advances 11(48) (2021) 30132-30147.
[18] C. Dominguez, R. Boelens, A.M. Bonvin, HADDOCK: a protein− protein docking approach based on biochemical or biophysical information, Journal of the American Chemical Society 125(7) (2003) 1731-1737.
[19] P.I. Koukos, M.F. Reau, A.M. Bonvin, Shape-restrained modelling of protein-small molecule complexes with HADDOCK, bioRxiv (2021).
[20] A. Khan, M.T. Khan, S. Saleem, M. Junaid, A. Ali, S.S. Ali, M. Khan, D.-Q. Wei, Structural Insights into the mechanism of RNA recognition by the N-terminal RNA-binding domain of the SARS-CoV-2 nucleocapsid phosphoprotein, Computational and Structural Biotechnology Journal (2020).
[21] R.A. Laskowski, PDBsum: summaries and analyses of PDB structures, Nucleic acids research 29(1) (2001) 221-222.
[22] R. Salomon-Ferrer, A.W. Götz, D. Poole, S. Le Grand, R.C. Walker, Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald, Journal of chemical theory and computation 9(9) (2013) 3878-3888.
[23] R. Salomon‐Ferrer, D.A. Case, R.C. Walker, An overview of the Amber biomolecular simulation package, Wiley Interdisciplinary Reviews: Computational Molecular Science 3(2) (2013) 198-210.

[24] A. Khan, T. Zia, M. Suleman, T. Khan, S.S. Ali, A.A. Abbasi, A. Mohammad, D.-Q. Wei, Higher infectivity of the SARS-CoV-2 new variants is associated with K417N/T, E484K, and N501Y mutants: An insight from structural data, Journal of Cellular Physiology n/a(n/a) (2021).
[25] D.R. Roe, T.E. Cheatham III, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, Journal of chemical theory and computation 9(7) (2013) 3084-3095.
[26] T. Hou, J. Wang, Y. Li, W. Wang, Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations, Journal of chemical information and modeling 51(1) (2011) 69-82.
[27] A. Khan, W. Heng, Y. Wang, J. Qiu, X. Wei, S. Peng, S. Saleem, M. Khan, S.S. Ali, D.-Q. Wei, In silico and in vitro evaluation of kaempferol as a potential inhibitor of the SARS-CoV-2 main protease (3CLpro), Phytotherapy research: PTR.
[28] A. Khan, M. Junaid, C.-D. Li, S. Saleem, F. Humayun, S. Shamas, S.S. Ali, Z. Babar, D.-Q. Wei, Dynamics insights into the gain of flexibility by Helix-12 in ESR1 as a mechanism of resistance to drugs in breast cancer cell lines, Frontiers in Molecular Biosciences 6 (2020) 159.
[29] A. Khan, Z. Rehman, H.F. Hashmi, A.A. Khan, M. Junaid, A.M. Sayaf, S.S. Ali, F.U. Hassan, W. Heng, D.-Q. Wei, An Integrated Systems Biology and Network-Based Approaches to Identify Novel Biomarkers in Breast Cancer Cell Lines Using Gene Expression Data, Interdisciplinary Sciences: Computational Life Sciences (2020) 1-14.
[30] A. Khan, M. Khan, S. Saleem, Z. Babar, A. Ali, A.A. Khan, Z. Sardar, F. Hamayun, S.S. Ali, D.-Q. Wei, Phylogenetic analysis and structural perspectives of RNA-dependent RNA-polymerase inhibition from SARs-CoV-2 with natural products, Interdisciplinary Sciences: Computational Life Sciences 12(3) (2020) 335-348.
[31] A. Khan, M. Junaid, A.C. Kaushik, A. Ali, S.S. Ali, A. Mehmood, D.-Q. Wei, Computational identification, characterization and validation of potential antigenic peptide vaccines from hrHPVs E6 proteins using immunoinformatics and computational systems biology approaches, PloS one 13(5) (2018).
[32] I. Hussain, N. Pervaiz, A. Khan, S. Saleem, H. Shireen, D.-Q. Wei, V. Labrie, Y. Bao, A.A. Abbasi, Evolutionary and structural analysis of SARS-CoV-2 specific evasion of host immunity, Genes & Immunity (2020) 1-11.
[33] S. Wold, K. Esbensen, P. Geladi, Principal component analysis, Chemometrics and intelligent laboratory systems 2(1-3) (1987) 37-52.
[34] K. Pearson, LIII. On lines and planes of closest fit to systems of points in space, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 2(11) (1901) 559-572.
[35] M.A. Balsera, W. Wriggers, Y. Oono, K. Schulten, Principal component analysis and long time protein dynamics, The Journal of Physical Chemistry 100(7) (1996) 2567-2572.
[36] M. Ernst, F. Sittel, G. Stock, Contact-and distance-based principal component analysis of protein dynamics, The Journal of chemical physics 143(24) (2015) 12B640_1.
[37] S. Lata, M. Akif, Probing structural basis for enhanced binding of SARS‐CoV‐2 P. 1 variant spike protein with the human ACE2 receptor, Journal of Cellular Biochemistry.
[38] T.N. Starr, A.J. Greaney, S.K. Hilton, D. Ellis, K.H. Crawford, A.S. Dingens, M.J. Navarro, J.E. Bowen, M.A. Tortorici, A.C. Walls, Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding, Cell 182(5) (2020) 1295-1310. e20.
[39] I. Celik, R. Yadav, Z. Duzgun, S. Albogami, A.M. El-Shehawi, Fatimawali, R. Idroes, T.E. Tallei, T.B. Emran, Interactions of the Receptor Binding Domain of SARS-CoV-2 Variants with hACE2: Insights from Molecular Docking Analysis and Molecular Dynamic Simulation, Biology 10(9) (2021) 880.
[40] A. Bornot, C. Etchebest, A.G. de Brevern, Predicting protein flexibility through the prediction of local structures, Proteins 79(3) (2011) 839-852.
[41] D. Chen, N. Oezguen, P. Urvil, C. Ferguson, S. Dann, T. Savidge, Regulation of protein-ligand binding affinity by hydrogen bond pairing. Sci. Adv. 2 (3), e1501240, 2016.

[42] J.D. Chodera, D.L. Mobley, Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design, Annual review of biophysics 42 (2013) 121-142.
[43] R. Patil, S. Das, A. Stanley, L. Yadav, A. Sudhakar, A.K. Varma, Optimized hydrophobic interactions and hydrogen bonding at the target-ligand interface leads the pathways of drug-designing, PloS one 5(8) (2010) e12029.
[44] T.S. Olsson, J.E. Ladbury, W.R. Pitt, M.A. Williams, Extent of enthalpy–entropy compensation in protein–ligand interactions, Protein Science 20(9) (2011) 1607-1618.