Complete Docking Report on SARS-CoV-2 Mutation

Docking Study SARA-CoV-2 protein with Ligand:

Molecular docking was performed to obtain more insights about interactions between the protein Spike Glycoprotein (7KDJ-Mutant) (PDBID: D614G), Spike Glycoprotein (PDBID: G476S), Spike Glycoprotein (PDBID: V483G), Spike Glycoprotein (PDBID: E484K), Spike Glycoprotein (PDBID: P681H), Spike Glycoprotein (PDBID: K417N), Spike Glycoprotein (PDBID: N501Y), Spike Glycoprotein (PDBID: N440K) with nine ligands. Molecular docking study were carried out by AutoDock 4.2 software, using the implemented empirical free energy function and the Lamarckian Genetic Algorithm (LGA). Molecular docking score of Spike Glycoprotein (7KDJ-Mutant)(PDBID: D614G), Spike Glycoprotein (7KDJ-Mutant)(PDBID: D614G),Spike Glycoprotein (PDBID: G476S),Spike Glycoprotein (PDBID: V483G),Spike Glycoprotein (PDBID: E484K), Spike Glycoprotein (PDBID: P681H), Spike Glycoprotein (PDBID: K417N), Spike Glycoprotein (PDBID: N501Y), Spike Glycoprotein (PDBID: N440K) were found to be -8.57, -7.76, -8.00, -7.18, -6.94, -7.54, 7.88 and -7.96 respectively.

1] D614G Mutation:

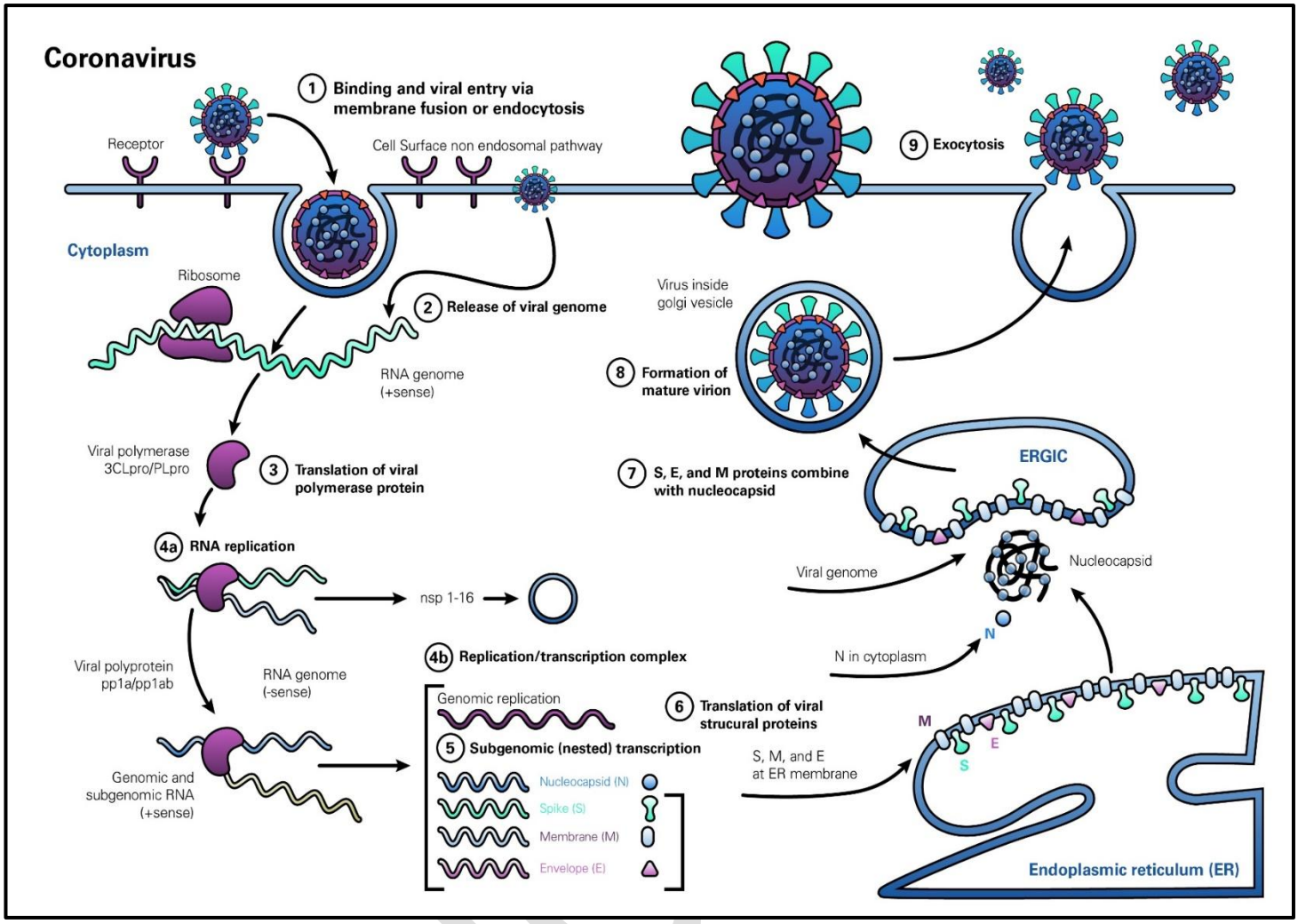
The spike aspartic acid-614 to glycine (D614G) substitution is prevalent in global severe acute respiratory syndrome coronavirus 2 SARS-CoV-2 strains. The variant exhibits more efficient infection, replication, and competitive fitness in primary human airway epithelial cells but maintains similar morphology and in vitro neutralization properties, compared with the ancestral wild-type virus. Infection of human angiotensin-converting enzyme 2 (ACE2) transgenic mice and Syrian hamsters with both viruses resulted in similar viral titers in respiratory tissues and pulmonary disease. However, the D614G variant transmits significantly faster and displayed increased competitive fitness than the wild type virus in hamsters. The data show that the D614G substitution enhances SARS-CoV-2 infectivity, competitive fitness, and transmission in primary human cells and animal models.

The D614G mutation is associated with the B.1 lineage of SARS-CoV-2 which now dominates the global pandemic, based upon global SARS-CoV-2 genome sequences shared via GISAID Retrospectively sampled viruses suggest this mutation was present in Guangzhou, Sichuan, and Shanghai Provinces, China in late January. In Europe, the 614G variant was first observed in genomes sampled on January 28 in a small outbreak in Bavaria, Germany, which was initiated by a visitor from Shanghai and subsequently controlled through public health efforts. It is therefore likely that the D614G mutation occurred in China before being introduced on multiple occasions to European countries where it increased in frequency. This scenario is consistent with the rapid increase in February and March of European virus genomes that carry the 614G variant. In the United

Kingdom, the first observation of a genome carrying the D614G mutation was in a sample collected on February 28 from a patient in Scotland who had recently traveled through Italy.

The D614G mutation in SARS-CoV-2, a non-synonymous mutation resulting in a replacement of aspartic acid with glycine at position 614 of the virus's spike protein (D614G). The trimeric spike protein, composed of subunits S1 and S2, is a large glycoprotein that mediates cell entry and has been studied extensively in other corona viruses, including SARS-CoV and Midde East respiratory syndrome (MERS). SARS-CoV-2 spike protein binds to angiotensin-converting enzyme 2 (ACE2) to gain cell entry, hence mutations in this gene have the potential to alter receptor binding affinity and infectivity, as well as viral immune evasion and immunogenicity.

During the first wave of COVID-19 cases 71 percent of the SARS-CoV-2 particles identified in patients in Houston had the D614G mutation. However, in the second wave of the outbreak during the summer, this variant had 99.9 percent prevalence. The researchers say this finding mirrors a trend observed around the world; a study published in July based on more than 28,000 genome sequences found that variants carrying the D614G mutation became the globally dominant form of SARS-CoV-2 in about a month.

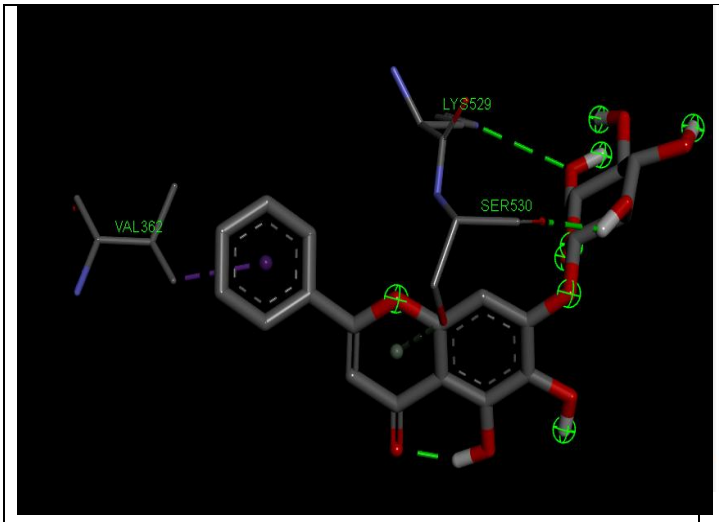


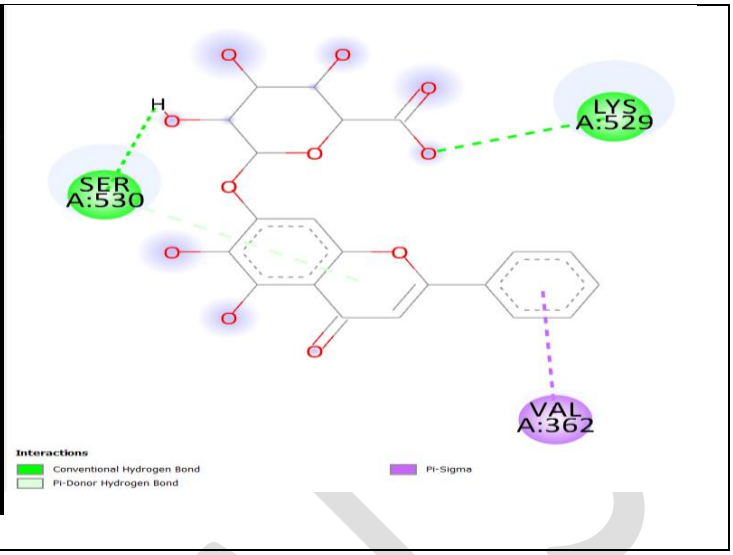
SN	Protein Info	PDB ID	Compounds Name								
			Comp 1	Comp 2	Comp 3	Comp 4	Comp 5	Comp 6	Comp 7	Comp 8	Comp 9
	SARS-CoV-2 Mutation										
1	**Spike Glycoprotein (7KDJ-Mutant)**	**D614G**	-7.08	-7.06	-6.41	-2.87	-4.34	-4.83	-5.40	-5.14	-8.57

SARS-CoV-2 Mutation

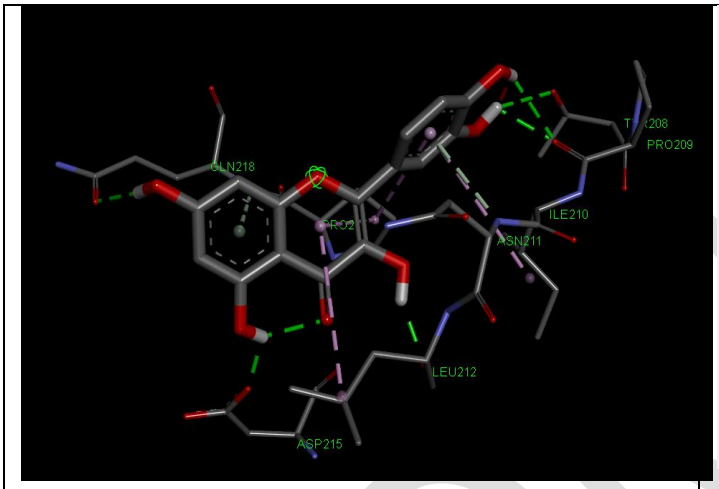
Docking Study 1: Spike (Glycoprotein (7KDJ-Mutant)) 7KDJ with Ligand Baicalin, Quercetin, Luteolin, Rutin, Hesperidin, Curcumin, EGCG, Piperine, Glycyrrhizic

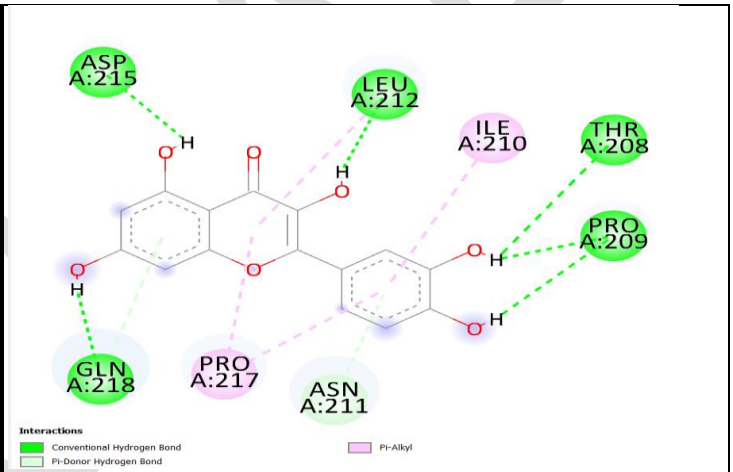
1.Baicalin



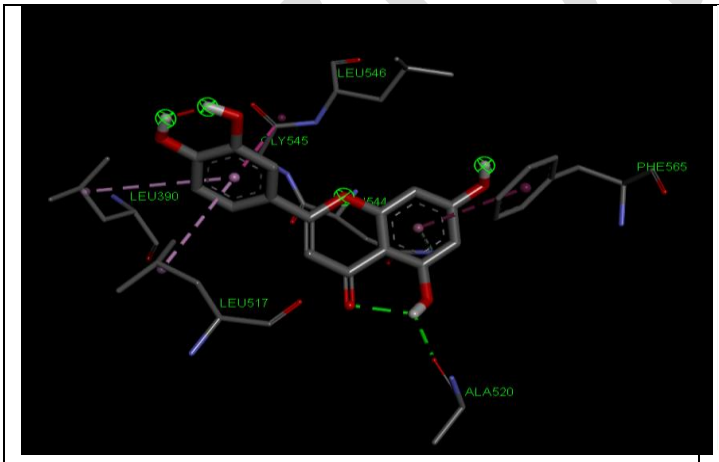


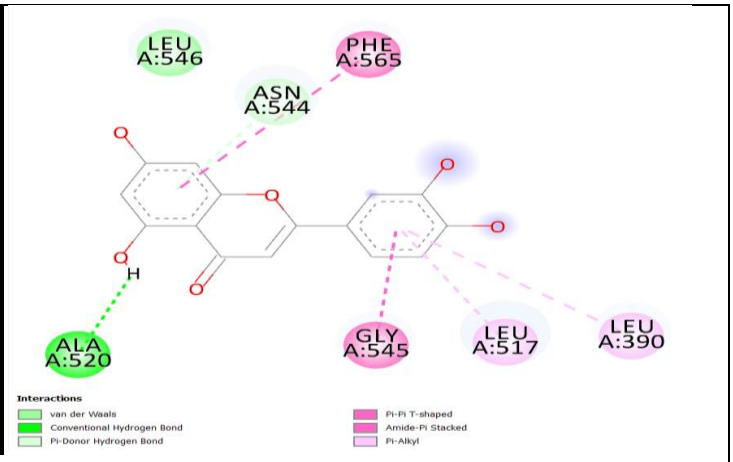
2.Quercitin



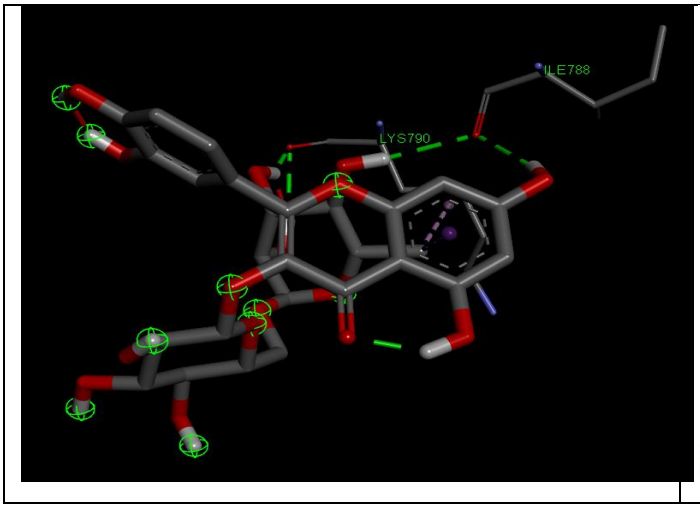


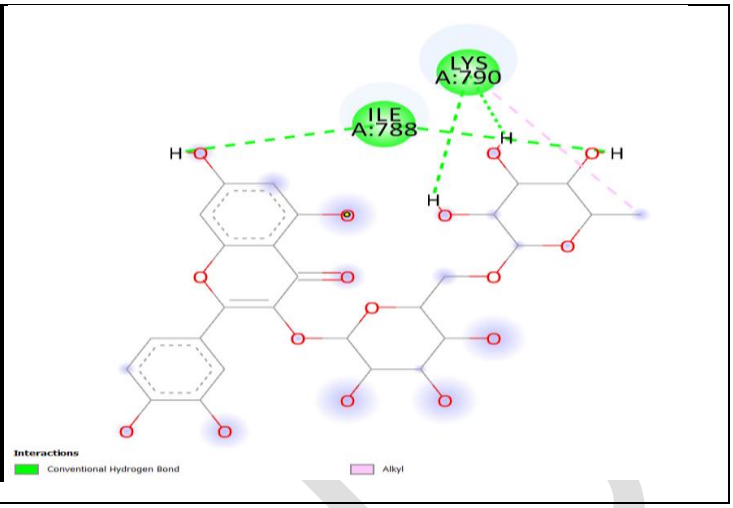
3.Leuteolin



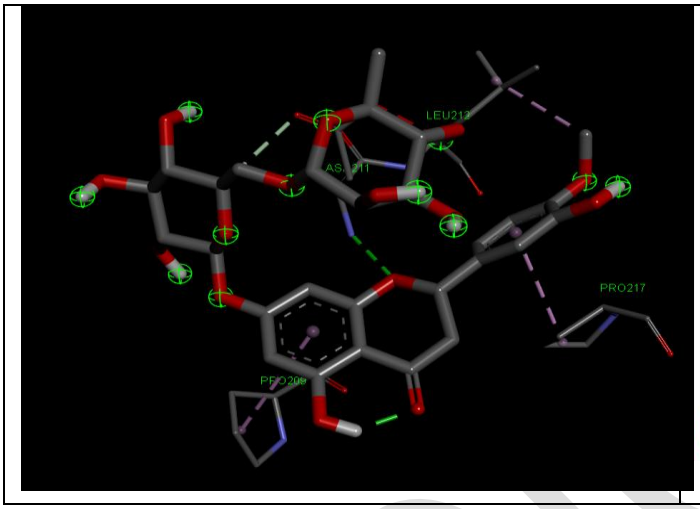


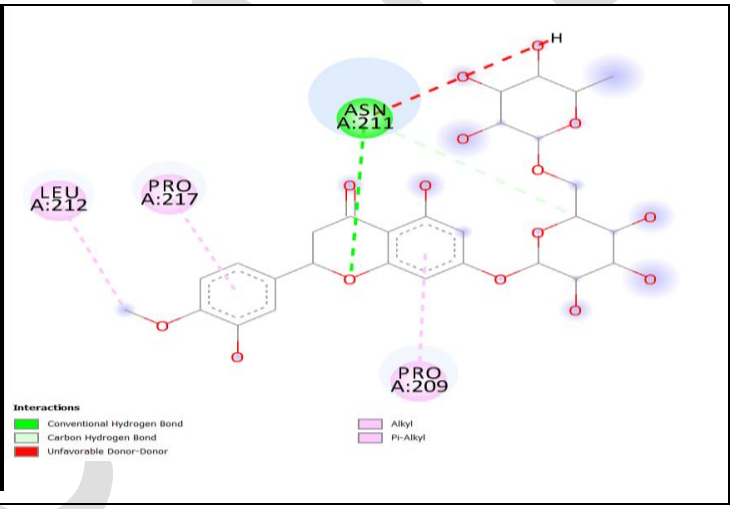
4.Rutin



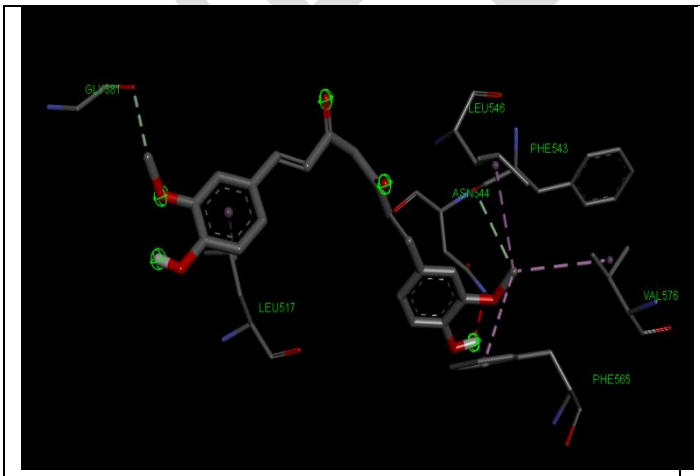


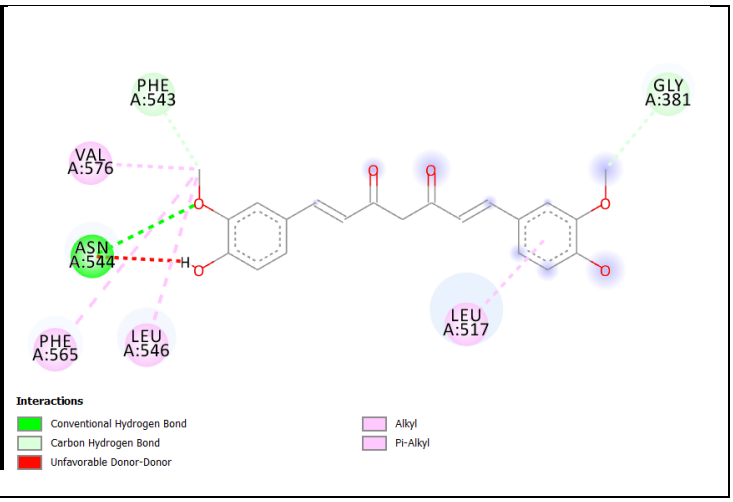
5. Hesperidin



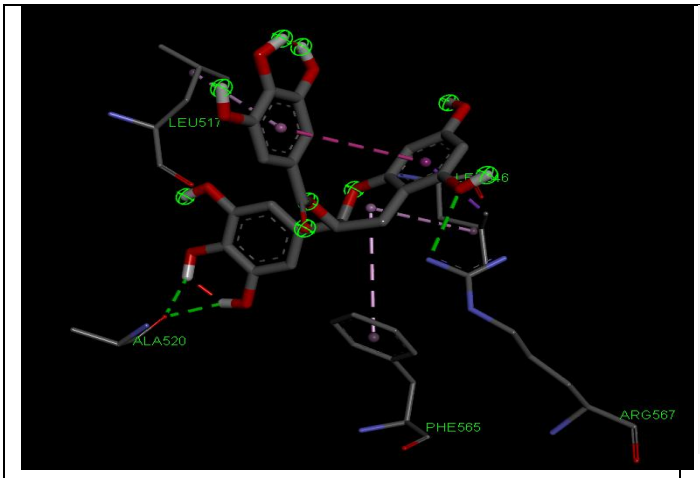


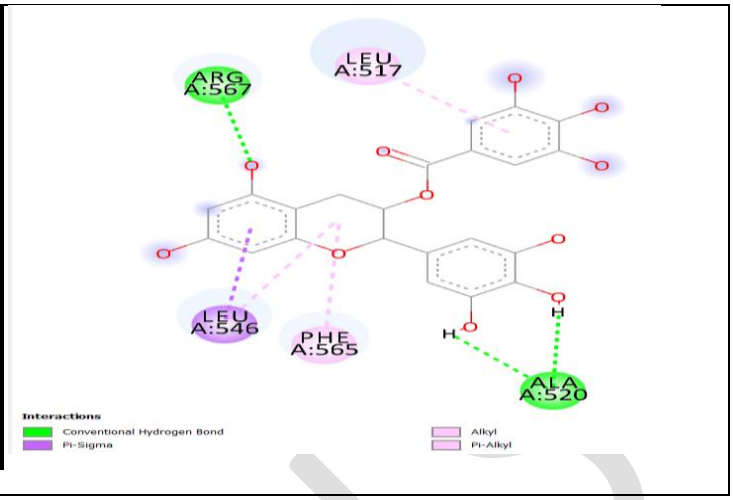
6. Curcumin



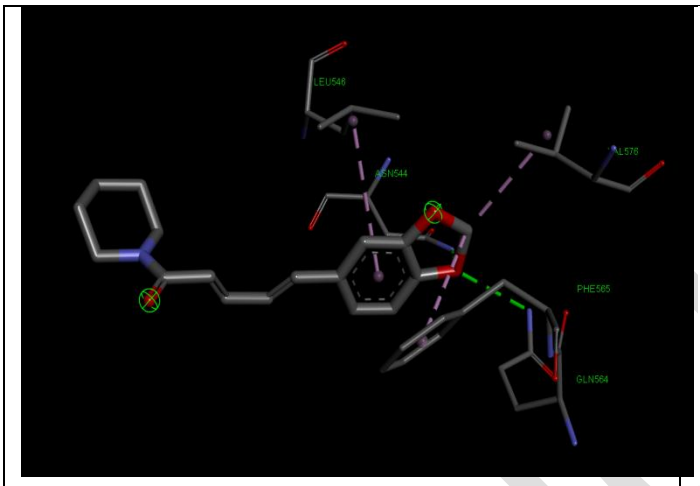


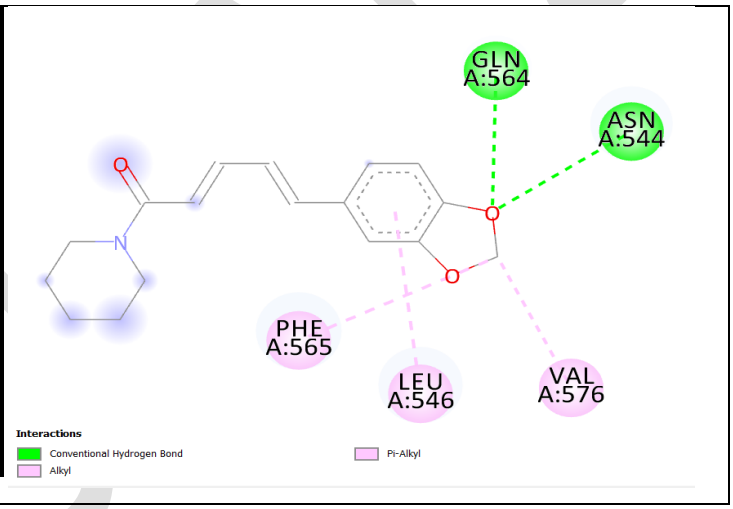
7. EGCG



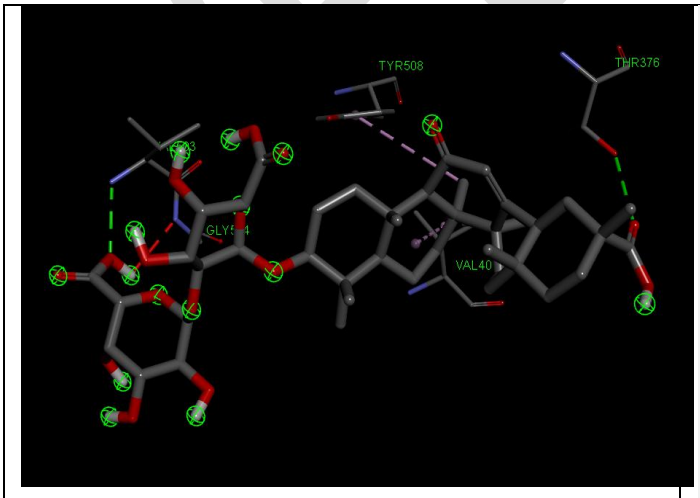


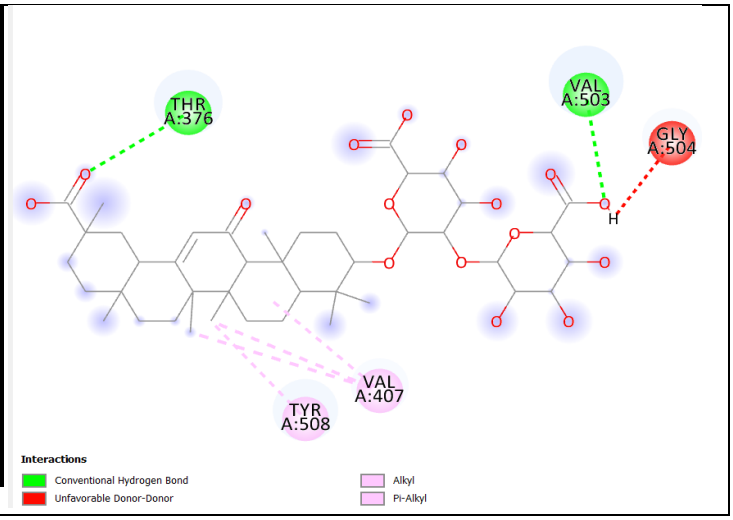
8.Piperine





9. Glycyrrhizic





2] G476S Mutation:

The substitution of Gly-476 to Ser-476 in the SARS-CoV-2 Receptor-binding domain (RBD) largely affected the structural dynamics of the S-protein leading to significant influence on the interactions with ACE-2 and neutralizing antibodies. Structural properties of the S-protein such as conformation changes, residual fluctuations and residue surface area largely varied between the wild-type and G476S variant, especially in the RBD. Analyses of the interaction energies between S-protein and ACE-2 suggest that the G476S variant may have enhanced interactions with ACE-2 compared to the wild-type. The G476S variant was found to have weaker interactions with the neutralizing antibody H014 compared to the wild-type.

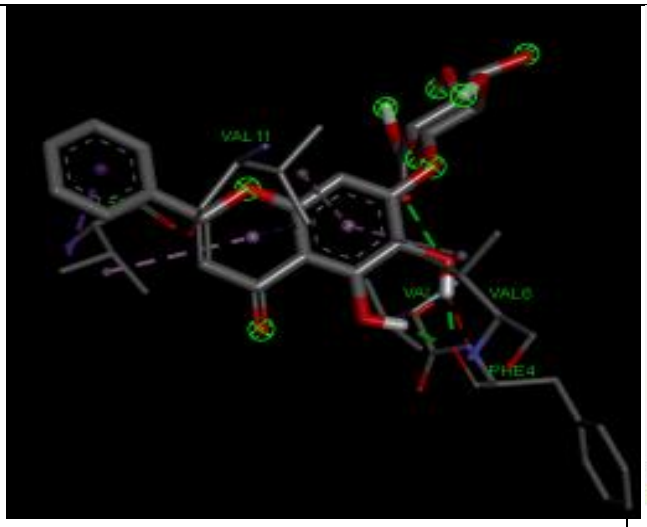
Using integrative computational analyses investigated the structural dynamics of the G476S variant of SARS-CoV-2 and assessed the influence of this mutation on the interactions with ACE-2 and neutralizing antibodies. The substitution of Gly-476 to Ser-476 in the RBD of SARS-CoV-2 S-protein has serious implications with respect to the infectivity and antigenicity of the virus. In the case of viral infectivity, which is initiated by an interaction with the host receptor ACE-2, the G476S variant showed stronger interactions with the ACE-2 compared to the wild-type. On the neutralization of the virus by antibodies, the results suggest that the influence is dependent on the epitopes recognized by the antibody. The surveillance of the G476S mutations should be increased to track the distribution and spread of the G476S variant.

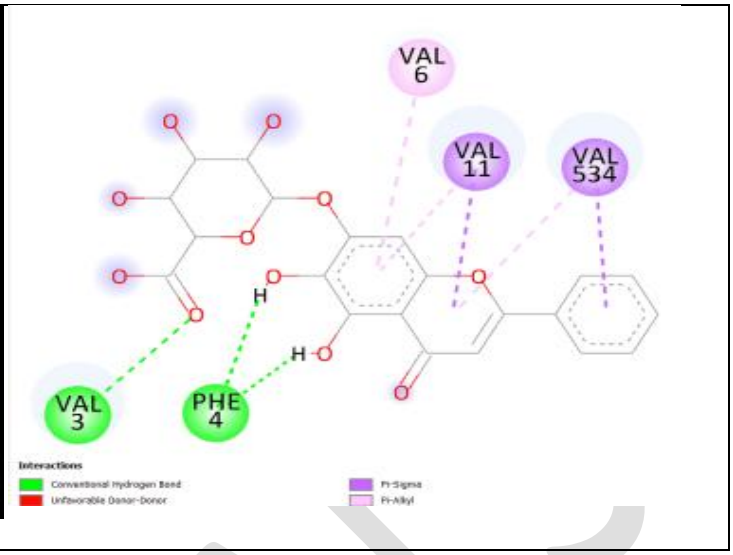
Epitope mapping analysis revealed G476S mutation as antigenic determinants and thus the mutations are important while designing a therapeutic vaccine or chimeric antibody. The findings will help in further understanding the role of such arising mutations in modulating immunogenicity, viral tropism and pathogenesis of the disease, which in lieu will help in designing vaccine more precisely to mitigate pandemic COVID-19.

SN	Protein Info	PDB ID	Compounds Name								
			Comp 1	Comp 2	Comp 3	Comp 4	Comp 5	Comp 6	Comp 7	Comp 8	Comp 9
	SARS-CoV-2 Mutation										
3	**Spike Glycoprotein**	**G476S**	-7.43	-7.40	-7.01	-4.18	-5.06	-6.48	-6.80	-6.60	-7.76

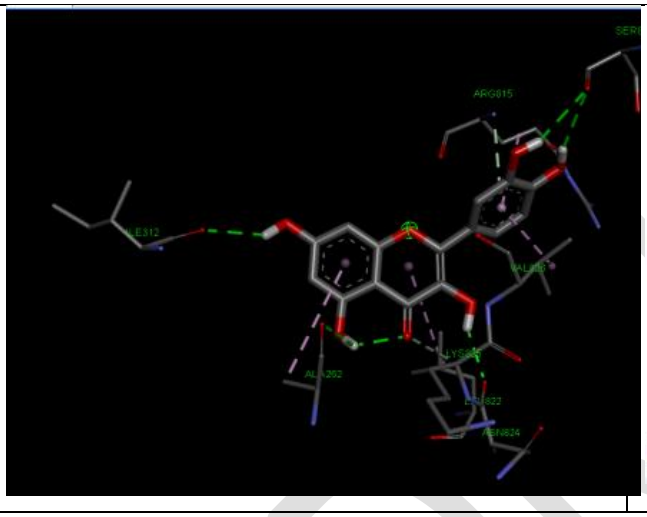
Docking Study 2: Spike Gycoprotein PROGS (G476S) with Ligand Baicalin, Quercetin, Luteolin, Rutin, Hesperidin, Curcumin, EGCG, Piperine, Glycyrrhizic

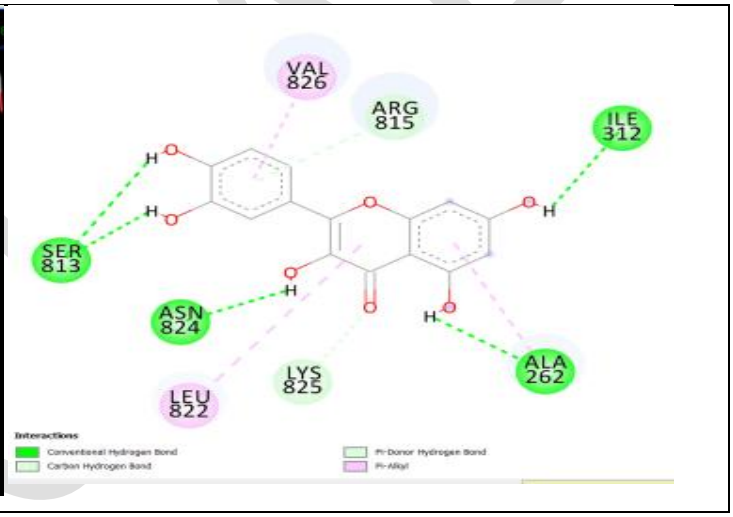
1.Baicalin



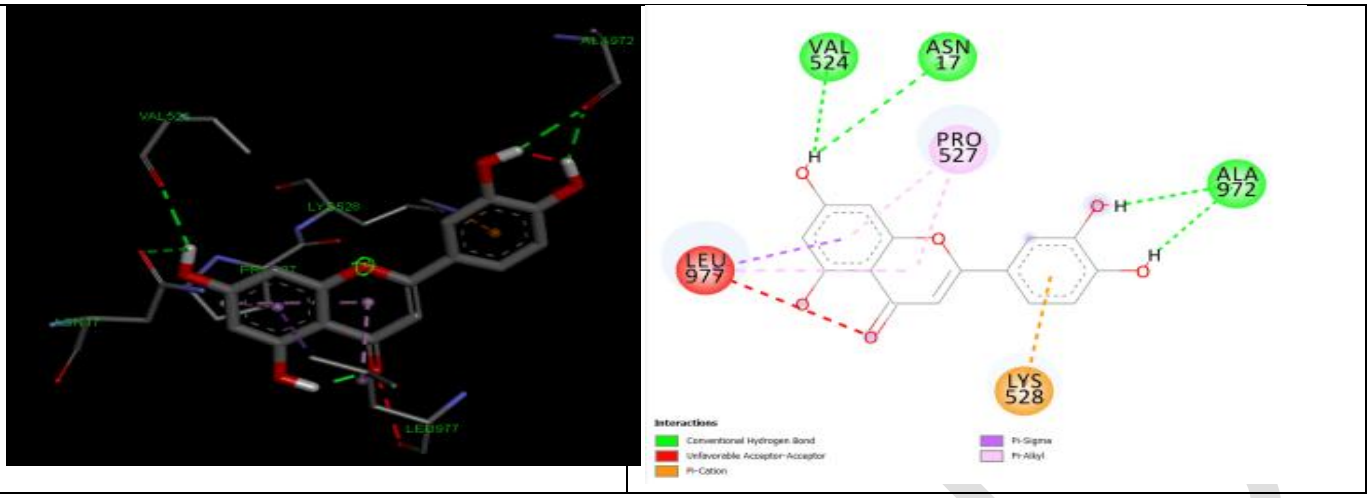


2.Quercitin

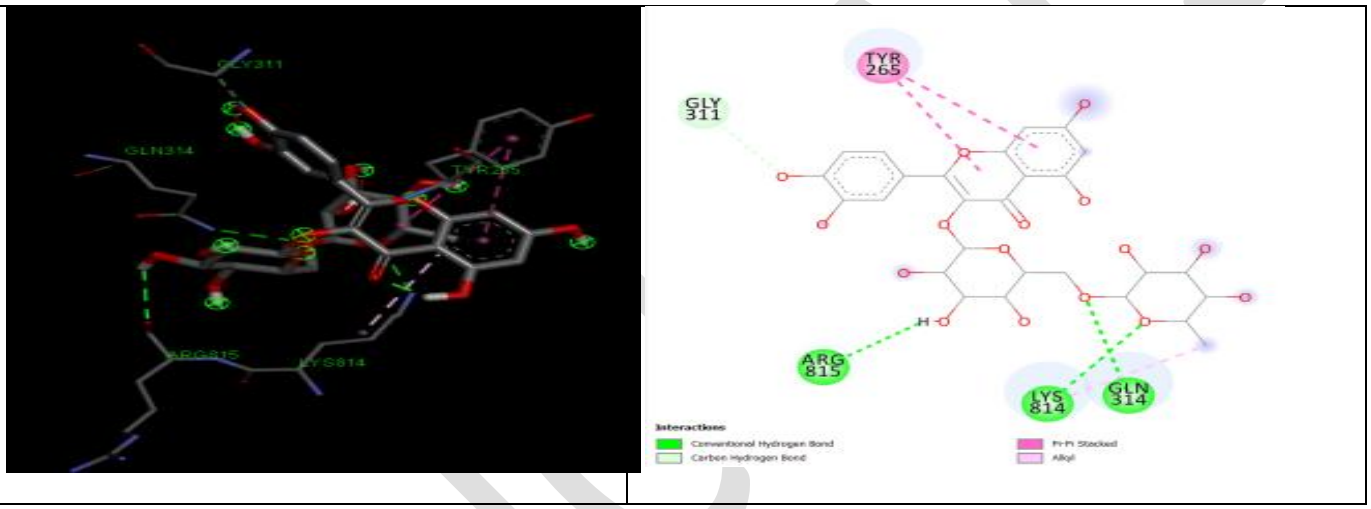




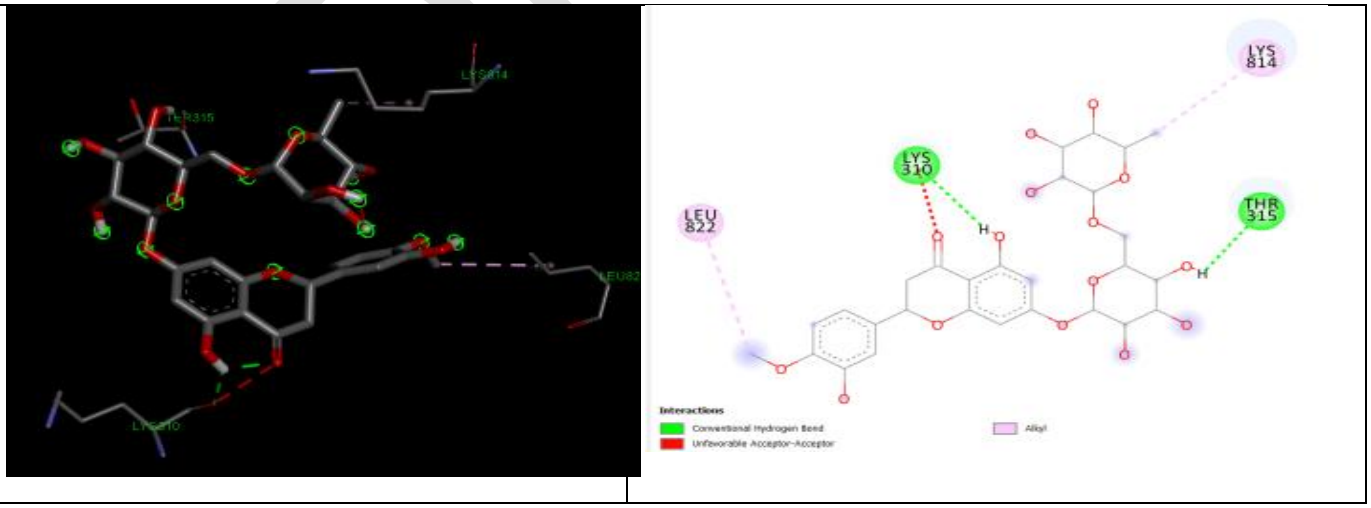
3.Leuteolin



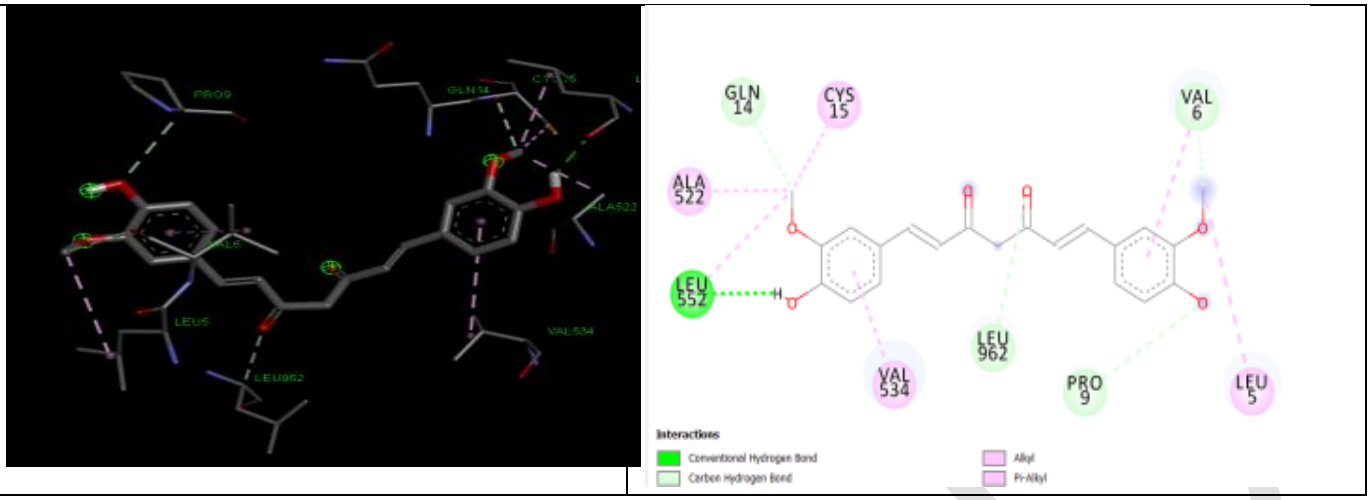
4.Rutin



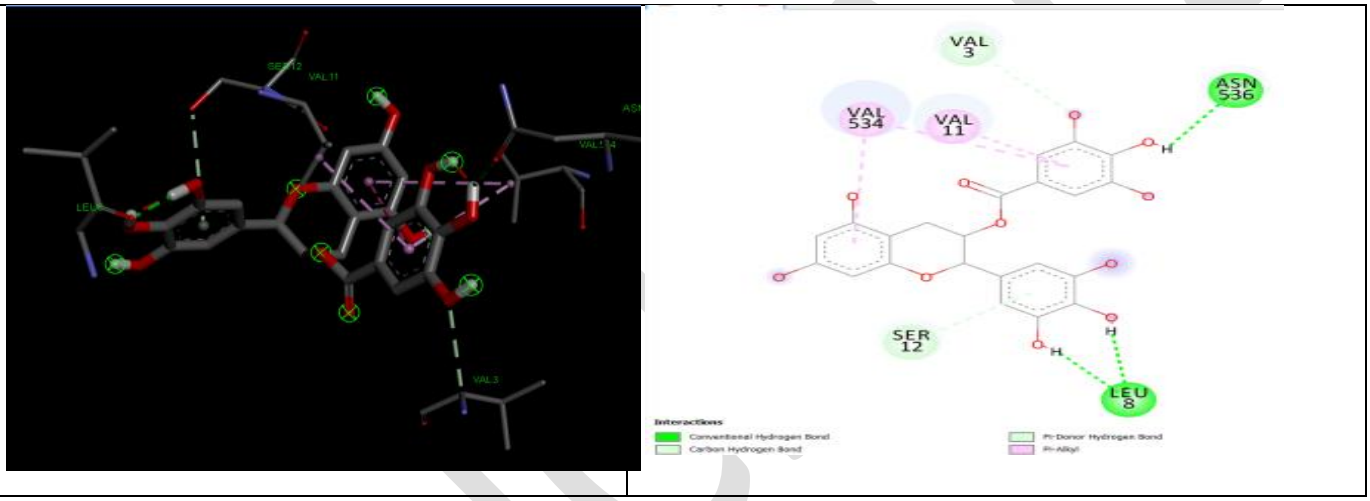
5. Hesperidin



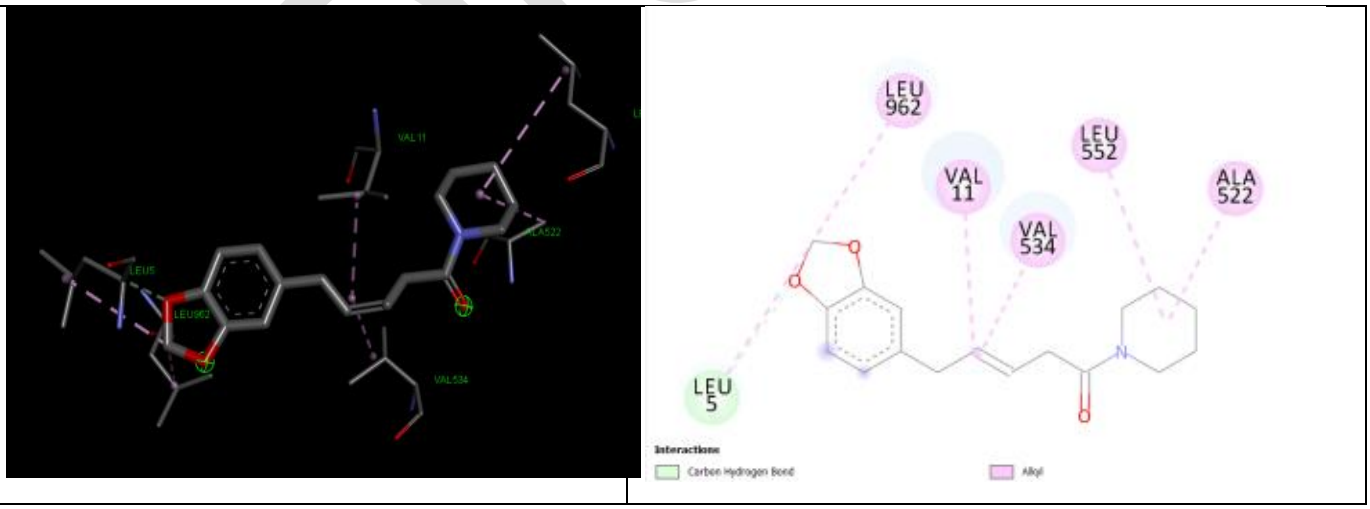
6. Curcumin



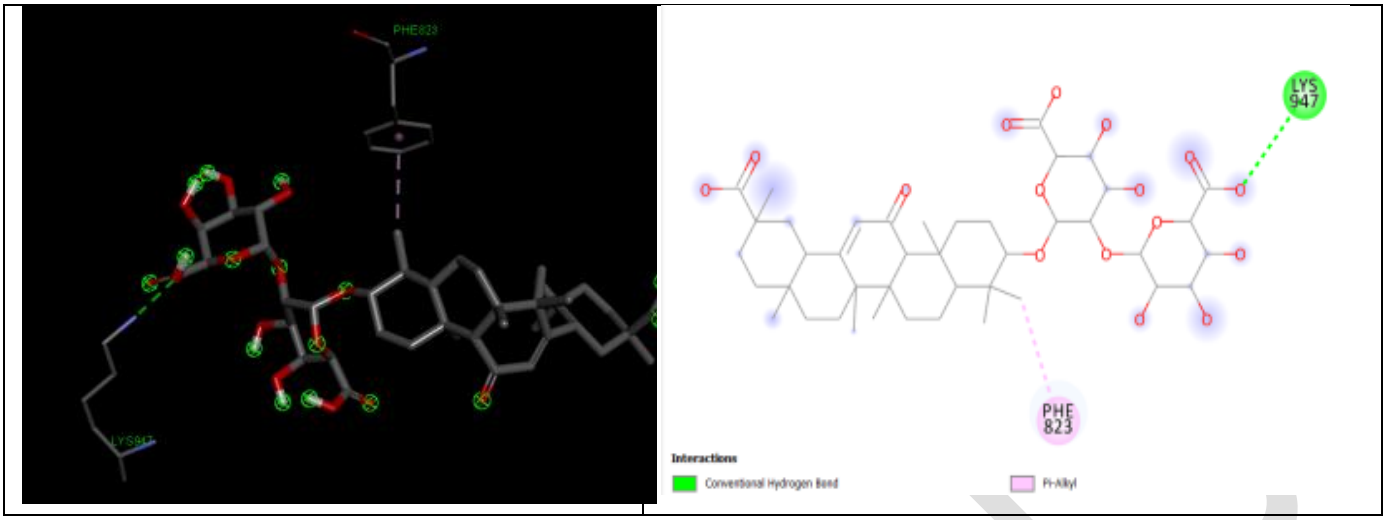
7. EGCG



8.Piperine



9. Glycyrrhizic

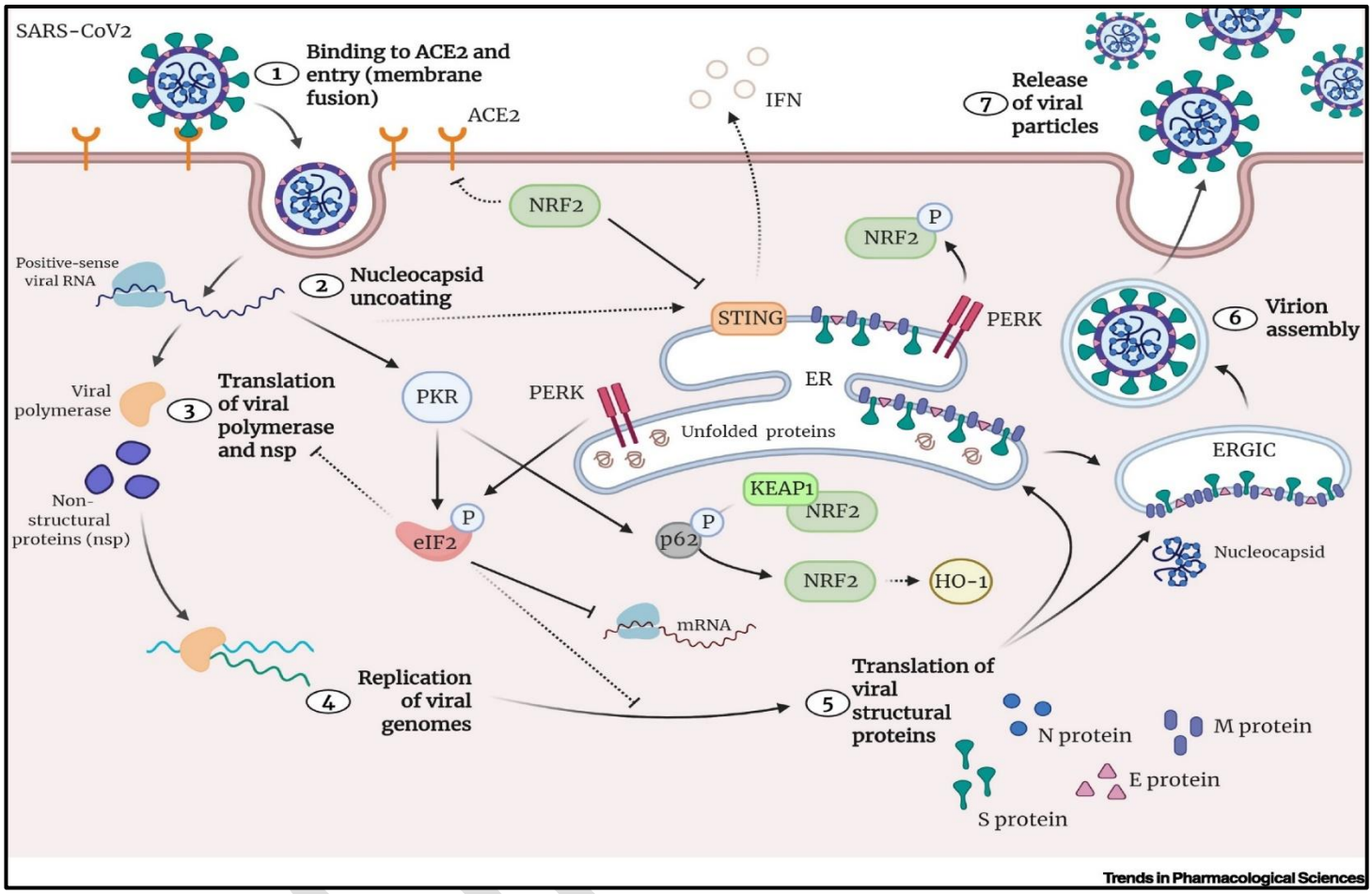


4] <u>V483G Mutation</u>:

V483A is an important amino acid residue in the RBM region of the spike glycoprotein, where the Valine at position 483 has changed to Alanine, making the viral genome a unique mutant strain. V483A is one of the few mutations that have the potential to change the protein secondary structure and relative solvent accessibility in the RBM region. The RBM makes the contacts between the SARS-CoV-2 and the human ACE2 receptor acting as the core binding site. Furthermore, the RNA replication rate causes the virus to mutate at a faster rate evading host immunity, thereby posing strong drug resistance. This mutagenic capability of the virus has become the leading cause of its evolution and genomic variation. Interestingly MD simulation supports strong favorable interaction of ACE2 with RBD region containing V483A mutation. Radius of gyration analysis also showed high degree of compactness in V483A. The landscape plot and Gibbs free energy also support findings. Overall, study indicates that V483G in the RBD region can enhance its binding with the human ACE2 receptor. V483A mutation led to enhanced and broadens the virus host cell entry and transmission of the disease. . Further epitope mapping analysis revealed the mutation as antigenic determinants and thus the mutations are important while designing a therapeutic vaccine or chimeric antibody. The findings will help in further understanding the role of such arising mutations in modulating immunogenicity, viral tropism and pathogenesis of the disease, which in lieu will help in designing vaccine more precisely to mitigate pandemic COVID-19.

Researchers and health officials are warning that that a mutant strain of the SARS-CoV-2 coronavirus, the V483G strain is fast making appearances in various countries including India, Brazil, The Middle-East, UK and even in the United States. According to genomic experts and immunologist, the V483 strain is resistant to the neutralizing properties of antibodies and is also deemed to be even more infectious than the D614G strains. Of the various emerging mutated strains, a mutation that is worrying and concerning that has occurred in the viral genome is the V483A mutation, which is a part of the receptor binding motif (RBM), present in the S1 domain

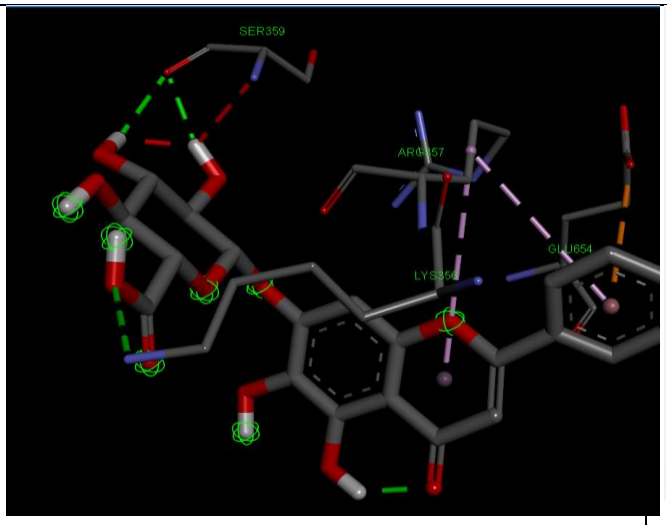
of the spike protein. This V483A mutant virus is becoming popular in North America with 36 cases detected in random sequencing studies recently so far and considering that many sequencing studies have not been conducted, its prevalence could be far more extensive.

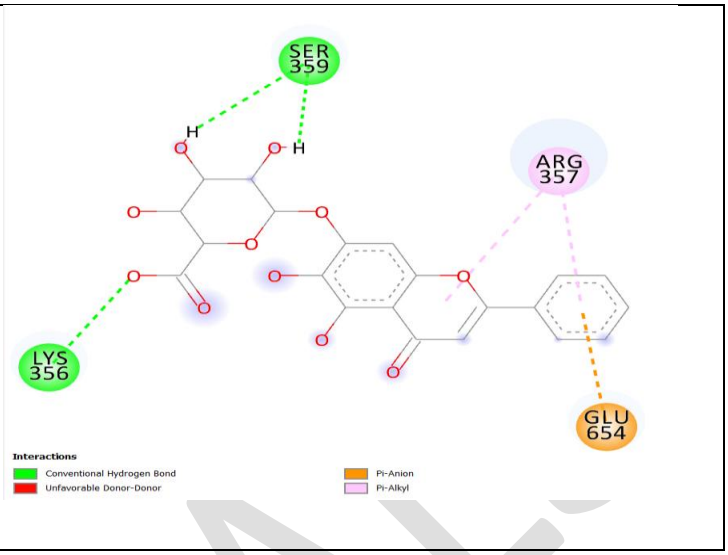


SN	Protein Info	PDB ID	Compounds Name								
			Comp 1	Comp 2	Comp 3	Comp 4	Comp 5	Comp 6	Comp 7	Comp 8	Comp 9
	SARS-CoV-2 Mutation										
4	Spike Glycoprotein	V483G	-7.70	-7.29	-7.54	-4.35	-5.59	-6.97	-6.36	-7.72	-8.00

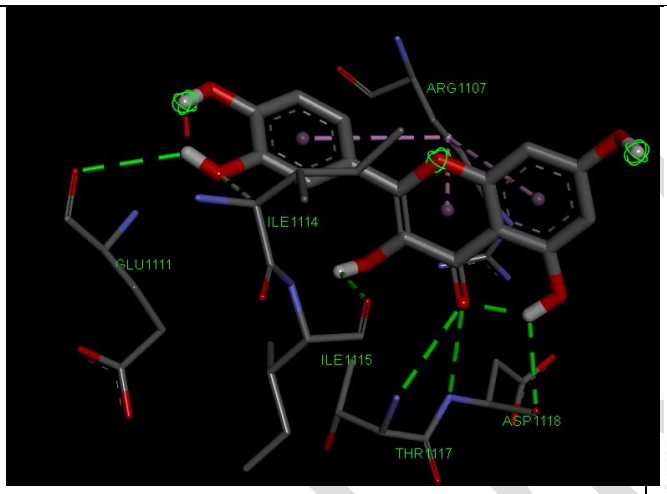
Docking Study 3: Spike Gycoprotein PROV (V483G) with Ligand Baicalin, Quercetin, Luteolin, Rutin, Hesperidin, Curcumin, EGCG, Piperine, Glycyrrhizic

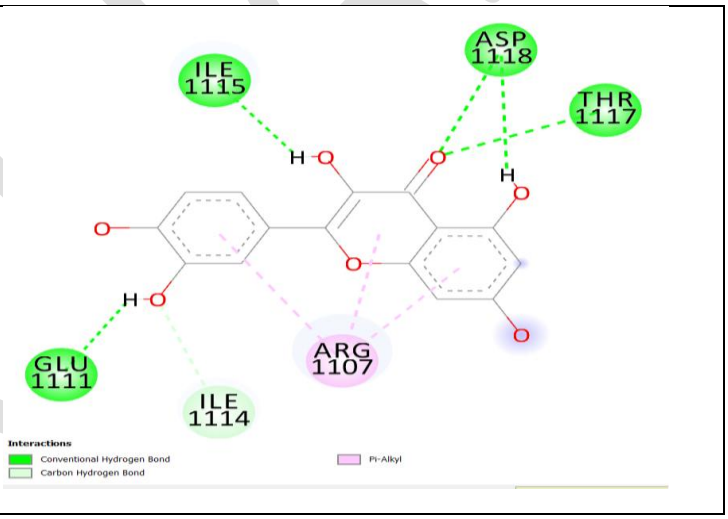
1.Baicalin



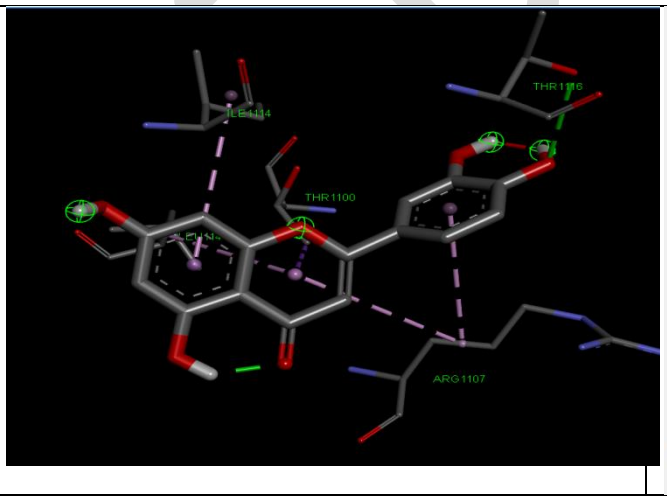


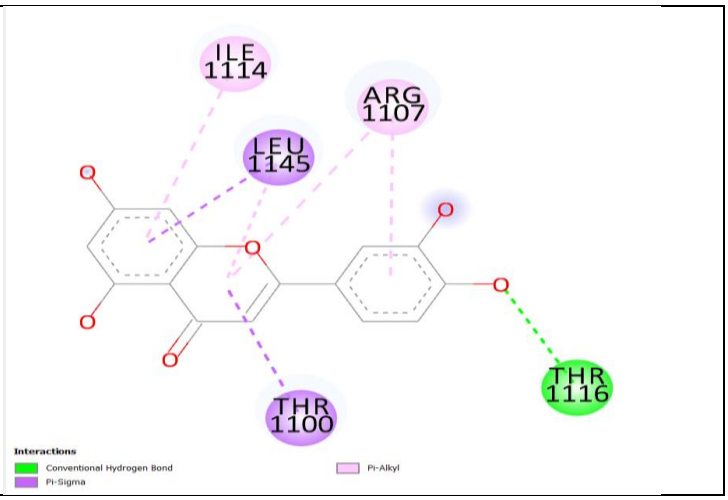
2.Quercitin



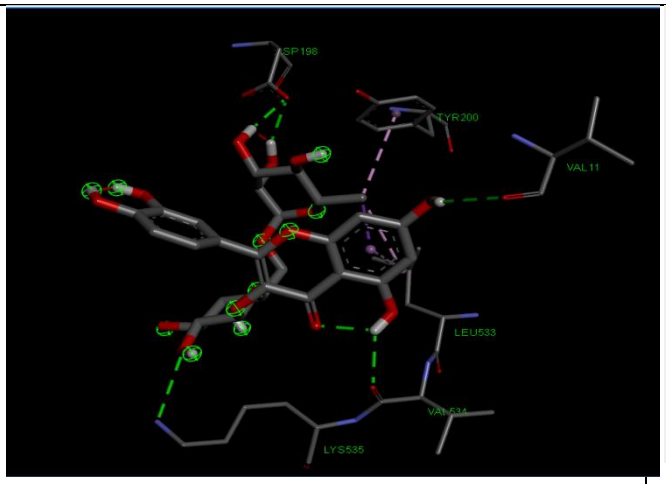


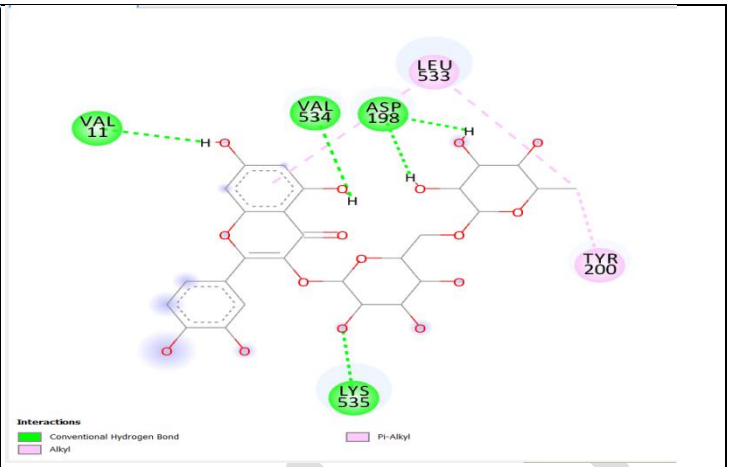
3.Leuteolin



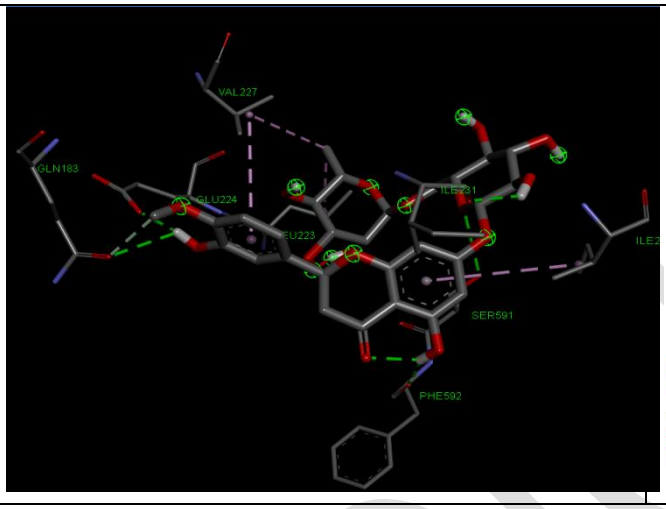


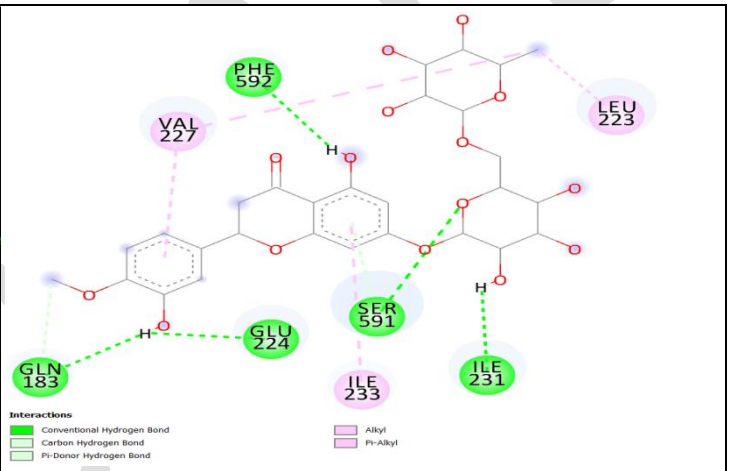
4.Rutin



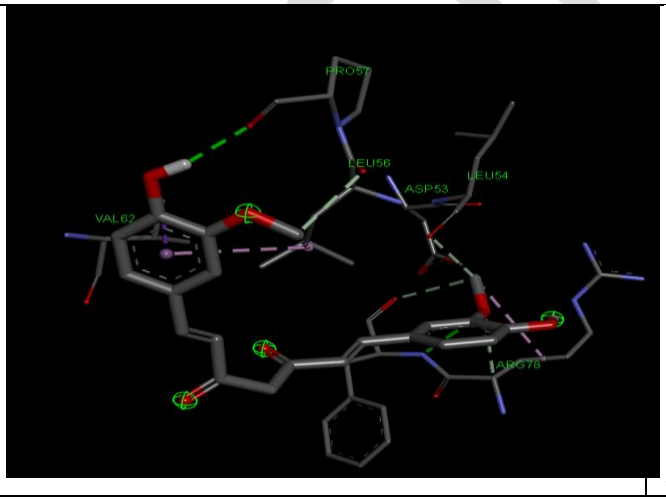


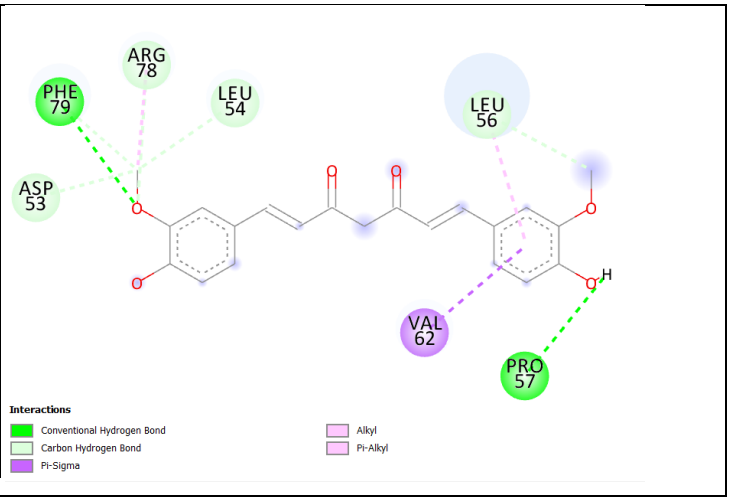
5. Hesperidin



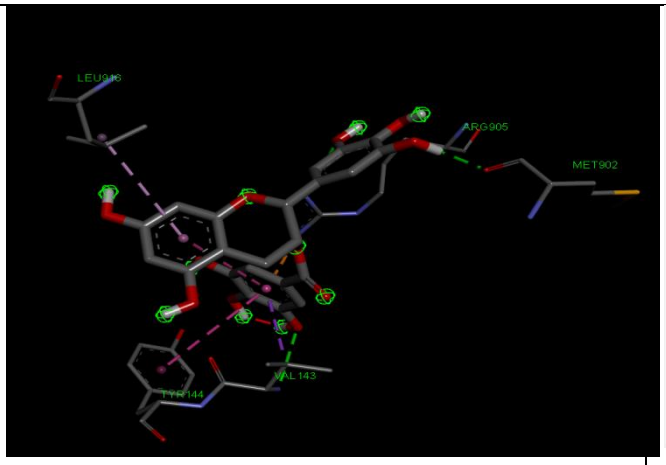


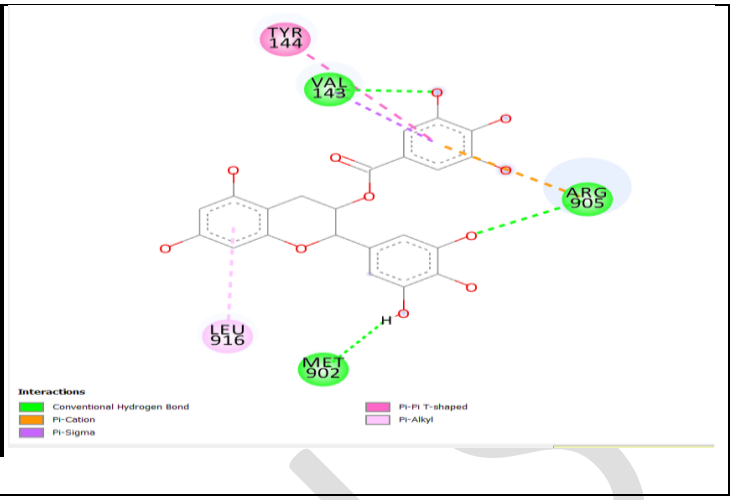
6. Curcumin



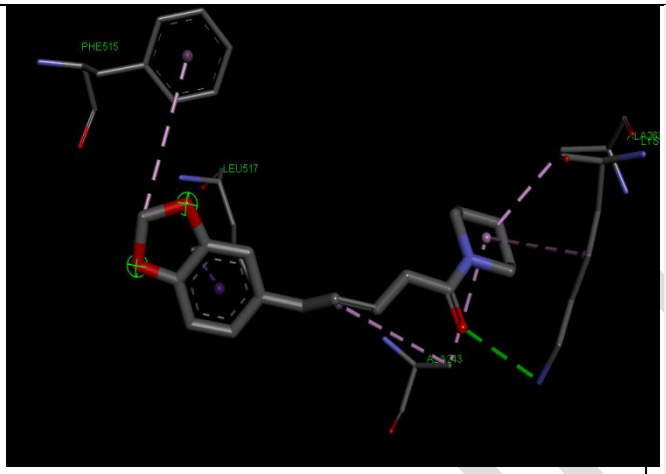


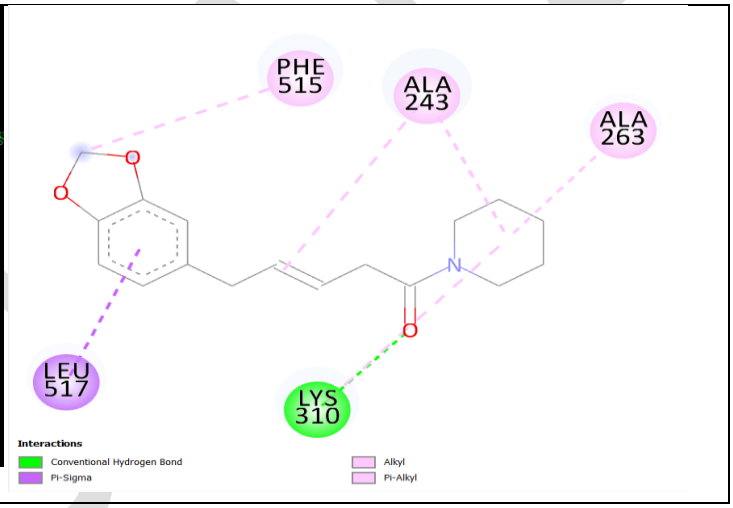
7. EGCG



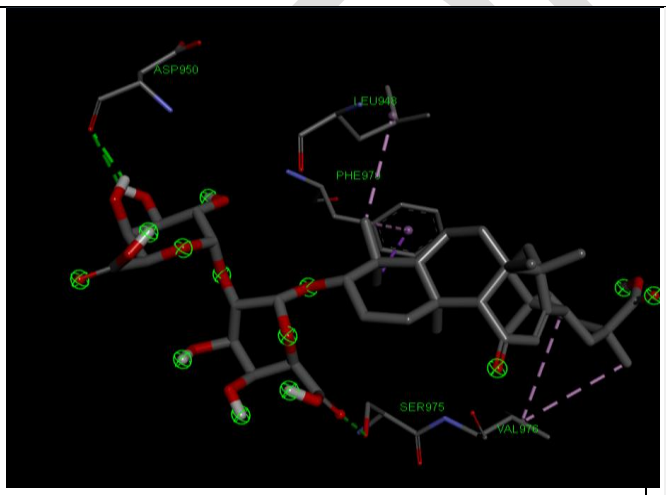


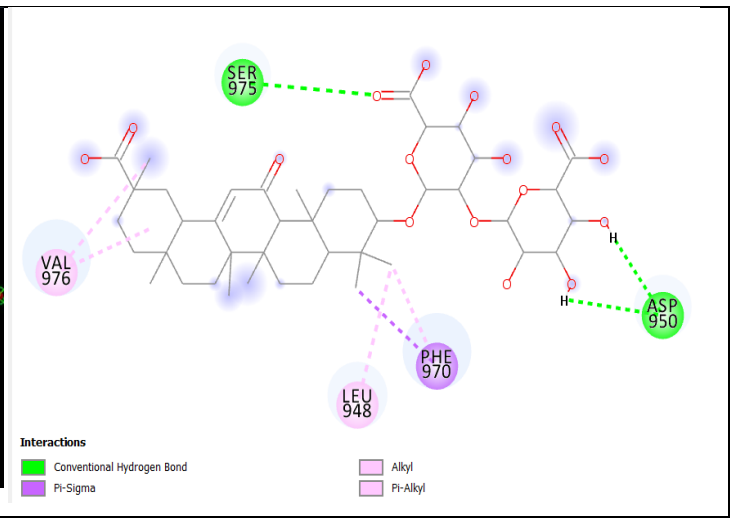
8.Piperine





9. Glycyrrhizic



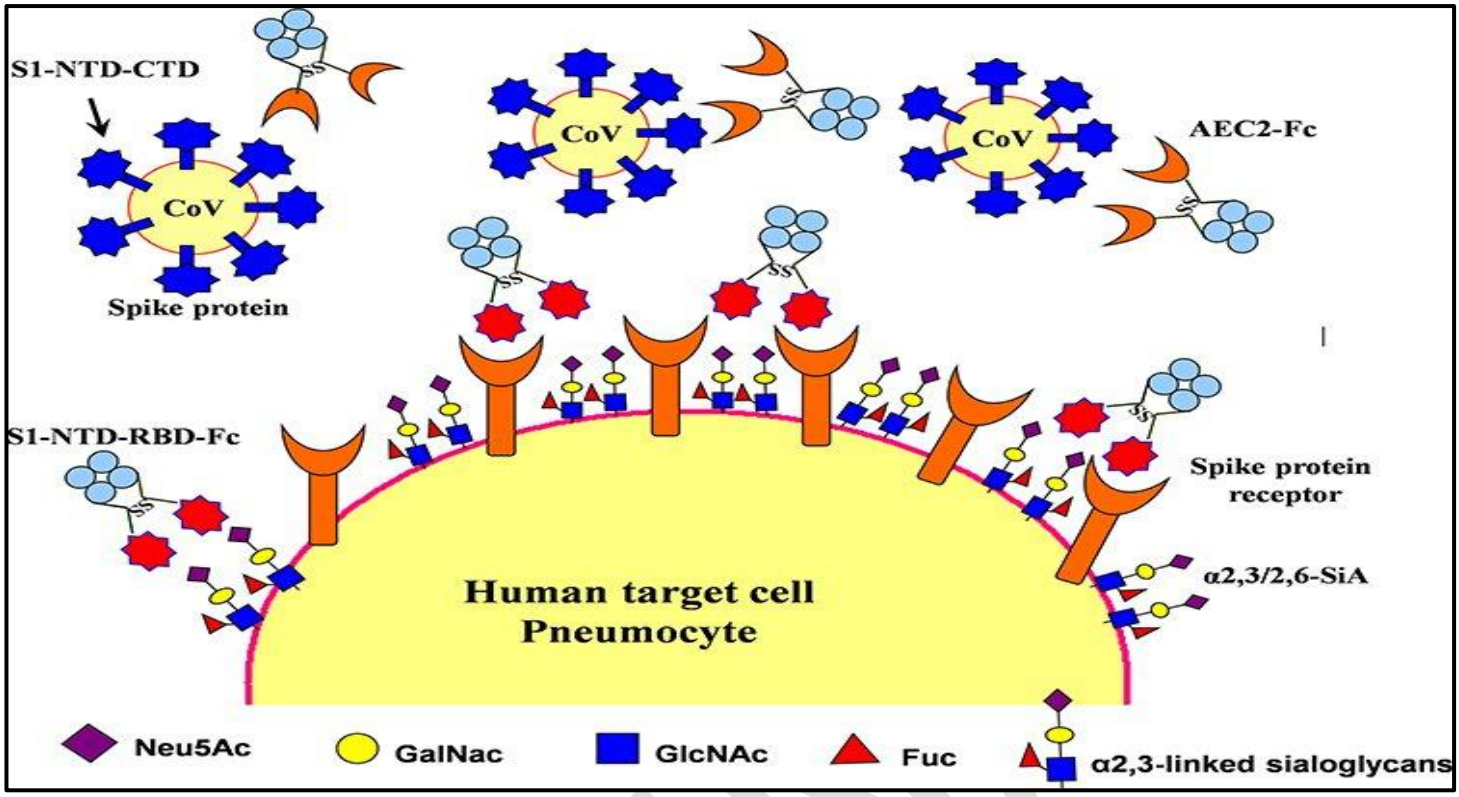


5] E484K Mutation:

Sometimes nicknamed Eric or Eek, E484K is a new mutation. The mutation changes the spike protein that the virus uses to enter human cells. It can make it harder for the immune system to recognize and fight the virus, if the body has been trained to fight the virus from earlier vaccines. E484K is called an "escape mutant" because it's been shown it might be able to escape some of the antibodies produced by the vaccine. First it was the South African strain. But recently it has also been detected in the UK strain, with samples found in discovered in south-west England.

The E484K mutation is present in the P.1 variant in Brazil. In the E484K mutation, a negatively charged amino acid (glutamic acid) is substituted with a positively charged amino acid (lysine). Thus, it can be expected that the mutation has a significant impact on viral sustainability and adaptive evolution. The structural analysis revealed that a new site for ACE2 binding (amino acid 75) is generated because of the E484K mutation. This appears to create a significantly stronger interaction between ACE2 and the native binding site located at the RBD and ACE2 interface (amino acid 501). The E484K mutation could potentially increase the infectivity and immune evasion potency of SARS-CoV-2.

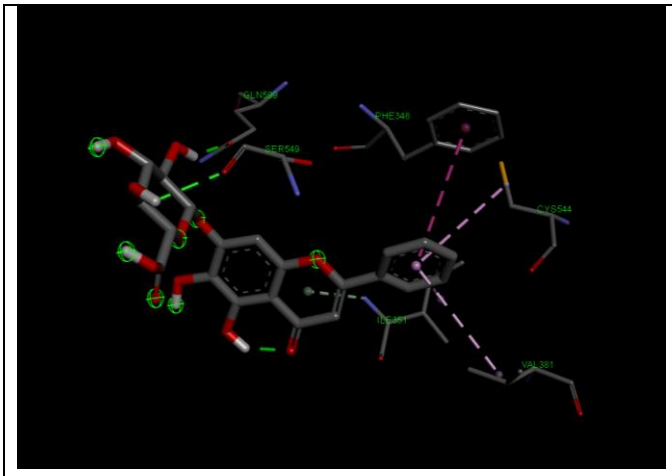
A highly diverse range of genetic mutations was observed in all Brazilian lineages with E484K mutation. On average, about 19 and 30 mutations were observed in B.1.1.33 and P.1 lineages, respectively. Further genomic analysis of the most recently emerged lineage P.2 indicated that both P.1 and P.2 lineages are evolving rapidly and have been circulating in Brazil for a more extended period. The study findings revealed that SARS-CoV-2 variants with E484K mutation are widely distributed in many regions in Brazil. The mutation was introduced in Brazil in October 2020. Because the E484K mutation is found in different viral lineages simultaneously with other mutations, the scientists suggest that this particular amino acid substitution may act as a common driving force for viral evolution in different genetic variants of SARS-CoV-2.

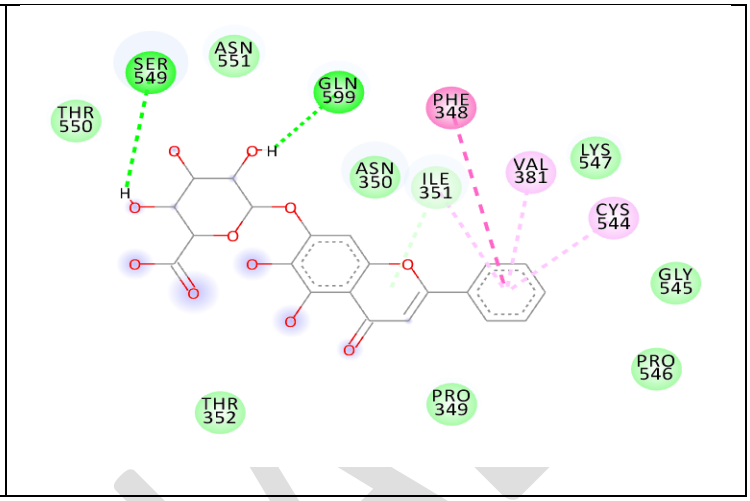


B	Protein Info	PDB ID	Compounds Name								
			Comp 1	Comp 2	Comp 3	Comp 4	Comp 5	Comp 6	Comp 7	Comp 8	Comp 9
	SARS-CoV-2 Mutation										
5	Spike Glycoprotein	E484K	-7.18	-6.56	-6.83	-3.13	-4.41	-4.72	-5.62	-6.92	-6.79

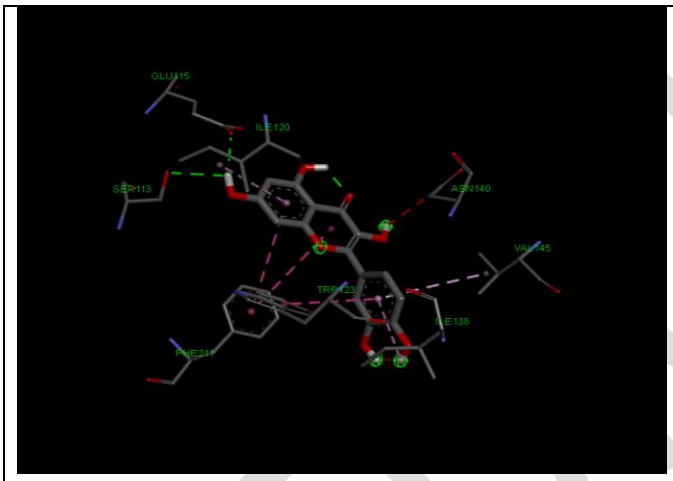
Docking Study 4: Spike Glycoprotein E484K with Ligand Baicalin, Quercetin, Luteolin, Rutin, Hesperidin, Curcumin, EGCG, Piperine, Glycyrrhizic

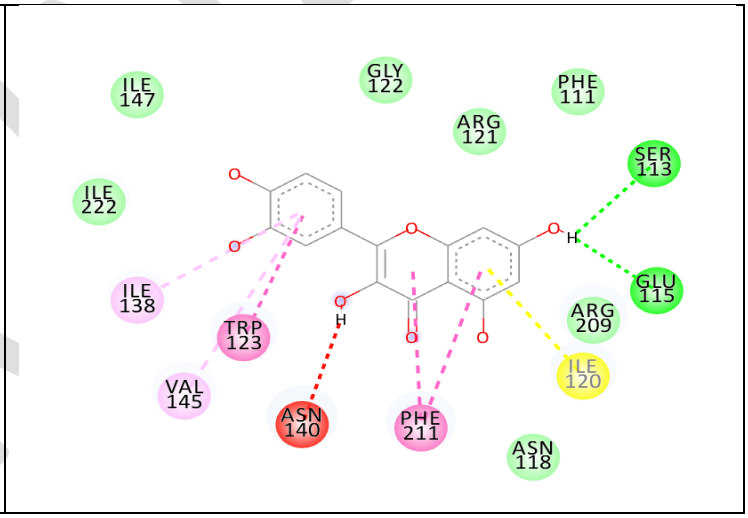
1.Baicalin



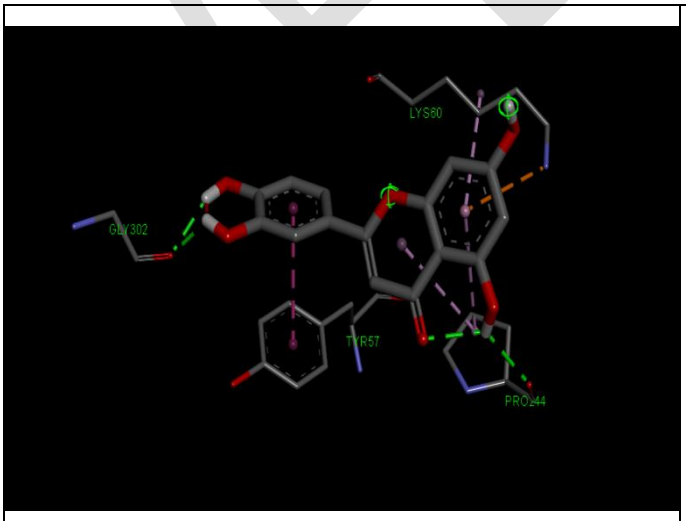


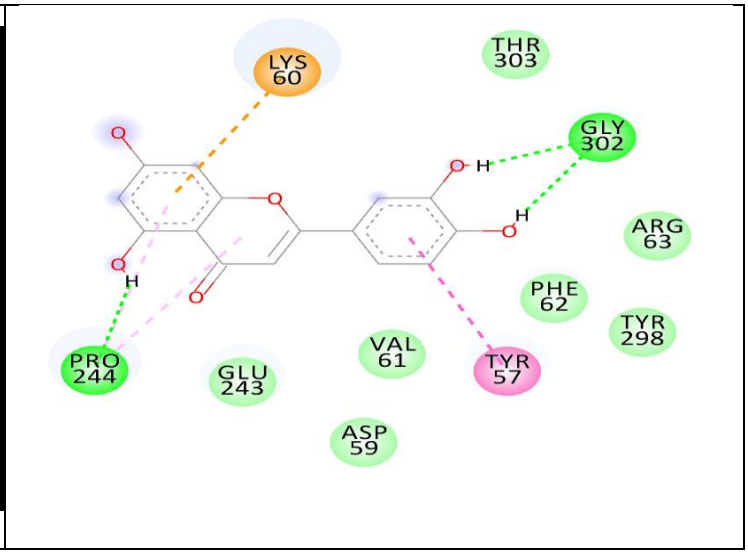
2.Quercitin



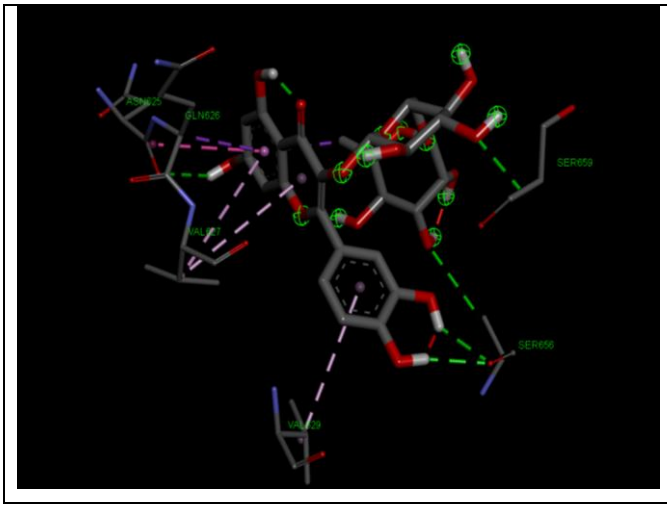


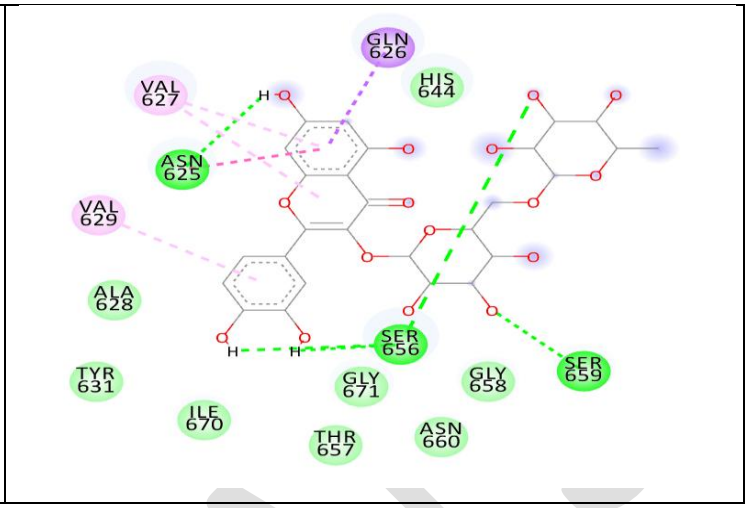
3.Leuteolin



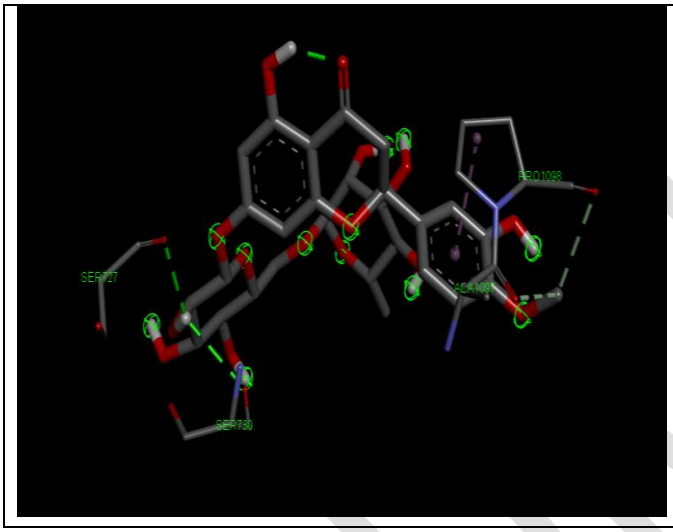


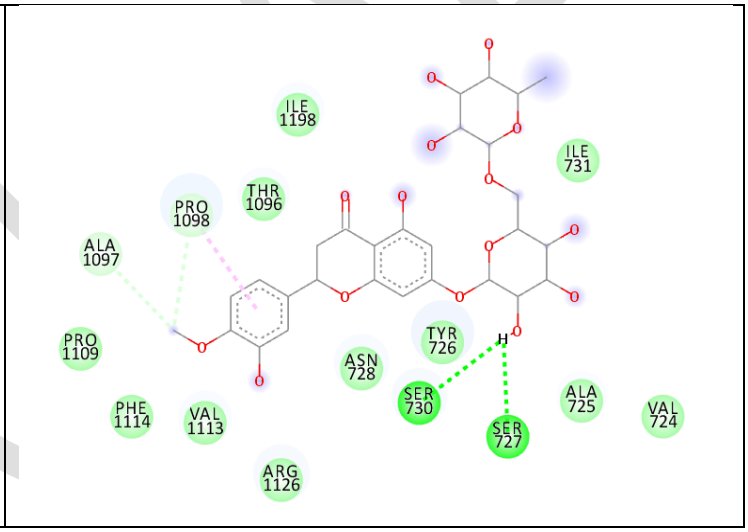
4.Rutin



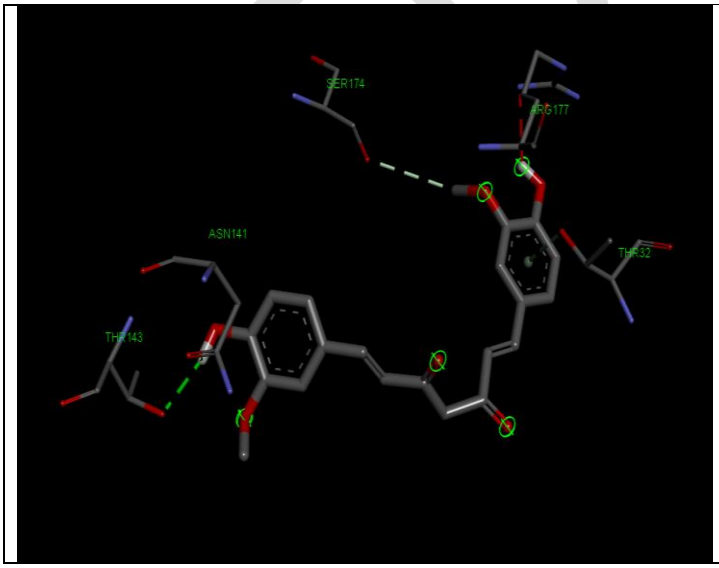


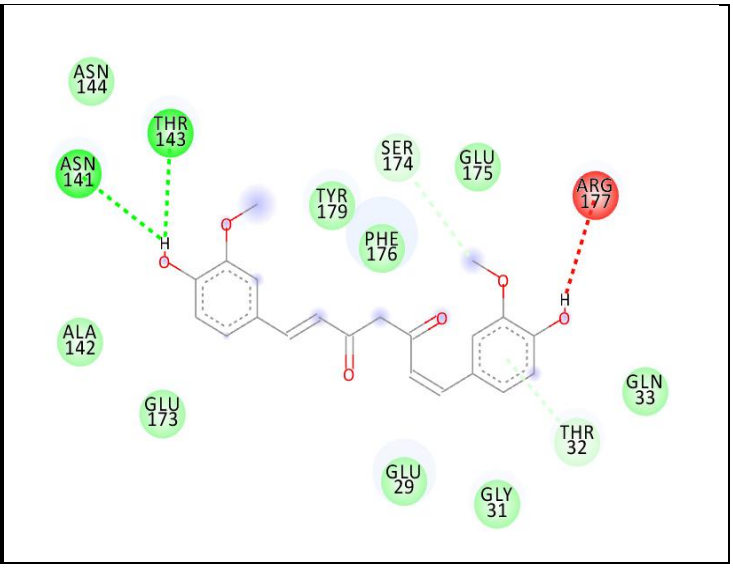
5. Hesperidin



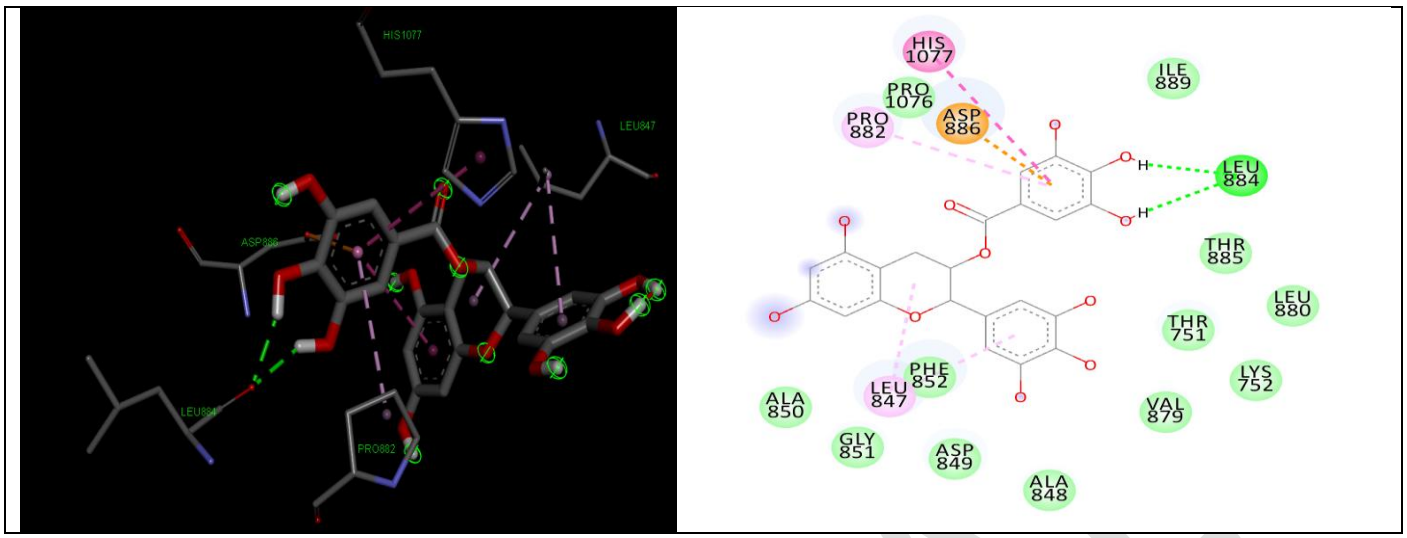


6. Curcumin

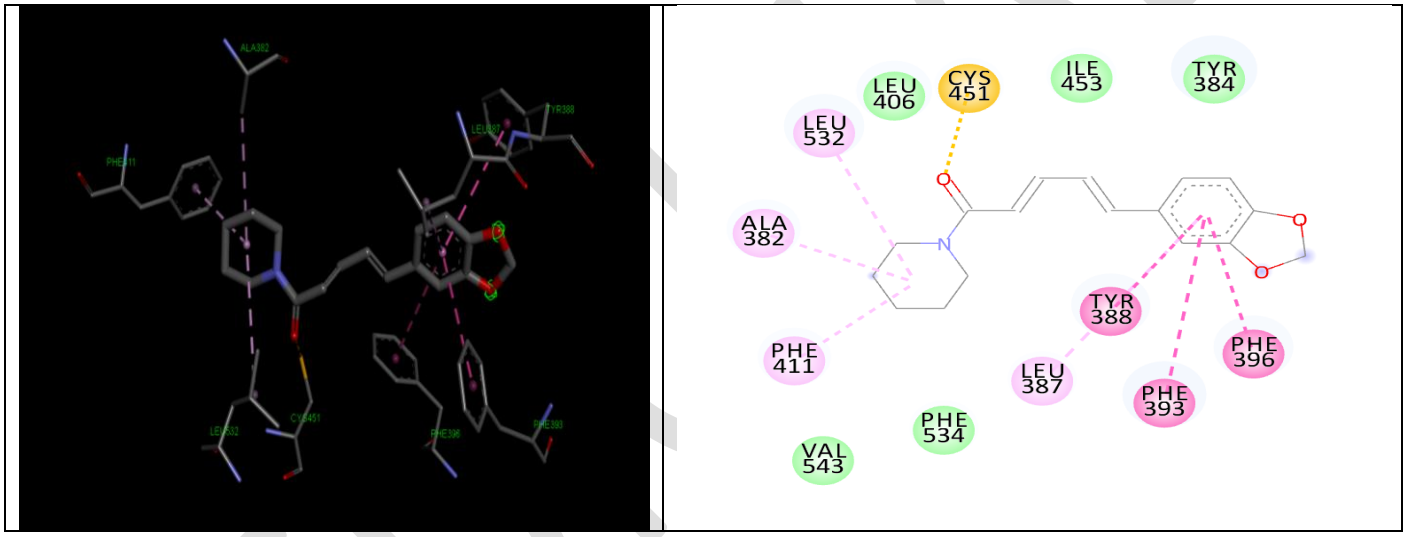




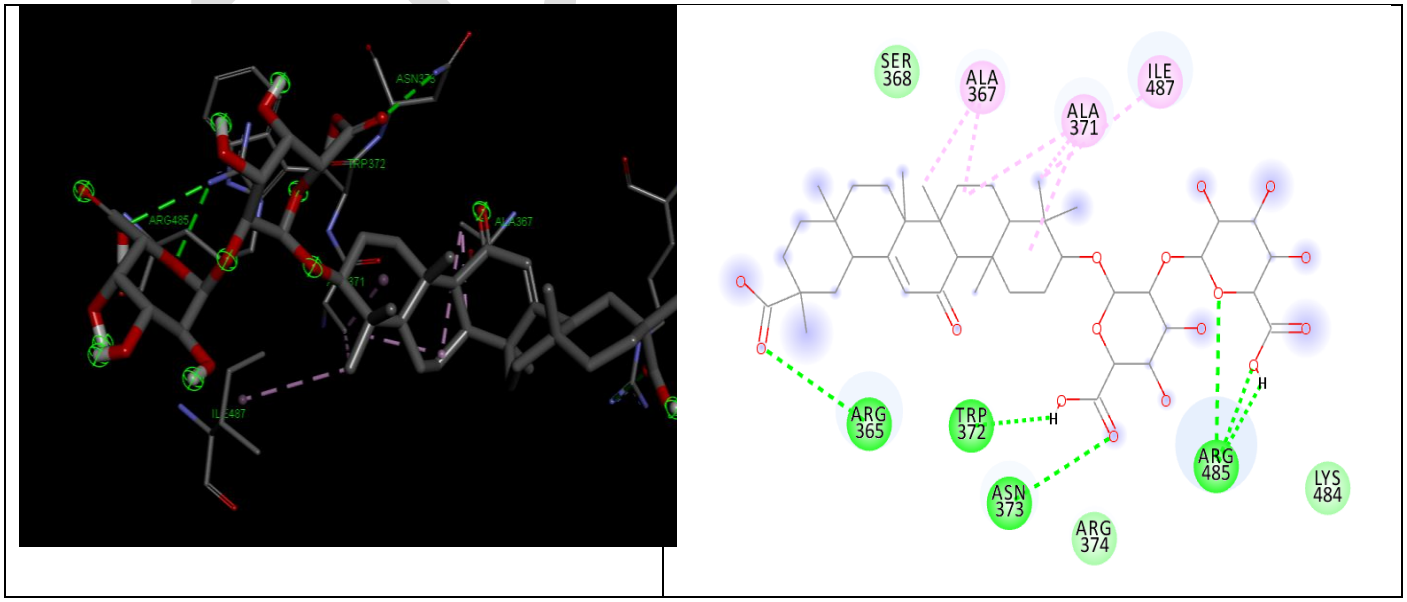
7. EGCG



8.Piperine



9. Glycyrrhizic



6] <u>P681H Mutation</u>:

B.1.1.7 viruses have also been shown to have a **P681H** mutation in the cleavage site of spike protein. This location is one of the residues that make up the furin cleavage site between S1 and S2 in spike. The S1/S2 furin cleavage site has been shown in animal models to promote viral entry into respiratory epithelial cells and transmission (Hoffmann et al. 2020; Peacock et al. 2020; Zhu et al. 2020). The spike proteins of this lineage have also been shown to have a deletion at amino acids. This mutation in the receptor binding domain of spike is a recurrent deletion that has been found in various lineages associated with SARS-CoV-2 (McCarthy et al. 2020; Kemp et al. 2020). Outside of spike, a Q27 stop mutation truncates the ORF8 protein of the virus, rendering the protein inactive. An ORF8 deletion at amino acid 382 has a mild effect on virus replication in human airway cells (Gamage et al. 2020). The B.1.1.7. lineage also has five synonymous mutations in ORF1ab and one synonymous mutation in the M gene.

Hawaiian SARS-CoV-2 strains that were deposited in the GenBank in March 2020 clustered with sequences from Wuhan, China, Sweden, and the state of New York (USA). Moreover, phylogenetic tree results suggest that the virus has been brought to Hawaii from many sources. Thirteen single nucleotide polymorphisms were decoded across 13 unique SARS-CoV-2 genomes within the S gene region – with one non-synonymous mutation (P681H) detected in the two Hawaii strains.

The P681H mutation is shared in VOC-202012/01, but has emerged spontaneously several times earlier also, and there is no evidence to indicate it contributing to increased transmission of the virus in Nigeria. P681H is immediately juxtaposed to the amino acid 682-685, furin cleavage site, identified at the S1/S2 linkage site, which has been predicted to enhance systemic infection based on bioinformatics analysis, and increased membrane fusion in laboratory experiments. The relevance of this to human infection is not known. This lineage has also been indirectly associated with higher virus load in samples tested by an assay using RT-qPCR and increased transmissibility. It has been hypothesized (but not proven) that this lineage may have resulted from the transmission of the virus from a chronically infected individual.

The P681H mutation is also characteristic of the new SARS-CoV-2 variants from the United Kingdom and Nigeria. SARS-CoV-2 sequences by the P681H mutation to create a ratio of sequences containing the P681H mutation to all sequences reported in the GISAID database for a given month. Inclusion criteria were for sequences providing a full month, day, and year. The D614G mutation underwent assessment in the same manner for comparison. All prevalence data converted into ratio underwent a logarithmic transformation.

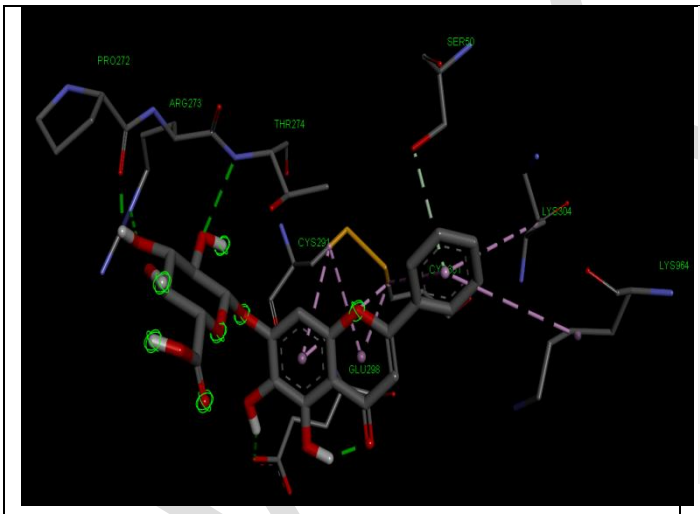
This finding will help in further understanding the role of such arising mutations in modulating immunogenicity, viral tropism and pathogenesis of the disease, which in lieu will help in designing vaccine more precisely to mitigate pandemic COVID-19.

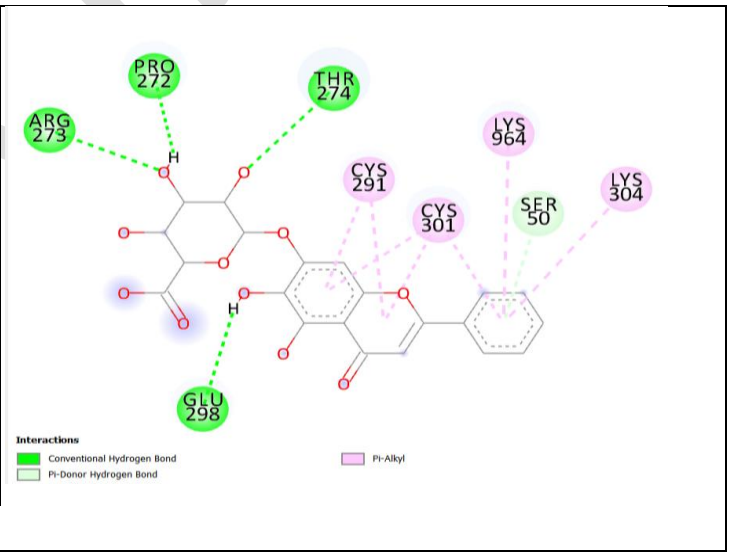
As per the following table comparison between wild type and mutant spike glycoproteins shows approximate same binding energy. Comp1 and Comp2 shows good binding energy as compare to wild type. As per the result and analysis except Comp4 others shows good result and useful for the inhibition of particular protein.

SN	Protein Info	PDB ID	Compounds Name								
			Comp 1	Comp 2	Comp 3	Comp 4	Comp 5	Comp 6	Comp 7	Comp 8	Comp 9
	SARS-CoV-2 Mutation										
6	Spike Glycoprotein	P681H	-6.80	-6.17	-6.92	-4.76	-4.43	-6.07	-6.24	-6.03	-6.94

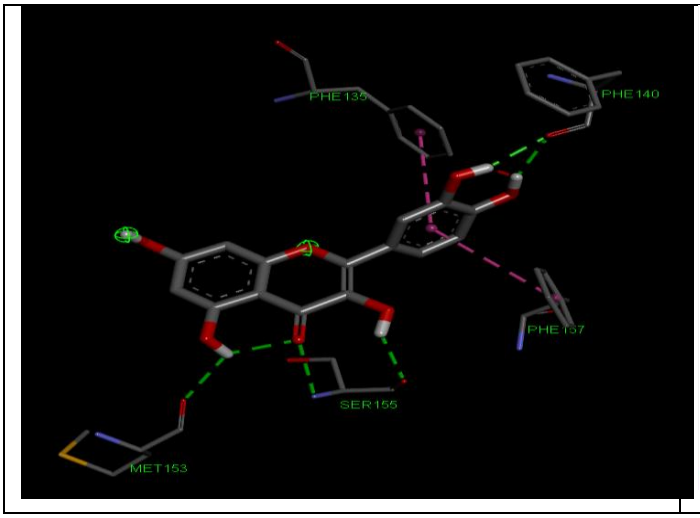
Docking Study 5 : SARS-CoV-2 Spike glycoprotein PROPH (P681H) with Ligand Baicalin, Quercetin, Luteolin, Rutin, Hesperidin, Curcumin, EGCG, Piperine, Glycyrrhizic

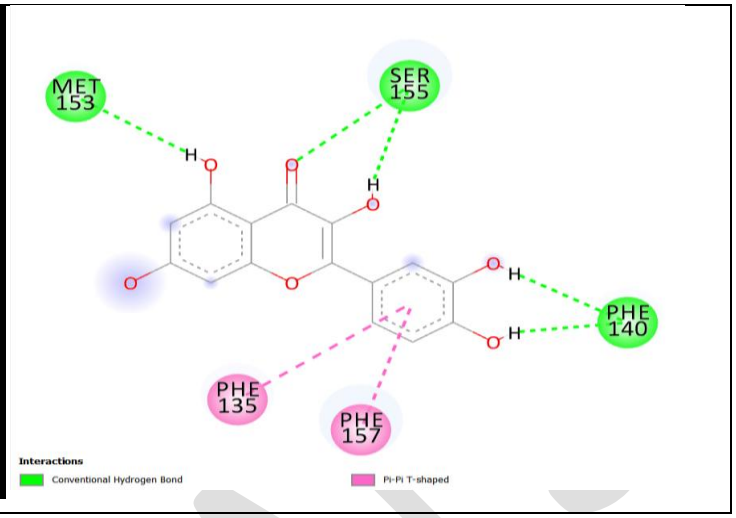
1.Baicalin



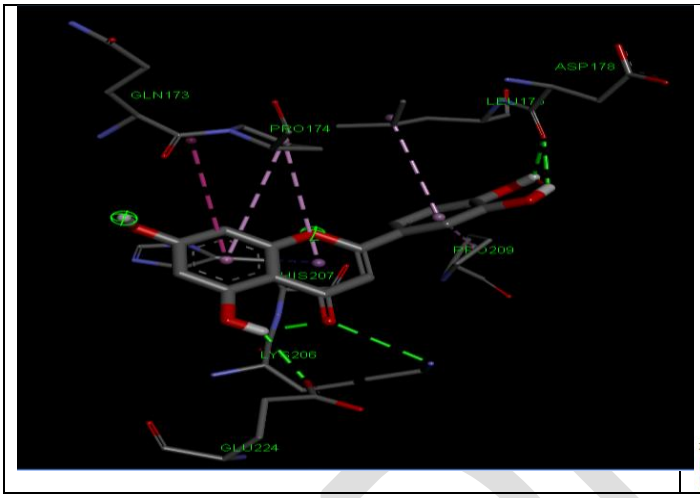


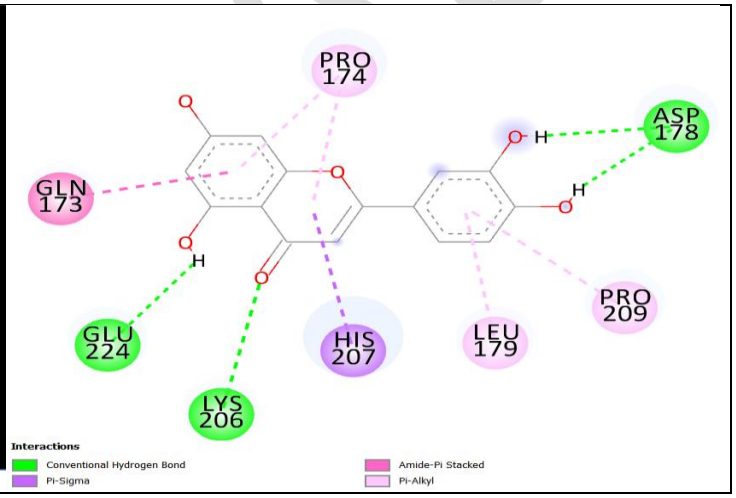
2.Quercitin



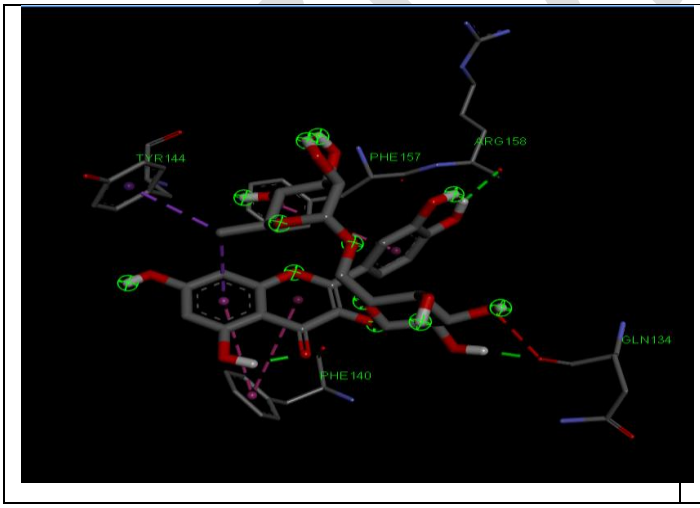


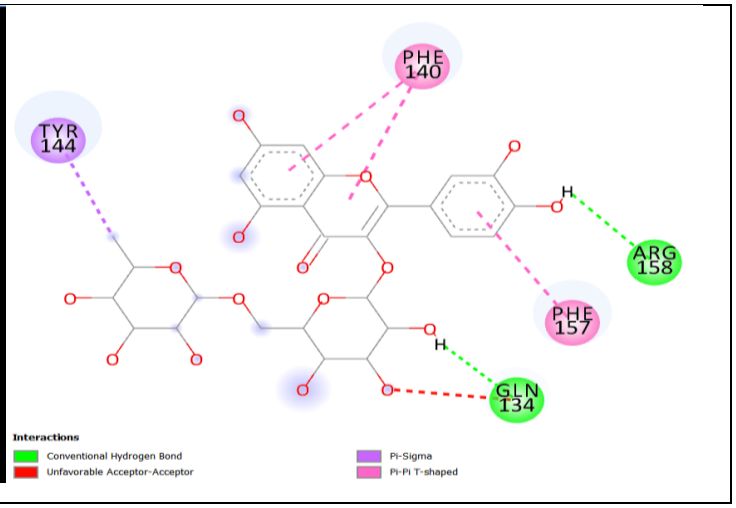
3.Leuteolin



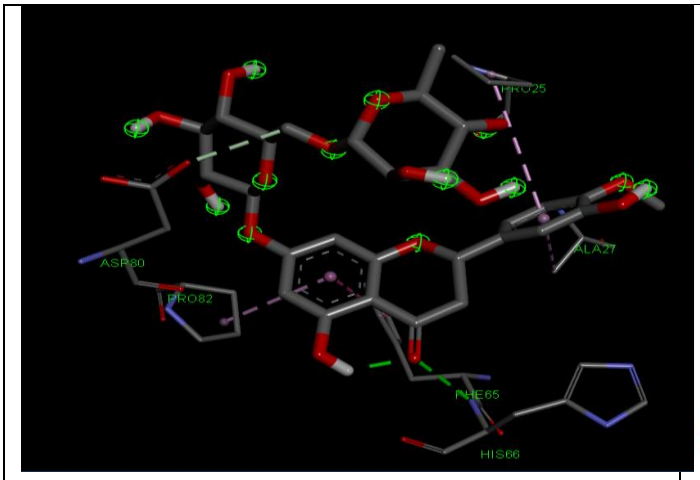


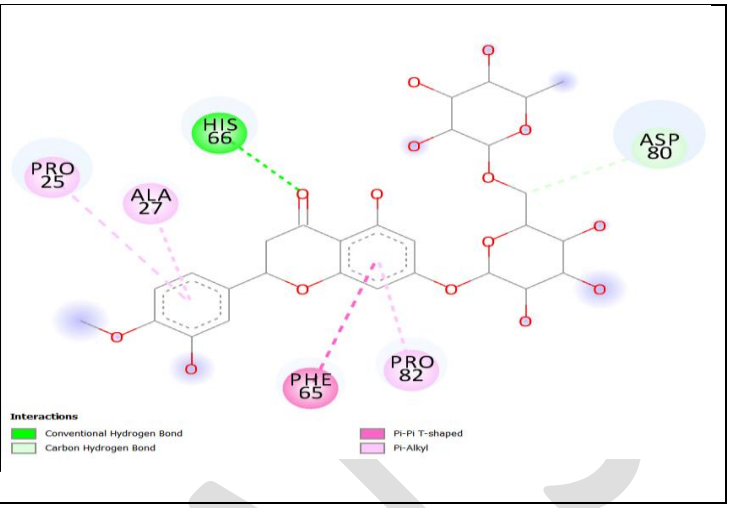
4.Rutin



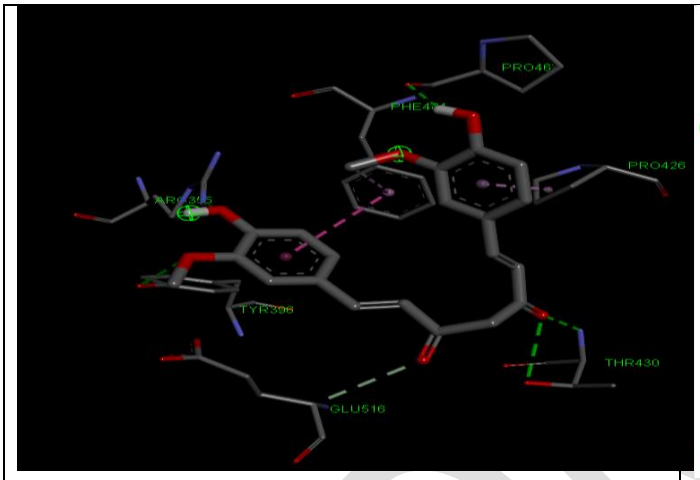


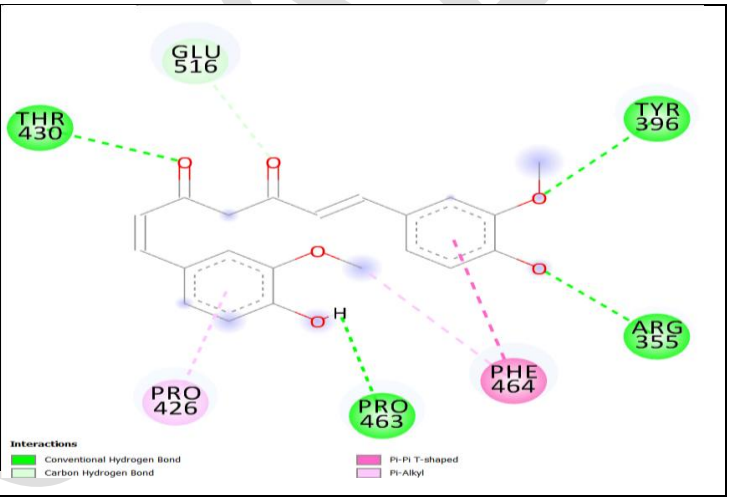
5. Hesperidin



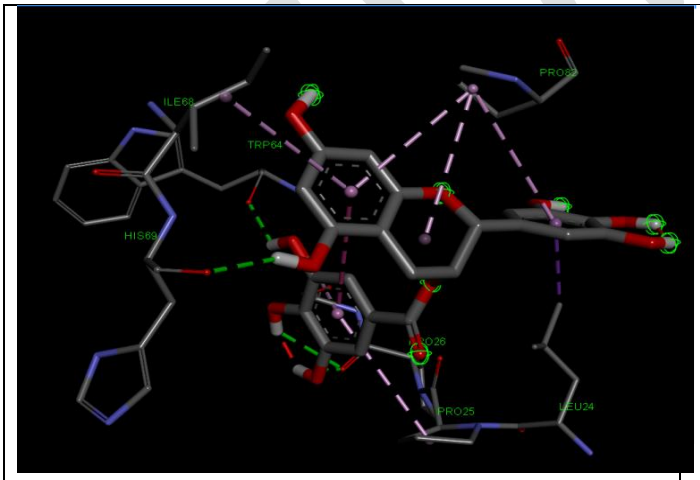


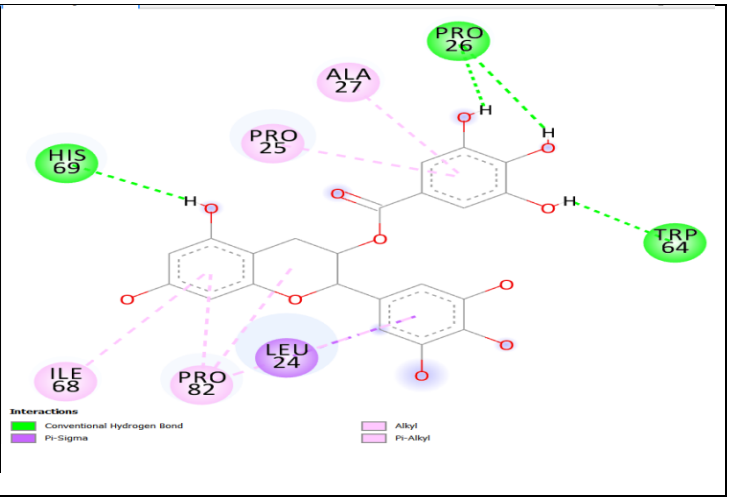
6. Curcumin



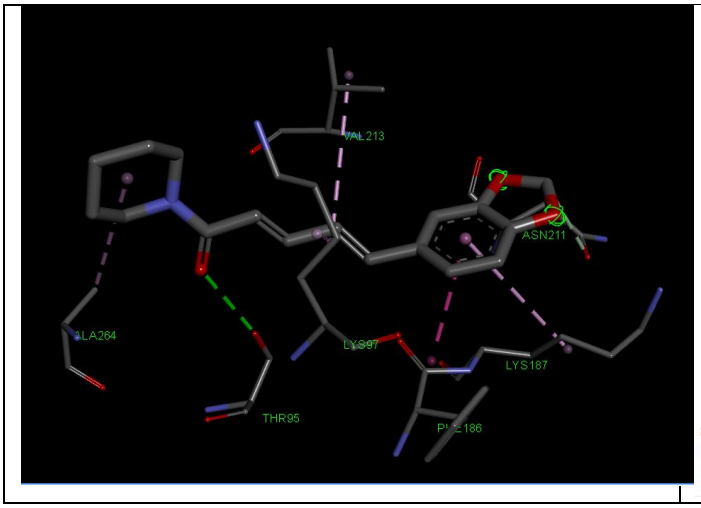


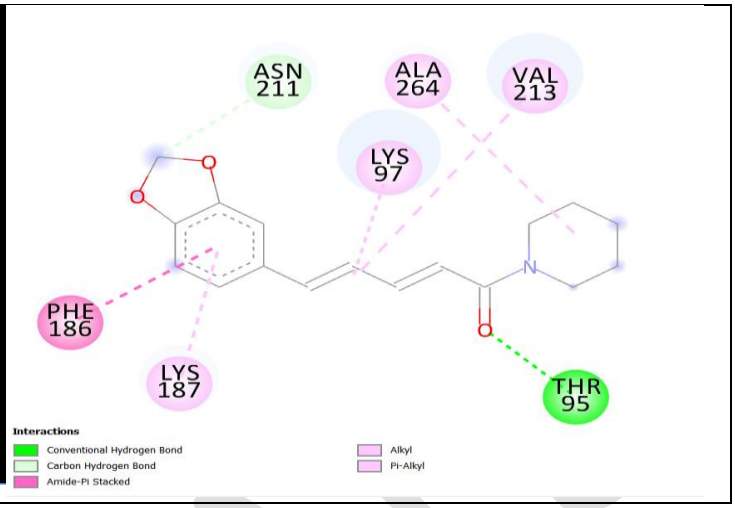
7. EGCG



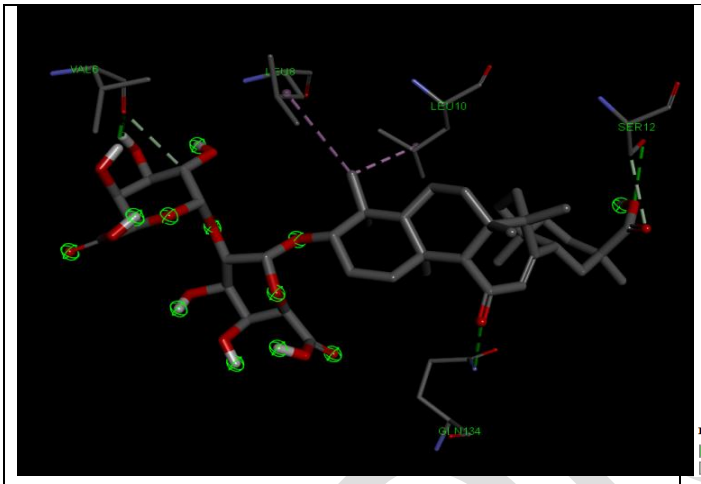


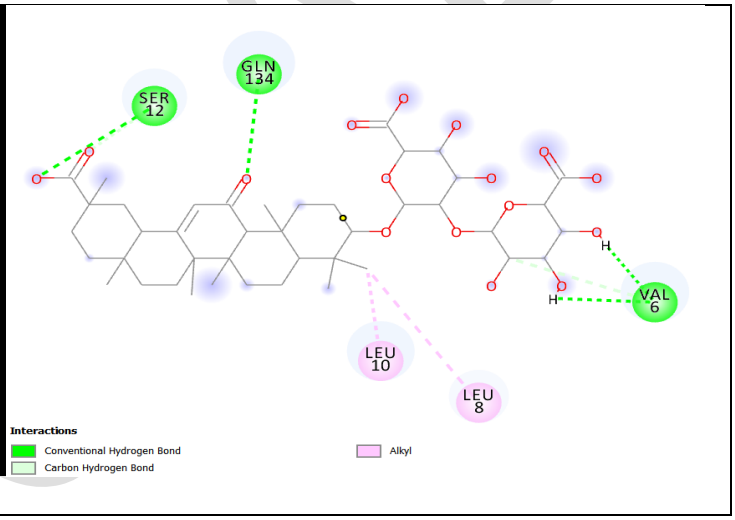
8.Piperine





9. Glycyrrhizic





7] K417N Mutation:

The accompanying K417N and E484K mutations in the South African strain provide a counterpoise to the spike's increased affinity due to the N501Y mutation. They prevent the formation of two salt bridges that help to form and stabilize the RBD-ACE2 complex. This reduces ACE2 binding affinity. Thus, this strain is less infectious and less rapidly spreading than the UK strain. This is even though both share the latter substitution.

The stronger binding to ACE2 caused by the substitution N501Y, the South African 501.V2 variant that has undergone two additional mutations (K417N and E484K) is unlikely to exhibit an increased infectivity and likely transmissibility.”

Deep mutational scanning suggests that the K417N mutation has minimal impact on binding affinity to hACE235. The spike RBD is the main target of neutralizing antibodies (NAbs) elicited during SARSCoV-2 infection. NAbs to the RBD can be broadly divided into four main classes. Of these, class 1 and class 2 antibodies appear to be most frequently elicited during SARS-CoV2 infection, and their epitopes directly overlap the hACE2 binding site. Class 1 antibodies have a VH3-53 restricted mode of recognition centred on spike residue K417. The K417N mutation would abolish key interactions with class 1 NAbs, and likely contributes toward immune evasion at this site.

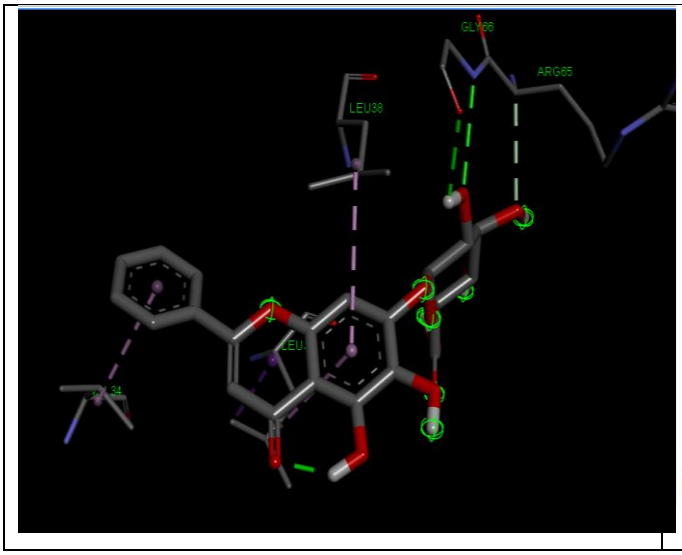
Single Spike RBD-RBM amino acid substitutions, which were 310 found in other new SARS-CoV-2 Spike variants, namely the 501.V2 variant from South Africa 311 (K417N).

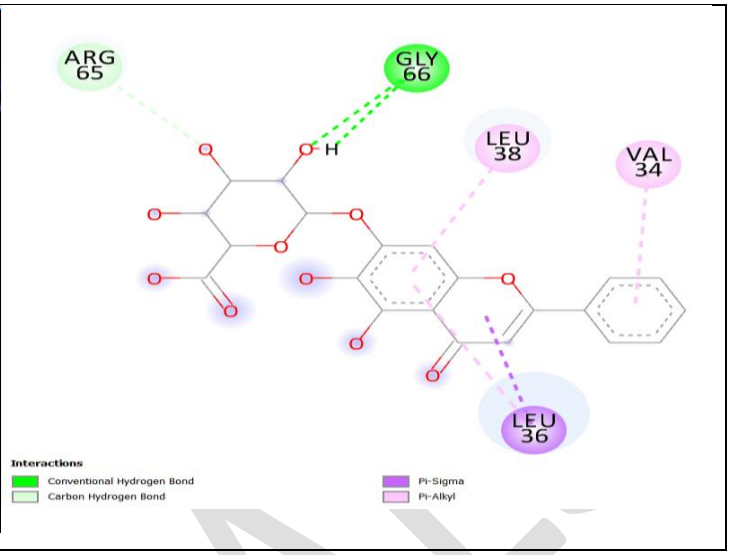
Results predicted that the combined effect of the three amino acid 322 substitutions N501Y, K417N and E484K was less than that of the single N501Y (Fig. 5A) in terms 323 of increased computed affinity for ACE2 (from -50.26 to -56.37 Kcal/mol as compared to -67.49 324 Kcal/mol of N510Y). According to research, the addition of K417N mutation led to a dramatic decrease of the STE90-C11 antibody binding to virus's S1 RBD. The default FEP sampling protocol calculated a ΔΔG value of 5.74 kcal/mol. In the case when both N501Y and K417N mutations are present at the same time this value further increased to 8.61 kcal/mol but this simulation was not well converged and after its extension to 15 ns we obtained a value of 5.83 kcal/mol. Thus, it seems that the effect of these mutations is not additive and only K417N mutation can abolish the interaction with STE90-C11 antibody. These results also suggest that even well tolerated to mutations, antibodies eventually would be resisted to this variant of SARS-CoV-2. 4). The K417N mutation also increases the S1 RBD binding to ACE2, however by only -0.39 kcal/mol.

SN	Protein Info	PDB ID	Compounds Name								
			Comp 1	Comp 2	Comp 3	Comp 4	Comp 5	Comp 6	Comp 7	Comp 8	Comp 9
	SARS-CoV-2 Mutation										
7	Spike Glycoprotein	K417N	-7.26	-7.35	-7.54	3.81	-6.65	-6.89	-6.89	-7.38	-7.28

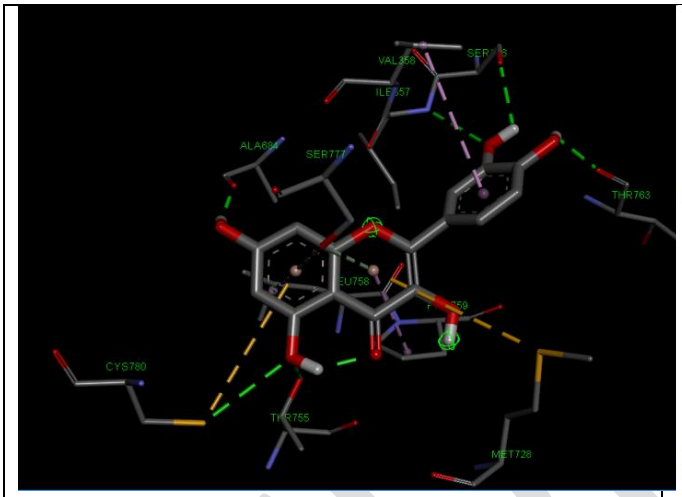
Docking Study 9: Spike Glycoprotein PROPH (K417N) with Ligand Baicalin, Quercetin, Luteolin, Rutin, Hesperidin, Curcumin, EGCG, Piperine, Glycyrrhizic

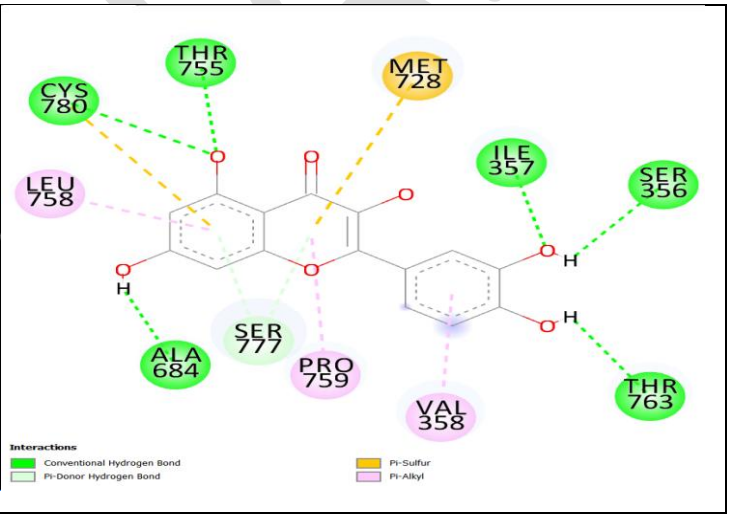
1.Baicalin



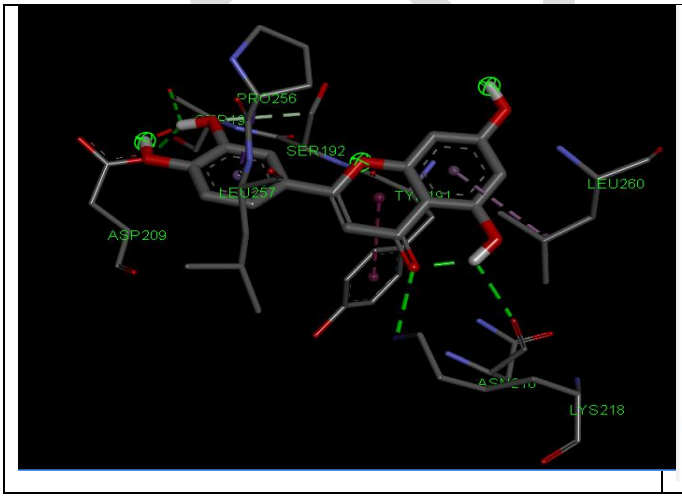


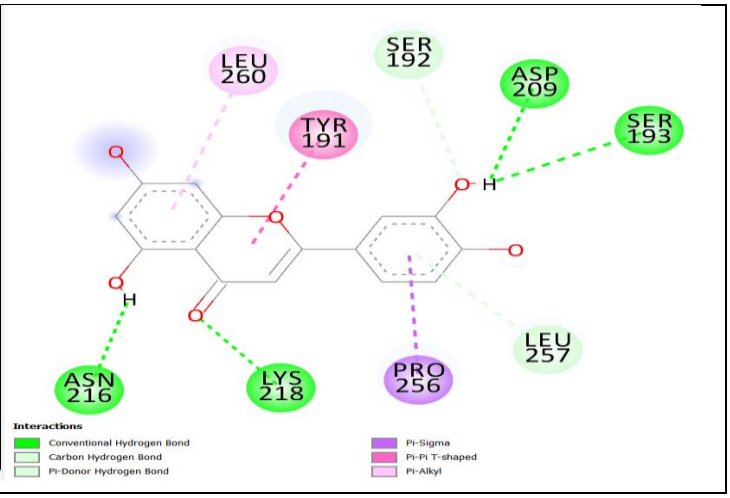
2.Quercitin



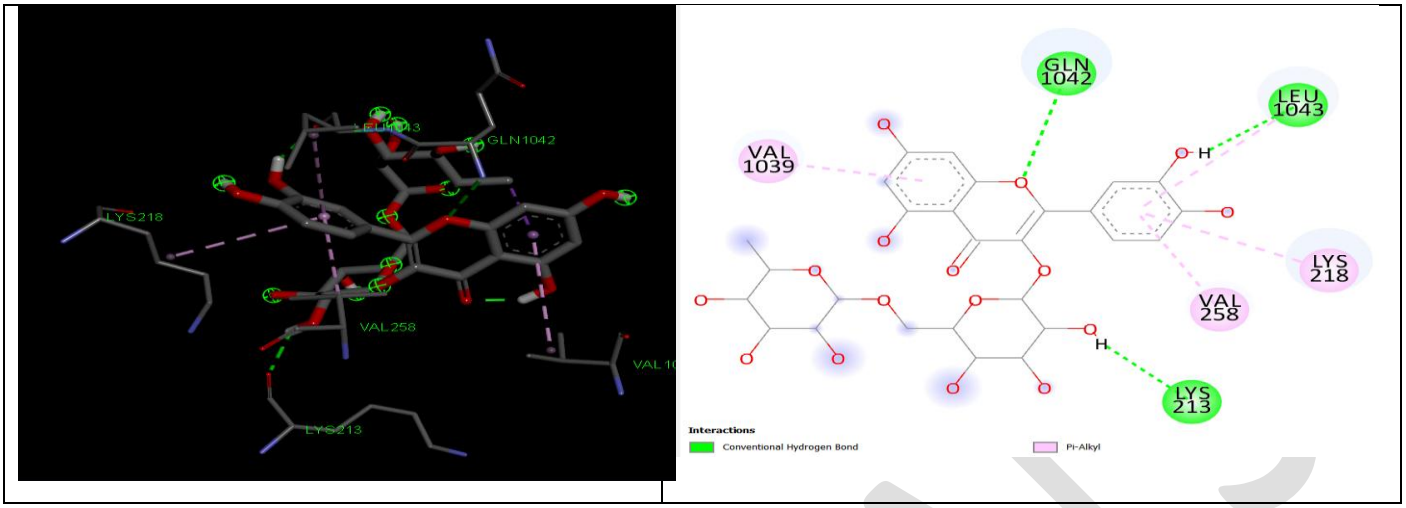


3.Leuteolin

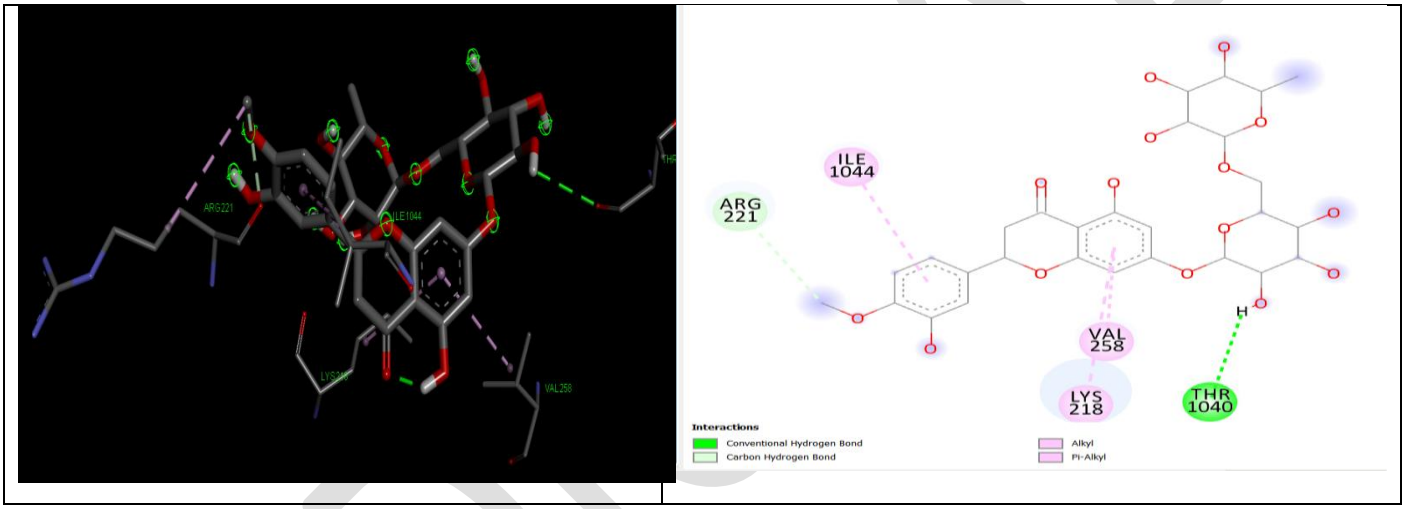




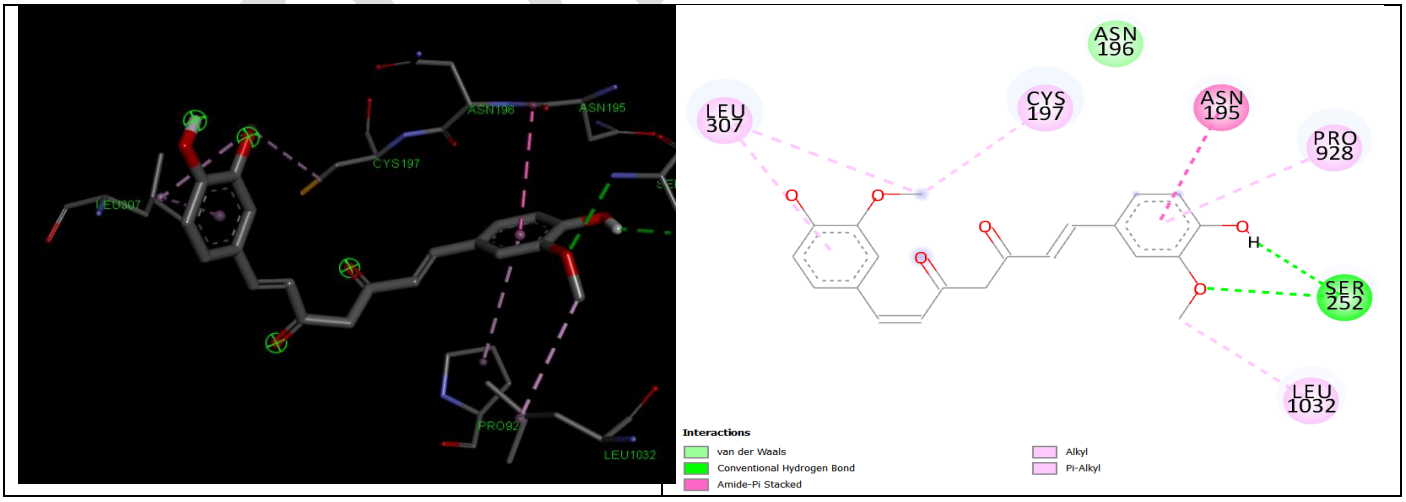
4.Rutin



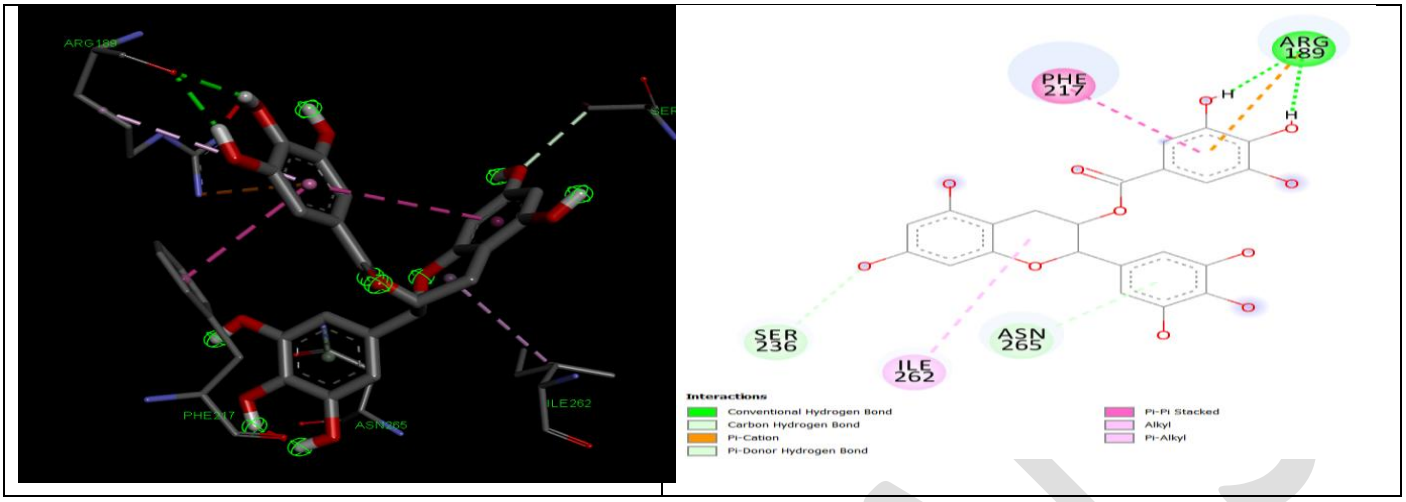
5. Hesperidin



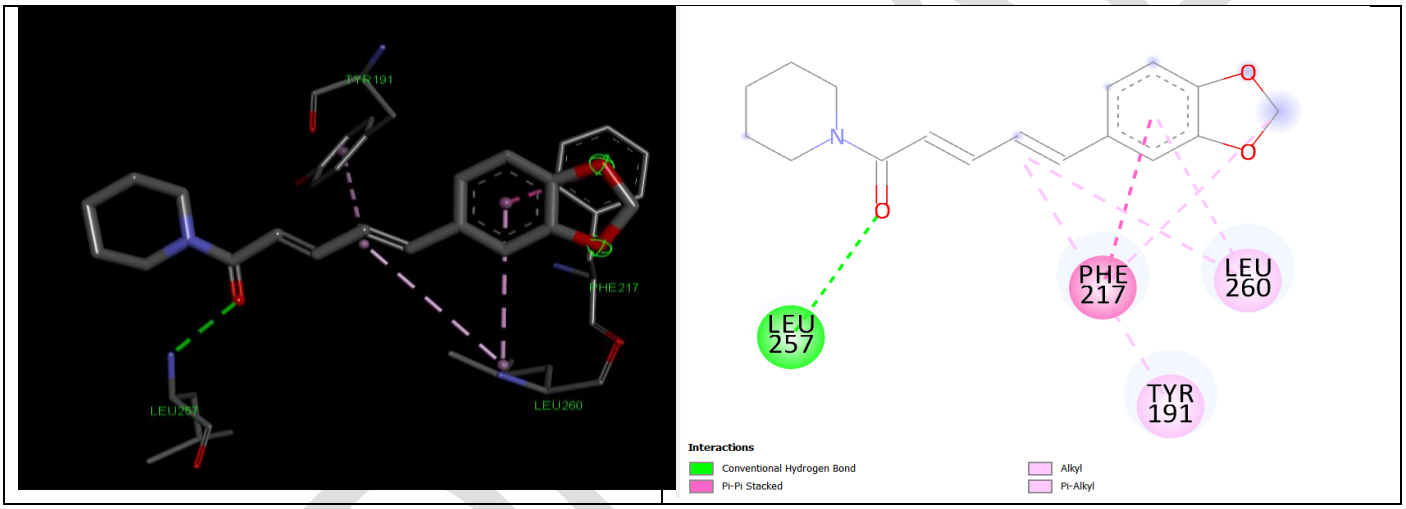
6. Curcumin



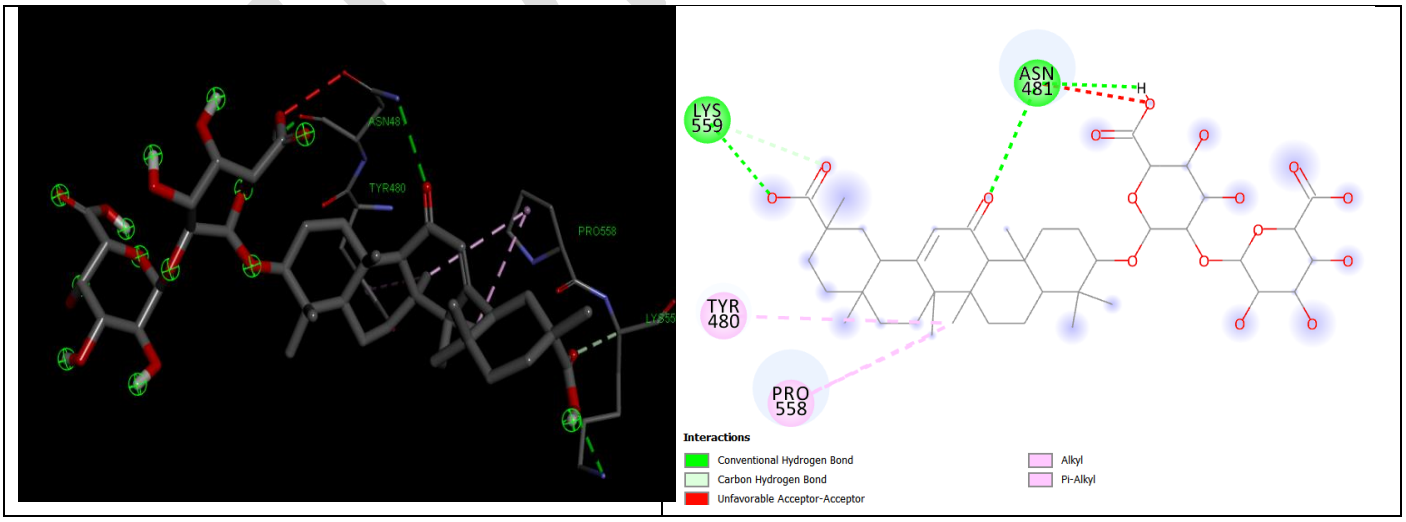
7. EGCG



8.Piperine



9. Glycyrrhizic



8] N501Y Mutation:

New severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) lineages carrying the amino acid substitution N501Y in the receptor-binding domain (RBD) of the spike protein have spread rapidly in the United Kingdom (UK) during late autumn 2020. The most concerning mutation is N501Y, which co-occurs with several mutations of potential biological importance, including P681H and deletion of the amino acid at the 69th and 70th residues (Δ69/Δ70) on the spike protein. It has recently been reported that this variant is rising in frequency in the South-east of England so fast as to raise the suspicion that it has increased transmissibility.

Researchers after study investigate this aspect of the virus in terms of increased viral load. The researchers sequenced all positive samples from four Lighthouse laboratories in the UK, using their quantitative sequencing approach. This yields the number of unique mapped reads, which bears a correlation with and therefore acts as a proxy for the viral load.

They found that the logarithm of unique mapped reads was negatively correlated with the Ct values obtained from the polymerase chain reaction (PCR) testing. They selected the presence of Y501 as a marker of the new variant. From 88 samples that showed this mutation, they considered only the samples taken in the period from October 31 to November 13, 2020. This showed that the number of unique mapped reads in the Y501 variant was more significant than in the N501 variant, indicating that the median viral loads are increased by about three times for the Y501 variant.

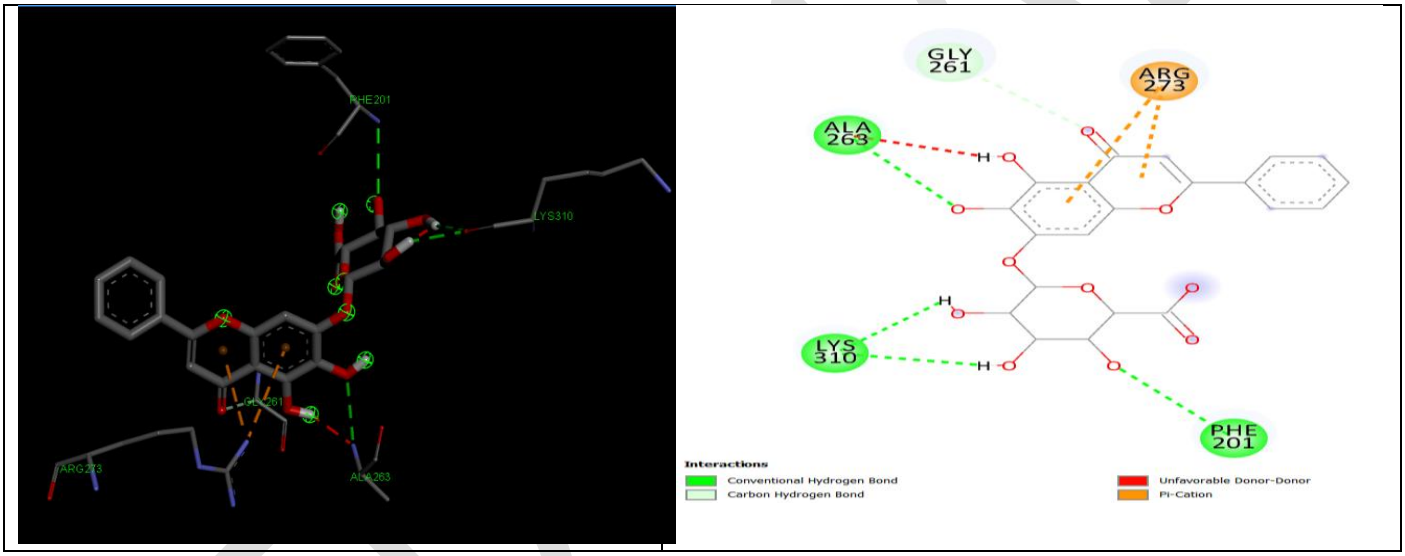
The mutation N501 at the receptor binding domain of the S protein results in the change of asparagine to tyrosine residue at position 501 (N501Y). This mutation is also present in the newly emerging SARS-CoV-2 variant viruses reported in the U.K. (20B/501Y.V1, B1.1.7 lineage) that is epidemiologically associated with high human to human transmission. The mutation is of particular concern because it is located in the viral receptor binding site for cell entry, increases binding to the receptor (angiotensin converting enzyme 2), and enables the virus to expand its host range to infect mice.

The reasons for the faster growth associated with the variant are not clear: it could be due to faster epidemic growth, demographic patterns, founder effects, or higher viral loads, among other biological mechanisms. The correlation with higher viral loads, in case, seems to suggest increased transmissibility of this virus, but further studies are required. Again, the N501Y mutation may not be the only reason for this expansion. There is more need to understand how viral levels are related to virulence since this may determine the infection's severity.

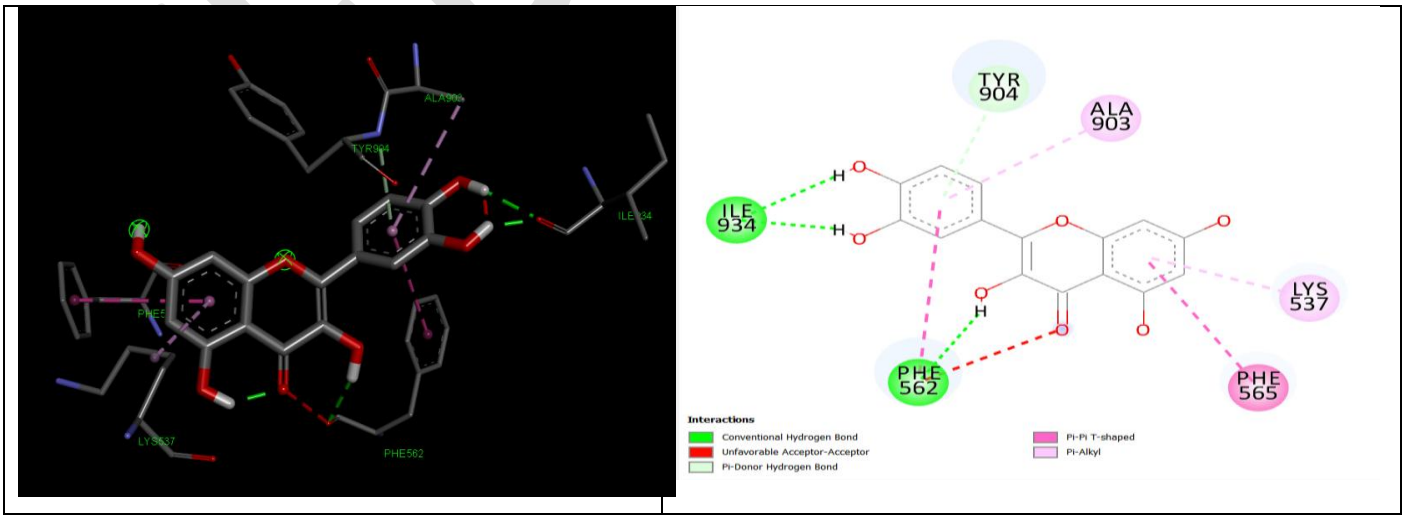
SN	Protein Info	PDB ID	Compounds Name								
			Comp 1	Comp 2	Comp 3	Comp 4	Comp 5	Comp 6	Comp 7	Comp 8	Comp 9
	SARS-CoV-2 Mutation										
8	Spike Glycoprotein	N501Y	-7.39	-7.35	-7.06	-4.75	-5.72	-6.43	-7.25	-6.86	-7.88

Docking Study 7: spike glycoprotein PRONY (N501Y) with Ligand Baicalin, Quercetin, Luteolin, Rutin, Hesperidin, Curcumin, EGCG, Piperine, Glycyrrhizic

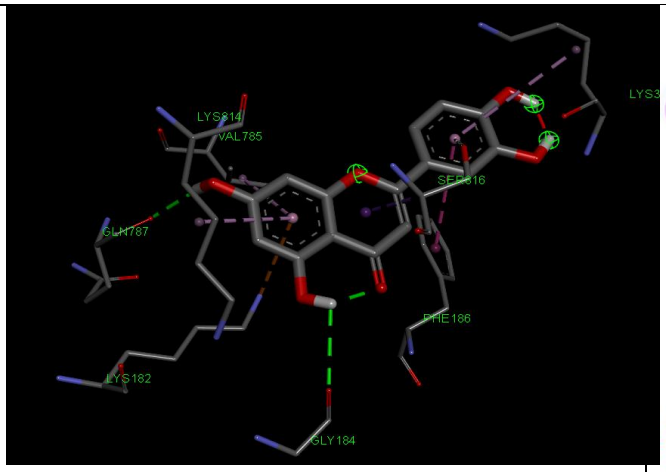
1.Baicalin

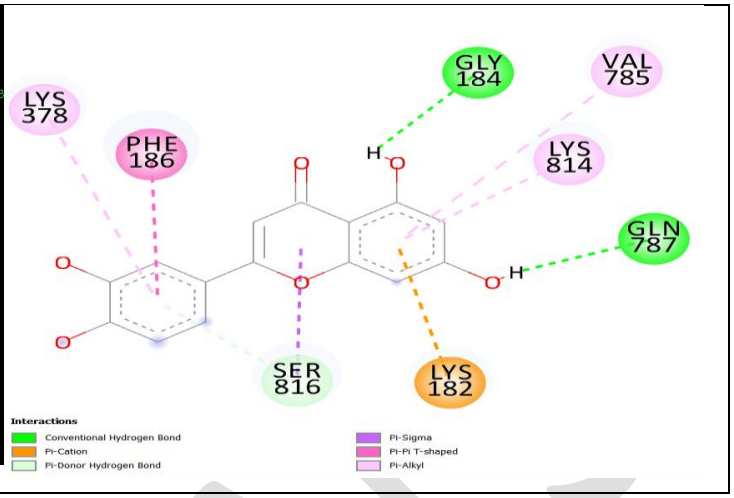


2.Quercitin

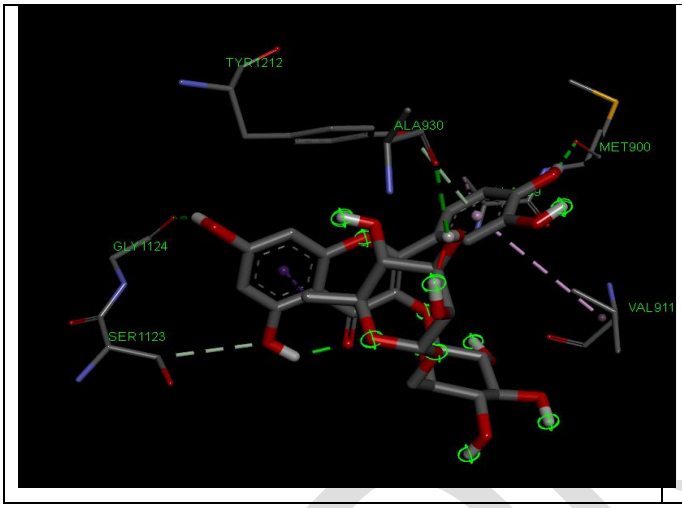


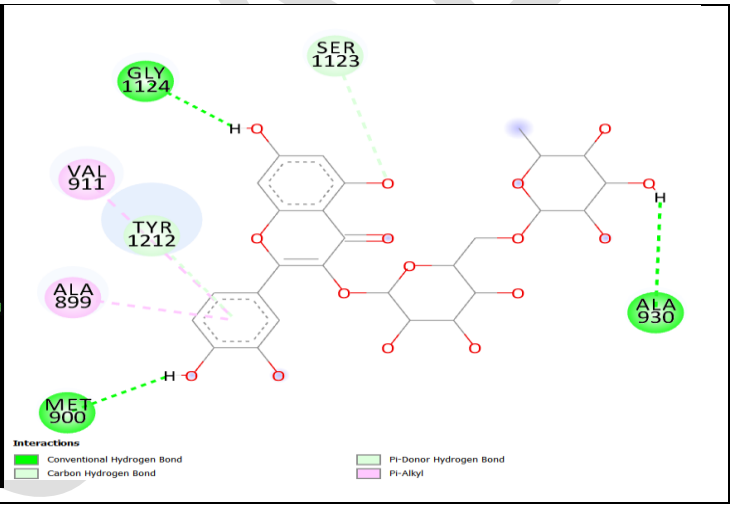
3.Leuteolin



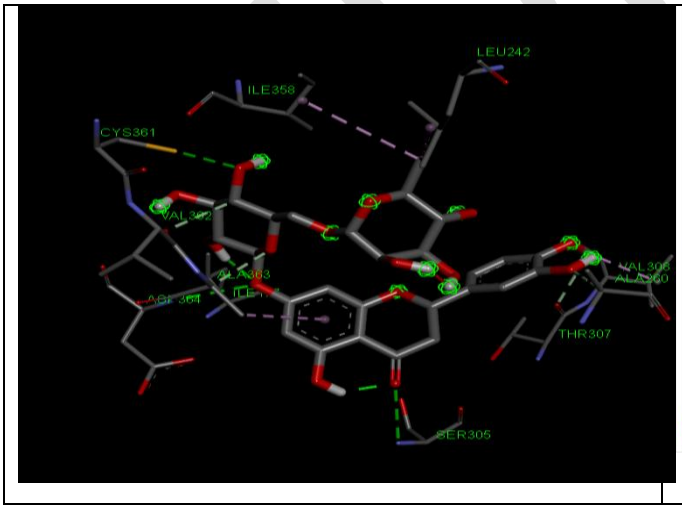


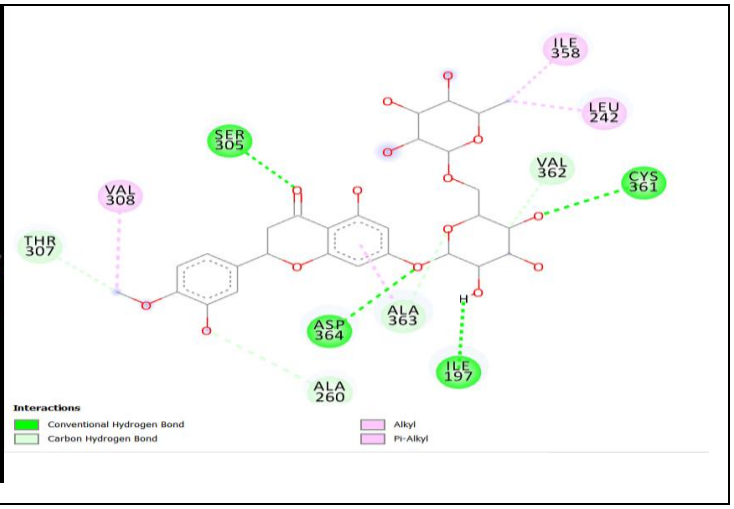
4.Rutin



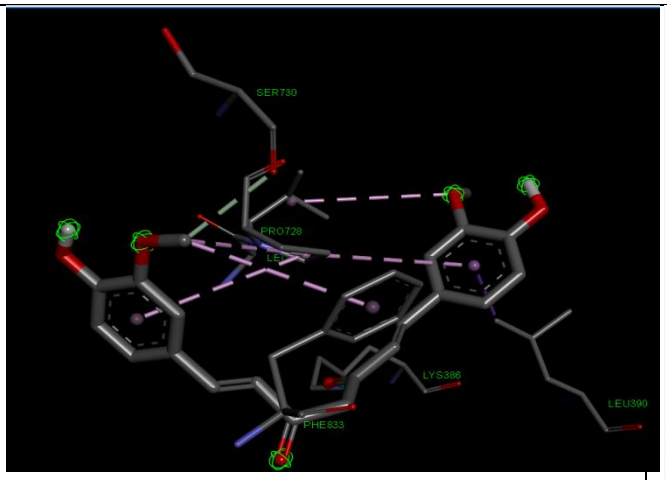


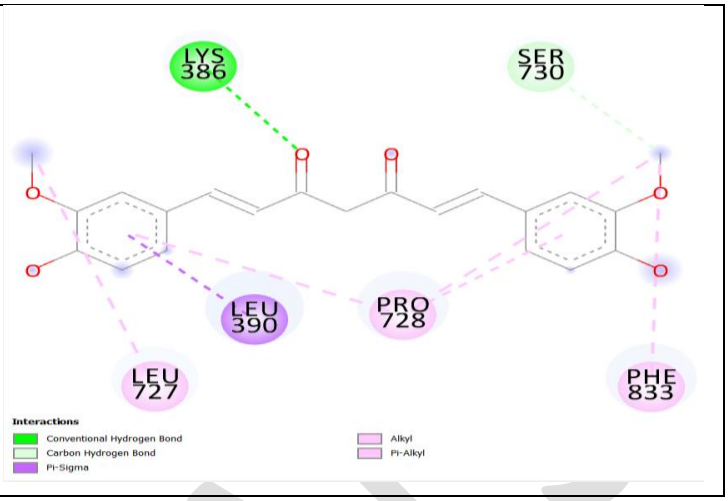
5. Hesperidin



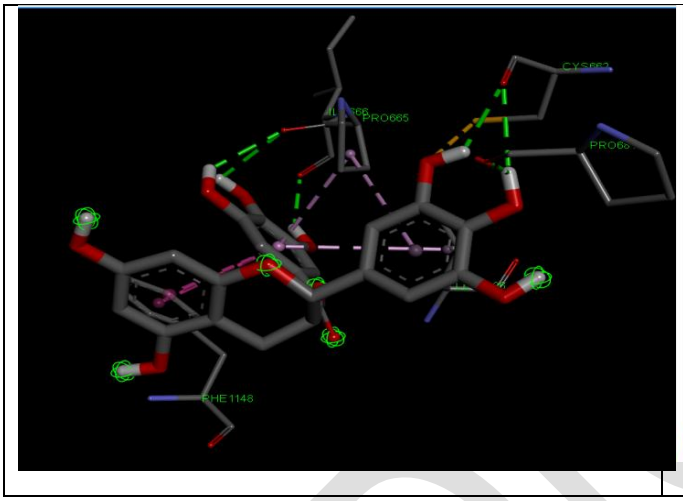


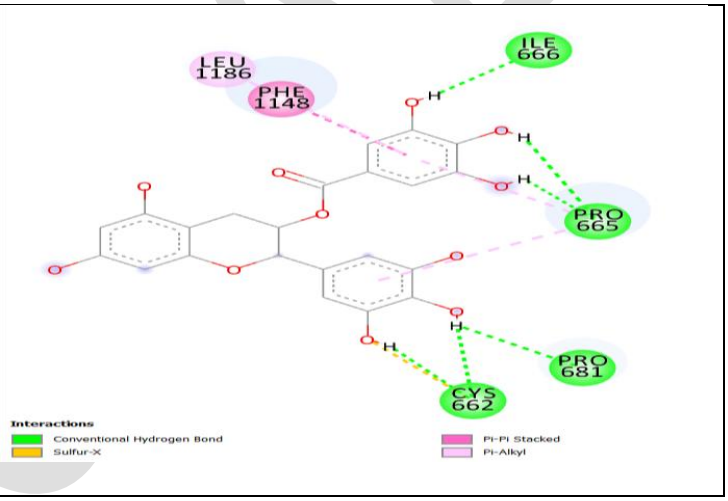
6. Curcumin



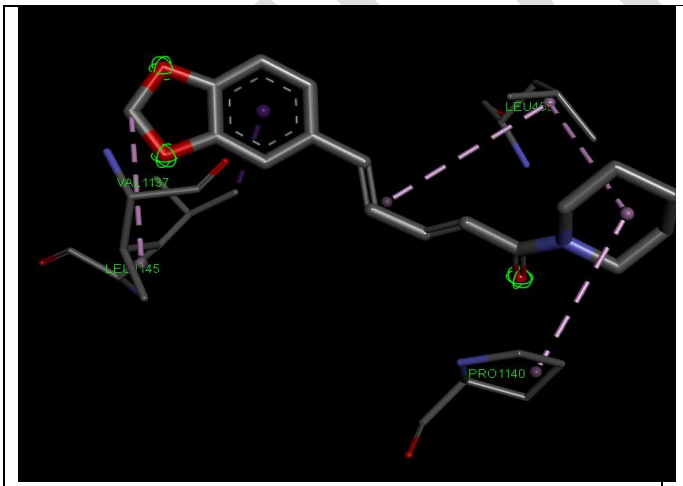


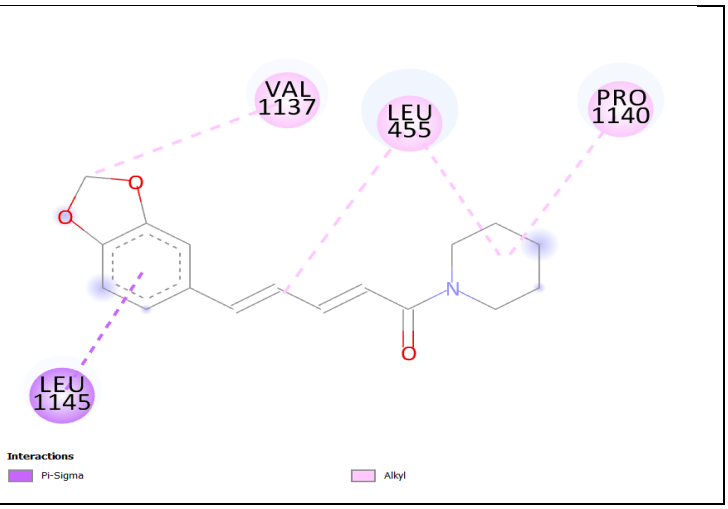
7. EGCG



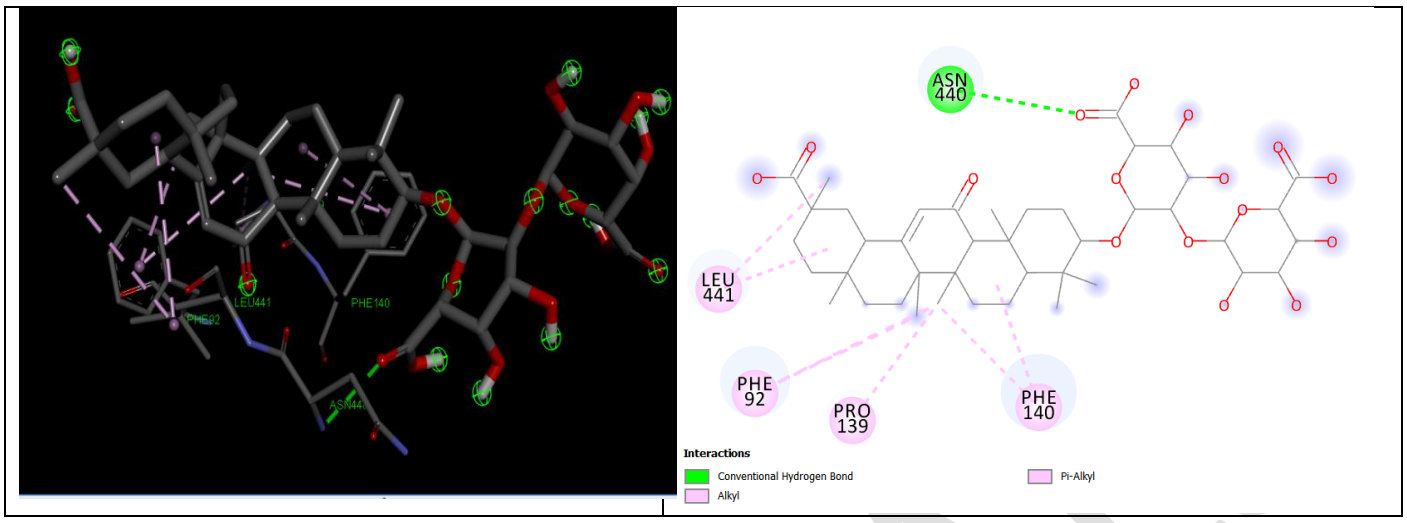


8.Piperine





9. Glycyrrhizic



9] N440K Mutation:

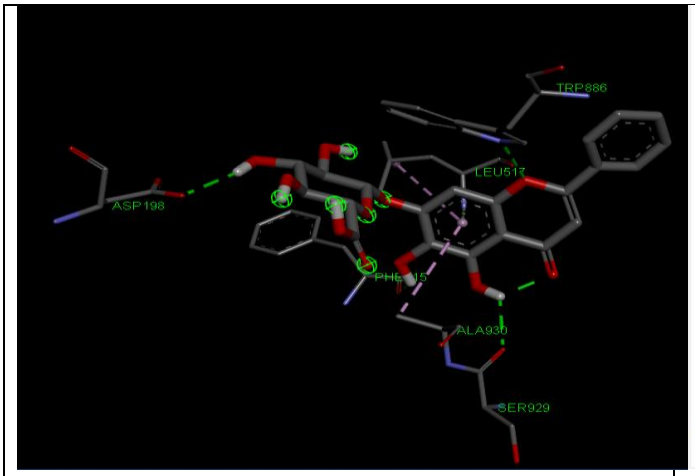
The N440K mutation is associated with greater binding affinity with human host receptors and is associated with greater infectivity ad transmission capability.The genetically tweaked variant with a mutation named N440K has been found in nearly 34% of the 272 SARS-CoV-2 genomes analyzed from Andhra Pradesh in December 2020. The variant has also been seen in Karnataka, Maharashtra and Telangana.For India, N440K high frequency variant was observed out of 19 immune escape variants. This mutation is located at C135 interaction interface. In wild type strain, N440 forms a strong H-bond network with D54 and weak H-bond networks with P52 and R55 of C135 antibody. Interestingly, it is observed 100% co-occurrence on N440K mutation along C64F mutation in membrane glycoprotein of SARS-CoV-2 apart from globally dominant D614G (S protein) and P323L (ORF1ab).

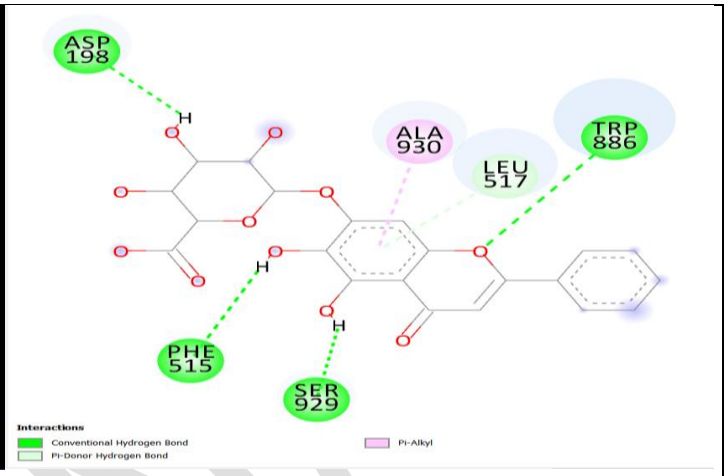
The calculated binding-free energy of mutant RBDs of spike protein complexed with human ACE2 revealed only four RBD mutant types (D364Y, N440K, N450K, S477R) displaying a much lower binding-free energy (ΔG), indicating a significantly higher affinity for the ACE2, which could influence the pathogenicity of SARS-CoV-2.

SN	Protein Info	PDB ID	Compounds Name								
			Comp 1	Comp 2	Comp 3	Comp 4	Comp 5	Comp 6	Comp 7	Comp 8	Comp 9
	SARS-CoV-2 Mutation										
9	Spike Glycoprotein	N440K	-7.55	-7.96	-7.21	-4.91	-6.31	-6.98	-7.07	-7.40	-7.69

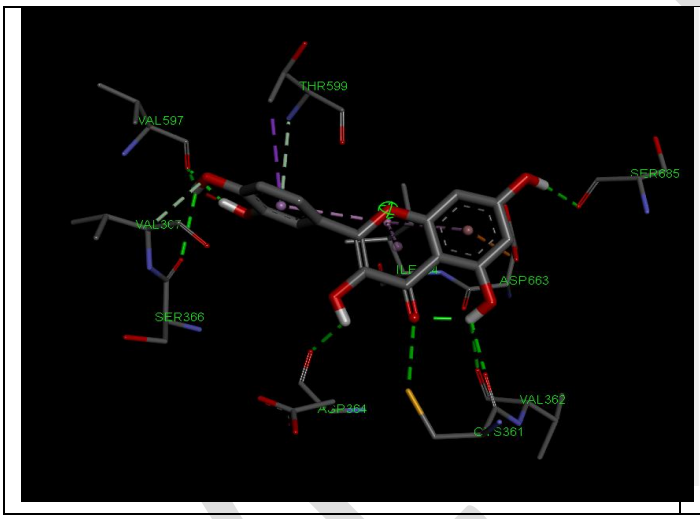
Docking Study 8 Spike Glycoprotein PRONK (N440K) with Ligand Baicalin, Quercetin, Luteolin, Rutin, Hesperidin, Curcumin, EGCG, Piperine, Glycyrrhizic

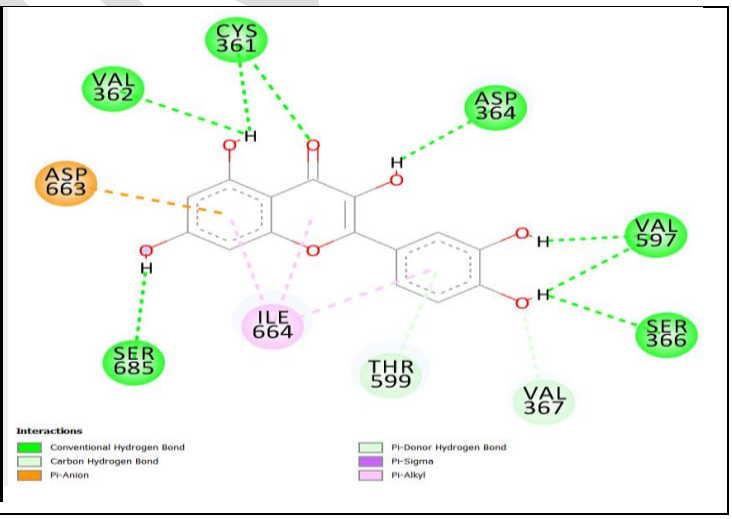
1.Baicalin



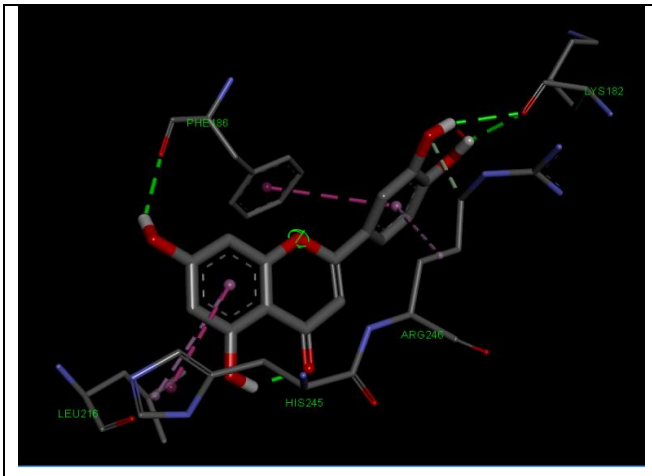


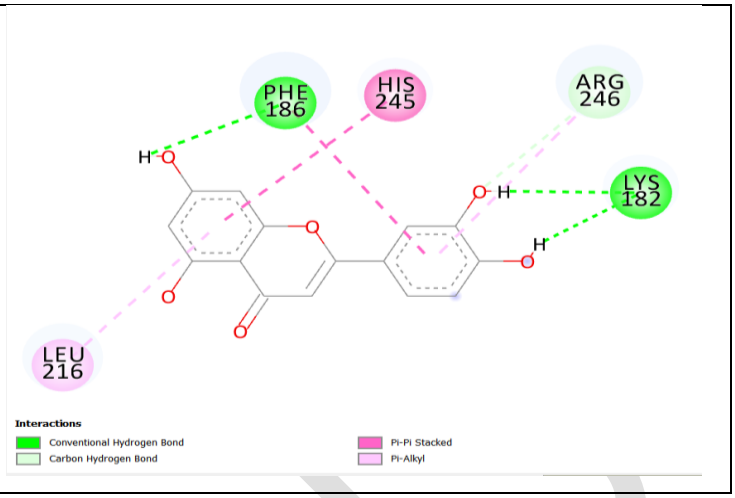
2.Quercitin



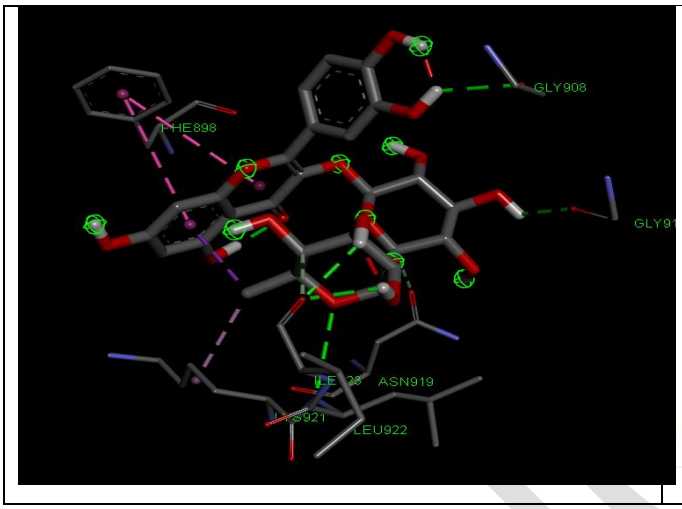


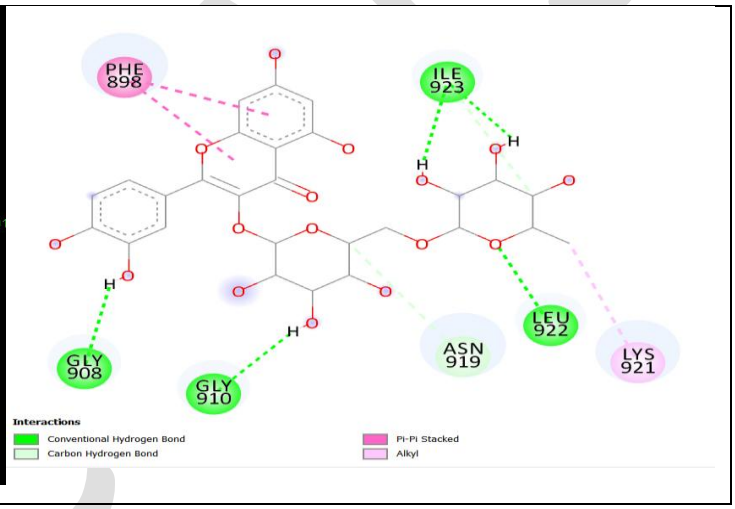
3.Leuteolin



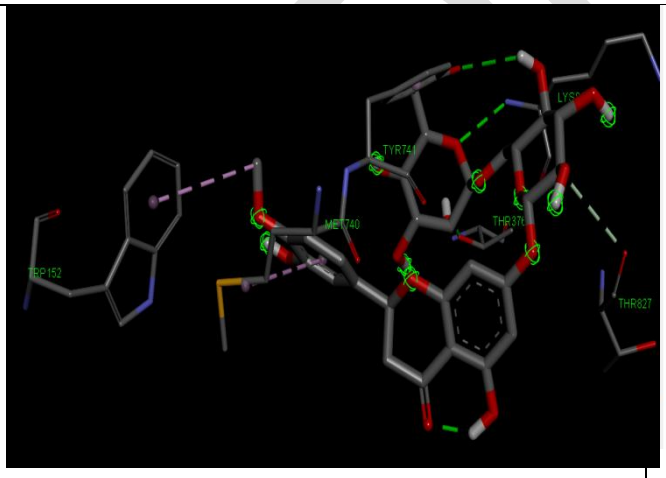


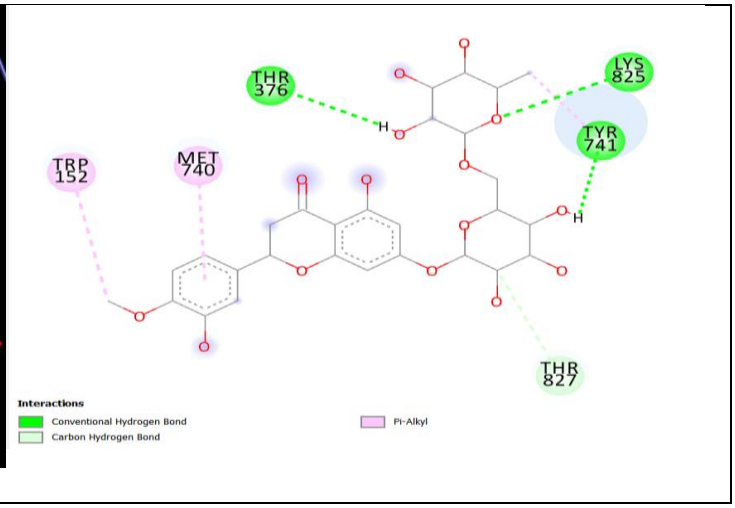
4.Rutin



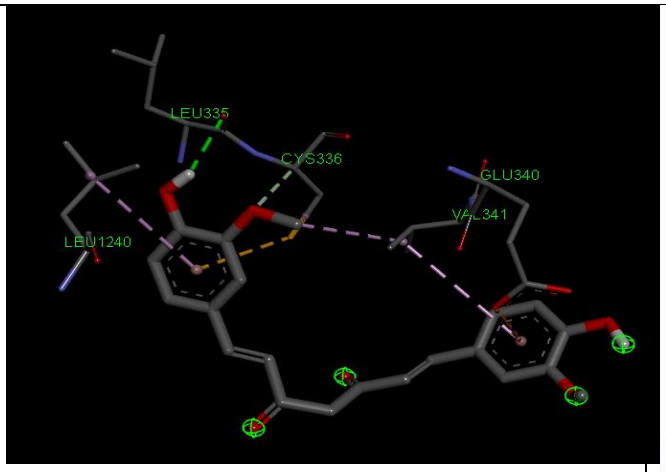


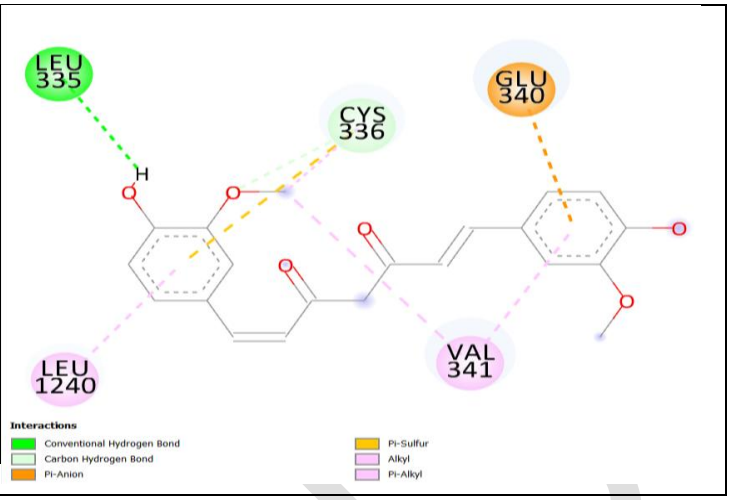
5. Hesperidin



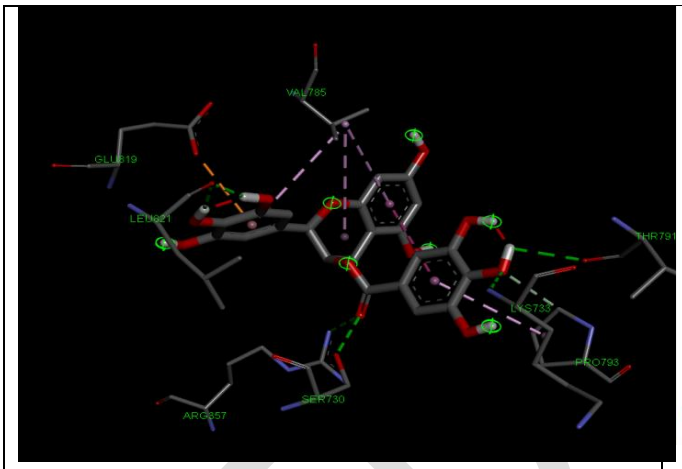


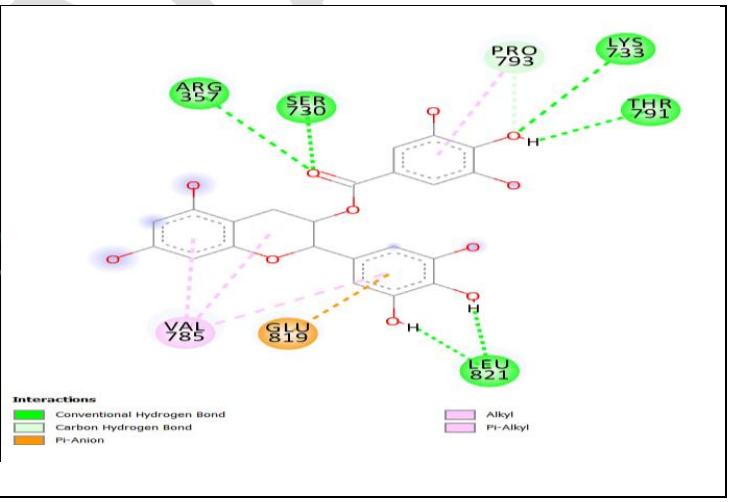
6. Curcumin



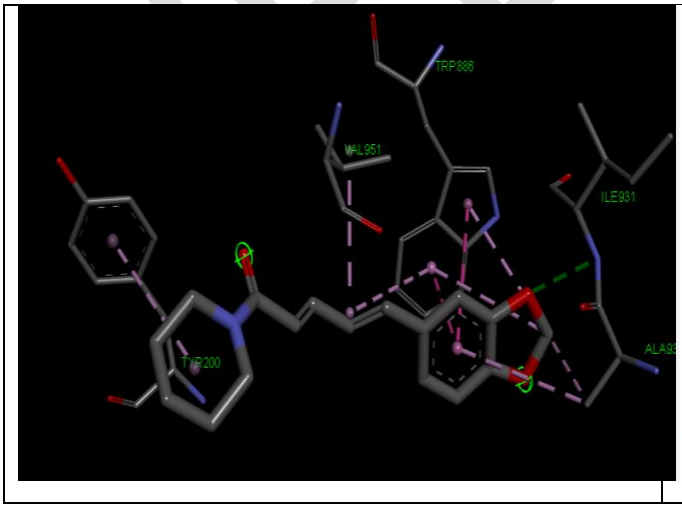


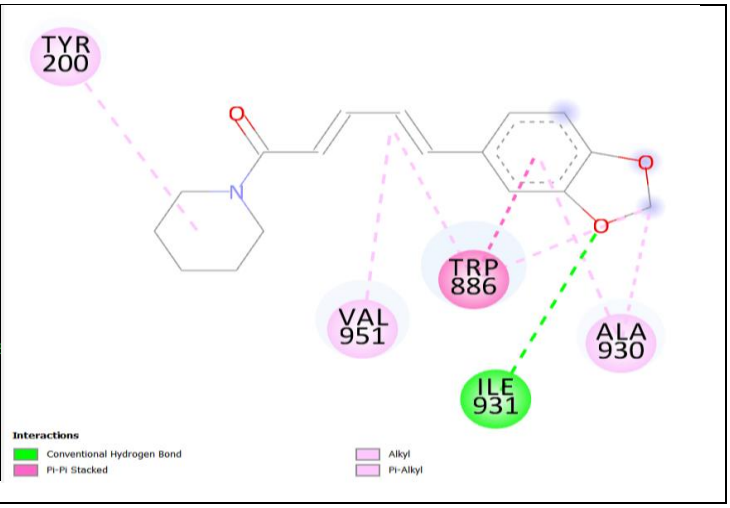
7. EGCG



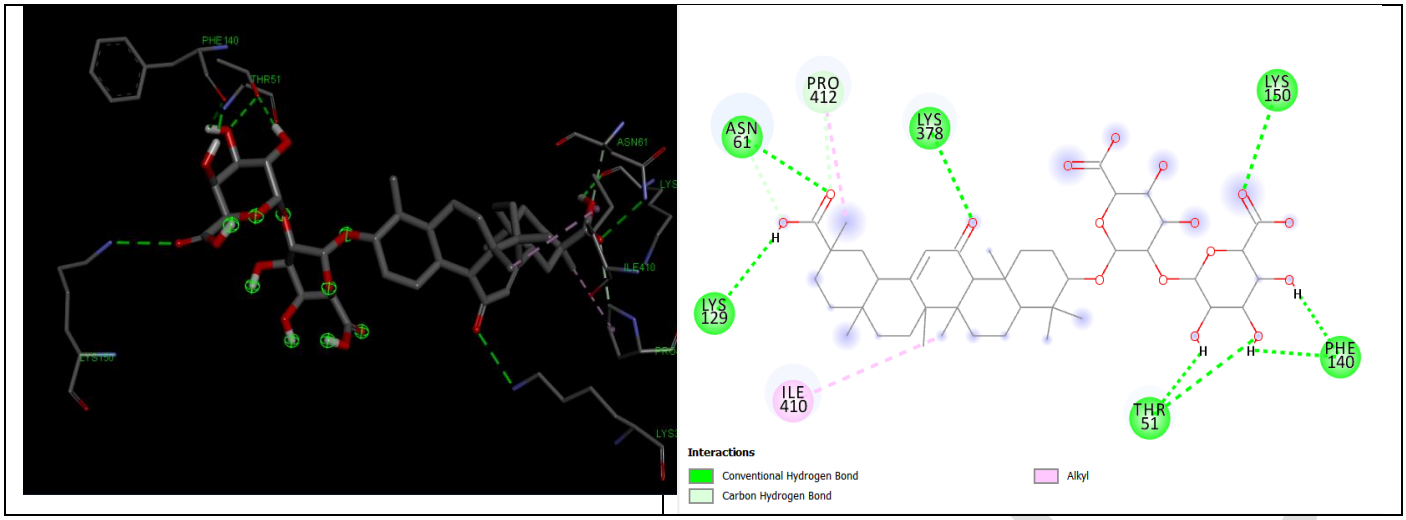


8.Piperine





9. Glycyrrhizic



Experimental:

Ligand preparation:

The structure of the nine ligands were downloaded from PUBCHEM database converted into 3D before analysis. The 3D structure later converted from SDF to PDB format using PyMOL software. The metals were also removed from the ligands using PyMOL software for appropriate docking study. The prepared ligands were saved in PDB format for further docking studies. On the basis of energy minimization the drug binds to effectors/receptors in the most stable form that is the minimum energy form. The active compounds were subjected to conformational analysis and energy minimization using Monte Carlo conformational search. Low energy conformers of all the structures were generated, which was utilized further for analysis.(Kim et al. 2016)

Protein Preparation:

The crystal structure of target proteins was retrieved from Protein Data Bank (PDB) with PDB IDs and was carried further for more studies of docking process. The crystal structure of (PDBID: D614G), (PDBID: G476S), (PDBID: V483G), (PDBID: E484K), (PDBID: P681H), (PDBID: K417N), (PDBID: N501Y), (PDBID: N440K)receptor subunit downloaded from PDB database.(Berman et al. 2002) All the water molecules were removed from the crystal structure of: 5RHC, 6VYB, 6X6P, 6M3M, 6W9C, 7BV2, 6WX4 receptor.

Molecular docking

Molecular Docking is an important component of computer-assisted drug discovery. It helps in predicting the intermolecular framework formed between a protein and ligand and outputs the appropriate binding between the molecules. AutoDock tool was used for the protein synthesis and Grid generation.

Docking was performed by AutoDock 4.2.6 program, using the implemented empirical free energy function and the Lamarckian Genetic Algorithm (LGA).Polar hydrogens were added into the structure and Gasteiger charges were computed and applied accordingly. Missing residues in the proteins were also added at the time of preparation. Molecular docking study of all 9 ligand and the all compounds D614G, G476S, V483G, E484K, P681H, K417N, N501Y, N440Kwere executed with AutoDock 4.2 software. The grid maps were calculated using AutoGrid. In all dockings grid-point spacing of 1.000 Å was applied. Using the gradient optimization algorithm and an empirical scoring function, the molecular docking was conducted to generate the best binding affinity or fitness of protein-ligand binding poses between compounds as D614G, G476S, V483G, E484K, P681H, K417N, N501Y, N440K receptor and 9 ligand. The best binding conformations of ligands were selected and analyzed using AutoDock 4.2 software as well as in Discovery Studio 4.0.(Yuan, Chan, and Hu 2017)

The best conformation with the lowest docked energy was chosen from the docking search. Number of torsions are choosen from 0-6, and if any ligand shows more than 6 it is adjusted to 6. Hydrogen bond interactions are also calculated and mentioned, presence of H-bonds depicts stable interaction between ligand and protein. Discovery studio 2020 Client and Chimera softwares are used to depict Hydrogen bonds, 2-D images and protein-ligand interactions images for a good visualization of the docking.Essentially, the aim of molecular docking is to give a prediction of the ligand-receptor complex structure using computation methods. Binding energy is a measure of the affinity of ligand-protein complex, or is the difference between the energy of complex and the sum of energies of each molecule separately(Gaillard 2018). The binding energy of protein which is greater than -5 is choose. The protein N440K with comp1binding energy is -7.08, comp 2 with -7.06,comp3 with -6.41, comp7 with -5.40,comp8 with -5.14,comp9 with -8.57.The protein N440K with comp1binding energy is -7.43,comp 2 with -7.40,comp3 with -7.01, comp5 with -5.06,comp6 with -6.48, comp7 with -6.80,comp8 with -6.60,comp9 with -7.76.The protein N440K with comp1binding energy is -7.70,comp 2 with -7.29,comp3 with -7.54, comp5 with -5.59,comp6 with -6.97, comp7 with -6.36,comp8 with -7.72,comp9 with -8.00.The protein N440K with comp1binding energy is -7.18,comp 2 with -6.56,comp3 with -6.83, comp7 with -5.62,comp8 with -6.92,comp9 with -6.79.The protein N440K with comp1binding energy is -6.80,comp 2 with -6.17,comp3 with -6.92, comp6 with -6.07, comp7 with -6.24,comp8 with -6.03,comp9 with -6.94.The protein N440K with comp1binding energy is -7.26,comp 2 with -7.35,comp3 with -7.54, comp5 with -6.65,comp6 with -6.89, comp7 with -6.89,comp8 with -7.38,comp9 with -7.28.The protein N440K with comp1binding energy is -7.39,comp 2 with -7.35,comp3 with -7.06, comp5 with -6.31,comp6 with -6.98, comp7 with -7.07,comp8 with -7.40,comp9 with -7.69.The protein N440K with comp1binding energy is -7.55,comp 2

with -7.96,comp3 with -7.21, comp5 with -6.31,comp6 with -6.98, comp7 with -7.07,comp8 with -7.40,comp9 with -7.69.

References:-

Berman, Helen M. et al. 2002. "The Protein Data Bank." *Acta Crystallographica Section D: Biological Crystallography* 58(6 I): 899–907.

Gaillard, Thomas. 2018. "Evaluation of AutoDock and AutoDock Vina on the CASF-2013 Benchmark." *Journal of Chemical Information and Modeling* 58(8): 1697–1706.

Kim, Sunghwan et al. 2016. "PubChem Substance and Compound Databases." *Nucleic Acids Research* 44(D1): D1202–13.

Yuan, Shuguang, H. C.Stephen Chan, and Zhenquan Hu. 2017. "Using PyMOL as a Platform for Computational Drug Design." *Wiley Interdisciplinary Reviews: Computational Molecular Science* 7(2): 1–10.