

# **Identification of Potential Binders of the SARS-Cov-2 Spike Protein via Molecular Docking, Dynamics Simulation and Binding Free Energy Calculation**

**Chirag N. Patel<sup>1,¶</sup>, Sivakumar Prasanth Kumar<sup>2,¶</sup>, Himanshu A. Pandya<sup>1</sup> and Rakesh M. Rawal<sup>2</sup> \***

<sup>1</sup>Department of Botany, Bioinformatics, and Climate Change Impacts Management, University School of Sciences, Gujarat University, Ahmedabad, 380009, India.

<sup>2</sup>Department of Life Sciences, University School of Sciences, Gujarat University, Ahmedabad, 380009, India.

**¶Equal first authors**

**\*Corresponding author:** rakeshrawal@gujaratuniversity.ac.in and +91-0-9925244855

**ORCID profiles:**

Chirag N. Patel, <http://orcid.org/0000-0003-0777-7720>

Sivakumar Prasanth Kumar, <https://orcid.org/0000-0003-3461-2489>

Himanshu A. Pandya, <https://orcid.org/0000-0003-4536-9541>

Rakesh M. Rawal, <https://orcid.org/0000-0002-7985-1187>

# Identification of Potential Binders of the SARS-Cov-2 Spike Protein via Molecular Docking, Dynamics Simulation and Binding Free Energy Calculation

## ABSTRACT

The pandemic outbreak of COVID-19 virus (SARS-CoV-2) has become critical global health issue. The biophysical and structural evidence shows that SARS-CoV-2 spike protein possesses higher binding affinity towards angiotensin-converting enzyme 2 (ACE2) and hemagglutinin-acetylcysteine (HE) glycoprotein receptor. Hence, it was selected as a target to generate the potential candidates for the inhibition of HE glycoprotein. The present study focuses on extensive computational approaches which contains molecular docking, ADMET prediction followed by molecular dynamics simulations and free energy calculations. Furthermore, virtual screening of NPACT compounds identified 3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2-yl]benzo[7]annulen-6-one, Silymarin, Withanolide D, Spirosolane and Oridonin were interact with high affinity. The ADMET prediction revealed pharmacokinetics and drug-likeness properties of top-ranked compounds. Molecular dynamics simulations and binding free energy calculations affirmed that these five NPACT compounds were robust HE inhibitor.

**Key words:** COVID-19 virus (SARS-CoV-2), angiotensin-converting enzyme 2 (ACE2) and hemagglutinin-acetylcysteine (HE) glycoprotein, molecular docking, Molecular dynamics simulations, NPACT compounds

# **Identification of Potential Binders of the SARS-Cov-2 Spike Protein via Molecular Docking, Dynamics Simulation and Binding Free Energy Calculation**

## **INTRODUCTION**

Coronaviruses (CoV) is a (+) RNA containing virus which appears as an oval-shaped envelop with spike like protrusions [1] (+) strand RNA [2]. The initiation of CoV was first emerged in Middle East and it causes respiratory tract disease called as Middle East Respiratory Syndrome Coronavirus (MERS-CoV) which was caused by new form of SARS-CoV (Severe Acute Respiratory Syndrome) [3, 4]. Coronavirus disease 2019 (COVID-19) is a pandemic and caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). CoV infection develops asymptotically as fever, cough, and severe shortness of breathing, nausea, vomiting and diarrhoea symptoms [5]. The envelop of SARS-CoV composed of various types of proteins envelop, membrane, nucleocapsid, replicase, spike glycoprotein, and other accuracy proteins 3a, 6, 7a and 7b [6-12] .

Recent reports by Vankadari and Wilce, 2020 and Wrapp *et al.*, 2020 glycosylation process occurs between the spike protein of coronaviruses and human host cell [13]. This spike protein recruits S1 (N-terminal half) and S2 (C-terminal half) fusion peptides [14] along with hemagglutinin-acetyesterase (HE) glycoprotein to interact with angiotensin-converting enzyme 2 (ACE2) [2], a surface receptor protein expressed in lungs, heart, kidneys, and intestine [15-20] cell adhesion and virulence [21, 22]. Recent studies showed that the ectodomain of the SARS-CoV-2 with ~15 nM affinity passes ACE2 receptor [23]. Numerous reports suggest that Zhou *et al.*, 2020, Walls *et al.*, 2020, Letko, M., Marzi, A. & Munster, V, 2020 and Hoffmann, M. *et al.*, 2020 SARS-CoV-2 utilizes ACE2 for the inserting viral particles into human target ACE2 cells [24-27]. After establishing connection for sugar elements on cell membranes, the hemagglutinin–acetyesterase (HE) of SARS-CoV-2 inserts messenger RNA and performs replication process [28] and it causes various diseases leads to death. World Health Organization declared (WHO) announced that the COVID-19 has become pandemic and due to its infection 38,62,676 confirmed cases, 2,65,961 confirmed deaths in more than 215 countries, areas, or territories with cases [29].

To combat the first line of infection developed due to CoV2 spike and ACE2 complex formation, computational strategies are needed to prevent this complex formation by developing small molecules which can target at its complex interface site. In this study, we

employed molecular docking approach to screen NPACT (Naturally occurring Plant-based Anti-cancer Compound activity-Target database) compounds and the best-scoring molecules were validated using molecular dynamics simulations analysis to better understand the interactions and conformational changes [30, 31].

## **MATERIALS AND METHODS**

Different molecular modeling techniques were employed in this study viz, molecular docking, ADMET prediction, molecular dynamics simulations, and binding free energy calculations to obtain novel leads from NPACT compounds [30, 32]. NPACT compounds have diverse pharmacological effects with natural drug-like characteristics. These compounds possess absorption, distribution, metabolism, excretion and toxicity (ADMET) properties which leads to determine the potential drug candidates for drug discovery process [33].

### **Molecular data set and its preparation**

The coronavirus belongs to +single-stranded RNA harbouring virus family and possesses hemagglutinin–acetylerase (HE) glycoprotein which is responsible for the connection with sugar moieties on cell membranes of receptor [28, 34-36]. This HE acquires flexibility to bind with O-acetylated sialic acids which demolishes receptor and its membrane by working as HE fusion protein [37]. The x-ray diffracted structure of coronavirus hemagglutinin-esterase in complex with 4,9-O-diacetyl sialic acid (PDB ID:3CL5) was chosen as a receptor for the study which containing the 4,9-O-diacetyl sialic acid as inbound ligand with single chain and 377 amino acids length and 1.80 Å resolution [2, 38]. In order to perform, Natural product-based drug discovery 1574 natural compounds were retrieved from NPACT database [33, 39]. These protein and natural compounds were prepared using yasara parameters like removing crystallographic waters, adding polar hydrogens and ions followed by atom typed using the Amber03 force field and geometry optimized using the steepest gradient approach (100 iterations) [40, 41].

### **Molecular docking**

Molecular docking was implemented on coronavirus hemagglutinin-esterase and 1574 NPACT compounds using YASARA software version 19.12.14 with AutoDock Vina algorithm and AMBER03 force field [40-42]. The docking scores were calculated using the following empirical equation:

$$\Delta G = \Delta G_{vdW} + \Delta G_{Hbond} + \Delta G_{elec} + \Delta G_{tor} + \Delta G_{desol}$$

Where  $\Delta G_{vdW}$  = van der Waals term for docking energy;  $\Delta G_{Hbond}$  = H bonding term for docking energy;  $\Delta G_{elec}$  = electrostatic term for docking energy;  $\Delta G_{tor}$  = torsional free energy term for compound when the compound transits from unbounded to bounded state;  $\Delta G_{desolv}$  = desolvation term for docking energy

### **Lipinsky's rule and admet prediction**

SwissADME webserver was used for the absorption, distribution, metabolism, elimination, toxicity, prediction of top scoring compounds. From this server, the Lipinski rule of 5 [43], GI absorption, BBB permeant, cytochrome inhibition were calculated which enhance the probability these compounds are more potent as a drug [44, 45].

### **Assessment of the stability of the identified leads with protein**

#### ***Molecular dynamics simulations study***

The physical movements of atoms and molecules of protein–ligand docked complex was identified through molecular dynamics simulation. Top scoring 5 compounds were chosen for the MD simulation at 10 ns time interval each. Before proceeding the MD simulation the docked complexes were prepared using steepest descent minimization, YASARA energy minimization module (YASARA Biosciences, GmbH) using AMBER (Assisted model building with energy refinement) force field [46, 47]. The pre-defined parameters of MD simulation were applied which includes AMBER03 force field, temperature of 298 K, pressure at 1 bar, coulomb electrostatics at a cut off of 7.86, 0.9% NaCl, solvent density 0.997, pH 7.0, 1-fs time steps, periodic boundaries, all atoms mobile [31, 48] . After the completion of MD simulation, structural level integrity and conformational changes which have been occurred in the protein–ligand docked complex were identified through interaction profiles and superposition study [49, 50].

#### ***Binding free energy calculations***

The single trajectory approach was used for the binding free energy calculation using yasara Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA approaches) [51-55] YASARA AMBER14 with the “single trajectory approach”. The MM/PBSA calculations, contains 10 ns MD simulations of all 15 docked protein-ligand complexes. The free energy values were calculated using the following equations:

$$\Delta G_{bind} = \Delta G_{complex(minimized)} - [\Delta G_{ligand(minimized)} + \Delta G_{receptor(minimized)}]$$

and

$$\Delta G_{bind} = \Delta G_{MM} + \Delta G_{PB} + \Delta G_{SA-TAS}$$

Where  $\Delta G_{TDS}$  is the conformation entropic contribution, and  $\Delta G_{MM}$  is the molecular mechanics interaction energy (electrostatic + van der Waals interaction) between protein and ligand.  $\Delta G_{PB}$  and  $\Delta G_{SA}$  depict the polar solvation energy and the nonpolar solvation energy, respectively [56].

## RESULTS AND DISCUSSION

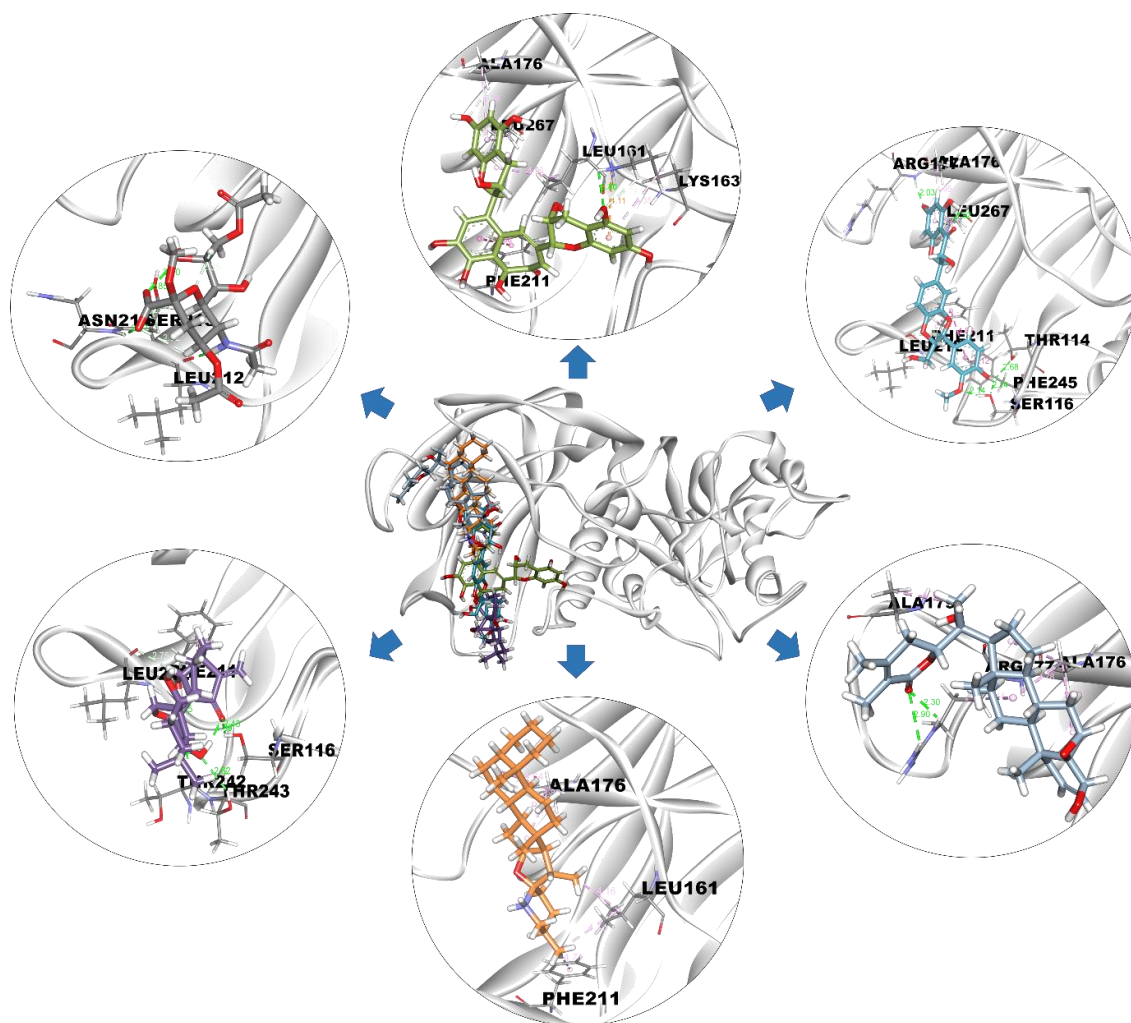
### Molecular docking and its validation

The molecular docking analysis was carried out for the identification of potential NPACT compounds which can inhibit the HE glycoprotein. The docking procedure was first justified by redocking the co-crystal ligand 4,9-O-diacetyl sialic acid into CoV-2 hemagglutinin-esterase protein and calculate the RMSD between bound and docked conformation of co-crystal ligand. A RMSD of 1.84 Å was achieved which depicted that the robustness of the docking procedure to develop physically related dock poses as close as possible. The redocked pose possessed the binding energy of 5.233 kcal/mol with 3 hydrogen bonds and 0.1945 kcal/mol/Atom efficiency with competitive inhibition. The co-crystal ligand showed 3 hydrogen bonds with Leu 212, Ser 213 and Asn 214. This co-crystal ligand was bounded beside the hydrophobic pocket having Tyr184, Phe211, Leu266, and Leu267 which were maintained by Asp220, Ser221, Gln222, Ser263, Glu265, Leu267. This 9-O-acetyl moiety plays an essential role to inhibit the HE glycoprotein. With this prerequisite approach, 1574 lead-like NPACT compounds were docked within the active site of coronavirus hemagglutinin-esterase complex (3cl5).

**Table-1 Binding energy, Hydrogen bonds, Efficiency and contacting receptor residues of selected ligands.**

Name	Binding energy [kcal/mol]	Hydrogen Bonds	Dissoc. constant [pM]	Efficiency [kcal/(mol *Atom)]	Contacting receptor residues
4,9-O-diacetyl sialic acid	5.23	3	145935232	0.1945	Thr 114, Thr 159, Leu 161, Ala 176, Arg 177, Tyr 184, Phe 211, Leu

					212, Ser 213, Asn 214, Phe 245, Leu 266, Leu 267
3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2-yl]benzo[7]annulen-6-one	7.66	1	2427424	0.1868	Cys 113, Thr 114, Thr 115, Thr 159, Leu 161, Lys 163, Ser 164, Ala 176, Arg 177, Tyr 184, Phe 211, Leu 212, Ser 213, Asn 214, Leu 267
Silymarin	7.33	4	4208301	0.2095	Thr 114, Ser 116, Gly 117, Thr 159, Leu 161, Ala 176, Arg 177, Tyr 184, Phe 211, Leu 212, Thr 242, Thr 243, Phe 245, Leu 267
Withanolide D	7.29	1	4502224	0.2145	Ala 156, Gln 157, Thr 159, Leu 161, Ala 176, Arg 177, Glu 178, Ala 179, Phe 181, Tyr 184, Leu 267
Spirosolane	7.25	0	4776200	0.2503	Ser 155, Ala 156, Gln 157, Thr 159, Leu 161, Ala 176, Arg 177, Ala 179, Tyr 184, Phe 211, Ser 213, Leu 266, Leu 267
Oridonin	7.22	3	5101171	0.2777	Thr 114, Ser 116, Phe 207, Lys 210, Phe 211, Leu 212, Thr 242, Thr 243, Phe 245



**Figure-1** Docked poses of 4,9-O-diacetyl sialic acid, 3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2-yl]benzo[7]annulen-6-one, Silymarin, Withanolide D, Spirosolane and Oridonin.

Based on the binding energy of co-crystal ligand, the threshold was set to 6. The top ranged compounds were chosen which ranges from 7.22 to 7.66 kcal/mol. In addition, binding energy, hydrogen bonds, efficiency and interactions with receptor residues depicted that the top ranked molecules were more potent than the co-crystal ligand binding (**Figure-1**). The result in **Table-1** demonstrated that compound 3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2-yl]benzo[7]annulen-6-one showed highest binding energy 7.66 kcal/mol, followed by Silymarin, Withanolide D, Spirosolane and Oridonin. Thr 114, Thr 159, Leu 161, Ala 176, Arg 177, Tyr 184, Phe 211, Leu 212, Ser 213, Asn 214, Leu 267 were found as most shared key residues in all compounds accounting hydrogen bonding and high efficiency which may aid the competitive inhibition with low molecular weight. These results represented that these candidates can be potential drug candidate which can inhibit the coronavirus



hemagglutinin–acetylcholinesterase (AChE) glycoprotein progression. Furthermore, conformational changes, stability and MM/PBSA assessment were carried out for all these protein-ligand complexes.

### Lipinsky's rule and ADMET prediction

SwissADME server was used for the ADMET prediction of top-most ranked compounds which calculates pharmacokinetic properties, druglike nature and medicinal chemistry friendliness also. **Table-2** shows drug-likeness properties of Lipinski's rule-of-five, in which Oridonin, Spirosolane and Withanolide D were fulfilled all the properties within the range. 3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2-yl]benzo[7]annulen-6-one and Silymarin exceeded the molecular weight and TPSA ranges. However, all the compounds passed the remaining ADMET prediction (ESOL Solubility (mg/ml, GI absorption, BBB permeant, Pgp substrate, CYP1A2 inhibitor, CYP2C19 inhibitor, CYP2C9 inhibitor, CYP2D6 inhibitor, CYP3A4 inhibitor, log Kp (cm/s)) shown in **Table-3**.

**Table-2 Molecular pharmacokinetic properties calculated for the top-ranked compounds with SwissADME webserver.**

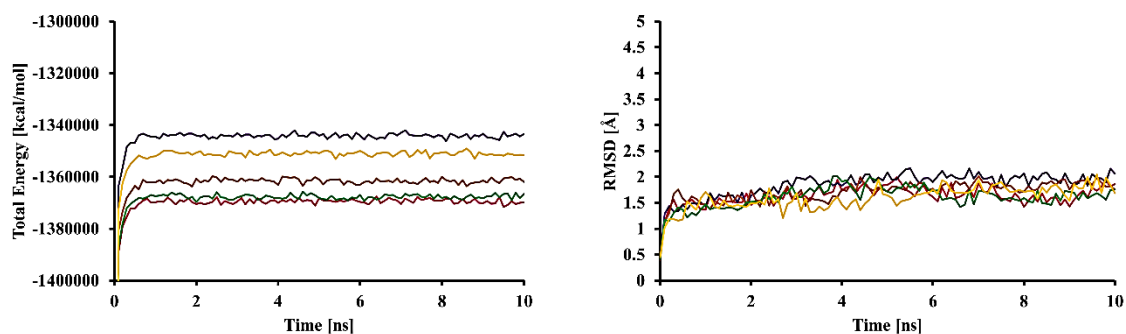
Molecule	3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2-yl]benzo[7]annulen-6-one	Oridonin	Silymarin	Spirosolane	Withanolide D
MW	566.51	364.43	482.44	400.66	470.6
Fraction Csp3	0.28	0.85	0.24	1	0.79
Rotatable bonds	2	0	4	0	2
HBA	12	6	10	1	6
HBD	9	4	5	1	2
TPSA	217.6	107.22	155.14	25.84	96.36
XLOGP3	0.8	0.08	1.9	7.92	3.12
ESOL Log S	-4.05	-2.15	-4.14	-7.31	-4.59

**Table-3 Predicted ADMET properties of top-ranked compounds.**

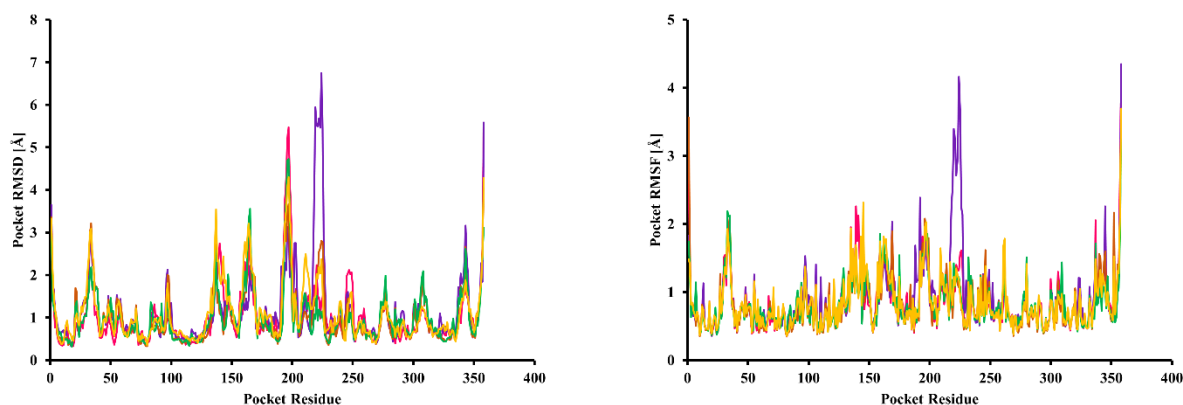
Molecule	3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2-yl]benzo[7]annulen-6-one	Oridonin	Silymarin	Spirosolane	Withanolide D
ESOL Solubility (mg/ml)	5.06E-02	2.58E+00	3.46E-02	1.95E-05	1.21E-02
GI absorption	Low	High	Low	High	High
BBB permeant	No	High	Low	Yes	No
Pgp substrate	No	No	No	No	Yes
CYP1A2 inhibitor	No	Yes	No	No	No
CYP2C19 inhibitor	No	No	No	No	No
CYP2C9 inhibitor	No	No	No	No	No
CYP2D6 inhibitor	No	No	No	No	No
CYP3A4 inhibitor	No	No	Yes	No	No
log Kp (cm/s)	-9.19	-8.47	-7.89	-3.12	-6.96

**Molecular dynamics simulations**

To investigate the dynamic properties of the HE glycoprotein with top-ranked NPACT compounds necessary for structural changes related to inhibition mechanism. MD simulations of this protein and top-ranked 5 compounds were carried out for the 10 ns simulation time, respectively. During this time interval the total energy of all 5 protein-ligand complexes were remained from -1340000 to -1380000 kcal/mol (**Figure-2**) while RMSD values lied under 2.3 Å which depicted less fluctuations. While pocket residues against pocket RMSD and pocket RMSF plots (**Figure-3**) suggested fewer fluctuations in Silymarin compound.

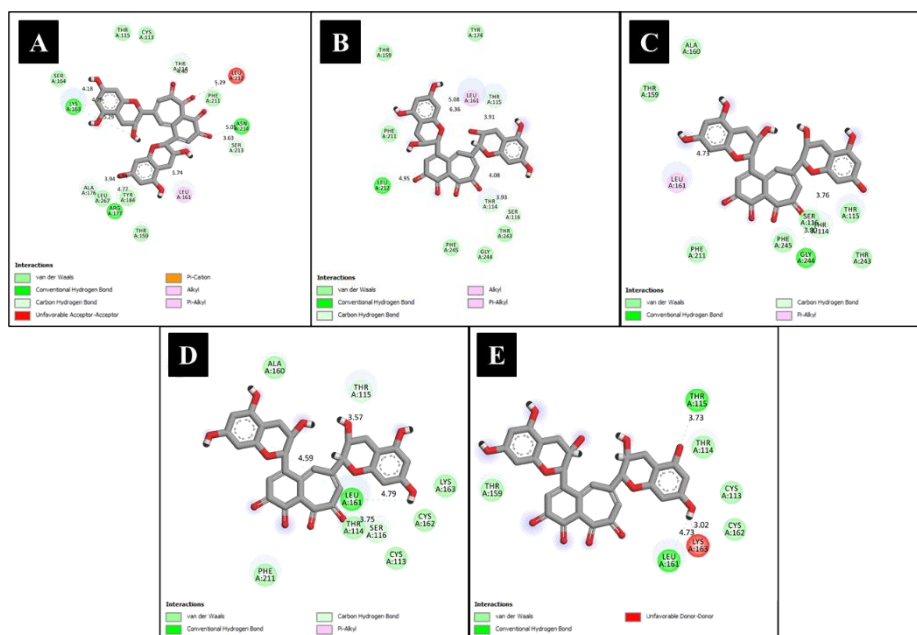


**Figure-2 Molecular Dynamics Simulation analysis: Time vs. Total energy and Time vs. RMSD graph. Color denotation: Magenta-3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2-yl]benzo[7]annulen-6-one; Purple-Silymarin; Green-Withanolide D; Brown-Spirosolane; Yellow-Oridonin.**

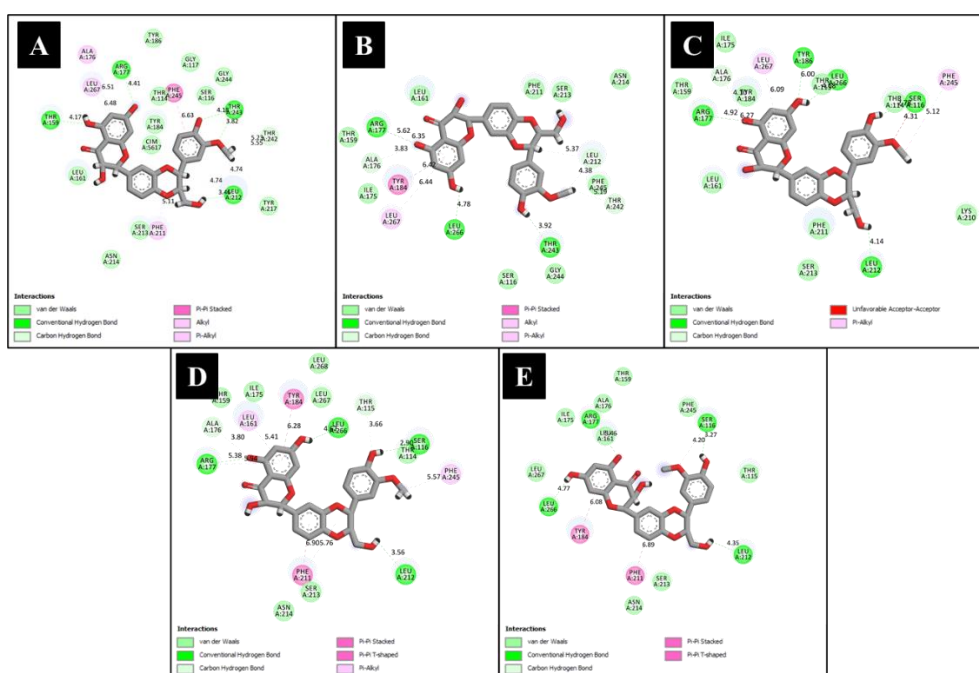


**Figure-3 Stability analysis: Pocket residue vs. Pocket RMSD and Pocket residues vs. Pocket RMSF.**

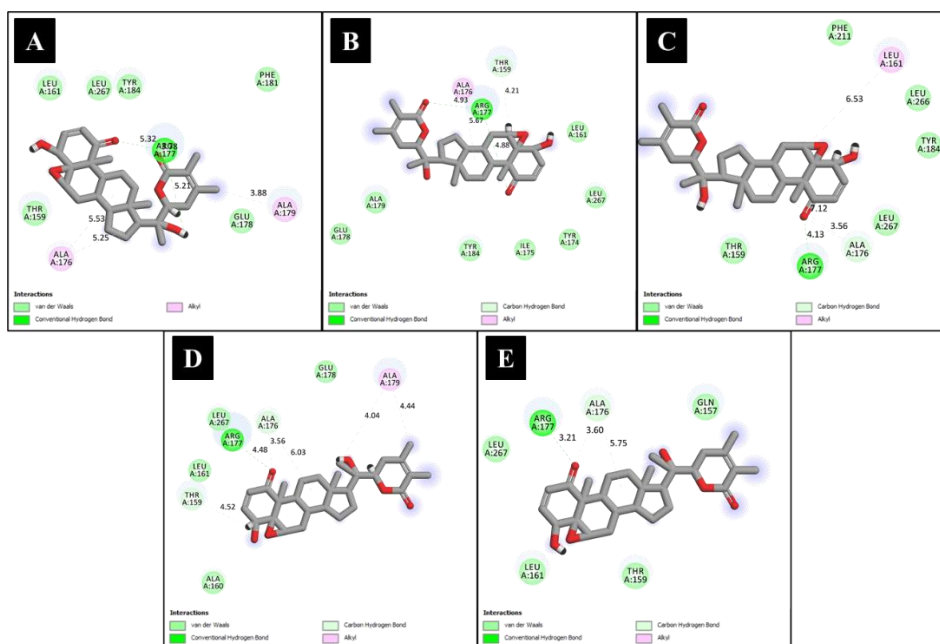
**Figure-S1 to Figure-S5** represents the conformational changes captured during the 10 ns time interval of all 5 docked complexes. **Figure-S1** depicts the changes were observed after the 2.5 ns in CoV-2 HE with 3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2 yl]benzo[7]annulen-6-one which remained stable with minor changes during the whole event. The CoV-2 HE and Silymarin complex possessed the variation at 2.5 ns simulation time which followed by 5 ns as well as 7 ns and 10 ns time interval with conformation changes (**Figure-S2**). However, major change was identified in CoV-2 HE with Withanolide D after 2.5 ns to 7 ns time intervals followed by stabilized complex till 10 ns (**Figure-S3**). CoV-2 HE with Spirosolane and Oridonin complexes persisted no changes during the MD simulation event (**Figure-S4** and **Figure-S5**).



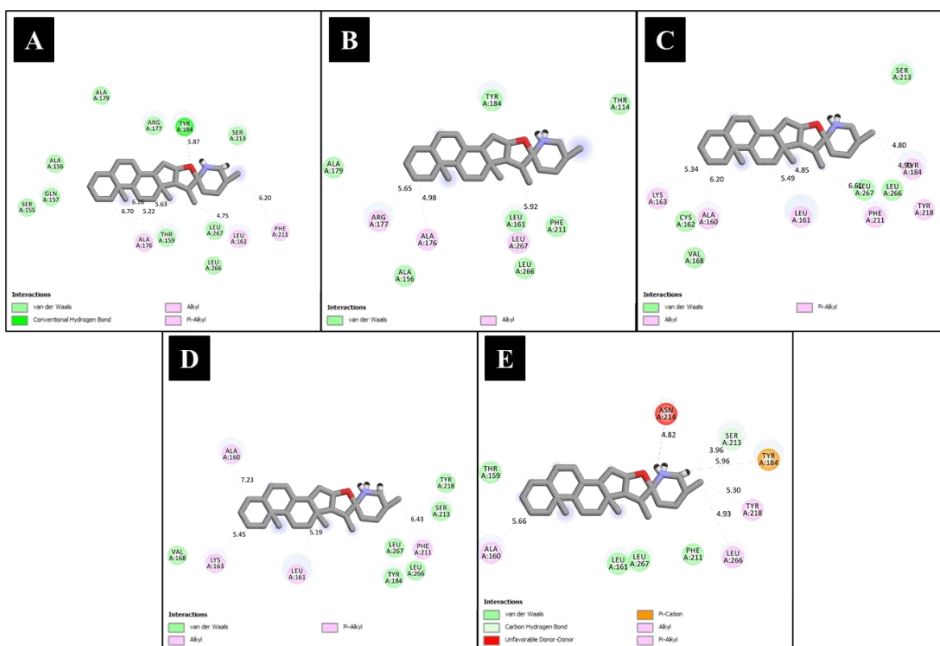
**Figure-S1 Graphical representation of the interactions between Coronavirus Hemagglutinin-Esterase with 3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2-yl]benzo[7]annulen-6-one simulated molecule by Discovery studio visualiser ((A) 0 nanoseconds, (B) 2.5 nanoseconds, (C) 5 nanoseconds, (D) 7.5 nanoseconds, (E) 10 nanoseconds).**



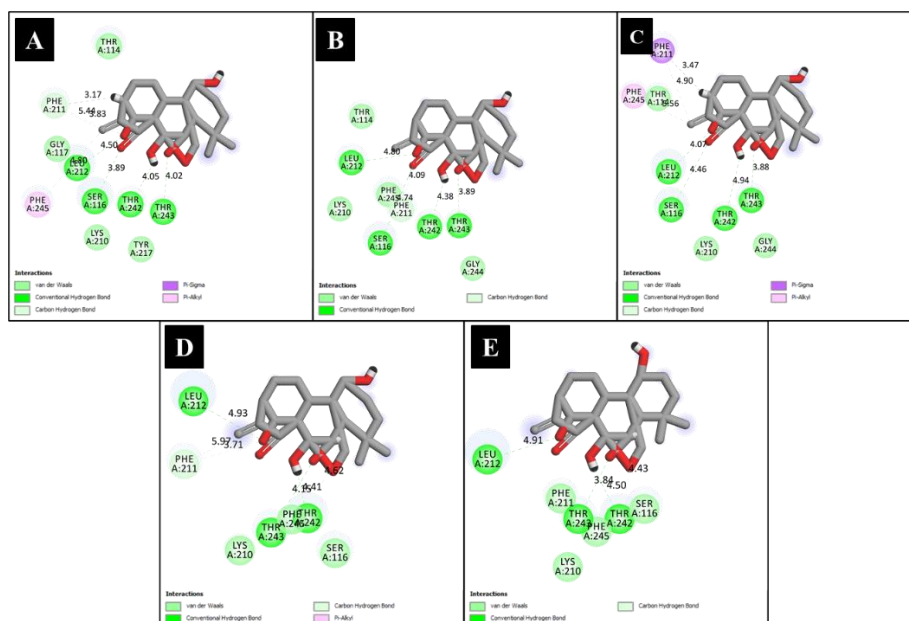
**Figure-S2 Graphical representation of the interactions between Coronavirus Hemagglutinin-Esterase with Silymarin simulated molecule by Discovery studio visualiser ((A) 0 nanoseconds, (B) 2.5 nanoseconds, (C) 5 nanoseconds, (D) 7.5 nanoseconds, (E) 10 nanoseconds).**



**Figure-S3 Graphical representation of the interactions between Coronavirus Hemagglutinin-Esterase with Withanolide D simulated molecule by Discovery studio visualiser ((A) 0 nanoseconds, (B) 2.5 nanoseconds, (C) 5 nanoseconds, (D) 7.5 nanoseconds, (E) 10 nanoseconds).**



**Figure-S4 Graphical representation of the interactions between Coronavirus Hemagglutinin-Esterase with Spirosolane simulated molecule by Discovery studio visualiser ((A) 0 nanoseconds, (B) 2.5 nanoseconds, (C) 5 nanoseconds, (D) 7.5 nanoseconds, (E) 10 nanoseconds).**



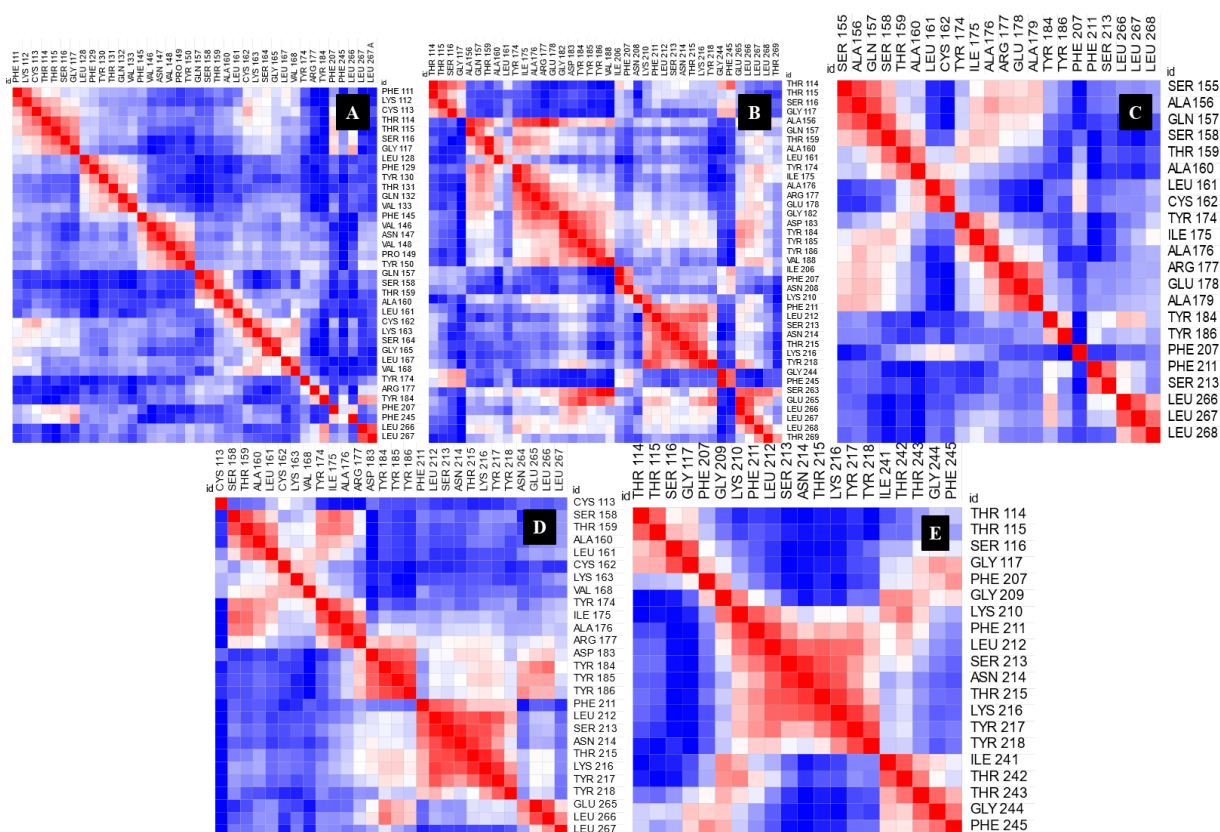
**Figure-S5 Graphical representation of the interactions between Coronavirus Hemagglutinin-Esterase with Oridonin simulated molecule by Discovery studio visualiser ((A) 0 nanoseconds, (B) 2.5 nanoseconds, (C) 5 nanoseconds, (D) 7.5 nanoseconds, (E) 10 nanoseconds).**

During the whole MD simulation events for all docked complexes some stable contacts were identified using 2D interaction diagrams from Discovery studio visualiser. For 3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2yl]benzo[7]annulen-6-one Thr 114 and Thr 115 possessed hydrogen bonding while Leu 161 bounded with Pi-Alkyl interactions. However, Silymarin showed Arg 177, Leu 212 Thr242 and Thr 243 amino acids having hydrogen bonding and Pi-Alkyl with Leu 267 amino acid. With respect to Withanolide, three amino acids were interacted through Alkyl (Ala 176 and Ala 179) and hydrogen bonding (Arg 177). Spirosolane owned the same number of the stable interactions as Withanolide with Ala 160, Tyr 184 (hydrogen bond) and Tyr 218 (Alkyl). Oridonin represented more influence about stable interactions with 6 stable contacts which were identified as five hydrogen bonds (Ser 116, Phe 211, Leu 212, Thr 242 and Thr 243) and one Pi-Alkyl bond interaction with Phe 245.

### Dynamic cross correlation matrix to study residue motions of cross-terms

Dynamic cross correlation matrix depicted the correlative motions of various pocket residues. **Figure-4** shows the residual correlative motion of all simulated protein-ligand complexes. This correlated motions among pocket residues were represented in the form of heatmaps with deep color intensity. The red color exhibited the positive correlation in the residue motions whereas blue color indicated the negative or anti-correlated movements of pocket residues. Among all

DCCM plots the common positive correlated residues were Phe 111, Lys 112, Cys 113, Thr 114, Thr 115, Ser 116, Gly 117, Leu 128, Phe 129, Tyr 130, Thr 131, Gln 132, Val 133, Phe 145, Val 146, Asn 147, Val 148, Pro 149, Tyr 150, Ser 155, Ala 156, Gln 157, Ser 158, Thr 159, Ala 160, Leu 161, Cys 162, Lys 163, Leu 167, Val 168, Tyr 174, Ile 175, Ala 176, Arg 177, Glu 178, Ala 179, Gly 182, Asp 183, Tyr 184, Tyr 185, Tyr 186, Val 188, Ile 206, Phe 207, Asn 208, Gly 209, Lys 210, Phe 211, Leu 212, Ser 213, Asn 214, Thr 215, Lys 216, Tyr 217, Tyr 218, Ile 241, Thr 242, Thr 243, Gly 244, Phe 245, Ser 263, Glu 265, Leu 266, Leu 267, Leu 268, Thr 269 and they were very essential for the generation of positive correlation.

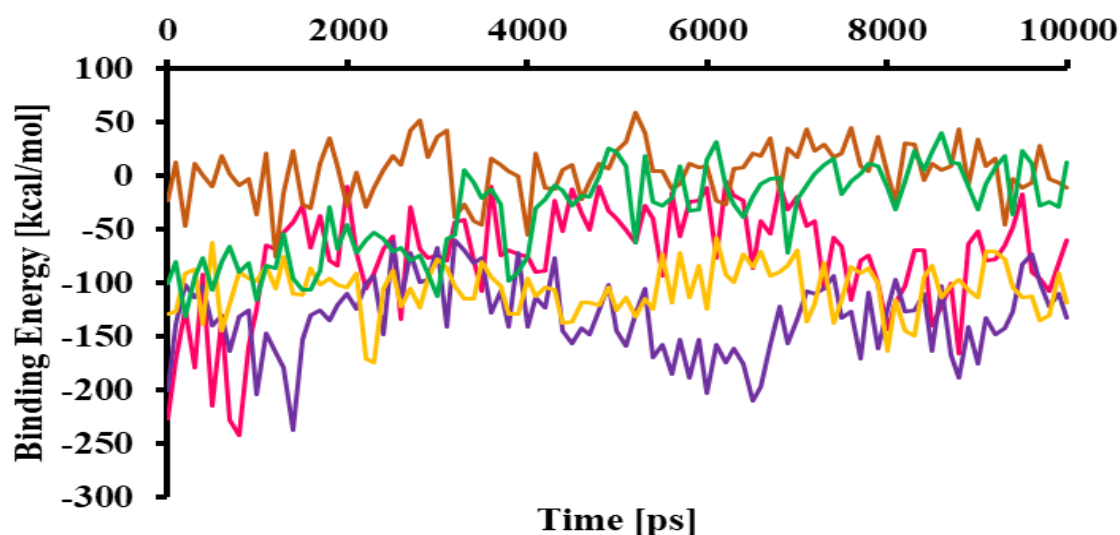


**Figure-4** Dynamic cross correlation map analysis of (A) 3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2-yl]benzo[7]annulen-6-one; (B) Silymarin; (C) Withanolide D; (D) Spirosolane and (E) Oridonin.

### Binding free energy evaluation

MM/PBSA calculation of each complexes were performed using MD simulation in YASARA software which pertained 10 ns time interval with 100 snapshots. **Figure-5** and **Table-5** depicts the results of MM-/PBSA calculation of all complexes. CoV-2 HE protein and Spirosolane complex obtained highest average positive binding energy of 4.386 kcal/mol, while CoV-2 HE protein and Silymarin showed the less binding energy of -132.237 kcal/mol. According to the

MM/ PBSA binding energy more positive binding energy indicates strong binding. However, the energy ranges from -250 kcal/mol to 70 kcal/mol with the indication of fluctuation between 0 to 80 ps and stabilized after it. This decomposition of different binding free energies deciphered partial contribution towards the inhibitory activity of top-ranked compounds.



**Figure-5** Binding free energy calculations of coronavirus hemagglutinin-esterase and top ranked compounds. Color denotation: Magenta-32,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2-yl]benzo[7]annulen-6-one; Purple-Silymarin; Green-Withanolide D; Brown-Spirosolane; Yellow-Oridonin.

**Table-4** The Binding free energy of binding for a series of prioritized hits using MM-/PBSA approach.

No.	Prioritized molecules	Binding free energy (MM/PBSA) (kcal/mol)		
		Initial	Final	Average
1	3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2-yl]benzo[7]annulen-6-one	-226.896	-60.142	-74.136
2	Silymarin	-200.094	-131.884	-132.237
3	Withanolide D	-100.228	12.539	-33.177
4	Spirosolane	-22.802	-10.833	4.386
5	Oridonin	-129.055	-118.204	-105.539

## CONCLUSION

The absence of effective therapeutic drug or vaccine for COVID-19 virus (SARS-CoV-2) which became pandemic outbreak, the spike protein of SARS-CoV-2 was targeted. The



prominent interaction between Human ACE2 and CoV2 HE proteins, which formed the primary event of viral invasion and infection, is the key focus of this research. The present computational study proposed the identification of robust natural candidates for the inhibition of coronavirus hemagglutinin-esterase complex. The combined computational molecular interaction approach was employed on natural repositories and 3,4,5-Trihydroxy-1,8-bis [(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2yl]benzo[7]annulen-6-one, Silymarin, Withanolide D, Spirosolane and Oridonin were revealed promising binding affinity. The molecular dynamics simulations study of the docked complexes offered binding stability while correlative motions of all simulated protein-ligand complexes were obtained from Dynamic cross correlation matrix. Furthermore, the MM/PBSA calculations were applied to calculate the binding free energy which revealed the major contribution of hydrogen bonds with selected target. These computational results provide an insight into the coronavirus hemagglutinin-esterase complex inhibition. Although, they were at preliminary level which requires the experimental confirmation through *in vitro* and *in vivo* studies towards the effective computer-guided drug discovery.

## CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest.

## ACKNOWLEDGMENTS

This work is supported by the Financial Assistance Programme – Department of Science and Technology [Grant number GSBTM/MD/JDR/1409/2017-18] and Gujarat Council on Science and Technology [GUJCOST/Supercomputer/2019-20/1359]. The authors gratefully acknowledge the Department of Botany, Bioinformatics and Climate Change Impacts Management, Gujarat University for providing an opportunity to access the bioinformatics research facilities.

## REFERENCES

- [1] Pene, F., Merlat, A., Vabret, A., Rozenberg, F., Buzyn, A., Dreyfus, F., et al. Coronavirus 229E-related pneumonia in immunocompromised patients. *Clinical infectious diseases*. 2003, 37, 929-32.
- [2] Zeng, Q., Langereis, M.A., van Vliet, A.L., Huizinga, E.G., de Groot, R.J. Structure of coronavirus hemagglutinin-esterase offers insight into corona and influenza virus evolution. *Proceedings of the National Academy of Sciences*. 2008, 105, 9065-9.

- [3] Coleman, C.M., Frieman, M.B. Coronaviruses: important emerging human pathogens. *Journal of virology*. 2014, 88, 5209-12.
- [4] Zaki, A.M., Van Boheemen, S., Bestebroer, T.M., Osterhaus, A.D., Fouchier, R.A. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *New England Journal of Medicine*. 2012, 367, 1814-20.
- [5] Porcheddu, R., Serra, C., Kelvin, D., Kelvin, N., Rubino, S. Similarity in case fatality rates (CFR) of COVID-19/SARS-COV-2 in Italy and China. *The Journal of Infection in Developing Countries*. 2020, 14, 125-8.
- [6] Huang, C., Ito, N., Tseng, C.-T.K., Makino, S. Severe acute respiratory syndrome coronavirus 7a accessory protein is a viral structural protein. *Journal of virology*. 2006, 80, 7287-94.
- [7] Huang, C., Peters, C., Makino, S. Severe acute respiratory syndrome coronavirus accessory protein 6 is a virion-associated protein and is released from 6 protein-expressing cells. *Journal of virology*. 2007, 81, 5423-6.
- [8] Schaecher, S.R., Mackenzie, J.M., Pekosz, A. The ORF7b protein of severe acute respiratory syndrome coronavirus (SARS-CoV) is expressed in virus-infected cells and incorporated into SARS-CoV particles. *Journal of virology*. 2007, 81, 718-31.
- [9] Shen, S., Lin, P.-S., Chao, Y.-C., Zhang, A., Yang, X., Lim, S.G., et al. The severe acute respiratory syndrome coronavirus 3a is a novel structural protein. *Biochemical and biophysical research communications*. 2005, 330, 286-92.
- [10] Nieto-Torres, J.L., DeDiego, M.L., Álvarez, E., Jiménez-Guardeño, J.M., Regla-Nava, J.A., Llorente, M., et al. Subcellular location and topology of severe acute respiratory syndrome coronavirus envelope protein. *Virology*. 2011, 415, 69-82.
- [11] Peiris, J., Lai, S., Poon, L., Guan, Y., Yam, L., Lim, W., et al. Coronavirus as a possible cause of severe acute respiratory syndrome. *The Lancet*. 2003, 361, 1319-25.
- [12] Chu, D., Peiris, J., Chen, H., Guan, Y., Poon, L.L. Genomic characterizations of bat coronaviruses (1A, 1B and HKU8) and evidence for co-infections in *Miniopterus* bats. *Journal of General Virology*. 2008, 89, 1282-7.
- [13] Xiong, X., Tortorici, M.A., Snijder, J., Yoshioka, C., Walls, A.C., Li, W., et al. Glycan shield and fusion activation of a deltacoronavirus spike glycoprotein fine-tuned for enteric infections. *Journal of virology*. 2018, 92, e01628-17.
- [14] Sainz, B., Rausch, J.M., Gallaher, W.R., Garry, R.F., Wimley, W.C. Identification and characterization of the putative fusion peptide of the severe acute respiratory syndrome-associated coronavirus spike protein. *Journal of virology*. 2005, 79, 7195-206.

- [15] Donoghue, M. Hsieh F Baronas E, Godbout K, Gosselin M, Stagliano N, et al. A novel angiotensin converting enzyme related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. *Circ Res.* 2000, 87, E1-E9.
- [16] Zhang, H., Kang, Z., Gong, H., Xu, D., Wang, J., Li, Z., et al. The digestive system is a potential route of 2019-nCov infection: a bioinformatics analysis based on single-cell transcriptomes. *BioRxiv.* 2020.
- [17] Wan, Y., Shang, J., Graham, R., Baric, R.S., Li, F. INFORMATION: READ & LOOK AT ALL BELOW!
- [18] Hofmann, H., Hattermann, K., Marzi, A., Gramberg, T., Geier, M., Krumbiegel, M., et al. S protein of severe acute respiratory syndrome-associated coronavirus mediates entry into hepatoma cell lines and is targeted by neutralizing antibodies in infected patients. *Journal of virology.* 2004, 78, 6134-42.
- [19] Hofmann, H., Pyrc, K., van der Hoek, L., Geier, M., Berkhout, B., Pöhlmann, S. Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. *Proceedings of the National Academy of Sciences.* 2005, 102, 7988-93.
- [20] Hofmann, H., Simmons, G., Rennekamp, A.J., Chaipan, C., Gramberg, T., Heck, E., et al. Highly conserved regions within the spike proteins of human coronaviruses 229E and NL63 determine recognition of their respective cellular receptors. *Journal of virology.* 2006, 80, 8639-52.
- [21] Song, W., Gui, M., Wang, X., Xiang, Y. Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. *PLoS pathogens.* 2018, 14, e1007236.
- [22] Millet, J.K., Kien, F., Cheung, C.-Y., Siu, Y.-L., Chan, W.-L., Li, H., et al. Ezrin interacts with the SARS coronavirus Spike protein and restrains infection at the entry stage. *PLoS One.* 2012, 7.
- [23] Kirchdoerfer, R.N., Wang, N., Pallesen, J., Wrapp, D., Turner, H.L., Cottrell, C.A., et al. Stabilized coronavirus spikes are resistant to conformational changes induced by receptor recognition or proteolysis. *Scientific reports.* 2018, 8, 1-11.
- [24] Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *nature.* 2020, 579, 270-3.
- [25] Walls, A.C., Park, Y.-J., Tortorici, M.A., Wall, A., McGuire, A.T., Veessler, D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell.* 2020.

- [26] Letko, M., Marzi, A., Munster, V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol* 2020; 5: 562-9. PUBMED.
- [27] Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*. 2020.
- [28] Holmes, K.V. SARS-associated coronavirus. *New England Journal of Medicine*. 2003, 348, 1948-51.
- [29] Organization, W.H. Coronavirus disease (COVID-2019) situation reports (<https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/>). Copyright© 2020 Massachusetts Medical Society. 2020.
- [30] Patel, C.N., Kumar, S.K.P., Pandya, H.A., Modi, K.M., Patel, D.P., Gonzalez, F.J. Retrieval of promiscuous natural compounds using multiple targets docking strategy: A case study on kinase polypharmacology. In: 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE, 2017, pp. 288-91.
- [31] Kumar, S.P., Patel, C.N., Jha, P.C., Pandya, H.A. Molecular dynamics-assisted pharmacophore modeling of caspase-3-isatin sulfonamide complex: Recognizing essential intermolecular contacts and features of sulfonamide inhibitor class for caspase-3 binding. *Computational biology and chemistry*. 2017, 71, 117-28.
- [32] Patel, C.N., George, J.J., Modi, K.M., Narechania, M.B., Patel, D.P., Gonzalez, F.J., et al. Pharmacophore-based virtual screening of catechol-o-methyltransferase (COMT) inhibitors to combat Alzheimer's disease. *Journal of Biomolecular Structure and Dynamics*. 2018, 36, 3938-57.
- [33] Mangal, M., Sagar, P., Singh, H., Raghava, G.P., Agarwal, S.M. NPACT: naturally occurring plant-based anti-cancer compound-activity-target database. *Nucleic acids research*. 2013, 41, D1124-D9.
- [34] Wagner, R., Matrosovich, M., Klenk, H.D. Functional balance between haemagglutinin and neuraminidase in influenza virus infections. *Reviews in medical virology*. 2002, 12, 159-66.
- [35] Yuan, P., Thompson, T.B., Wurzburg, B.A., Paterson, R.G., Lamb, R.A., Jardetzky, T.S. Structural studies of the parainfluenza virus 5 hemagglutinin-neuraminidase tetramer in complex with its receptor, sialyllactose. *Structure*. 2005, 13, 803-15.
- [36] De Groot, R.J. Structure, function and evolution of the hemagglutinin-esterase proteins of corona-and toroviruses. *Glycoconjugate journal*. 2006, 23, 59-72.

- [37] Tortorici, M.A., Walls, A.C., Lang, Y., Wang, C., Li, Z., Koerhuis, D., et al. Structural basis for human coronavirus attachment to sialic acid receptors. *Nature structural & molecular biology*. 2019, 26, 481-9.
- [38] Brian, D.A., Hogue, B.G., Kienzle, T.E. The coronavirus hemagglutinin esterase glycoprotein. In: *The Coronaviridae*, Springer, 1995, pp. 165-79.
- [39] Chen, Y., de Bruyn Kops, C., Kirchmair, J. Data resources for the computer-guided discovery of bioactive natural products. *Journal of chemical information and modeling*. 2017, 57, 2099-111.
- [40] Krieger, E., Darden, T., Nabuurs, S.B., Finkelstein, A., Vriend, G. Making optimal use of empirical energy functions: force-field parameterization in crystal space. *Proteins: Structure, Function, and Bioinformatics*. 2004, 57, 678-83.
- [41] Patel, C.N., Narechania, M.B. Targeting epidermal growth factor receptors inhibition in non-small-cell lung cancer: a computational approach. *Molecular Simulation*. 2018, 44, 1478-88.
- [42] Patel, C.N., Kumar, S.P., Modi, K.M., Soni, M.N., Modi, N.R., Pandya, H.A. Cardiotonic steroids as potential Na<sup>+</sup>/K<sup>+</sup>-ATPase inhibitors—a computational study. *Journal of Receptors and Signal Transduction*. 2019, 39, 226-34.
- [43] Lipinski, C.A. Lead-and drug-like compounds: the rule-of-five revolution. *Drug Discovery Today: Technologies*. 2004, 1, 337-41.
- [44] Daina, A., Michielin, O., Zoete, V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific reports*. 2017, 7, 42717.
- [45] Karim, A., Mishra, A., Newton, M.H., Sattar, A. Efficient toxicity prediction via simple features using shallow neural networks and decision trees. *ACS Omega*. 2019, 4, 1874-88.
- [46] Krieger, E., Vriend, G. New ways to boost molecular dynamics simulations. *Journal of computational chemistry*. 2015, 36, 996-1007.
- [47] Ahmad, S., Raza, S., Uddin, R., Azam, S.S. Binding mode analysis, dynamic simulation and binding free energy calculations of the MurF ligase from *Acinetobacter baumannii*. *Journal of Molecular Graphics and Modelling*. 2017, 77, 72-85.
- [48] Tietze, D., Kaufmann, D., Tietze, A.A., Voll, A., Reher, R., König, G., et al. Structural and dynamical basis of G protein inhibition by YM-254890 and FR900359: an inhibitor in action. *Journal of Chemical Information and Modeling*. 2019, 59, 4361-73.
- [49] Bhoi, M.N., Borad, M.A., Jethava, D.J., Acharya, P.T., Pithawala, E.A., Patel, C.N., et al. Synthesis, biological evaluation and computational study of novel isoniazid containing

- 4H-Pyrimido [2, 1-b] benzothiazoles derivatives. *European journal of medicinal chemistry*. 2019, 177, 12-31.
- [50] Gmelch, T.J., Sperl, J.M., Sieber, V. Molecular dynamics analysis of a rationally designed aldehyde dehydrogenase gives insights into improved activity for the non-native cofactor NAD<sup>+</sup>. *ACS Synthetic Biology*. 2020.
- [51] Borah, P., Sanjeev, A., Mattaparthi, V.S.K. Computational investigation on the effect of Oleuropein aglycone on the  $\alpha$ -Synuclein aggregation. *Journal of Biomolecular Structure and Dynamics*. 2020, 1-12.
- [52] Wang, J., Hou, T., Xu, X. Recent advances in free energy calculations with a combination of molecular mechanics and continuum models. *Current Computer-Aided Drug Design*. 2006, 2, 287-306.
- [53] Wang, W., Donini, O., Reyes, C.M., Kollman, P.A. Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions. *Annual review of biophysics and biomolecular structure*. 2001, 30, 211-43.
- [54] Kollman, P.A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of chemical research*. 2000, 33, 889-97.
- [55] Massova, I., Kollman, P.A. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspectives in drug discovery and design*. 2000, 18, 113-35.
- [56] Razzaghi-Asl, N., Mirzayi, S., Mahnam, K., Sepehri, S. Identification of COX-2 inhibitors via structure-based virtual screening and molecular dynamics simulation. *Journal of Molecular Graphics and Modelling*. 2018, 83, 138-52.

## Figure Legends

**Figure-1** Docked poses of 4,9-O-diacetyl sialic acid, 3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2-yl]benzo[7]annulen-6-one, Silymarin, Withanolide D, Spirosolane and Oridonin.

**Figure-2** Molecular Dynamics Simulation analysis: Time vs. Total energy and Time vs. RMSD graph. Color denotation: Magenta-3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2-yl]benzo[7]annulen-6-one; Purple-Silymarin; Green-Withanolide D; Brown-Spirosolane; Yellow-Oridonin.

**Figure-3** Stability analysis: Pocket residue vs. Pocket RMSD and Pocket residues vs. Pocket RMSF.

**Figure-4** Dynamic cross correlation map analysis of (A) 3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2-yl]benzo[7]annulen-6-one; (B) Silymarin; (C) Withanolide D; (D) Spirosolane and (E) Oridonin.

**Figure-S1** Graphical representation of the interactions between Coronavirus Hemagglutinin-Esterase with 3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2-yl]benzo[7]annulen-6-one simulated molecule by Discovery studio visualiser ((A) 0 nanoseconds, (B) 2.5 nanoseconds, (C) 5 nanoseconds, (D) 7.5 nanoseconds, (E) 10 nanoseconds).

**Figure-S2** Graphical representation of the interactions between Coronavirus Hemagglutinin-Esterase with Silymarin simulated molecule by Discovery studio visualiser ((A) 0 nanoseconds, (B) 2.5 nanoseconds, (C) 5 nanoseconds, (D) 7.5 nanoseconds, (E) 10 nanoseconds).

**Figure-S3** Graphical representation of the interactions between Coronavirus Hemagglutinin-Esterase with Withanolide D simulated molecule by Discovery studio visualiser ((A) 0 nanoseconds, (B) 2.5 nanoseconds, (C) 5 nanoseconds, (D) 7.5 nanoseconds, (E) 10 nanoseconds).

**Figure-S4** Graphical representation of the interactions between Coronavirus Hemagglutinin-Esterase with Spirosolane simulated molecule by Discovery studio visualiser ((A) 0 nanoseconds, (B) 2.5 nanoseconds, (C) 5 nanoseconds, (D) 7.5 nanoseconds, (E) 10 nanoseconds).

**Figure-S5** Graphical representation of the interactions between Coronavirus Hemagglutinin-Esterase with Oridonin simulated molecule by Discovery studio visualiser ((A) 0 nanoseconds, (B) 2.5 nanoseconds, (C) 5 nanoseconds, (D) 7.5 nanoseconds, (E) 10 nanoseconds).

**Figure-5** Binding free energy calculations of coronavirus hemagglutinin-esterase and top ranked compounds. Color denotation: Magenta-3,4,5-Trihydroxy-1,8-bis[(2R,3R)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-2-yl]benzo[7]annulen-6-one; Purple-Silymarin; Green-Withanolide D; Brown-Spirosolane; Yellow-Oridonin.

**Table Legends**

**Table-1** Binding energy, Hydrogen bonds, Efficiency and contacting receptor residues of selected ligands.

**Table-2** Molecular pharmacokinetic properties calculated for the top-ranked compounds with SwissADME webserver.

**Table-3** Predicted ADMET properties of top-ranked compounds.

**Table-4** The Binding free energy of binding for a series of prioritized hits using MM-/PBSA approach.