

Drug and drug-like molecule binding to interface of SARS-CoV-2 S-protein:human ACE2 complex: A density functional theory study

Cherumuttathu H. Suresh

The active site of SARS-CoV-2 S-protein:human ACE2 complex (**M**) for ligand (**L**) binding is modelled using a two layer ONIOM(B3LYP/6-31G*:PM7) method in conjunction with full density functional theory (DFT) calculations. The **ML** complexes are located for thirty-five **L** molecules which include sixteen traditional herbal isolates (THI) and several FDA approved drugs. Previously, from a docking study of around 9000 molecules at the S-protein:human ACE2 interface, Smith and Smith observed the best docking score (-7.0 to -7.7 kcal/mol) for these ligands. The **Ls** showed close similarity in many structural and functional motifs such as six-membered double ring fused (drf) moiety, drf to phenyl ring connection by single bond, six-membered O- and N-heterocycles, and presence of hydroxyl, carbonyl, carboxylate, ester, amino and imino units. The interface cavity is highly polar with dipole moment 44.4 D and $\sim 1 \text{ nm}^3$ in size ($\mu = 44.4 \text{ D}$). Hence, most of the ligands show an increasing trend in binding energy (E_b) with increases in μ . The traditional herbal isolate (THI) fustin (**5**) is the weakest binding with E_b -17.5 kcal/mol, whereas more than five times better E_b is observed for the strongest binding m-digallic acid (THI, **13**) and adenosine 3',5'-bisphosphate (**35**). Similarly, the THIs myricetin (**4**) and glucogallin (**14**) and approved drugs such as sapropterin (**18**), tetrahydrobiopterin (**19**), protirelin (**30**) and fidarestat (**32**) show E_b values better than -60.0 kcal/mol. The effective binding of these ligands occurs due to multiple noncovalent interactions (total around 6 - 8) between the functional groups on **L** and moieties of S-protein and ACE2 receptors such as arginine, histidine, tyrosine, lysine, carboxylate and amide. The THI myricetin (**4**) and the sugar substitute aspartame (**28**) show E_b -67.7 and -50.7 kcal/mol, respectively while the new ligand **36**, designed by utilizing the most binding structural motifs of **4** and **28**, shows the best E_b -91.7 kcal/mol. Such a hybrid ligand design strategy is promising for the development of new drugs from THI leads and existing drugs. The ONIOM-linked DFT study is effective, affordable and reliable for a quantum chemical treatment of the drug-receptor binding process which sheds lights upon the preserved features of the receptor cavity that could lead to a rational design approach to COVID-19 drug development.

Introduction

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2 or COVID-19) is a positive strand RNA virus which has been causing one of the deadliest pandemics of the world with infections above 33 million globally and deaths around 1.0 million to date within a span of eight months from late December 2019.¹⁻³ The RNA genome of COVID-19 is one among the largest and it encodes four major structural proteins, viz. the spike (S) protein, nucleocapsid (N) protein, membrane (M) protein, and the envelope (E) protein.⁴ The S-protein is responsible for facilitating entry of the virus into the target cell due to strong S-protein–host cell protein interaction⁵ and it has been an important target for coronavirus vaccine and antiviral development. Very recently Xu et al.⁶ used the computer-guided homology modelling method to the crystal structure SARS coronavirus S-protein (PDB accession: 6ACD) to develop the 3-D structural model for SARS-CoV-2. Previous studies have shown that the spike protein (S-protein) of SARS-CoV-2 shows

strong binding affinity to the angiotension-converting enzyme 2 (ACE2) which acts as its receptor.^{5,7,8} Xu et al.⁶ used the structural superimposition and molecular rigid docking approaches⁹ to develop the 3-D complex structure of the S-protein binding to human ACE2 (Fig. 1a) and reported the binding free energy -50.6 kcal/mol for the complex. Very recently, Smith and Smith¹⁰ followed the work of Xu et al. to generate the coordinates for the S-protein-ACE2 complex and used the model to enact an ensemble docking virtual high throughput screening study to identify small-molecules which bind to either the isolated S-protein at its host receptor region or to the S-protein-human ACE2 interface region. On the basis of docking scores, they reported the binding affinities of over 9000 drugs, metabolites, and natural products (and their isomers) with respect to the S-protein and the S-protein:ACE2 receptor complex. The massive amount of docked structural data obtained using Autodock Vina¹¹ is available from the work of Smith and Smith¹⁰ which is very useful for further research in this area.

Numerous studies have appeared in the recent literature on the modelling of drug-receptor interactions related to SARS-CoV-2

^aChemical Sciences and Technology Division, CSIR–National Institute for Interdisciplinary Science and Technology (CSIR–NIIST), Thiruvananthapuram 695019, India..

Electronic Supplementary Information (ESI) available: [Suppinfo1_optimized_structures.xlsx (Coordinates of optimized structures of the reduced model of the S-protein:ACE2 interface complex (M), all ligands (Ls) and all ML complexes) and suppinfo2_energies.xlsx (Energy parameters of M, L and ML complexes)]. See DOI: 10.1039/x0xx00000x

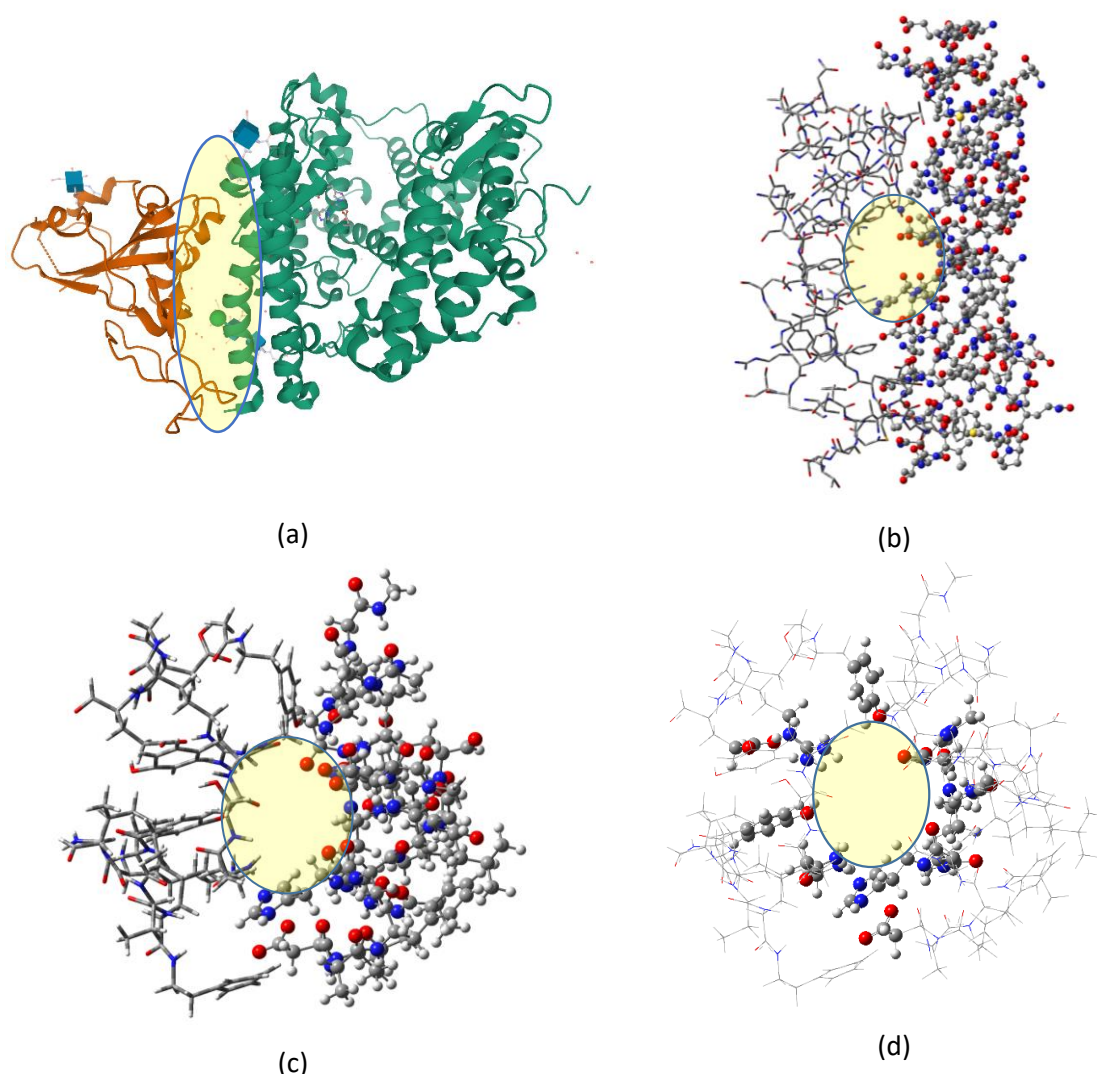


Fig. 1 (a) S-protein of SARS-CoV-2 (orange) with human ACE2 receptor (green) complex (b) Interface model of S-protein:ACE2 complex used for docking studies by Smith and Smith. The stick model is for S-protein and the rest is for ACE2 (c) The reduced interface model of the docked region of S-protein:ACE2 complex (M) used for DFT studies. (d) Two layer ONIOM(B3LYP/6-31G*:PM7) model used for optimizing M. The wireframe model is in semiempirical layer PM7 and the rest is in DFT layer. Docked molecules are located at the circled region.

using computational approaches such as molecular docking, molecular dynamics simulation, and other computer-aided drug-design (CADD) techniques,^{3,12-27} Very recently, Hagar et al. have reported the molecular docking and DFT calculations on some antiviral N-heterocycles.²⁸ The present study focuses on understanding the electronic features of the binding region of the S-protein:ACE2 interface model for small molecules using a density functional theory (DFT) method in conjunction with a hybrid method that combines DFT and a semiempirical approach. Further, the molecules listed by Smith and Smith¹⁰ as the most binding at the interface are selected to obtain structural and energetic features at the DFT level. On the basis of the non-covalent binding features of the small molecule at the interface, new design strategies are applied to improve the binding energy of novel derivatives.

Computational Methodology

The interface model (Fig. 1b) reported by Smith and Smith¹⁰ for the SARS-CoV-2 S-protein:human ACE2 complex (Fig. 1a) is used to get a reduced model (M) for DFT studies (Fig. 1c). Since M contains a total of 611 atoms (302 without hydrogen), a full-fledged optimization of the structure using a reliable DFT method is cumbersome. Among the several less expensive and accurate methods devised for the determination of the structure and properties of large molecular systems, the multi-layer hybrid method known as ONIOM method, developed by Morokuma and co-workers²⁹ is widely applied for modelling molecular systems such as homogeneous catalysts, nanomaterials and biological macromolecules. Here we used a two layer ONIOM method as implemented in Gaussian 16 suite of programs³⁰ to optimize the reduced model (M). In the

ONIOM scheme, the high layer is treated with the B3LYP/6-31G* level DFT³¹ while the low layer is treated with the semiempirical PM7 method³² (Fig. 1d). All the atoms in the low layer are frozen to avoid the collapse of the structure while atoms in the high layer is fully optimized under the ONIOM(B3LYP/6-31G*:PM7) scheme. Further, the structure obtained using this constrained optimization scheme (full model) is subjected to single point calculation at B3LYP/6-31G* to obtain the total energy of **M** (E_1). The ONIOM(B3LYP/6-31G*:PM7) optimization is also performed on the docked structure of the small molecule complex with the reduced interface model of the SARS-CoV-2 S-protein:human ACE2 complex (ML). The small molecule complex of the reduced model is prepared from the reported structures by Smith and Smith¹⁰. We have selected the best thirty five Ls from their database with docking score in the range -7.0 to -7.7 kcal/mol for this study. For the optimized small molecule complex ML, B3LYP/6-31G* single point calculation is performed to obtain the total energy E_2 . All Ls were optimized at B3LYP/6-31G* level to obtain their energy E_0 . The binding energy (E_b) is calculated as $E_b = E_2 - (E_0 + E_1)$.

Further, the molecular electrostatic potential (MESP) calculation³³⁻³⁵ is done at B3LYP/6-31G* level to understand the electron distribution features of **M**. The MESP analysis is useful for the interpretation of the noncovalent bonding features of molecular complexes.

Results and discussion

Reduced interface model (**M**) for SARS-CoV-2 S-protein:human ACE2 complex.

Two views of the ONIOM(B3LYP/6-31G*:PM7) level optimized geometry of **M** is given in Fig. 2. The size of the cavity wherein the small molecule can bind is approximately 10.9 x 10.6 x 11.1 Å in dimension. Though the total charge of the high layer in ONIOM (sphere model in Fig. 2) is zero, it contains three carboxylate ($-\text{COO}^-$) moieties, two positively charged arginine moieties ($-\text{CN}_3\text{H}_5^+$) and one ammonia ($-\text{NH}_3^+$) moiety. The anionic carboxylate moieties lie close to the cationic moieties. The high layer of the cavity also consists of two phenolic moieties from the tyrosine units (both belong to the S-protein), a portion from a histidine unit and a portion from a proline unit (both from ACE2).

With the ONIOM optimized geometry, a single point DFT calculation at B3LYP/6-31G* level is done for **M** to obtain its electron density distribution. The MESP on planes passing through the middle region of the cavity is shown in Fig. 3 (for the view given in Fig. 2a). These pictures show that the bottom region of the cavity (proline portion is at this region) is electron rich (blue area) while the top region (histidine portion is at this region) is electron deficient (red area). This suggests a highly polar binding domain for the interface. The high dipole moment 44.4 D observed for this binding cavity with a top to bottom direction (Fig. 2b) indicates its natural tendency to attract polar molecules for strong binding.

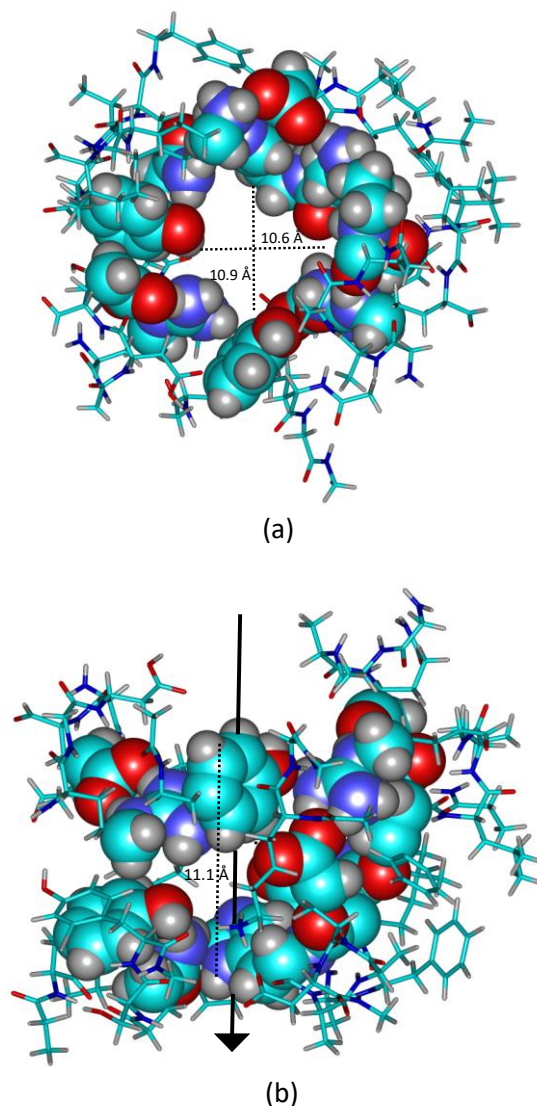


Fig. 2 Optimized geometry at ONIOM(B3LYP/6-31G*:PM7) level for the interface region of SARS-CoV-2 S-protein:human ACE2 complex (**M**). (a) Top view through the cavity and (b) side view. Approximate dimension of the cavity is given in Å. The dipole moment (μ) vector (44.4 D) is given in Fig. b.

Selected small molecules for binding

The thirty five top scoring ligands reported by Smith and Smith¹⁰ fall in the category of approved FDA drugs, NPC approved drugs, traditional herbal isolates (THI), Indian approved chemicals (IAC) and other small molecules which are depicted in Fig. 4. Each molecule is identified with a number as well as another number-label used by Smith and Smith¹⁰ along with the corresponding popular/approved/scientific name. The molecular diagrams of them (**1 – 35**) are created with ChemCraft software by visualizing the corresponding optimized structures at B3LYP/6-31G* level. The initial orientation of the molecules for optimization is obtained from the docked geometry reported by Smith and Smith¹⁰. Further, all of them are taken in the neutral state by attaching appropriate number of hydrogen atoms to the main block atoms.

Molecules, **1** – **11** show very close structural similarity among themselves and all are THI. They all have a structural moiety composed of a phenyl ring fused to a six-membered O-

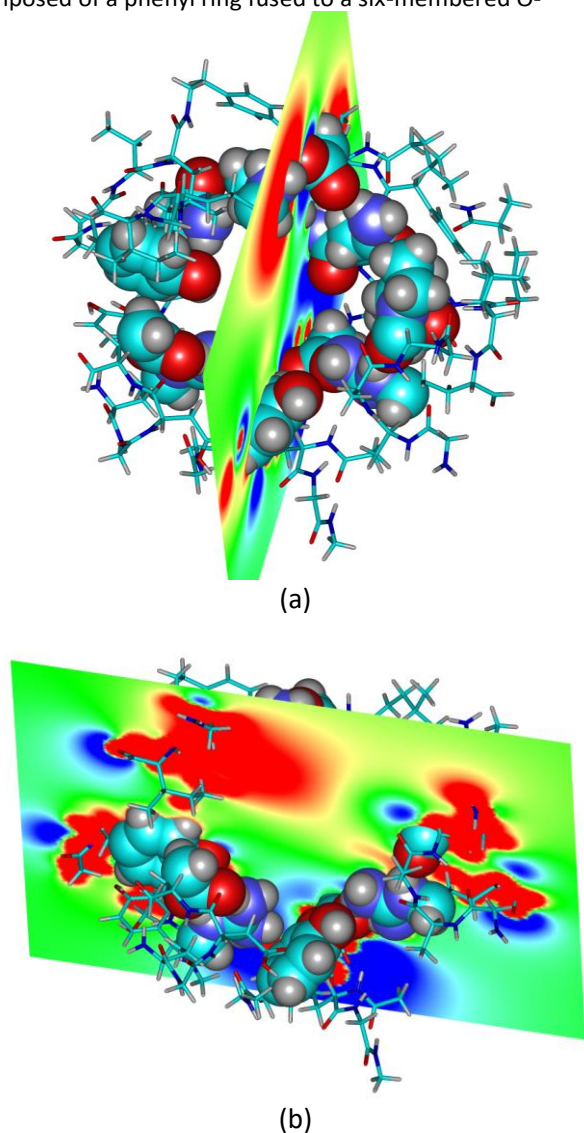


Fig. 3 Molecular electrostatic potential (MESP) map plotted on two planes passing approximately through the center region of the cavity. From blue to red region, MESP is at -30.0 to $+30.0$ kcal/mol.

hetrocyclic ring (saturated/unsaturated). In all cases, this six-membered double ring-fused (drf1) structural moiety is connected to a phenyl ring at the third position from the fusion point. Further, all are rich with hydroxy substitution at multiple points on the rings whereas except **6** and **11**, all contain carbonyl group/s too. The molecule **12** (THI - shikonin) is also similar to **1** – **11** except for the phenyl ring. Similarly, except for the drf1 structure, THI **13** (m-digallic acid) and THI **14** (glocogallin) show good similarity to **1** – **12** due to their hydroxy rich character. In the cases of **15** – **19**, the drf structure (drf2) unlike **1** – **12** is composed of at least one six-membered N-heterocycle. Also a single bond connects the drf2 unit to a small structural unit containing N- or O- or both N- and O-hetero

atoms. Structures **20**, **21** and **22** are also similar to **15** – **19**, the main difference being the drf unit is composed of five and six-membered rings (drf3). In the case of **23**, the FDA approved drug ofatumumab, a triple ring-fused (trf1) structural moiety is present which can also be considered as fusion between a drf3 type structure to a six-membered O-heterocycle. The trf1 in **23** is connected to a five-membered ring through a double bond whereas only single bond connection can be seen between drf unit and other moiety in **1** – **22**, except **20**. Structures **24** – **29** contain one aromatic ring (phenyl, furan or pyridine moiety) which is connected to a small unit containing multiple hetero O- and N-atoms *via* a single bond. The small unit is composed of ester, acid, carbonyl, hydroxy, nitro, amino and imino functional groups. The **30** and **31** contain saturated five-membered N-heterocycles as well as functional groups typically found in other ligands (carbonyl, hydroxy, amino, etc). Compounds **32**, **33** and **34** are characterized with the presence of a spiro carbon center as well as phenyl unit and O- or N-heterocycle. **32** is the only one having a fluorine substitution on the phenyl ring while **34** is unique due to the presence of a hypervalent sulphur center. The **35** is perhaps the odd one out considering the phosphate-sugar-base connectivity which is typically seen in drugs targeting RNA polymerase.

Binding energy

The binding energy calculated at DFT level for all the molecules is given in Table 1. The weakest binding is observed for **5**, the traditional herbal isolate (THI) fustin with E_b -17.5 kcal/mol, whereas the strongest binding is observed for another THI **13**, m-digallic acid with E_b -89.1 kcal/mol. Adenosine 3',5'-bisphosphate (**35**) also showed strong binding, E_b -87.9 kcal/mol. The E_b data for most of them can be clubbed in the range -30.0 to -50.0 kcal/mol (Fig. 5). It is clear that the docking score in the narrow range -7.0 to -7.7 kcal/mol (Table 1) is not sensitive enough to distinguish the binding behaviour of these molecules at the interface. Table 1 also reports the dipole moment value for the optimized structure of all the thirty five ligands. The correlation plot in Fig. 6 suggests that for most of the molecules, the binding interactions become stronger with increase in the dipole moment. This result supports the highly polarized nature of the binding pocket which is clearly brought out from the MESP plot given in Fig. 3. Contrary to this observation, the strongest binding m-digallic acid has the lowest dipole moment 0.65 D. Though it appears counter intuitive, the structure of **13** (Fig. 4) suggests that its two phenyl units are substituted with several polar hydroxy groups and the low dipole moment is due to the nearly perpendicular orientation of the phenyl units. The molecules **14**, **30** and **32** are also examples for this observation as they show high binding energy and relatively low dipole moment. In **14**, the sugar unit is nearly perpendicular to the phenyl unit while in **32**, the spiro carbon forces the ring moieties in a perpendicular orientation. In **30**, the three N-heterocycles occupy a triangular orientation which contributes to the low dipole moment. It appears that more than the total dipole moment, the locally polarized nature of the binding units is important for effective binding.

 1 (2462) luteolin (THI)	 2 (2828) quercetin (THI)	 3 (2526) eriodictyol (THI)	 4 (3223) myricetin (THI)
 5 (2542) fustin (THI)	 6 (2600) leucopelargonidin (THI)	 7 (3224) gossypetin (THI)	 8 (2156) naringenin (THI)
 9 (1885) hydrangenol (THI)	 10 (2463) scutellarein (THI)	 11 (3362) leucodelphinidin (THI)	 12 (2502) shikonin (THI)
 13 (3347) m-digallic acid (THI)	 14 (3594) glucogallin (THI)	 15 (1447) pemirolast (FDA)	 16 (2185) demethyl-coclaurine (THI)
 17 (2873) alpha-dichroine (THI)	 18 (1617) sapropterin (FDA)	 19 (1585) tetrahydrobiopterin (FDA)	 20 (1560) carbazochrome (NPC)
 21 (2102) vidarabine (FDA)	 22 (1531) adrenochrome guanylylhydrazone	 23 (1629) ofatumumab (FDA)	 24 (1933) benserazide (FDA)
 25 (1135) Isoniazid pyruvate (FDA)	 26 (1564) nitrofurantoin (FDA)	 27 (1148) phenformin (IAN)	 28 (2712) aspartame (IAN)
 29 (1476) acitazanolast (NPC)	 30 (4303) protirelin (FDA)	 31 (2954) vildagliptin (IAN)	 32 (2357) fidarestat (NPC)

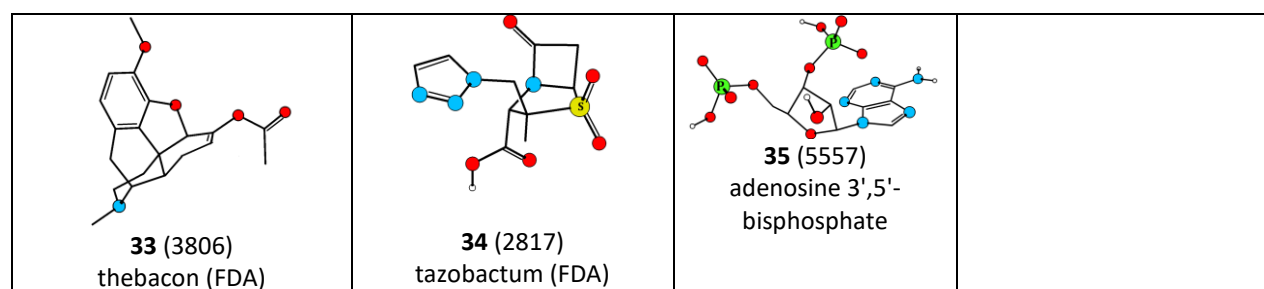


Fig. 4. The selected molecules (1 - 35) for the binding study with the interface of SARS-CoV-2 S-protein:human ACE2 complex. The label used by Smith and Smith (in bracket) as well as the common names of the molecules are also shown with their category.

Table 1. Binding energy of all the selected molecules in kcal/mol.

Molecule	Smith & Smith label	Name	Classification	Docking score	Dipole moment (D)	Binding Energy (kcal/mol)
1	2462	luteolin	THI	-7.4	4.94	29.4
2	2828	quercetin	THI	-7.3	7.00	45.5
3	2526	eriodictyol	THI	-7.1	3.96	22.8
4	3223	myricetin	THI	-7.7	3.57	67.7
5	2542	fustin	THI	-7.6	2.20	17.5
6	2600	leucopelargonidin	THI	-7.5	4.88	37.6
7	3224	gossypetin	THI	-7.5	3.83	37.1
8	2156	naringenin	THI	-7.3	2.58	22.5
9	1885	hydrangenol	THI	-7.2	4.20	24.5
10	2463	scutellarein	THI	-7.0	4.11	28.6
11	3362	leucodelphinidin	THI	-7.7	5.34	35.9
12	2502	shikonin	THI	-7.3	2.52	20.2
13	3347	m-digallic acid	THI	-7.5	0.65	89.1
14	3594	glucogallin	THI	-7.3	1.77	62.1
15	1447	pemirolast	FDA	-7.4	3.94	34.1
16	2185	demethyl-coclaurine	THI	-7.0	2.74	27.2
17	2873	alpha-dichroine	THI	-7.3	1.78	22.4
18	1617	sapropterin	FDA	-7.1	6.78	61.4
19	1585	tetrahydrobiopterin	FDA	-7.3	7.60	63.3
20	1560	carbazochrome	NPC	-7.2	4.01	31.4
21	2102	vidarabine	FDA	-7.1	4.99	41.1
22	1531	adrenochrome guanyldihydrazone	Others	-7.4	4.51	47.8
23	1629	ofatumumab	FDA	-7.6	9.12	42.0
24	1933	benserazide	FDA	-7.4	6.12	36.9
25	1135	isoniazid pyruvate	FDA	-7.3	4.87	30.5
26	1564	nitrofurantoin	FDA	-7.2	3.58	48.8
27	1148	phenformin	IAN	-7.0	3.60	49.3
28	2712	aspartame	FDA	-7.4	5.07	50.7
29	1476	acitazanolast	NPC	-7.2	4.42	40.2
30	4303	protirelin	FDA	-7.3	3.59	62.0
31	2954	vildagliptin	IAN	-7.0	6.48	30.7
32	2357	fidarestat	NPC	-7.3	4.38	67.5
33	3806	thebacon	FDA	-7.4	5.71	21.5
34	2817	tazobactam	FDA	-7.1	3.91	40.2
35	5557	adenosine 3',5'-bisphosphate	Others	-7.3	12.47	87.9

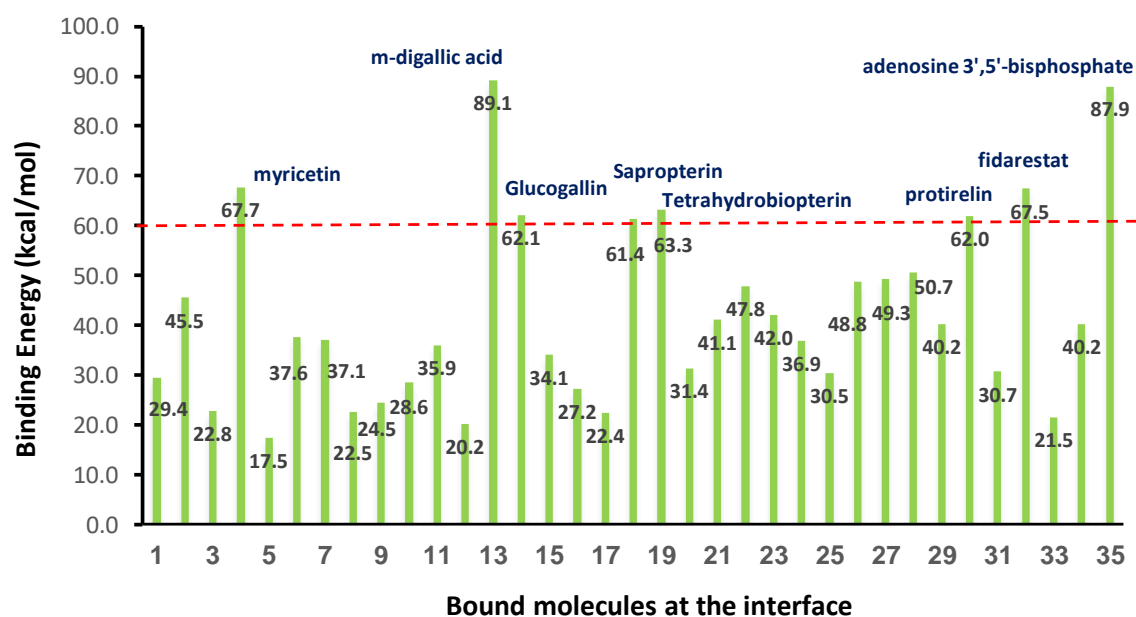


Fig. 5. Binding energy profile for all the molecules at the interface of SARS-CoV-2 S-protein:human ACE2 complex.

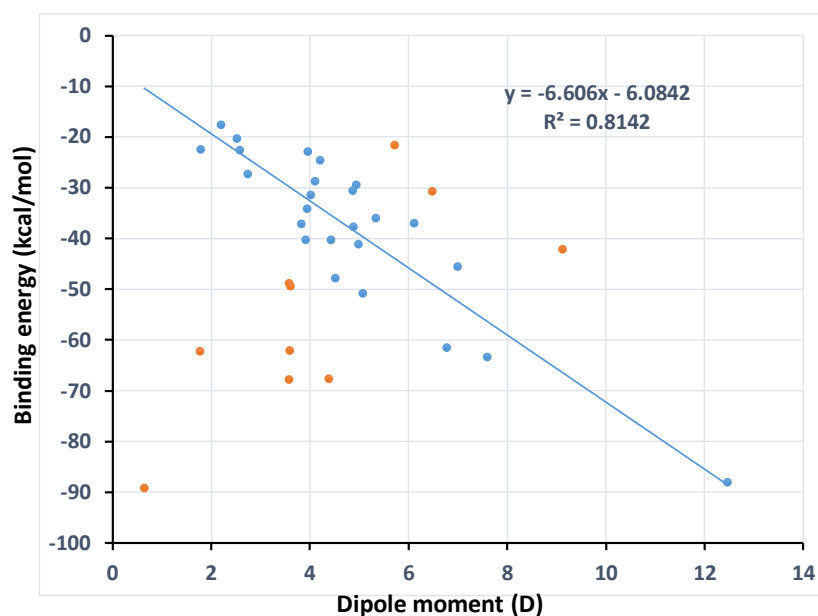


Fig. 6. Correlation between dipole moment of molecules and their binding energy at the interface of SARS-CoV-2 S-protein:human ACE2 complex obtained at B3LYP/6-31G* level DFT.

Noncovalent bonding

Here we select eight molecules, viz. **4** (myricetin), **13** (m-digallic acid), **14** (glucogallin), **18** (sapropterin), **19** (tetrahydrobiopterin), **30** (protirelin), **32** (fidarestat) and adenosine 3',5'-bisphosphate (**35**) which show binding energy values better than -60.0 kcal/mol for an in-depth analysis of the noncovalent binding features of the binding pocket. Fig. 7 depicts the entrapped structure of these molecules within the

interface region of SARS-CoV-2 S-protein:human ACE2 complex (**ML**). In all the cases, both S-protein portion as well as ACE2 portion are interacting with the molecule and they all fit well at the void space (approximately 1 nm³ size).

The ONIOM(B3LYP/6-31G*:PM7) level optimized geometries can provide accurate information about the specific noncovalent interactions between the small molecules (L) and

the binding pocket of M. Fig. 8 shows the detailed information about the interactions for the seven complexes. In all cases,

around L, six to eight interactions can be identified which support their high binding behavior in the cavity. These

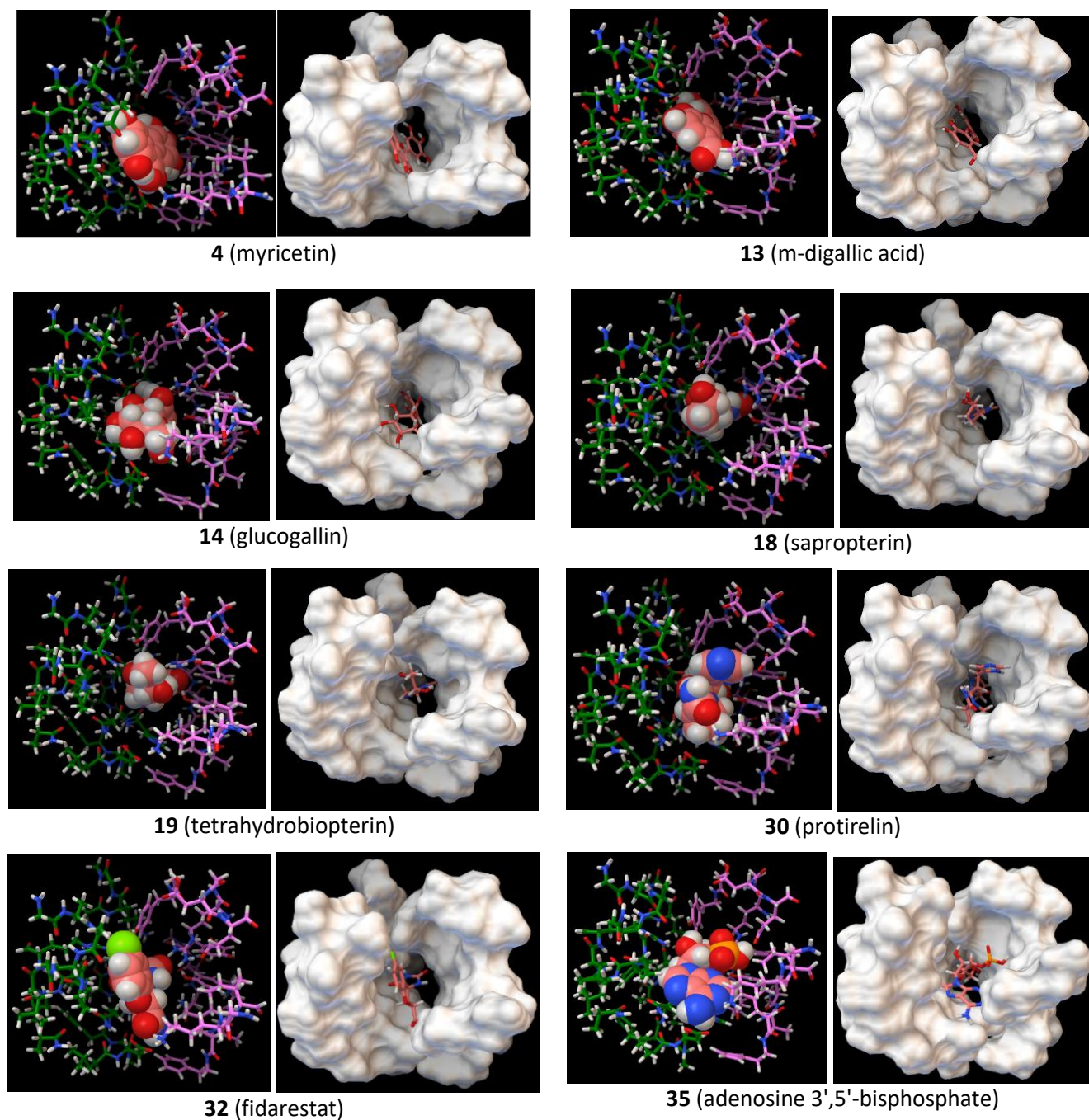


Fig. 7 The entrapped molecules in the cavity. The green and pink stick models represent the S-protein and ACE2 portions, respectively. To show the fitting of the molecule in the cavity, the van der Waals surface model is also depicted for each complex. ONIOM(B3LYP/6-31G*:PM7) results.

interactions mainly arise from arginine, histidine, tyrosine and lysine moieties as well as from carboxylate and amide units of the interface. Noncovalent bonding with the arginine moiety is found in all cases. The **19** (tetrahydrobiopterin) and **32** (fidarestat) show interactions from two arginine moieties. The THI molecules myricetin, m-digallic acid and glucogallin are devoid of interactions with tyrosine moiety whereas FDA drugs except fidarestat shows tyrosine-drug interactions.

Myricetin, aspartame and a hybrid design

As seen in Fig. 5, only one molecules fall in the binding energy range 50 – 60 kcal/mol, the FDA approved sugar substitute aspartame (**28**). It shows seven significant hydrogen bonding interactions, viz. HO...HN and CO...HN from acid moiety, CO...HO, CO...HN and NH...OC from amid group and two CO...NH from ester moiety (Fig. 8). The E_b -50.7 kcal/mol observed for

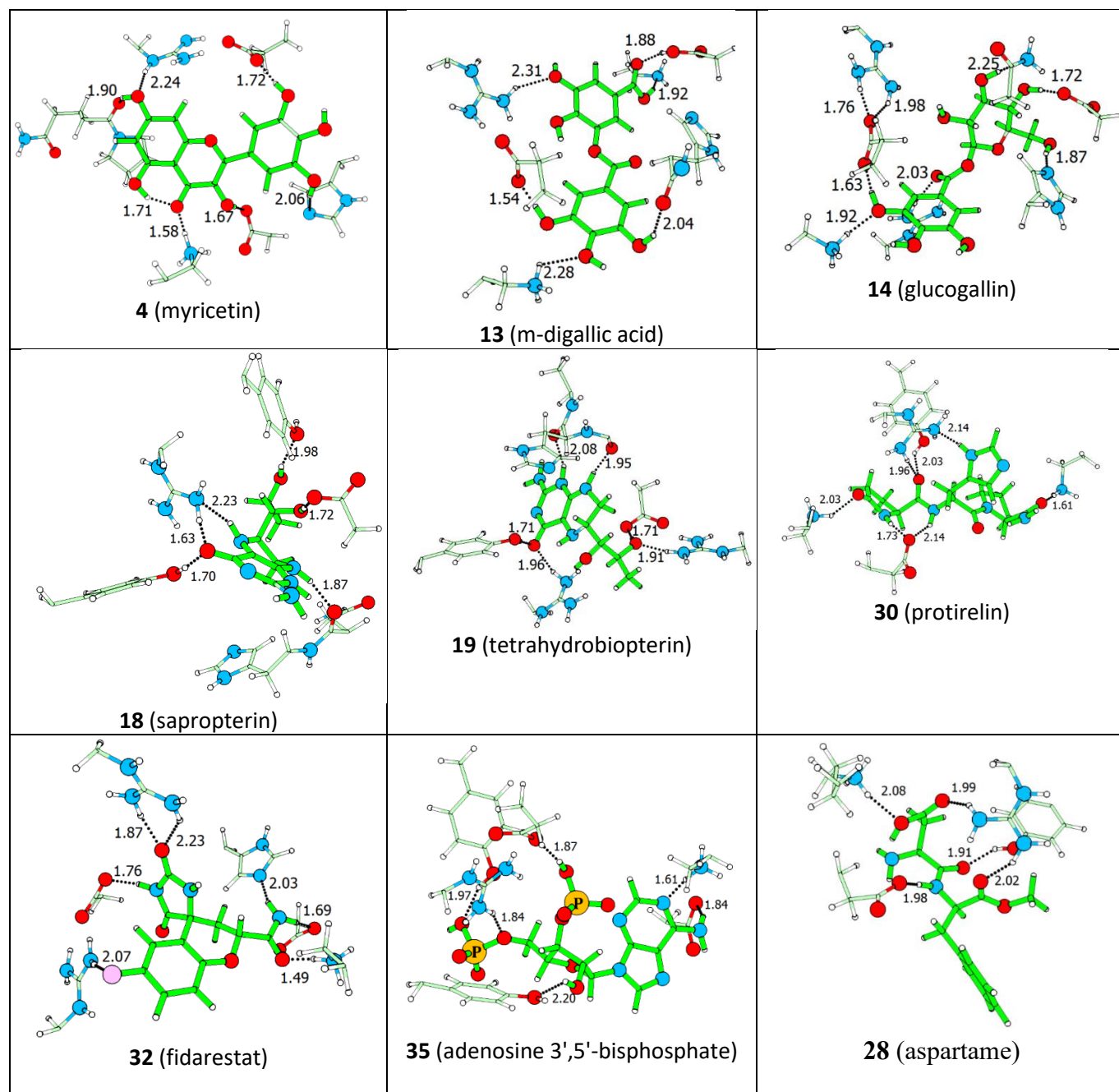


Fig. 8. Significant hydrogen bonding interactions seen in ML complexes. The L is represented with thick green bonds. Only the interacting portions on M is shown. Distances in Å.

aspartame can be mainly attributed to the collective strength of these hydrogen bonds. A conspicuous fact is that its benzyl moiety is devoid of any significant interactions and suggests that by modifying the benzyl moiety, improvement in E_b can be achieved. In order to test this hypothesis, we have designed a hybrid structure **36** composed of portions of the THI myricetin and aspartame (Fig. 9). In this design, the benzyl moiety and ester group of aspartame is replaced with the drf1 unit of myricetin. It may be noted that in ML complex of myricetin, among the seven hydrogen bonds (CO...HN, OH...O and OH...N types), five are with the double ring fused (drf1) moiety (Fig. 8).

The **36** shows nine significant interactions with the interface and their collective strength gives the best E_b value -91.7 kcal/mol. Here the arginine moiety is strongly connected to the molecule through three hydrogen bonds. The hybrid design indeed strengthens the overall interaction between the ligand and receptor. Though some of the THIs show strong interaction in the binding domain of SARS-CoV-2 S-protein:human ACE2 complex (M), their modification through the incorporation of N-heteroatoms may be considered for improving the drug-like characters.

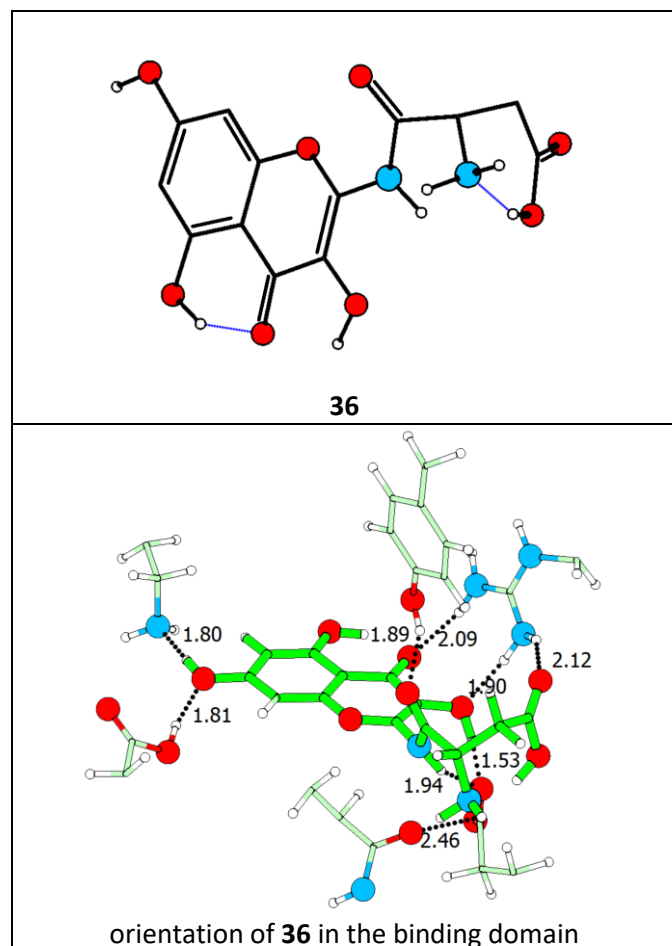


Fig. 9 Molecular drawing for myricetin-aspartame hybrid **36** (top) and its interactions with interface region of SARS-CoV-2 S-protein:human ACE2 complex (bottom). Distances in Å.

Conclusions

The binding of thirty five drug and drug-like molecules (**L**) at the active site of the SARS-CoV-2 S-protein:human ACE2 interface complex (**M**) is successfully modelled using DFT method in conjunction with a two layer ONIOM method. The **Ls** showed close structural similarity in terms of the presence of a double ring fused structure (drf1), flexible single bond between drf1 and a phenyl ring, and presence of several hydroxyl and carbonyl groups. In the case of drug molecules, drf-type structure composed of aromatic ring and N-heterocycles is observed in majority of them which was connected by single bond to other small structural units composed of typical functional groups such as carbonyl, hydroxyl, amino and imino. The **Ls** showed typically 10 – 12 Å length and their binding occurs at a highly polarized interface cavity having dimensions slightly higher than 1 nm. The binding affinity of the ligands typically increases with increase in their dipole moment.

The **Ls** showed multiple noncovalent interactions with both S-protein and ACE2 units. Though the docking score reported by Smith and Smith is nearly same for all **Ls**, the DFT results show that the binding energy of the two most strongly binding **Ls** is five times better than the least binding **L**. Among the most

binding eight molecules at the interface, three are THIs, viz. myricetin (**4**), m-digallic acid (**13**) and glucogallin (**14**) and others are approved drugs such as sapropterin (**18**), tetrahydrobiopterin (**19**), protirelin (**30**), fidarestat (**32**) and adenosine 3',5'-bisphosphate (**35**). All these molecules showed six to eight interactions around them connecting with moieties such as arginine, histidine, tyrosine, lysine, carboxylate and amide from both S-protein and ACE2 residues. The high binding energy is attributed to the cumulative effect of these interactions.

Aspartame, a ligand with moderate binding energy showed rich noncovalent interactions at the interface through its amino acid and amide units while a strongly interacting THI myricetin has significant interactions at the interface through its drf1 structural motif. The newly designed ligand **36** is emerged as the most strongly binding **Ls** at the interface due to a hybrid design strategy that joins the most binding moiety of a THI with the most binding portions of aspartame. The incorporation of N-centers in oxygen-rich THIs can improve their drug-like features.

In summary, the two layer ONIOM optimization strategy is an effective and affordable way to model the active site of drug-receptor complexes. The inner layer of ONIOM optimized structures provide the specific binding features of the molecules while the all electron DFT calculation provides a more reliable quantum chemical treatment of the drug-receptor binding process. The conserved features of the binding domain can be clearly brought out from such studies with structural and energetic details of noncovalent binding features. Several promising ligands for the inhibition of the activity of S-protein on the ACE2 receptors is obtained in this study and the new quantum chemical strategy to design molecular systems with improved binding affinity could be effective for drug development.

Conflicts of interest

"There are no conflicts to declare".

Acknowledgements

The research is supported by CSIR-NIIST COVID-19 project fund. The author is also grateful to Director CSIR-NIIST Dr. A. Ajayaghosh and Dr. S. Savithri, Chief Scientist, CSIR-NIIST for the encouragement.

References

- (1) Lai, C. C.; Shih, T. P.; Ko, W. C.; Tang, H. J.; Hsueh, P. R. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *Int. J. Antimicrob. Agents* **2020**, *55*, 105924.
- (2) Nishiura, H.; Linton, N. M.; Akhmetzhanov, A. R. Serial interval of novel coronavirus (COVID-19) infections. *Int. J. Infect. Dis.* **2020**, *93*, 284-286.

- (3) de Oliveira, O. V.; Rocha, G. B.; Paluch, A. S.; Costa, L. T. Repurposing approved drugs as inhibitors of SARS-CoV-2 S-protein from molecular modeling and virtual screening. *J. Biomol. Struct. Dyn.* **2020**, 1-10, doi: 10.1080/07391102.07392020.01772885.
- (4) Belouzard, S.; Millet, J. K.; Licitra, B. N.; Whittaker, G. R. Mechanisms of Coronavirus Cell Entry Mediated by the Viral Spike Protein. *Viruses* **2012**, 4, 1011-1033.
- (5) Li, W. H.; Moore, M. J.; Vasilieva, N.; Sui, J. H.; Wong, S. K.; Berne, M. A.; Somasundaran, M.; Sullivan, J. L.; Luzuriaga, K.; Greenough, T. C.; Choe, H.; Farzan, M. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **2003**, 426, 450-454.
- (6) Xu, X. T.; Chen, P.; Wang, J. F.; Feng, J. N.; Zhou, H.; Li, X.; Zhong, W.; Hao, P. Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. *Sci. China Life Sci.* **2020**, 63, 457-460.
- (7) Li, W. H.; Zhang, C. S.; Sui, J. H.; Kuhn, J. H.; Moore, M. J.; Luo, S. W.; Wong, S. K.; Huang, I. C.; Xu, K. M.; Vasilieva, N.; Murakami, A.; He, Y. Q.; Marasco, W. A.; Guan, Y.; Choe, H. Y.; Farzan, M. Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. *Embo J.* **2005**, 24, 1634-1643.
- (8) Prabakaran, P.; Mao, X. D.; Dimitrov, D. S. A model of the ACE2 structure and function as a SARS-CoV receptor. *Biochem. Biophys. Res. Co.* **2004**, 314, 235-241.
- (9) Li, F.; Li, W. H.; Farzan, M.; Harrison, S. C. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* **2005**, 309, 1864-1868.
- (10) Smith, M.; Smith, J. C. Repurposing Therapeutics for COVID-19: Supercomputer-Based Docking to the SARS-CoV-2 Viral Spike Protein and Viral Spike Protein-Human ACE2 Interface. *ChemRxiv* **2020**, <https://doi.org/10.26434/chemrxiv.11871402.v4>.
- (11) Trott, O.; Olson, A. J. Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, 31, 455-461.
- (12) Bhatnager, R.; Bhasin, M.; Arora, J.; Dang, A. M. S. Epitope based peptide vaccine against SARS-COV2: an immune-informatics approach. *Biochem. Biop. Res. Co.* **2020**, <https://doi.org/10.1080/07391102.2020.1787227>.
- (13) Bongini, P.; Trezza, A.; Bianchini, M.; Spiga, O.; Niccolai, N. A possible strategy to fight COVID-19: Interfering with spike glycoprotein trimerization. *Biochem. Biop. Res. Co.* **2020**, 528, 35-38.
- (14) Havranek, B.; Islam, S. M. An in silico approach for identification of novel inhibitors as potential therapeutics targeting COVID-19 main protease. *J. Biomol. Struct. Dyn.* **2020**, 1-12, doi: 10.1080/07391102.07392020.01776158.
- (15) Kandeel, M.; Al-Nazawi, M. Virtual screening and repurposing of FDA approved drugs against COVID-19 main protease. *Life Sci.* **2020**, 251, 117627.
- (16) Macchiagodena, M.; Pagliai, M.; Procacci, P. Identification of potential binders of the main protease 3CL(pro) of the COVID-19 via structure-based ligand design and molecular modeling. *Chem. Phys. Lett.* **2020**, 750, DOI: 10.1016/j.cplett.2020.137489.
- (17) Onawole, A. T.; Sulaiman, K. O.; Kolapo, T. U.; Akinde, F. O.; Adegoke, R. O. COVID-19: CADD to the rescue. *Virus Res.* **2020**, 285, 198022.
- (18) Shah, B.; Modi, P.; Sagar, S. R. In silico studies on therapeutic agents for COVID-19: Drug repurposing approach. *Life Sci.* **2020**, 252, 117652.
- (19) Wang, J. M. Fast Identification of Possible Drug Treatment of Coronavirus Disease-19 (COVID-19) through Computational Drug Repurposing Study. *J. Chem. Inf. Model.* **2020**, 60, 3277-3286.
- (20) Elfiky, A. A. Anti-HCV, nucleotide inhibitors, repurposing against COVID-19. *Life Sci.* **2020**, 248, 117477.
- (21) Elmezayen, A. D.; Al-Obaidi, A.; Sahin, A. T.; Yeleki, K. Drug repurposing for coronavirus (COVID-19): in silico screening of known drugs against coronavirus 3CL hydrolase and protease enzymes. *J. Biomol. Struct. Dyn.* **2020**, 1-13, doi: 10.1080/07391102.07392020.01758791.
- (22) Gentile, D.; Patamia, V.; Scala, A.; Sciortino, M. T.; Piperno, A.; Rescifina, A. Putative Inhibitors of SARS-CoV-2 Main Protease from A Library of Marine Natural Products: A Virtual Screening and Molecular Modeling Study. *Marine Drugs* **2020**, 18, doi: 10.3390/md18040225.
- (23) Ibrahim, I. M.; Abdelmalek, D. H.; Elshahat, M. E.; Elfiky, A. A. COVID-19 spike-host cell receptor GRP78 binding site prediction. *J. Infection* **2020**, 80, 554-562.
- (24) Islam, R.; Parves, M. R.; Paul, A. S.; Uddin, N.; Rahman, M. S.; Al Mamun, A.; Hossain, M. N.; Ali, M. A.; Halim, M. A. A molecular modeling approach to identify effective antiviral phytochemicals against the main protease of SARS-CoV-2. *J. Biomol. Struct. Dyn.* **2020**, 1-12, doi: 10.1080/07391102.07392020.01761883.

- (25) Jin, Z. M.; Du, X. Y.; Xu, Y. C.; Deng, Y. Q.; Liu, M. Q.; Zhao, Y.; Zhang, B.; Li, X. F.; Zhang, L. K.; Peng, C.; Duan, Y. K.; Yu, J.; Wang, L.; Yang, K. L.; Liu, F. J.; Jiang, R. D.; Yang, X. L.; You, T.; Liu, X. C.; Yang, X. N.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W. Q.; Xiao, G. F.; Qin, C. F.; Shi, Z. L.; Jiang, H. L.; Rao, Z. H.; Yang, H. T. Structure of M-pro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020**, *582*, 289-293.
- (26) Mittal, L.; Kumari, A.; Srivastava, M.; Singh, M.; Asthana, S. Identification of potential molecules against COVID-19 main protease through structure-guided virtual screening approach. *J. Biomol. Struct. Dyn.* **2020**, *1-19*, doi: 10.1080/07391102.07392020.01768151.
- (27) Pant, S.; Singh, M.; Ravichandiran, V.; Murty, U. S. N.; Srivastava, H. K. Peptide-like and small-molecule inhibitors against Covid-19. *J. Biomol. Struct. Dyn.* **2020**, *1-10*, doi: 10.1080/07391102.07392020.01757510.
- (28) Hagar, M.; Ahmed, H. A.; Aljohani, G.; Alhaddad, O. A. Investigation of Some Antiviral N-Heterocycles as COVID 19 Drug: Molecular Docking and DFT Calculations. *Int. J. Mol. Sci.* **2020**, *21*, doi: 10.3390/ijms21113922.
- (29) Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z. F.; Liu, F. Y.; Li, H. B.; Ding, L. N.; Morokuma, K. The ONIOM Method and Its Applications. *Chem. Rev.* **2015**, *115*, 5678-5796.
- (30) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Wallingford, CT, 2016.
- (31) Tirado-Rives, J.; Jorgensen, W. L. Performance of B3LYP density functional methods for a large set of organic molecules. *J. Chem. Theor. Comput.* **2008**, *4*, 297-306.
- (32) Hostas, J.; Rezac, J.; Hobza, P. On the performance of the semiempirical quantum mechanical PM6 and PM7 methods for noncovalent interactions. *Chem. Phys. Lett.* **2013**, *568*, 161-166.
- (33) Remya, G. S.; Suresh, C. H. Quantification and classification of substituent effects in organic chemistry: A theoretical molecular electrostatic potential study. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20615-20626.
- (34) Bijina, P. V.; Suresh, C. H. Molecular Electrostatic Potential Reorganization Theory to Describe Positive Cooperativity in Noncovalent Trimer Complexes. *J. Phys. Chem. A* **2020**, *124*, 2231-2241.
- (35) Bijina, P. V.; Suresh, C. H.; Gadre, S. R. Electrostatics for probing lone pairs and their interactions. *J. Comput. Chem.* **2018**, *39*, 488-499.