



Targeting SARS-CoV-2 main protease: structure based virtual screening, in silico ADMET studies and molecular dynamics simulation for identification of potential inhibitors

Ankit Uniyal^{a†}, Manoj Kumar Mahapatra^{b†} , Vinod Tiwari^a, Rajat Sandhir^c and Rajnish Kumar^a 

^aDepartment of Pharmaceutical Engineering & Technology, Indian Institute of Technology (B.H.U.), Varanasi, India; ^bKanak Manjari Institute of Pharmaceutical Sciences, Rourkela, India; ^cDepartment of Biochemistry, Panjab University, Chandigarh, India

Communicated by Ramaswamy H. Sarma

ABSTRACT

COVID-19 pandemic has created a healthcare crisis across the world and has put human life under life-threatening circumstances. The recent discovery of the crystallized structure of the main protease (M_{pro}) from SARS-CoV-2 has provided an opportunity for utilizing computational tools as an effective method for drug discovery. Targeting viral replication has remained an effective strategy for drug development. M_{pro} of SARS-CoV-2 is the key protein in viral replication as it is involved in the processing of polyproteins to various structural and nonstructural proteins. Thus, M_{pro} represents a key target for the inhibition of viral replication specifically for SARS-CoV-2. We have used a virtual screening strategy by targeting M_{pro} against a library of commercially available compounds to identify potential inhibitors. After initial identification of hits by molecular docking-based virtual screening further MM/GBSA, predictive ADME analysis, and molecular dynamics simulation were performed. The virtual screening resulted in the identification of twenty-five top scoring structurally diverse hits that have free energy of binding (ΔG) values in the range of -26.06 (for compound AO-854/10413043) to -59.81 Kcal/mol (for compound 329/06315047). Moreover, the top-scoring hits have favorable ADME properties as calculated using in silico algorithms. Additionally, the molecular dynamics simulation revealed the stable nature of protein-ligand interaction and provided information about the amino acid residues involved in binding. Overall, this study led to the identification of potential SARS-CoV-2 M_{pro} hit compounds with favorable pharmacokinetic properties. We believe that the outcome of this study can help to develop novel M_{pro} inhibitors to tackle this pandemic.

ARTICLE HISTORY

Received 6 October 2020
Accepted 4 November 2020

KEYWORDS

Virtual screening; hit identification; SARS-CoV-2; main protease; in silico ADME; molecular dynamics; anti-viral

1. Introduction

SARS-CoV-2 is a novel coronavirus belonging to the coronaviridae family and is responsible for acute respiratory illness (COVID-19). World Health Organization has declared it as a global pandemic and categorized it into the emergency diseases. More than 32.7 million confirmed cases along with 991000 fatalities have been reported as on 28th September 2020 (Weekly Epidemiological Update, 2020). COVID-19 has created a substantial threat to human lives by putting the healthcare system in unprecedented situations. Scientists and researchers across the globe are trying to identify new therapies as well as repurposing the clinically established therapies for the management of COVID-19 (Guo et al., 2020; Liu et al., 2020). No specific vaccine and the antiviral agent has translated to the clinics so far for the effective control of COVID-19. The available drug regimens have limited efficacy thus imposing the need for comprehensive strategies and optimized therapeutics development (Sanders et al., 2020). At this unprecedented time to reduce the false-positive errors and to expedite the drug discovery process, artificial

intelligence could provide a real-time solution. Using the structure-based virtual and high throughput screening followed by pharmacological investigation may be an effective strategy for the development of therapeutics against SARS-CoV-2 (Adeshina et al., 2020).

The genome of SARS-CoV-2 is single-stranded positive-sense RNA having a length of $\sim 30,000$ nucleotides (Chen et al., 2020). This RNA is a 5'-cap structure with 3'-poly (A) tail. There are six open reading frames (ORFs) observed in the SARS-CoV-2 genome. The ORF1a and ORF1ab encode for polyproteins (pp1a and pp1ab) that undergoes the process of proteolysis and get processed into nonstructural proteins (NSP1-NSP16) (Gordon et al., 2020). Enzymes responsible for the proteolysis are the main protease (M_{pro}), also known as 3-chymotrypsin like cysteine protease ($3CL^{pro}$), and papain-like proteases. These NSPs form the replicase-transcriptase complex which assembles the subgenomic RNA which further promotes the synthesis of structural proteins namely envelope, spike, membrane and nucleocapsid (Chen et al., 2020; Gordon et al., 2020). Thus, M_{pro} is responsible for

charging the polyproteins with their respective functional entities and is crucial for maintaining the life cycle of COVID-19. The M_{pro} (NSP5) interacts with the epigenetic regulators i.e. histone deacetylase 2 (HDAC2) and tRNA methyltransferase (Gordon et al., 2020). Also, the trafficking across the mitochondria and endoplasmic reticulum are reported to be regulated by the NSP5 (Gordon et al., 2020).

Superimposed X-ray crystal structures of the main proteases of SARS-CoV-2, SARS-CoV and MERS-CoV display a high degree of sequence identity and conservation of active site residues. SARS-CoV-2 main protease has about 96% sequence identity with that of SARS-CoV main protease, while the rest 4% variable residues are not primarily involved in substrate binding (Zhang et al., 2020). M_{pro} is one of the nonstructural proteins which plays a pivotal role in the life cycle of the virus by processing various polyproteins that are translated from viral RNA. This fact along with the highly conserved nature of the main protease makes it a promising target for designing antiviral agents. The main protease is a cysteine protease that is biologically active in its dimeric form only. Electron density mapping of the individual chain confirmed the presence of 306 residues in each of the monomer, which in turn comprised of three domains and a catalytic dyad (His₄₁ and Cys₁₄₅). Residues from 8-101, 102-184, 201-303 form the domain-I, II and III respectively (Goyal & Goyal, 2020). Domain-III contains 5 antiparallel α helices which are arranged into a globular cluster and is attached to the domain-II with the help of a long loop of residues 185-200 (Jin et al., 2020). The catalytic dyad (His₄₁ and Cys₁₄₅) as well as the substrate-binding site, is situated inside the cleft present in between domain-I and II. In 6LU7 main protease, the co-crystallized ligand has shown to have anti-parallel alignment with the residues 164-168 on one side and with residues 189-191 on the other side. Several studies have reported the important residues like His₄₁, Met₄₉, Gly₁₄₃, Ser₁₄₄, His₁₆₃, His₁₆₄, Met₁₆₅, Glu₁₆₆, Leu₁₆₇, Asp₁₈₇, Arg₁₈₈, Gln₁₈₉, Thr₁₉₀, Ala₁₉₁, Gln₁₉₂ are responsible for substrate binding activity, whereas residues Arg₄, Ser₁₀, Gly₁₁, Glu₁₄, Asn₂₈, Ser₁₃₉, Phe₁₄₀, Ser₁₄₇, Glu₂₉₀, Arg₂₉₈ are responsible for dimerization of the main protease. Jin et al. (2020) have reported that important residues like His₄₁, Phe₁₄₀, Cys₁₄₅, His₁₆₃, Met₁₆₅, Glu₁₆₆, Leu₁₆₇, His₁₇₂, Asp₁₈₇ forming the catalytic binding pocket are highly conserved. The surface loop and the domain-III of M_{pro} contains the most variable regions, whereas the substrate-binding site is highly conserved among the main proteases found in all coronaviruses. Inhibition of the main protease leads to blocking of replication of the coronavirus. As the M_{pro} of SARS-CoV-2 is not homologous with any human proteases known so far, its inhibitors would be less likely to be toxic (Kim et al., 2016). Thus, targeting the M_{pro} enzyme of SARS-COV-2 could provide a strategic approach to curb the viral replication and ultimately effective management of the COVID-19 crisis (Ullrich & Nitsche, 2020).

Determination of the crystal structure of M_{pro} at high resolution (PDB ID:6YB7) provides a great opportunity for the use of structure based drug design tools to screen large

libraries of commercially available compounds to identify hit molecules which would be cost-effective in the development of new antiviral agents (Jin et al., 2020; Ren et al., 2013). A global program COVID MoonShot has been launched to crowdsource drug-design by medicinal chemists. The MoonShot database suggests that the crystallography studies have confirmed that M_{pro} has a solvent accessible and empty active binding site. At this collaborative organ, more than 1500 crystal experiments have been done yet and the data is made available in the public domain. Whereas, more than 500 compounds have been tested which demonstrates the drug-design-based effective screening of hit molecules. Multiple fragment libraries e.g.; DSI-poised library, York3D, and MiniFragments FragLites & Peplites, etc. have been created and screened against M_{pro} (Main protease structure and XChem fragment screen, 2020). Crystal soaking and co-crystallization are being used for the identification of potential covalent, non-covalent, and dimer targeting hit molecules.

In this study, we have performed a structure based virtual screening of a commercially available library of compounds against the active site binding pocket of SARS-CoV-2 M_{pro} . The top-ranking compounds were further subjected to MM/GBSA binding free energy calculation and in silico absorption, distribution, metabolism, excretion, and toxicity (ADMET) predictions for pharmacokinetic profiling of identified potential M_{pro} inhibitors. Finally, a short molecular dynamics simulation is performed in order to evaluate the stability of the enzyme-inhibitor complex.

2. Material and methods

The computational studies including molecular docking-based virtual screening, MM/GBSA, and ADMET calculations were performed using Schrödinger molecular modeling suite. Glide module implemented in Schrödinger molecular modeling suite was used for docking which approximates a complete systematic search of the conformational, orientational, and positional space of the docked ligand into the binding pocket (Friesner et al., 2004). MM/GBSA method was used to calculate binding free energy using the Prime module (Jacobson et al., 2002; Schrödinger Release 2020-3: Prime, Schrödinger, LLC, New York, NY, 2020). QikProp (Schrödinger Release 2020-3: QikProp, Schrödinger, LLC, New York, NY, 2020) was used for the prediction of ADME properties of the identified hit compounds. Molecular dynamics study was performed using Desmond (Bowers, 2006) module implemented in Schrödinger molecular modeling suite (Schrödinger Release 2020-3: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2020. Maestro-Desmond Interoperability Tools, Schrödinger, New York, NY, 2020).

2.1. Protein preparation and docking grid generation

Beginning from the outbreak of this pandemic, several attempts have been made on the front of structural biology to solve the three-dimensional structure of different drug targets to aid in the drug discovery process. The first three-dimensional crystal structure of M_{pro} was published in early

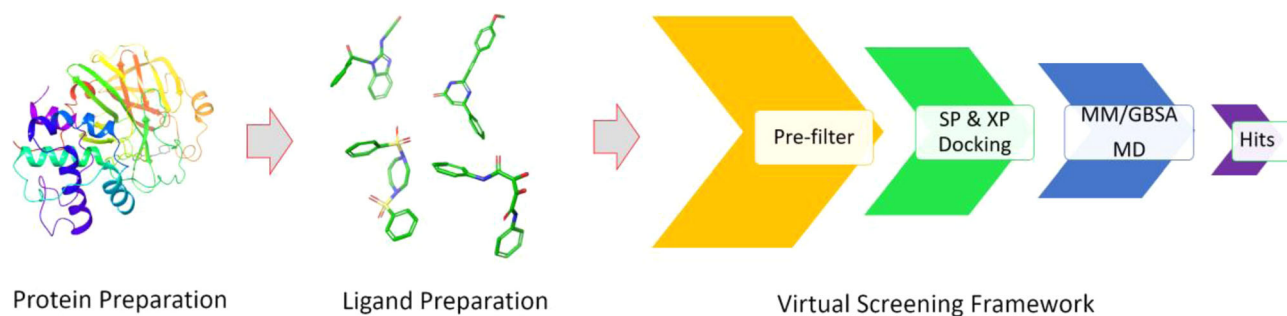


Figure 1. The overall workflow of virtual screening and analysis.

February 2020 (PDB ID: 6LU7) (Jin et al., 2020). For the current study, we have used the crystal structure of COVID-19 main protease in complex with a co-crystallized fragment 2-cyclohexyl-N-(3-pyridyl)acetamide (Z31792168) (PDB ID: 5R84) (<https://www.rcsb.org/structure/5R84>). As a standard practice, the protein structure was prepared using the protein preparation wizard of the Schrödinger suite before further use which includes the assignment of bond orders were, the addition of missing hydrogens, and removal of crystallization water molecules. Further, protonation state and tautomer assignment were carried out and the protein structure was submitted to a restrained molecular mechanics refinement using OPLS_2005 force field (Jorgensen & Tirado-Rives, 1988) until the heavy atom root square deviation value converged to 0.3 Å. To define the binding site for docking of database compounds, a docking grid was generated using centeroid of the co-crystallized ligand Z31792168 with a cutoff of docking ligands with a maximum length of 20 Å. The prepared grid was further used for structure-based virtual screening.

2.2. Database preparation

A large number of commercially available chemical libraries are present from where one can cherry pick the compounds for in vitro/in vivo testing based on the initial computational screening (Lyu et al., 2019). One such library is Asinex Gold & Platinum collection which contains 29174 compounds and provides diverse and cost-effective coverage of drug-like chemical space with most of them obeying Lipinski's Rule of 5 (Asinex, 2020). As a first step, the library was filtered on the basis of optimum physicochemical properties using a modified Lipinski filter (Lipinski et al., 1997), and PAINS substructure (Baell & Holloway, 2010) filter for removal of pan-assay interference compounds to increase the chance to obtain pharmacokinetically optimum hit compounds. LigPrep (Schrödinger Release 2020-3: LigPrep, Schrödinger, LLC, New York, NY, 2020) module was used for library preparation which performed several tasks such as generated stereoisomers, neutralized charged structures, and determined the most probable ionization state at pH 7.0 ± 2.0 , and generated tautomers. Stereoisomers were generated where the compounds were having chiral centers and a single lowest energy conformation obtained by ConfGen were retained which were further minimized using the Impref module by applying OPLS_2005 force field method. The prepared library

containing 132727 compounds was used for docking-based virtual screening in the next step. The co-crystallized ligand Z31792168 was prepared in a similar manner.

2.2.1. Structure-based virtual screening

The structure-based virtual screening was performed with a slight deviation from the earlier reported procedure (Yadav et al., 2016). Virtual screening workflow implemented in Maestro (Schrödinger Release 2020-3: Maestro, Schrödinger, LLC, New York, NY, 2020) interface of Schrödinger Suite was used and the overall workflow of virtual screening followed by other in silico studies has been depicted in Figure 1. The dataset prepared in the previous step was used in high throughput virtual screening (HTVS) and the output compounds from HTVS were ranked on the basis of Glide Score and 25% of top-scoring compounds further parsed to standard precision (SP) mode of docking. Further, the top-ranked 25% compounds were parsed to extra precision (XP) (Friesner et al., 2006) mode of docking which in the end provided 250 compounds. For the purpose of comparison, the co-crystallized ligand, Z31792168 was docked in a similar manner and the docking score and binding pose is compared with the top scoring hits. After manual inspection of the protein-ligand interaction of the top-scoring compounds, the top 25 compounds were chosen for further analysis. The output from the virtual screening utility containing all the thermodynamic information in the form of Glide Score was analyzed using Glide XP visualizer, which enables visualization of ligand-receptor interactions in an interactive manner.

2.3. Binding free energy calculations

The virtual screening workflow provided us with the top-scoring hit compounds but the XP docking is not able to estimate the binding free energy of the ligand-receptor complex. Therefore, we used Prime MM/GBSA approach to predict the binding free energy of the receptor-ligand complex (Lyne et al., 2006). The pose viewer file from the XP docking comprised of the receptor-ligand complex was used for calculation. Principally, the MM/GBSA approach employs molecular mechanics, the generalized Born model, and solvent accessibility method to predict the free energies from structural information circumventing the computational complexity of free energy simulations (Dill, 1997). The binding

free energy of each ligand was calculated using the following equation:

$$\Delta G_{bin} = \Delta E_{mm} + \Delta G_{sol} + \Delta G_{SA}$$

Where ΔE_{mm} is the difference in the minimized energies between the receptor-ligand complex and the sum of energies of the unliganded receptor and ligands. ΔG_{sol} is the difference in the GBSA solvation energy of the receptor-ligand complex and the sum of energies of the unliganded receptor and ligands. ΔG_{SA} is the difference in surface area energies for the receptor-ligand complex and the sum of energies of the unliganded receptor and ligands.

2.4. In silico ADMET prediction

Since non-optimal ADMET properties are a major cause of attrition in drug development, optimization of these properties at the early stage of drug discovery is highly encouraged (Alavijeh & Palmer, 2004). Therefore, the top twenty-five hits were subjected to the calculation of various physicochemically significant and pharmacokinetically relevant ADMET properties using the QikProp module implemented in Schrodinger molecular modeling suite. (Schrodinger Release 2020-3: QikProp, Schrodinger, LLC, New York, NY, 2020). Similarly, the ADMET calculations were performed for the co-crystallised ligand, Z31792168 and were compared with those of top scoring hits.

2.5. Molecular dynamics simulation

The general limitation of any grid-based docking algorithm is that it treats the receptor as a rigid entity and therefore it provides a still picture of the protein-ligand interaction. However, in the physiological system, this interaction is dynamic in nature. Therefore, to better understand the interaction between the protein-ligand complex of identified hits, we performed a molecular dynamics simulation of the top five identified hits using Desmond molecular dynamics program (Schrodinger Release 2020-3: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2020. Maestro-Desmond Interoperability Tools, Schrodinger, New York, NY, 2020). In general, the protein-ligand complex for each hit was placed into an orthorhombic box of size $10 \times 10 \times 10 \text{ \AA}$ and solvated explicitly with TIP3P. OPLS3e was used as a force field for the system. The system was neutralized by Na^+ and Cl^- ions at a final concentration of 0.15 M. The system was minimized and pre-equilibrated using the standard equilibration protocol implemented in Desmond which contains several steps before the final production run. Beginning with a Brownian Dynamics NVT simulation for 100 ps at 10 K temperature, with small timesteps and restraints on solute heavy atoms followed by NVT simulation for 12 ps at 10 K temperature with small timesteps and restraints on solute heavy atoms. Then a 12 ps NPT simulation at 10 K temperature with restraints on solute heavy atoms and a 12 ps NPT simulation with restraints on solute heavy atoms, and finally a 24 ps NPT simulation without any restraints. After all these equilibration steps, a final

production simulation was performed for 50 ns using NPT (normal Pressure and Temperature) ensemble at 300 K and 1.013 bars with the default setting of relaxation before simulation (Bowers, 2006). The Nose-Hoover Chain thermostat and Martyna–Tobias–Klein barostat were used to maintain the temperature and pressure, respectively (Martyna et al., 1994). The RESPA integrator was used with a time step of 2 fs during the simulation with a smooth PME method for the calculation of long-range electrostatic interaction. The energy and coordinates were saved in the trajectory at every 10 ps. The final simulation trajectory was analyzed using the simulation interaction diagram available in Maestro.

3. Results and discussion

3.1. Structure-based virtual screening

To date, several crystal structures of SRS-CoV-2 Mpro have been solved and deposited in protein data bank using x-ray crystallography and cryo-EM technologies. In this study, we have utilized viral M_{pro} structure in complex with a co-crystallized fragment 2-cyclohexyl-N-(3-pyridyl)acetamide (Z31792168) (PDB ID: 5R84). One of the important steps is proper curation and preparation of ligand libraries with diverse chemical scaffolds but with drug-like properties. The screening of dataset compounds was curated from Asinex website and the library was properly prepared properly as described in the methods section. The prepared library was then subjected to structure-based virtual screening workflow as implemented in the Schrodinger molecular modeling suite which finally led to the identification of top-scoring compounds hits. The docking score of the twenty-five top scoring hits compounds have been provided in Table 1 and it can be observed that the co-crystallised ligand Z31792168 has a Glide gscore of -4.81 as compared to the top twenty five hits which has a Glide gscore in the range of -8.09 to -6.37 . It indicates that the top scoring hits have better binding efficiency as compared to the co-crystallised ligand Z31792168. The 2D structures of the top twenty-five hits are shown in Figure 2 and the 2-D ligand interaction diagram and 3D pose of the ligand in the binding pocket of the enzyme obtained after final docking step for the cocrystallized ligand and the top five hits is shown in Figures 3–5. It can be observed from the structures that the top-scoring hits represent diverse chemical scaffolds like substituted xanthine, dibenzylamine, naphthyl sulfonamide, N,N-disubstituted pyrazine, indole, and benzamide, etc.

Docking studies revealed that all the top 5 hits shared the same binding cavity and have shown interactions with various catalytic site residues like His_41, Asn_142, Gly_143, Ser_144, Cys_145, Met_165, Glu_166 of the main protease. The amide nitrogen of the co-crystallized native ligand was found to interact with M^{pro} mainly through strong hydrogen bonding with Hie_164 amino acid residue and the pyridinyl nitrogen formed a hydrogen bond with Hie_163 residue (Figure 3(A)). Additionally, the pyridinyl ring is surrounded by Cys_145, Ser_144, Asn_142, Leu_141, and Phe_140. On the other hand, the cyclohexyl group is surrounded by Gln_189,

Table 1. Docking and MMGBSA dG binding scores of co-crystallized ligand Z31792168 and top twenty-five hits.

S. no.	Title	Glide gscore	Glide energy	Energy	XP GScore	MMGBSA dG bind	MMGBSA dG bind Coulomb	MMGBSA dG bind Covalent	MMGBSA dG bind Solv GB	MMGBSA dG bind vdW
1	Z31792168 (Co-crystallized ligand)	-4.81	-33.55	21.97	-4.81	-43.64	-8.24	1.19	14.24	-37.02
2	AG-690/11060013	-8.09	-55.52	13.27	-8.09	-54.50	-29.50	4.94	27.16	-41.91
3	AG-690/11203374_1	-7.87	-55.00	15.55	-7.87	-56.54	-22.55	1.02	23.58	-43.04
4	AG-690/11203374_2	-7.70	-56.30	16.82	-7.70	-51.93	-25.79	5.45	23.86	-41.41
5	AG-690/11203374_3	-7.44	-56.62	17.05	-7.44	-52.72	-17.42	4.40	21.30	-45.36
6	AH-034/04857012	-7.28	-47.08	40.38	-7.28	-34.03	4.41	6.27	24.64	-45.24
7	AG-690/11060018	-7.12	-48.54	14.54	-7.12	-40.20	-21.42	3.98	30.97	-38.95
8	AG-205/05184040	-6.99	-44.16	24.23	-6.99	-41.36	-7.47	1.87	19.17	-34.93
9	AO-365/11349014	-6.97	-44.70	28.83	-6.97	-43.94	-27.65	5.10	37.93	-41.41
10	AH-034/11365679	-6.93	-42.96	40.82	-6.93	-49.52	-9.85	5.09	25.70	-46.57
11	AN-329/06315047	-6.89	-49.15	47.06	-6.89	-59.81	-20.69	3.73	20.23	-44.16
12	AO-365/11349014	-6.66	-43.78	28.37	-6.66	-44.08	-31.43	4.11	41.19	-42.35
13	AN-329/09986025	-6.62	-39.13	42.51	-6.62	-44.20	-16.84	4.33	19.70	-34.56
14	AK-968/11492131	-6.60	-46.71	27.39	-6.60	-49.28	-16.93	5.53	23.15	-42.20
15	AE-641/00653027	-6.69	-37.00	39.04	-6.69	-42.15	-18.53	5.41	24.18	-39.27
16	AE-848/11243033	-6.56	-42.86	17.91	-6.56	-43.49	-6.54	5.30	13.82	-39.29
17	AG-205/08396013	-6.55	-38.71	19.20	-6.55	-40.86	-9.05	2.94	22.12	-40.18
18	AE-848/11243033	-6.54	-45.97	17.91	-6.54	-41.79	-19.42	4.27	23.03	-32.83
19	AE-641/00004064	-6.54	-40.19	37.01	-6.54	-41.59	-25.85	4.51	24.72	-28.55
20	AO-854/10413043	-6.50	-32.87	28.01	-6.50	-26.07	-12.62	3.65	26.72	-27.30
21	AK-778/10920048	-6.40	-39.01	28.60	-6.40	-46.57	-19.47	5.31	17.21	-36.37
22	AN-652/11380002	-6.38	-37.08	27.73	-6.38	-45.41	-23.92	3.40	23.00	-31.70
23	AG-690/08354009	-6.37	-40.55	27.95	-6.37	-47.09	-11.99	0.85	20.70	-39.52
24	AG-690/10772011	-6.30	-43.71	24.53	-6.30	-44.16	-14.10	2.01	19.37	-34.62
25	AN-652/11380002	-6.27	-32.81	27.24	-6.27	-41.92	-24.02	7.82	29.13	-33.63
26	AG-205/11132201	-6.30	-33.93	18.00	-6.30	-38.40	-15.08	1.70	17.15	-33.17

Arg₁₈₈, Asp₁₈₇, and Hie₄₁. Similar kinds of interactions were observed in the case of the compound AG-690/11060013 where the cyclic keto groups of the core nucleus formed hydrogen bonds with Gly₁₄₃ and Thr₂₆ (Figure 3(B)). Compound AG-690/11203374_1 and AG-690/11203374_2 interacted in a similar fashion as shown in Figure 4(A,B), respectively. with the active site of M_{pro} in a similar fashion. The core nucleus formed hydrogen bonds with the Thr₂₆ and Gly₁₄₃ while the p-chloro phenoxy side chain is surrounded by Ser₄₆, Cys₄₄ and Hie₄₁ amino acid residues. Figure 5(A,B) shows docked pose and ligand interaction diagram of hits AG-690/11203374_3 and AH-034/04857012, respectively. Compound AG-690/11203374_3 interacted well with the active site of M_{pro} through hydrogen bonds with Gly₁₄₃ and Thr₂₆. Additional interaction of the chloro group on phenyl sidechain was observed with Cys₄₄ in the form of a halogen bond. While the compound AH-034/04857012 formed a hydrogen bond with Glu₁₆₆. Overall, in contrast to the co-crystallized ligand, the hit molecules have shown additional favorable interactions with other important residues like His₄₁, Asn₁₄₂, Gly₁₄₃, Ser₁₄₄, Cys₁₄₅. which makes them potential ligand for this enzyme.

3.2. Binding free energy calculations

Further, in order to calculate the free energy of binding, the MMGBSA approach was used. The binding free energy scores of top hits is given in Table 1. From the table, it can be observed that the ΔG bind values were in the range of -26.06 (for AO-854/10413043) to -59.81 (for AN-329/06315047) Kcal/mol as compared to the co-crystallized

ligand Z31792168 which gave a ΔG bind value of -43.64 Kcal/mol.

3.3. In silico ADMET prediction

Incorporation of ADMET predictions during pre-clinical drug discovery can help to reduce the risk of failure of drug candidates in clinical trials and this motivated development of several tools for ADMET prediction. In this study, we performed *in silico* ADMET calculations of top twenty-five hits along with the co-crystallized ligand, Z31792168 using QikProp module and the predicted values of several important parameters along with their acceptable range are given in Table 2. A number of pharmaceutically significant Physico-chemical descriptors like SASA (total solvent accessible surface area), FOSA (saturated carbon and attached hydrogen representing a hydrophobic component of SASA), FISA (Hydrophilic component of SASA, i.e. N, O & H on heteroatoms), PISA (π component of SASA, i.e. C and attached H), QPlogPo/w (Predicted octanol/water partition coefficient), QPlogS (Predicted aqueous solubility), QPlogHERG (Predicted IC₅₀ for the blockage of HERG K⁺ channels), QPPCaco (Predicted apparent Caco-2 cell permeability for gut-blood barrier), QPlogBB (Predicted brain/blood partition coefficient) etc. were selected for this study. We found a slight deviation of PISA parameter for compound AH-034/04857012 and AH-034/11365679, whereas a little deviation of QPlogS has been found for compound AH-034/11365679. Apart from these two compounds, all other top twenty-five hits and the co-crystallized ligand, Z31792168 were having favorable pharmacokinetic parameters, which signifies their druggable potential for human use.

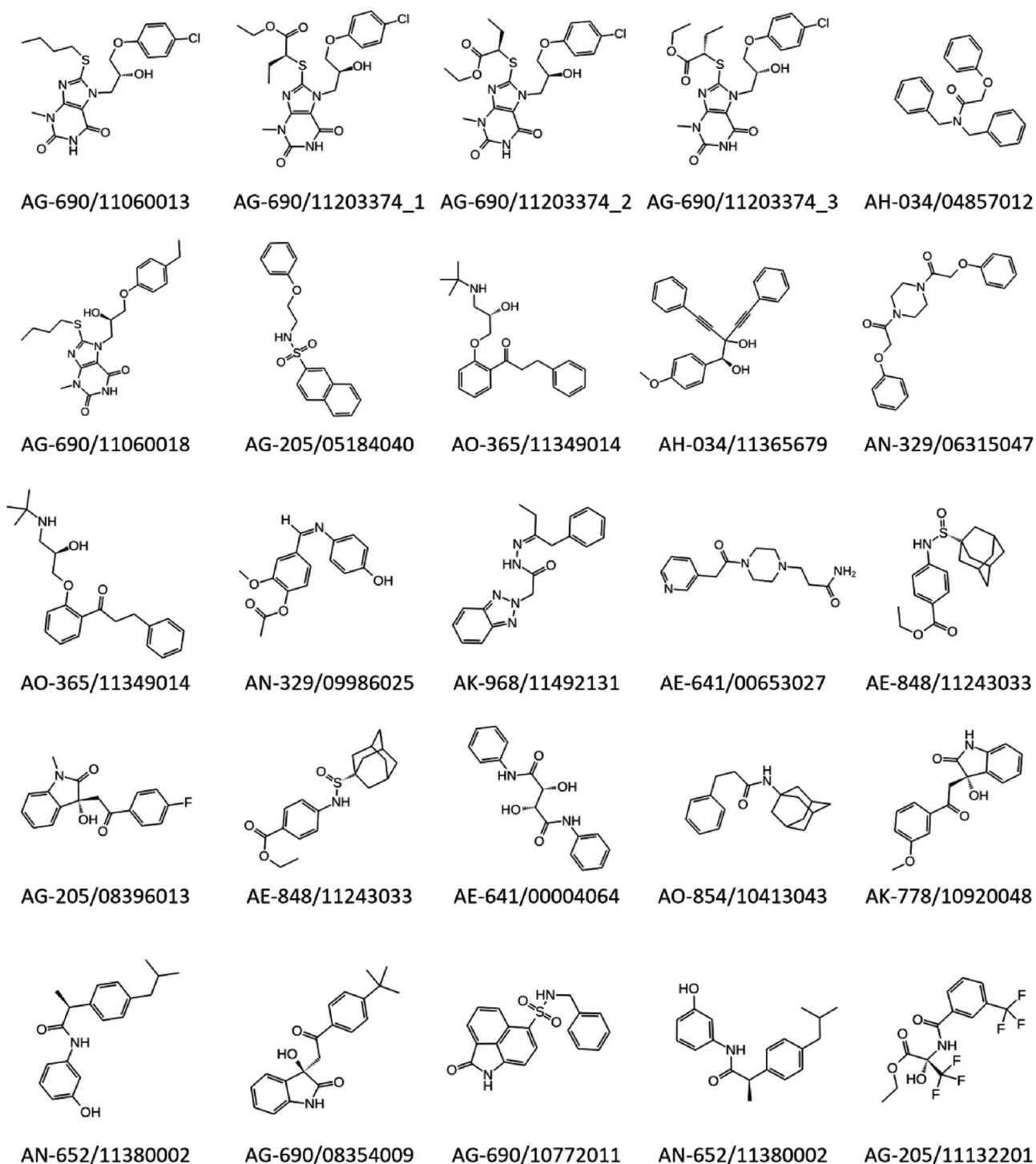


Figure 2. Two-dimensional structures of top twenty-five hits identified through structure-based virtual screening.

3.4. Molecular dynamics simulation

Various factors like glide score, MMGBSA energy and ADMET properties were considered for selecting the top 5 hits (AG-690/11060013, AG-690/11203374_1, AG-690/11203374_2, AG-690/11203374_3, AH-034/04857012), which were further utilized for simulation studies. As an additional measure, we performed 50 ns long molecular dynamics simulation of the ligand-protein complex for the top five hits obtained from the virtual screening experiment. The rationale for this was to understand the dynamics of ligand binding to the active

site of the enzyme. A molecular dynamics simulation study was necessary to examine the stability and dynamic fluctuations in the ligand-protein complex under a simulated biological environment.

Figure 6 depicts various MD trajectory data analysis for ligand AG-690/11060013. The RMSD plot (Figure 6(A)) indicated a stable ligand-protein complex throughout the entire simulation period which was determined from the RMSD values ranging from 0.5 to 2.5 Å for both the protein and ligand system. Additionally, the flexibility of the protein system was

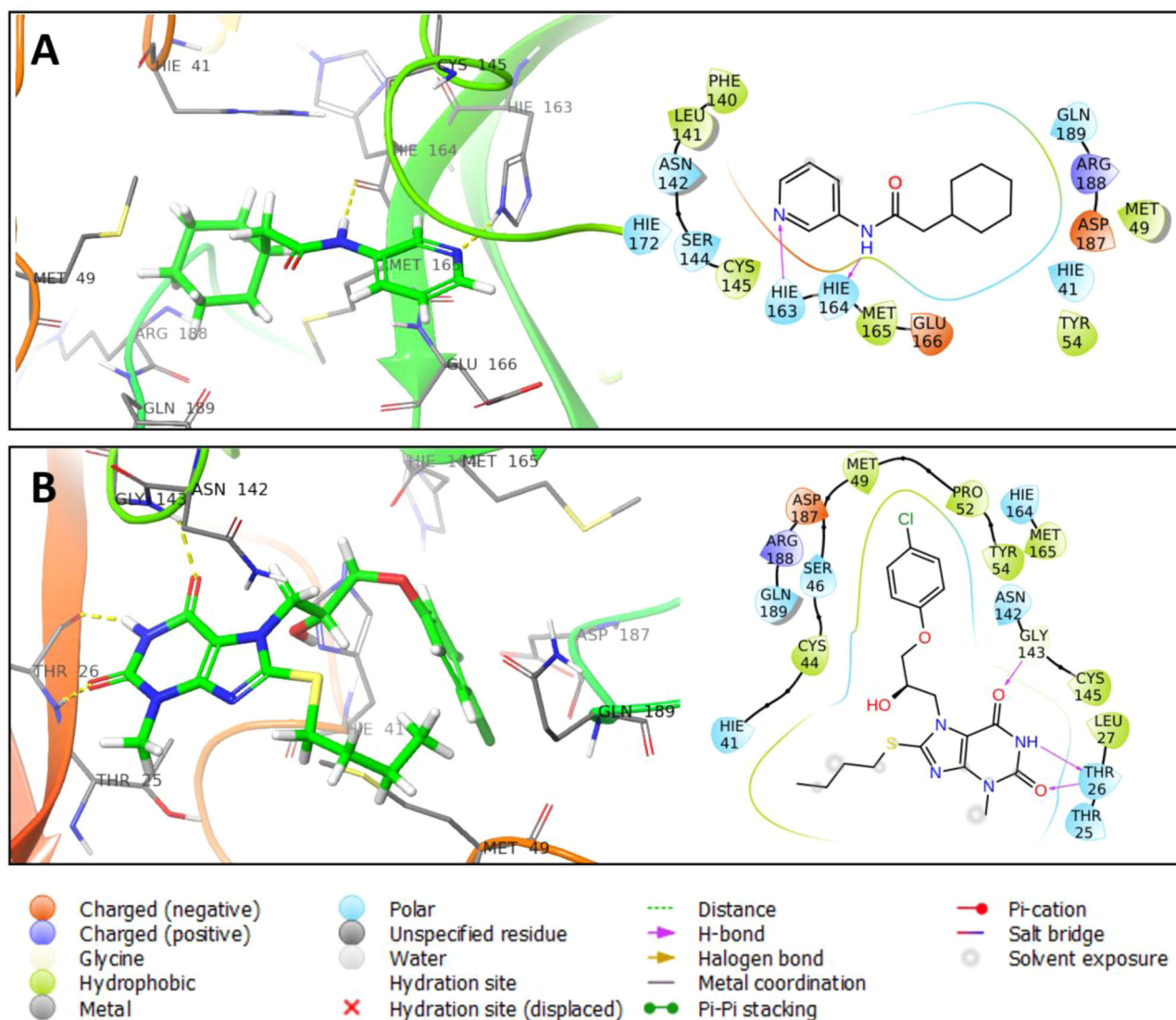


Figure 3. Three-dimensional binding pose and two-dimensional ligand interaction diagram of co-crystallized ligand, Z31792168 (A) and top hit AG-690/11060013 (B).

assessed during the entire simulation by calculating the RMSF of individual amino acid residues of the protein (Figure 6(B)) which ranged from 0.4 to 2.0 Å indicating a stable protein-ligand complex. The binding interactions between ligand and active site amino acid residues inside the binding pocket of M_{pro} were computed and represented in Figure 6(C). It was observed that the ligand interacted with the M_{pro} active site with THR₂₆, HIS₄₁, GLY₁₄₃ and GLU₁₆₆ mainly through hydrogen bonding. Hydrophobic interactions with HIS₄₁, MET₄₉ and MET₁₆₅ and various water bridges with residues SER₁₄₄ and CYS₁₄₅ also played a supportive role in binding the ligand with M_{pro} active site. RMSF plot of ligand (Figure 6(D)) suggested that the ligand is located in the binding pocket and bound to the amino acid residues at the active site. Figure 6(E) depicts specific interactions between the ligand and active site amino acid residues of M_{pro} . The xanthine core along with the side chain substitutions engaged the active site amino acid residues through various H-bonds, hydrophobic interactions, and water bridges. Finally, a timeline of the protein-ligand interactions was plotted (Figure 6(F)) in the form of hydrogen bonds,

hydrophobic interaction, ionic interaction, and water bridges throughout the 50 ns simulation period. The top panel indicates the total number of specific protein-ligand contacts and the bottom panel indicates residue level interaction of the ligand. Overall, the ligand interacted well with the M_{pro} binding pocket and a minimum of 4 contacts were present throughout the simulation period.

Figure 7 depicts all the analysis data of MD trajectory for ligand AG-690/11203374_1. It can be observed from the root mean square deviation (RMSD) plot (Figure 7(A)) that the protein as well ligand remain stable throughout the simulation period and the RMSD value was between 1.1 to 2.7 Å for both protein and ligand, indicating a stable complex between them. Additionally, in order to evaluate the flexibility of the protein structure throughout the simulation, we calculated root mean square fluctuation (RMSF) of individual amino acid residues of the protein (Figure 7(B)). The whole protein had an RMSF value between 0.4 to 2.0 Å except the terminal amino acids that showed moderate movement throughout the 50 ns long simulation. The interaction of the ligand with the individual amino acid residues in the binding

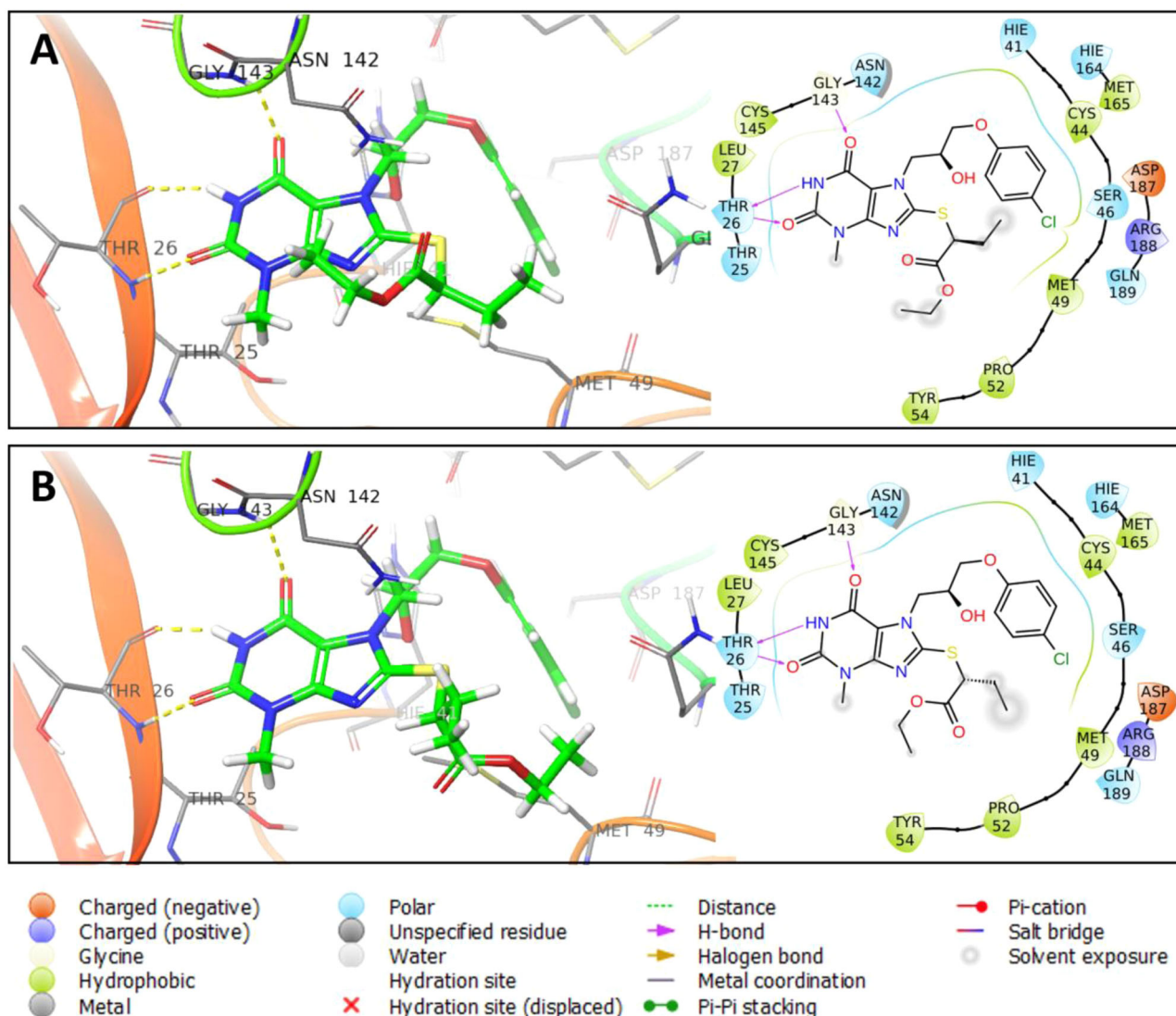


Figure 4. Three-dimensional binding pose and two-dimensional ligand interaction diagram of hits AG-690/11203374_1 (A) and AG-690/11203374_2 (B).

pocket of M_{pro} was quantified and represented in Figure 7(C). It is clear from the plot that the ligand interacted with the M_{pro} active site with THR_26 mainly through hydrogen bonding, with HIS_41 by hydrophobic interaction and with GLU_166 via both hydrogen bonding and water bridges. Hydrophobic interactions played a major role in the interaction of this ligand with M_{pro} active site. Hydrogen bonding is one of the major forces involved in protein-ligand interaction and it has an indispensable role in stabilizing the protein-ligand complexes (Du et al., 2016). RMSF plot of ligand (Figure 7(D)) indicated that the ligand largely stayed bound to the active site amino acids. Figure 7(E) shows specific interactions of the ligand with active site amino acid residues of M_{pro} . It is observed that both xanthine core, as well as side chain substitution, engaged the active site amino acid residues. Finally, a timeline of protein-ligand interactions in the form of hydrogen bonds, hydrophobic interaction, ionic interaction, and water bridges is plotted as shown in Figure 7(F) throughout the 50 ns simulation time. The top panel indicates the total number of specific protein-ligand contacts and the bottom panel indicates residue level interactions of the ligand. Overall, the ligand interacted well with the M_{pro}

binding pocket and a minimum of 4 contacts were present throughout the simulation period.

Figure 8 depicts the MD trajectory data for ligand AG-690/11203374_2. The RMSD plot (Figure 8(A)) indicated a stable ligand-protein complex throughout the entire simulation period which was ascertained from RMSD values ranging from 1 to 2.1 Å for protein and 0.37 to 2.1 Å for the ligand. The flexibility of the protein structure was evaluated throughout the simulation by calculating the RMSF of individual amino acid residues of the protein (Figure 8(B)). The entire protein system had an RMSF value between 0.3 to 1.7 Å except for the terminal amino acid residues which showed moderate movement throughout the 50 ns simulation. The interaction between ligand and individual amino acid residues inside the binding pocket of M_{pro} is quantified and represented in Figure 8(C). It can be seen from the plot that the ligand interacted with the M_{pro} active site with THR_26, HIS_41, ASN_142, GLY_143, and GLU_166 mainly through hydrogen bonding. Hydrophobic interactions with residues HIS_41 and MET_165 and water bridges with residues SER_144 and CYS_145 also played a supportive role in binding the ligand with M_{pro} active site. RMSF plot of ligand

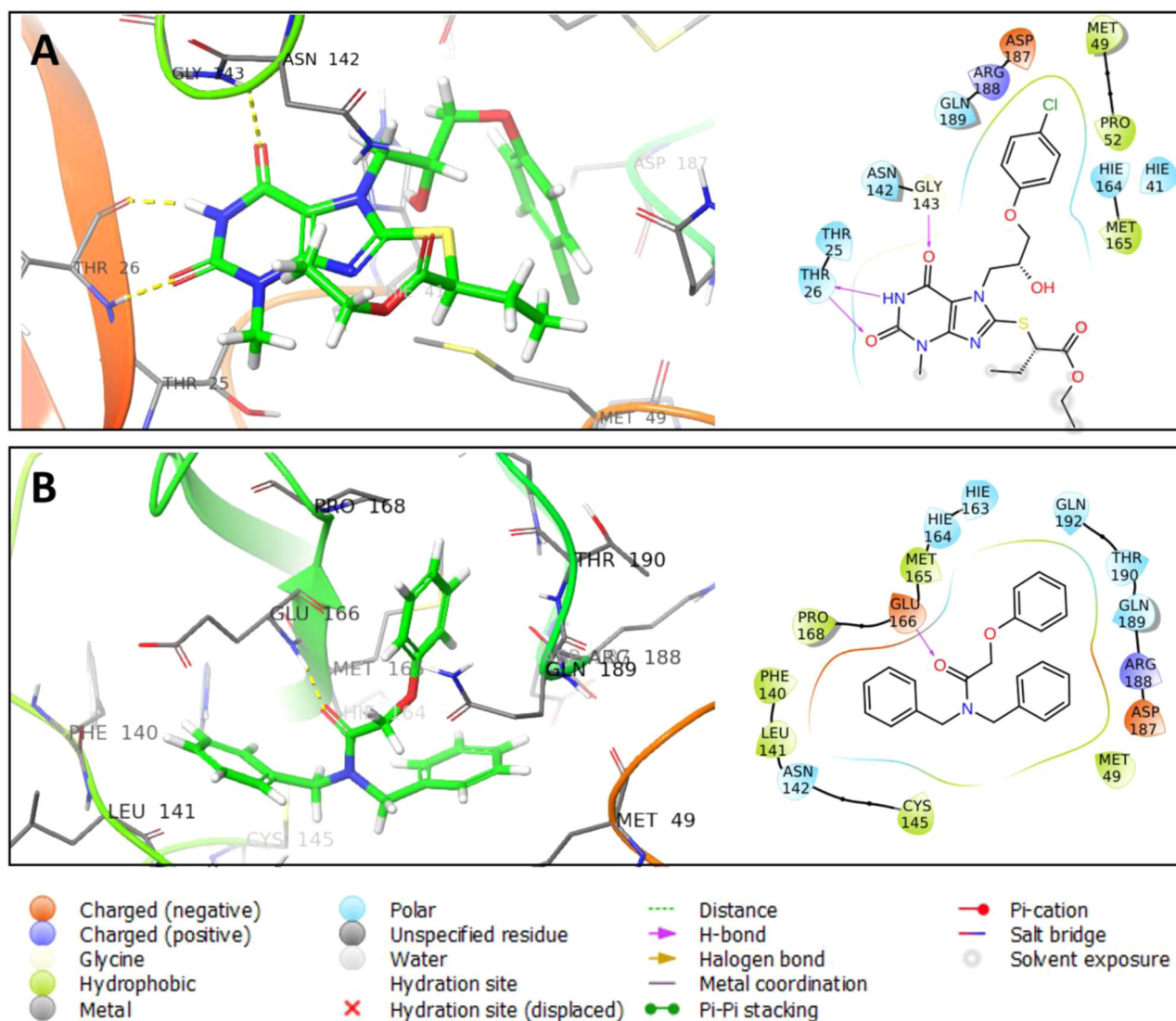


Figure 5. Three-dimensional binding pose and two-dimensional ligand interaction diagram of hits AG-690/11203374_3 (A) and AH-034/04857012 (B).

(Figure 8(D)) suggested that the ligand mostly stayed inside the binding pocket bound to the active site amino acid residues. Figure 8(E) depicts specific interactions between the ligand and active site amino acid residues of M_{pro} . The xanthine core along with the side chain substitutions engaged the active site amino acid residues through various H-bonds and hydrophobic interactions. Finally, a timeline of protein-ligand interactions in the form of hydrogen bonds, hydrophobic interaction, ionic interaction, and water bridges is shown in Figure 8(F) throughout the 50 ns simulation time. The top panel indicates the total number of specific protein-ligand contacts and the bottom panel indicates residue level interactions of the ligand. Overall, the ligand interacted well with the M_{pro} binding pocket and a minimum of 4 contacts were present throughout the simulation period.

Figure 9 depicts the MD trajectory data for ligand AG-690/11203374_3. The RMSD plot (Figure 9(A)) indicated a stable ligand-protein complex throughout the entire simulation period which was ascertained from the RMSD values varying from 0.5 to 2.4 Å for both the protein and ligand. The

flexibility of the protein structure was evaluated throughout the simulation by calculating the RMSF of individual amino acid residues of the protein (Figure 9(B)). The entire protein system had an RMSF value ranged from 0.4 to 2.0 Å except for the terminal amino acid residues which showed moderate movement throughout the 50 ns simulation duration. The interaction between ligand and individual amino acid residues inside the binding pocket of M_{pro} is quantified and represented in Figure 9(C). It was observed that the ligand interacted with the M_{pro} active site with amino acid residues THR_26, ASN_142, GLU_166, ASP_187, GLN_189, THR_190, and GLN_192 mostly through hydrogen bonding, which is one of the major forces involved in binding interactions as well as in stabilizing the protein-ligand complex. Hydrophobic interactions with HIS_41, MET_49 and MET_165 residues and water bridge with GLU_166 residue also played a supportive role in binding the ligand with M_{pro} active site. RMSF plot of ligand (Figure 9(D)) suggested that the ligand mostly stayed inside the binding pocket bound to the active site amino acid residues. Figure 9(E) depicted specific

Table 2. Qikprop calculated ADMET properties of co-crystallized ligand Z31792168 and top twenty-five hits.

S. no.	Title	MW	SASA	FOSA	FISA	PISA	volume	QLogPo/w	QLogS	QLogHERG	QPPCaco	QLogBB	OPPMDCK	logKp	QLogKhsa	PSA
1	Z31792168 (co-crystallized ligand)	218.3	495.37	259.21	66.43	169.73	821.84	2.51	-3.43	-4.66	2322.67	-0.22	1230.09	-1.86	-0.03	47.78
2	AG-690/11060013	438.93	658.20	306.36	136.00	129.63	1238.92	3.40	-4.33	-4.62	508.38	-1.09	706.36	-2.61	0.12	109.72
3	AG-690/11203374_1	496.97	725.49	363.78	160.11	125.73	1382.11	3.26	-4.60	-4.93	300.32	-1.47	351.01	-2.97	0.07	138.33
4	AG-690/11203374_2	496.97	739.19	366.98	166.01	124.38	1400.31	3.34	-4.85	-5.06	264.01	-1.55	329.16	-3.09	0.10	139.87
5	AG-690/11203374_3	496.97	725.35	342.66	169.91	133.24	1382.19	3.21	-4.61	-4.96	242.49	-1.56	291.77	-3.13	0.07	140.61
6	AH-034/04857012	331.41	619.21	75.62	17.88	525.70	1118.61	4.63	-4.27	-5.23	4870.40	-0.03	3868.09	0.68	0.24	30.62
7	AG-690/11060018	432.54	725.73	443.21	147.74	123.56	1342.12	3.63	-4.98	-5.18	393.44	-1.57	207.92	-2.75	0.31	110.32
8	AG-205/05184040	327.40	602.66	76.91	77.69	448.06	1062.47	3.60	-4.38	-6.33	1816.33	-0.54	943.00	-0.80	0.24	55.25
9	AO-365/11349014	355.48	688.47	307.37	62.03	319.07	1253.18	4.13	-3.72	-6.75	637.61	-0.32	336.49	-2.55	0.45	60.58
10	AH-034/11365679	368.43	737.96	106.52	63.65	567.79	1269.47	5.92	-6.87	-7.99	2468.03	-0.70	1313.52	0.17	0.96	48.29
11	AN-329/06315047	354.41	645.01	183.31	77.33	384.37	1149.55	1.81	-1.55	-3.48	610.72	-0.56	951.07	-1.02	-0.95	71.33
12	AO-365/11349014	355.48	704.63	329.39	57.43	317.82	1266.74	4.24	-4.00	-6.95	705.07	-0.29	375.13	-2.47	0.48	64.68
13	AN-329/09986025	285.30	552.31	207.88	117.31	227.12	938.79	2.54	-3.59	-5.21	764.58	-0.91	370.12	-2.31	-0.04	78.44
14	AK-968/11492131	321.38	593.14	165.05	109.93	318.17	1064.18	3.70	-4.32	-5.48	898.32	-0.87	440.57	-1.75	0.35	85.24
15	AE-641/00653027	276.34	531.15	216.15	155.11	159.89	926.39	-1.25	1.08	-2.45	24.60	-0.72	37.41	-5.40	-1.29	95.45
16	AE-848/11243033	347.47	638.94	417.68	76.09	128.45	1122.72	3.29	-3.50	-5.10	74.88	-0.46	1209.23	-1.99	0.18	62.73
17	AG-205/08396013	299.30	521.20	83.69	99.52	291.34	914.91	2.78	-3.16	-4.95	1127.58	-0.43	1014.48	-1.94	-0.17	71.55
18	AE-848/11243033	347.47	609.47	386.39	103.72	102.00	1096.79	2.93	-2.99	-4.54	40.96	-0.67	634.99	-2.59	0.13	65.20
19	AE-641/00004064	300.31	580.89	27.76	154.40	398.73	979.93	2.06	-3.49	-6.32	340.17	-1.38	154.23	-2.29	-0.24	109.94
20	AO-854/10413043	283.41	531.62	346.91	23.56	161.15	965.61	3.69	-3.80	-2.73	4236.29	0.12	3382.81	-1.00	0.28	31.16
21	AK-778/10920048	297.31	534.06	111.85	132.23	289.99	926.93	2.45	-3.44	-5.17	552.08	-0.93	260.31	-2.46	-0.02	92.31
22	AN-652/11380002	297.40	601.23	277.14	88.81	235.28	1058.29	4.18	-4.89	-5.32	1424.55	-0.65	725.20	-1.75	0.53	55.96
23	AG-690/08354009	323.39	601.69	220.72	134.00	246.97	1060.23	3.52	-4.88	-5.37	531.08	-1.02	249.62	-2.64	0.47	83.88
24	AG-690/10772011	338.38	549.84	33.12	176.94	339.78	977.69	1.60	-2.98	-5.31	207.98	-1.25	90.62	-3.20	-0.18	93.03
25	AN-652/11380002	297.40	588.04	262.38	91.15	234.51	1040.95	4.12	-4.66	-5.16	1353.80	-0.65	686.35	-1.80	0.49	56.03
26	AG-205/11132201	359.23	590.49	136.39	112.70	129.38	986.29	4.60	-6.21	-5.20	845.67	-0.36	5986.59	-2.66	0.59	83.88

MW (Molecular weight of the molecule; 130.0–725.0); SASA (Total solvent accessible surface area in square angstroms using a probe with a 1.4 Å radius; 300.0–1000.0); FOSA (Hydrophobic component of the SASA; 0.0–750.0); FISA (Hydrophilic component of the SASA; 7.0–330.0); PISA (π component of the SASA; 0.0–450.0); Volume (total solvent accessible volume: 500–2000); QLogPo/w (Predicted octanol/water partition coefficient: -2 to 6.5); QLogS (Predicted aqueous solubility: -6.5 to 0.5); QLogHERG (Predicted I_{C50} for blockage of HERG K^+ channels; concern < -5); QPPCaco (Predicted apparent Caco-2 cell, model for gut-blood barrier, permeability in nm²/sec: poor if < 25 and great if > 500); QLogBB (Predicted brain/blood partition coefficient: -3 to 1.2); QPPMDCK (Predicted apparent MDCK cell permeability in nm²/sec predicting non-active transport across blood brain barrier; < 25 poor and > 500 great); QLogKp (Predicted skin permeability; -8.0 to -1.0); QLogKhsa (Predicted human serum albumin binding: -1.5 to 1.5); PSA (Van der Waals surface area of polar nitrogen and oxygen atoms and carbonyl carbon atoms; 7.0–200.0).

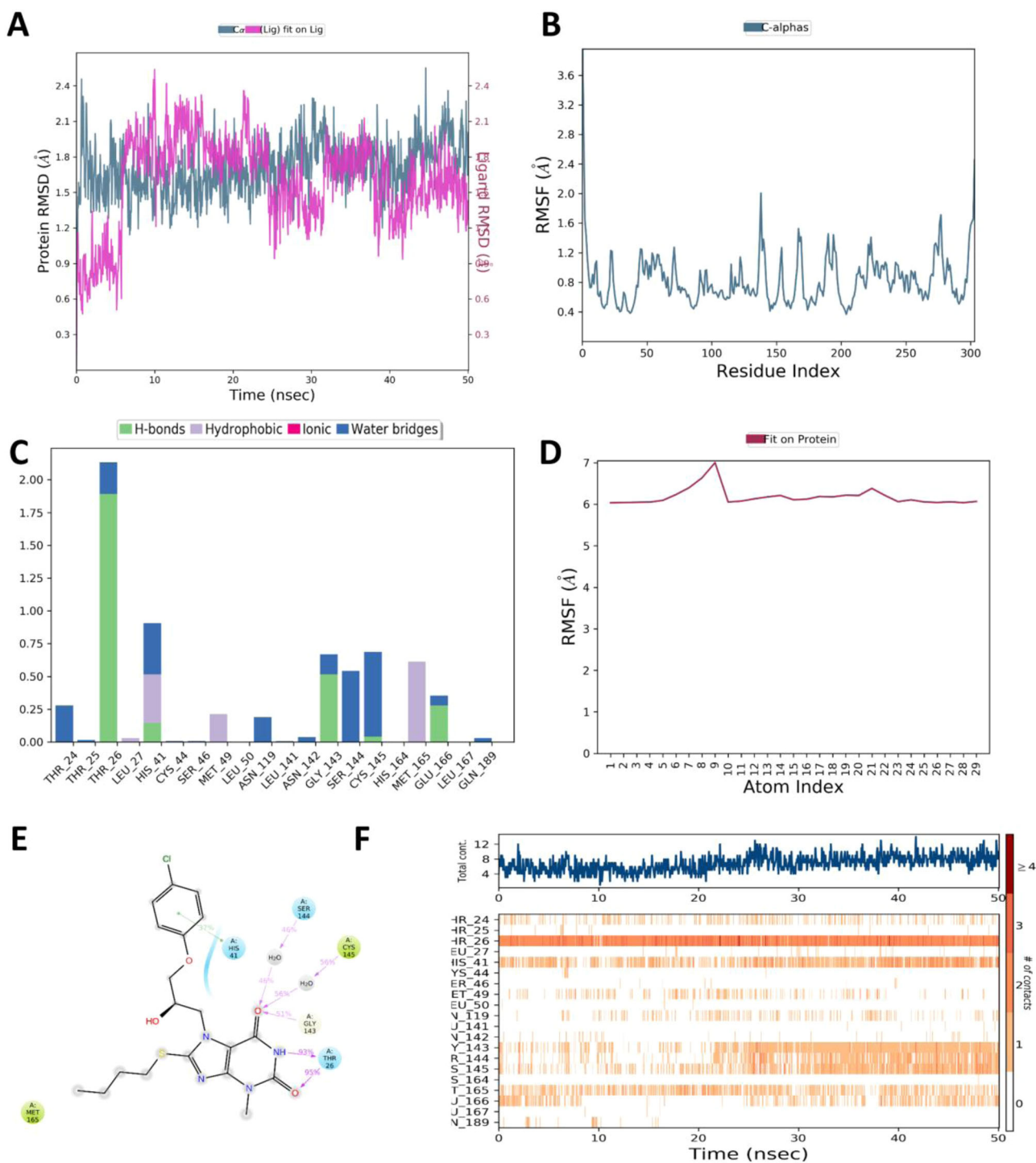


Figure 6. Molecular dynamics trajectory analysis of AG-690/11060013 (A) RMSD of the protein and ligand with respect to the first frame (B) Protein RMSF (C) Protein-ligand contacts histogram (D) Ligand RMSF (E) Ligand-protein contacts (F) Protein-ligand contacts.

interactions between the ligand and active site amino acid residues of M_{pro} . The xanthine core along with the side chain substitutions engaged the active site amino acid residues through various H-bonds and hydrophobic interactions. Finally, a timeline of protein-ligand interactions in the form of hydrogen bonds, hydrophobic interaction, ionic interaction and water bridges is plotted as shown in Figure 9(F) throughout the 50 ns simulation time. The top panel indicates the total number of specific protein-ligand contacts

and the bottom panel indicates residue level interaction of the ligand. Overall, the ligand interacted well with the M_{pro} binding pocket and a minimum of 4 contacts were present throughout the simulation period.

Figure 10 shows all the MD trajectory analysis data for the ligand AH-034/04857012. It was observed from the RMSD plot (Figure 10(A)) that the protein as well ligand remains stable throughout the simulation period and the RMSD value ranged from 0.87 to 2.12 Å for both the protein and ligand

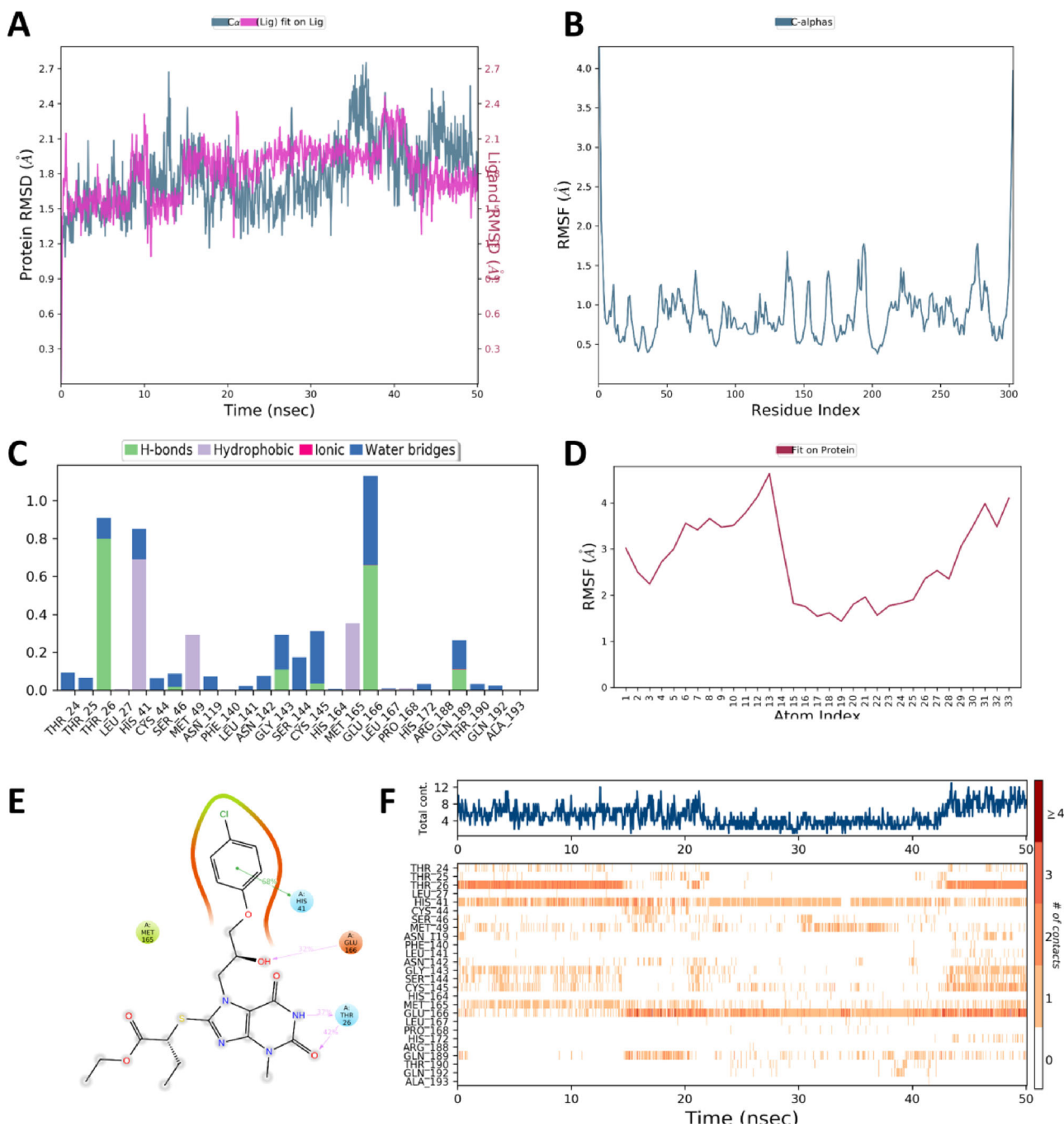


Figure 7. Molecular dynamics trajectory analysis of AG-690/11203374_1 (A) RMSD of the protein and ligand with respect to the first frame (B) Protein RMSF (C) Protein-ligand contacts histogram (D) Ligand RMSF (E) Ligand-protein contacts (F) Protein-ligand contacts.

indicating a stable complex between them. Furthermore, the flexibility of the protein structure was evaluated throughout the simulation by calculating the RMSF of individual amino acid residues of the protein (Figure 10(B)). The entire protein had an RMSF value ranging from 0.4 to 2.0 Å except for the terminal amino acids, which showed moderate movement throughout the 50 ns long simulation. The binding interactions of the ligand with active site amino acid residues inside the binding pocket of M_{pro} were quantified and represented in Figure 10(C). It was observed that the ligand interacted with the M_{pro} active site with GLU_166 and GLN_189 through hydrogen bonding and water bridge. Hydrophobic

interactions of the ligand with amino acid residues LEU_27, HIS_41, MET_49, MET_165, LEU_167, and PRO_168 played a major role in binding interactions with M_{pro} active site. RMSF plot of ligand (Figure 10(D)) indicated that the ligand largely stayed bound to the active site amino acids. Figure 10(E) shows specific interactions of the ligand with active site amino acid residues of M_{pro} . It can be observed that N-benzyl moiety and the carbonyl moiety present on phenoxyethyl side chain was engaged with the active site amino acid residues. Finally, a timeline of the protein-ligand interactions is plotted (Figure 10(F)) in the form of hydrogen-bonds, hydrophobic interaction, ionic interaction and water bridges

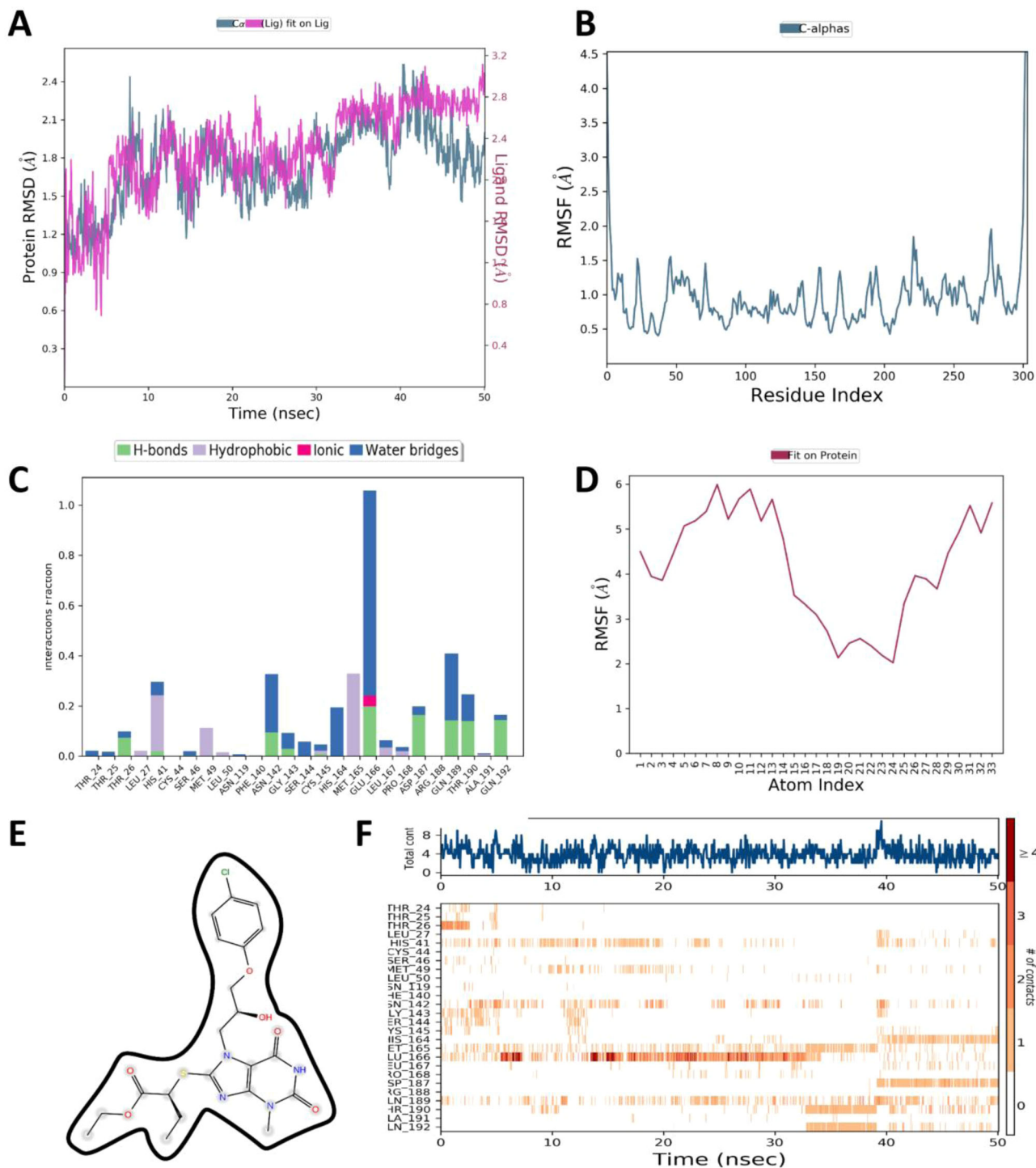


Figure 9. Molecular dynamics trajectory analysis of AG-690/11203374_3 (A) RMSD of the protein and ligand with respect to the first frame (B) Protein RMSF (C) Protein-ligand contacts histogram (D) Ligand RMSF (E) Ligand-protein contacts (F) Protein-ligand contacts.

4. Conclusions

Structure-based virtual screening based on molecular docking technique was utilized to identify twenty-five potential hit compounds against SARS-CoV-2 M_{pro} from commercially available small molecule dataset. Extra precision molecular docking analysis revealed that the top hits compounds were involved in moderate to strong protein-ligand interaction with the binding site residues of M_{pro} indicated

from their good XP-score. The dominant attractive forces for ligand binding inside the active site cavity were hydrogen bonding and hydrophobic interactions. In silico ADME studies revealed that the hit compounds were having optimal pharmacokinetic properties. Additionally, top-five hit compounds were subjected to molecular dynamics simulation studies to examine the stability and integrity of their complex with the enzyme, M_{pro} . The results from MD

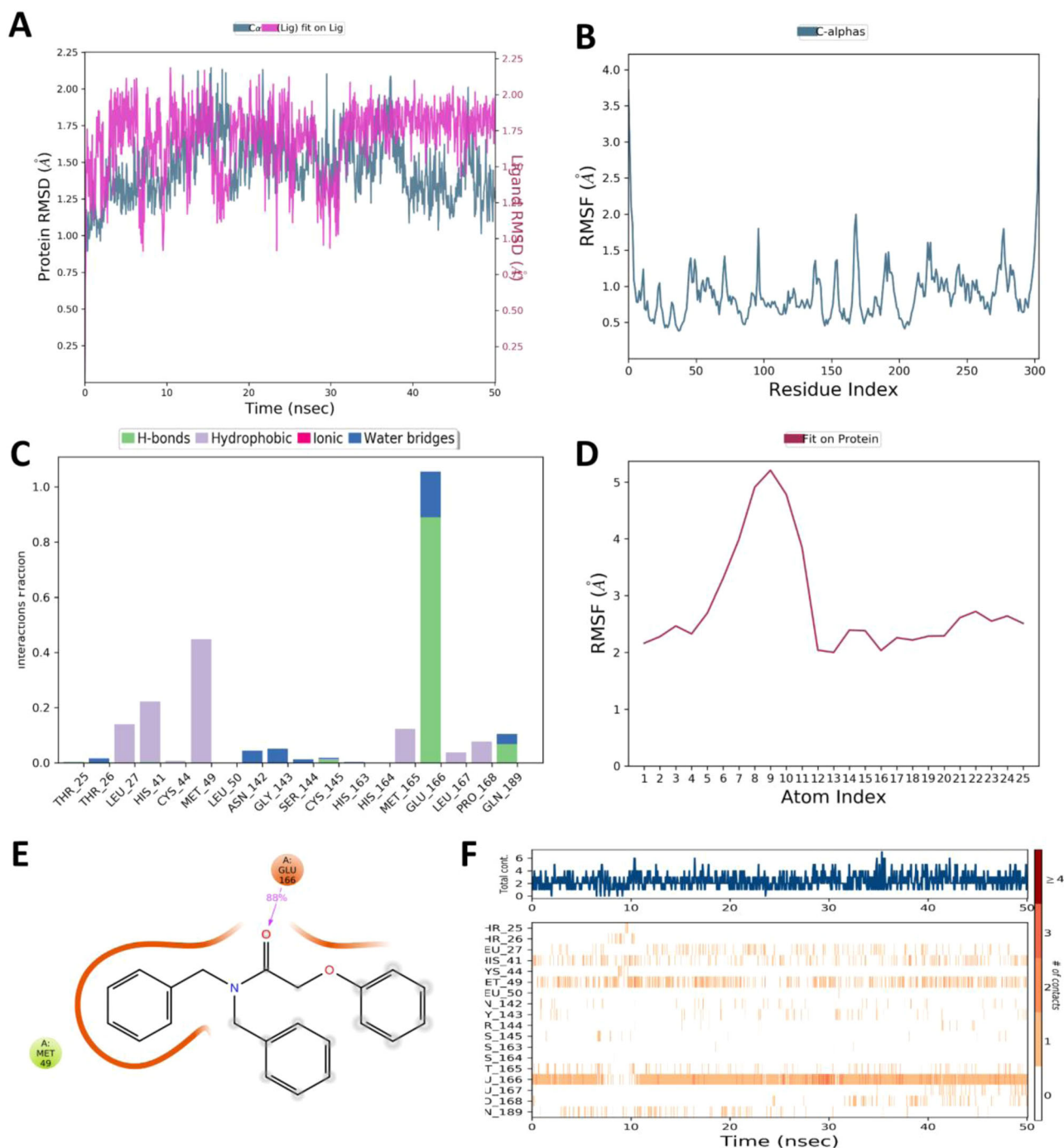


Figure 10. Molecular dynamics trajectory analysis of AH-034/04857012 (A) RMSD of the protein and ligand with respect to the first frame (B) Protein RMSF (C) Protein-ligand contacts histogram (D) Ligand RMSF (E) Ligand-protein contacts (F) Protein-ligand contacts.

simulation have shown good stability and less conformational change of the receptor-ligand complex during a 50 ns trajectory.

Overall results show that compound AG-690/11203374_1 and AG-690/11203374_2 fits snugly inside the binding cavity and have shown strong interactions with highly conserved residues like His_41, Cys_145, Glu_166. Among the top-ranked hits, compound AG-690/11203374_1 and AG-690/11203374_2 emerged as the best in silico hits based on the docking, MMGBSA, dynamics and ADMET studies. Although the biological efficacy has yet to be determined for the two in silico hits, but it is anticipated that they could serve as

suitable lead molecules for the development of new antiviral agents against this pandemic.

Acknowledgment

RK acknowledges the seed grant obtained from the Indian Institute of Technology (BHU) Varanasi, India.

Disclosure statement

The authors declare no conflict of interest.

Funding

Indian Institute of Technology (BHU)10.13039/501100002742The work is supported by seed grant from Indian Institute of Technology (BHU) awarded to RK.

ORCID

Manoj Kumar Mahapatra  <http://orcid.org/0000-0001-8147-5283>

Rajnish Kumar  <http://orcid.org/0000-0002-7021-6996>

References

- Adeshina, Y. O., Deeds, E. J., & Karanicolas, J. (2020). Machine learning classification can reduce false positives in structure-based virtual screening. *Proceedings of the National Academy of Sciences of the United States of America*, 117(31), 18477–18488. <https://doi.org/10.1073/pnas.2000585117>
- Alavijeh, M. S., & Palmer, A. M. (2004). The pivotal role of drug metabolism and pharmacokinetics in the discovery and development of new medicines. *IDrugs : The Investigational Drugs Journal*, 7(8), 755–763.
- Asinex. (2020). http://www.asinex.com/libraries-html/libraries_gold_platinum-html/
- Baell, J. B., & Holloway, G. A. (2010). New substructure filters for removal of Pan Assay Interference Compounds (PAINS) from screening libraries and for their exclusion in bioassays. *Journal of Medicinal Chemistry*, 53(7), 2719–2740. <https://doi.org/10.1021/jm901137j>
- Bowers, K. J. (2006). Scalable algorithms for molecular dynamics simulations on commodity clusters. SC '06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing (pp. 43–43), November 11–17. Institute of Electrical and Electronics Engineers (IEEE). <https://doi.org/10.1109/SC.2006.54>
- Chen, Y., Liu, Q., & Guo, D. (2020). Emerging coronaviruses: Genome structure, replication, and pathogenesis. *Journal of Medical Virology*, 92(4), 418–423. <https://doi.org/10.1002/jmv.25681>
- Dill, K. A. (1997). Additivity principles in biochemistry. *The Journal of Biological Chemistry*, 272(2), 701–704. <https://doi.org/10.1074/jbc.272.2.701>
- Du, X., Li, Y., Xia, Y.-L., Ai, S.-M., Liang, J., Sang, P., Ji, X.-L., & Liu, S.-Q. (2016). Insights into protein-ligand interactions: Mechanisms, models, and methods. *International Journal of Molecular Sciences*, 17(2), 144. <https://doi.org/10.3390/ijms17020144>
- Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P., & Shenkin, P. S. (2004). Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of Medicinal Chemistry*, 47(7), 1739–1749. <https://doi.org/10.1021/jm0306430>
- Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., Sanschagrin, P. C., & Mainz, D. T. (2006). Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of Medicinal Chemistry*, 49(21), 6177–6196. <https://doi.org/10.1021/jm051256o>
- Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., O'Meara, M. J., Rezelj, V. V., Guo, J. Z., Swaney, D. L., Tummino, T. A., Hüttenhain, R., Kaake, R. M., Richards, A. L., Tutuncuoglu, B., Foussard, H., Batra, J., Haas, K., Modak, M., ... Krogan, N. J. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*, 583(7816), 459–468. <https://doi.org/10.1038/s41586-020-2286-9>
- Goyal, B., & Goyal, D. (2020). Targeting the dimerization of the main protease of coronaviruses: A potential broad-spectrum therapeutic strategy. *ACS Combinatorial Science*, 22(6), 297–305. <https://doi.org/10.1021/acscombsci.0c00058>
- Guo, Y.-R., Cao, Q.-D., Hong, Z.-S., Tan, Y.-Y., Chen, S.-D., Jin, H.-J., Tan, K.-S., Wang, D.-Y., & Yan, Y. (2020). The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak - An update on the status. *Military Medical Research*, 7(1), 11. <https://doi.org/10.1186/s40779-020-00240-0>
- <https://www.rcsb.org/structure/5R84>. (2020).
- Jacobson, M. P., Friesner, R. A., Xiang, Z., & Honig, B. (2002). On the role of the crystal environment in determining protein side-chain conformations. *Journal of Molecular Biology*, 320(3), 597–608. [https://doi.org/10.1016/s0022-2836\(02\)00470-9](https://doi.org/10.1016/s0022-2836(02)00470-9)
- Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., Zhang, B., Li, X., Zhang, L., Peng, C., Duan, Y., Yu, J., Wang, L., Yang, K., Liu, F., Jiang, R., Yang, X., You, T., Liu, X., ... Yang, H. (2020). Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature*, 582(7811), 289–293. <https://doi.org/10.1038/s41586-020-2223-y>
- Jorgensen, W. L., & Tirado-Rives, J. (1988). The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, 110(6), 1657–1666. <https://doi.org/10.1021/ja00214a001>
- Kim, Y., Liu, H., Galasiti Kankanamalage, A. C., Weerasekara, S., Hua, D. H., Groutas, W. C., Chang, K.-O., & Pedersen, N. C. (2016). Reversal of the progression of fatal coronavirus infection in cats by a broad-spectrum coronavirus protease inhibitor. *PLoS Pathogens*, 12(3), e1005531. <https://doi.org/10.1371/journal.ppat.1005531>
- Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, 23(1–3), 3–25. [https://doi.org/10.1016/S0169-409X\(96\)00423-1](https://doi.org/10.1016/S0169-409X(96)00423-1)
- Liu, C., Zhou, Q., Li, Y., Garner, L. V., Watkins, S. P., Carter, L. J., Smoot, J., Gregg, A. C., Daniels, A. D., Jervey, S., & Albaiu, D. (2020). Research and development on therapeutic agents and vaccines for COVID-19 and related human coronavirus diseases. *ACS Central Science*, 6(3), 315–331. <https://doi.org/10.1021/acscentsci.0c00272>
- Lyne, P. D., Lamb, M. L., & Saeh, J. C. (2006). Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *Journal of Medicinal Chemistry*, 49(16), 4805–4808. <https://doi.org/10.1021/jm060522a>
- Lyu, J., Wang, S., Balius, T. E., Singh, I., Levit, A., Moroz, Y. S., O'Meara, M. J., Che, T., Alga, E., Tolmacheva, K., Tolmachev, A. A., Shoichet, B. K., Roth, B. L., & Irwin, J. J. (2019). Ultra-large library docking for discovering new chemotypes. *Nature*, 566(7743), 224–229. <https://doi.org/10.1038/s41586-019-0917-9>
- Main protease structure and XChem fragment screen. (2020). <https://www.diamond.ac.uk/covid-19/for-scientists/Main-protease-structure-and-XChem.html>
- Martyna, G. J., Tobias, D. J., & Klein, M. L. (1994). Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics*, 101(5), 4177–4189. <https://doi.org/10.1063/1.467468>
- Ren, Z., Yan, L., Zhang, N., Guo, Y., Yang, C., Lou, Z., & Rao, Z. (2013). The newly emerged SARS-like coronavirus HCoV-EMC also has an "Achilles' heel": Current effective inhibitor targeting a 3C-like protease. *Protein & Cell*, 4(4), 248–250. <https://doi.org/10.1007/s13238-013-2841-3>
- Sanders, J. M., Monogue, M. L., Jodlowski, T. Z., & Cutrell, J. B. (2020). Pharmacologic treatments for coronavirus disease 2019 (COVID-19): A review. *JAMA*, 323(18), 1824–1836. <https://doi.org/10.1001/jama.2020.6019>
- Schrödinger Release 2020-3. (2020). *Desmond Molecular Dynamics System*, D. E. Shaw Research, New York, NY, 2020. *Maestro-Desmond Interoperability Tools*. Schrödinger.
- Schrödinger Release 2020-3. (2020). *LigPrep*. Schrödinger, LLC.
- Schrödinger Release 2020-3. (2020). *Maestro*. Schrödinger, LLC.
- Schrödinger Release 2020-3. (2020). *Prime*. Schrödinger, LLC.
- Schrödinger Release 2020-3. (2020). *QikProp*. Schrödinger, LLC.
- Ullrich, S., & Nitsche, C. (2020). The SARS-CoV-2 main protease as drug target. *Bioorganic & Medicinal Chemistry Letters*, 30(17), 127377. <https://doi.org/10.1016/j.bmcl.2020.127377>
- Weekly epidemiological update. (2020). https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200928-weekly-epi-update.pdf?sfvrsn=9e354665_6

- Yadav, A., Kumar, R., Sunkaria, A., Singhal, N., Kumar, M., & Sandhir, R. (2016). Evaluation of potential flavonoid inhibitors of glyoxalase-I based on virtual screening and in vitro studies. *Journal of Biomolecular Structure and Dynamics*, 34(5), 993–1007. <https://doi.org/10.1080/07391102.2015.1064830>
- Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., Becker, S., Rox, K., & Hilgenfeld, R. (2020). Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science (New York, N.Y.)*, 368(6489), 409–412. <https://doi.org/10.1126/science.abb3405>