▍ Medicinal Chemistry & Drug Discovery

Identification of SARS-CoV-2 Main Protease Inhibitors Using Structure Based Virtual Screening and Molecular Dynamics Simulation of DrugBank Database

Pradip Debnath,*[a] Samhita Bhaumik,[b] Debanjan Sen,[c] Ravi K. Muttineni,[d] and Sudhan Debnath*[a]

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) is highly pathogenic to humans and has created an unprecedented global health care threat. Globally, intense efforts are going on to discover a vaccine or new drug molecules to control the COVID-19. However, till today, there is no effective therapeutics or treatment available for COVID-19. In this study, we aim to find out potential small molecule inhibitors for SARS-CoV-2 main protease (M^{pro}) from the known DrugBank database version 5.1.8. We applied structure-based virtual screening of the database containing 11875 numbers of drug candidates to identify potential hits for SARS-CoV-2 M^{pro} inhibitors. Seven potential inhibitors having admirable XP glide score ranging from −15.071 to −8.704 kcal/mol and good binding affinity with the active sites amino acids of M^{pro} were identified. The selected hits were further analyzed with 50 ns molecular dynamics (MD) simulation to examine the stability of protein-ligand complexes. The root mean square deviation and potential energy plot indicates the stability of the complexes during the 50 ns MD simulation. The MM-GBSA analysis also showed good binding energy of the selected hits (−83.2718 to −58.6618 kcal/mol). Further analysis revealed critical hydrogen bonds and hydrophobic interactions between compounds and the target protein. The compounds bind to biologically important regions of M^{pro}, indicating their potential to inhibit the functionality of this component.

1. Introduction

The present pandemic situation due to COVID-19 caused by SARS-CoV-2 is a huge challenge and burning issue to the research community, health workers, and government officials worldwide. Due to the high person-to-person dissemination capacity of SARS-CoV-2, the disease adversely affected 114.8 million people with 2.5 million deaths in 219 countries and territories around the world (https://www.worldometers.info/coronavirus/). Globally, coronavirus disease 2019 (COVID-19) has created a health disaster and economic crisis.[1,2] The disease forcibly affects economic growth worldwide and pushes millions of people towards unemployment. The SARS-CoV-2 was initially detected in late December 2019 in Wuhan City, Hubei Province, and caused an atypical pneumonia outbreak.[3–7] The type of pneumonia caused by SARS-CoV-2 is a highly infectious disease and has rapidly spread worldwide.[8] Observing the rapid spread of the virus worldwide, the World Health Organization (WHO) declared COVID-19 as a global public health emergency of international concern,[9] and it was announced a pandemic in March 2020. This urgent situation is pressing the world scientific community to respond with the development of a novel vaccine or small molecule therapeutics for SARS-CoV-2.[10–15] While drug repurposing[16–18] may be a short-term and non-specific solution to treat COVID-19 patients, the development of more targeted inhibitors is highly desirable.

SARS-CoV-2 is an envelope, single-strand, positive-sense RNA virus of approximately 30kb size, belonging to β-coronavirus under the sub-family *Coronavirinae* and family *Coronaviridae* of order *Nidovirales*.[19] It typically affects the respiratory tract of humans and leads to mild/severe respiratory tract infections.[20] In the last two decades, two highly pathogenic human coronaviruses (HCoVs), including Severe Acute Respiratory Syndrome (SARS)[21] in the year 2002 and Middle East Respiratory Syndrome (MERS)[22] in 2012, emerging from animal reservoirs, have also led to global epidemics with morbidity and mortality.

Potential therapies against coronaviruses can be divided into two categories depending on the target, one is acting on the human immune system and the other is on coronavirus itself.[10–15] Direct action on the virus includes inhibition of viral genome replication, blocking the critical enzymes essential for the replication cycle, and inhibiting the virus entry by blocking virus receptor-human cell complex formation or inhibiting the

[a] *Dr. P. Debnath, Dr. S. Debnath*
     Department of Chemistry, Maharaja Bir Bikram College, Agartala, Tripura, 799004, India
     E-mail: pradipchem78@gmail.com
            bcsdebnath@gmail.com
[b] *Dr. S. Bhaumik*
     Department of Chemistry, Women's College, Agartala, Tripura, 799001, India
[c] *Dr. D. Sen*
     BCDA College of Pharmacy & Technology, Jessore Road South, Hridaypur, Kolkata, West Bengal, 700127, India
[d] *Dr. R. K. Muttineni*
     Immunocure Discovery Solution Pvt. ltd. IKP, Hyderabad, 500078, India

ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry Europe
European Chemical
Societies Publishing

virus self-assembly process through the blocking of structural proteins.[23] Human CoVs genome has several conserved structural proteins such as- Spike (S) glycoprotein, envelope (E) protein, membrane (M) protein, and nucleocapsid (N) protein and at least four non-structural proteins (nsPs) such as- 3-chymotrypsin-like protease (3CL[pro]) also known as Main protease (M[pro]), papain-like protease (PL[pro]), helicase, and RNA-dependent RNA polymerase (RdRp). Protein sequence alignment analyses of SARS-CoV-2 indicated that catalytic sites of the four SARS-CoV-2 enzymes that could represent antiviral targets are highly conserved and shows a total of 79.9% nucleotide sequence identity with SARS-CoV.[24] Spike protein plays a crucial role in host cell attachment, accelerates cellular entry through the host cell receptor, and plays a vital role in viral pathogenesis. The non-structural proteins, 3CL[pro] and PL[pro], that are two important proteases, play a crucial role in the viral replication process through the extensive proteolysis of two replicase polyproteins, pp1a and pp1ab into 16 non-structural proteins (nsP1-nsP16).[25] These nonstructural proteins are assembled and form the replication-transcription complex which regulates the numerous functions of virus replication viz. replication of the viral genome, sub-genomic RNA processing, and packaging of the new virion.[26] Interrupting any replication process would become a potential molecular target to develop therapeutics against coronavirus. The non-structural proteins of SARS-CoV-2 such as- M[pro], PL[pro], and RdRp are the most striking targets for therapeutic agents and their X-ray crystallography structures are also available in the literature.[27–30] Among them, M[pro] is highly sensitive; therefore, it is recognized as a potential therapeutic target to develop antiviral agents against SARS-CoV-2.[31] This main protease enzyme of SARS-CoV-2 accelerates the proteolysis reaction and provides an extra domain to stabilize the dimer genome of the virus. Most of the earlier efforts to target SARS-CoV-2 resulted in the identification of several main protease (M[pro]) inhibitors targeting the catalytic dyad of the protein defined by His41 and Cys145 residues.[32–34] Thus, the inhibition of the function of this protein may reduce the rapid multiplication of this virus and can be a potentially effective target to combat COVID-19. So, the molecule has to be identified to target the active site of the main protease enzyme to stop the effective multiplication of this virus.

It is a well-known fact that drug discovery is a time-demanding process. On an average, the discovery of a new drug from initial findings to its marketing stage, takes months to years. For this reason, several research groups have paid their attention to drug repurposing strategies to find out the potential inhibitors for SARS-CoV-2 virus replication.[10–18,35–40] In this study, we aim to screen the DrugBank database (www.drugbank.ca, release version 5.1.8) using the computational approach with an intention to identify small molecules that exhibit inhibitory properties against the SARS-CoV-2 main protease enzyme to fasten the drug discovery process. So, this work has been designed to address the problem. To identify potential inhibitors, a structure-based drug design approach like virtual screening (VS) of DrugBank database-5.1.8 containing 11875 numbers molecules was performed using Glide.[41–43] The top-scoring compounds with the highest binding affinity

were subjected to 50 ns Molecular Dynamics (MD) simulations to determine the stability of the complex system. The results of the in silico binding studies revealed that the selected hits have a better binding affinity towards the active sites of the target protein (M[pro]) than the co-ligands (N3 and X77). Moreover, most of the inhibitors reported here are commercially available and therefore, easily accessible for further characterization and development by the scientific community.

2. Materials & methods

2.1. Collection of materials

Five different X-ray crystal structures of SARS-CoV-2 M[pro] in complex with respective co-ligand such as- 6LU7 (N3, 2.16 Å) [27a], 6 W79 (X77, 1.46 Å) [27b], 7 C6S (U5G, 1.60 Å) [27c], 7JQ4 (MPI7, 1.65 Å) [27d], 7JPZ (MPI1, 1.60 Å) [27e] were retrieved from RCSB Protein Data Bank (www.rcsb.org). The DrugBank database (www.drugbank.ca, release version 5.1.8) containing 11875 numbers molecules has been utilized for virtual screening (VS) and identification of potential SARS-CoV-2 main protease inhibitors. The database was prepared by using the Ligprep module of Schrodinger.[44–46] The molecular docking and dynamics part were conducted in AMD theadrepper 24 core cpu supported by 64gb ram integrated with nvidia tesla k20 gpu running over Ubuntu OS. Calculation of ki values and binding energy were performed using Autodock using Windows 10, OS architecture 64-bit, Core (TM) 2 Due CPU machine.

2.2. Protein preparation and grid generation

The X-ray crystal structure of M[pro] of SARS-CoV-2 was prepared using the 'Protein Preparation Wizard' in Maestro 12.2.[47–49] The prepared protein structures were used for the grid generation by using the 'Receptor Grid Generation' panel. During the protein preparation process, the protein structure bond order was assigned using the chemical compound database (CCD). In this process, the hydrogens were added, bond order with metals were assumed zero. Then, the missing side chains and loops were filled by using Prime.[50–52] Finally water molecules were removed (beyond 5.0 Å). The energy of the protein-coligand complex was minimized until the RMSD between the minimized structure and the starting structure reached to 0.3 Å, using the OPLS5 force field. The receptor grid box of 15 Å dimension from the center of each selected co-crystallized ligand of 6LU7, 6W79, 7C6S, 7JQ4, and 7JPZ was generated by selecting their co-ligands within the active site. For predicting K_i values using Auto dock 4.2,[53] the protein 6LU7 was imported in Autodock Tools 4, then water molecules as well as hetero atoms were removed, and polar hydrogens were added followed by computing Gasteiger and addition of Kollman charge. Finally, the system was saved in pdbqt format. The grid dimension of main protease was fixed by selecting active site amino acid residue information (41, 49, 140, 141, 143, 145, 163, 166, 189) and the grid center ($x = -10.88$, $y = 13.94$, $z = 68.21$) and grid size ($x = 58$, $y = 68$, $z = 70$) were fixed.

ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry Europe
European Chemical
Societies Publishing

2.3. Validation of docking protocol

The docking protocol was validated by re-docking the co-crystal ligand against the respective target protein. For this purpose, the co-crystallized ligand of 6LU7 (N3), 6 W79 (X77), 7 C6S (U5G), 7JQ4 (MPI7) and 7JPZ (MPI1)_were split from receptors and subjected to Extra Precision (XP) Glide docking with the corresponding receptor grid of protein. The Root Mean Square Deviation (RMSD) was measured by superimposing the best docked structural pose of co-ligand on its original X-ray crystallographic bound conformation. The RMSD value is used to measure the quality of reproduction of a crystallographic binding pose and docking accuracy of the docking method. The lower value of RMSD the better is the reproducibility of docking with respect to the original crystallographic bound conformation. The maximum acceptable range of RMSD value is < 3.0 Å.

2.4. Virtual screening (VS) of DrugBank database

The structure based VS of the DrugBank database was performed using the virtual screening workflow of Glide. In the VS workflow initially, the DrugBank molecules were filtered by Lipinski's rule of five followed by elimination of ligands having reactive functional groups and the removal of the duplicates using QikProp.[54–56] QikProp predicts the properties and descriptors of hits by comparing with 95% known drug molecules. After QikProp filtration, the ligands were prepared, and then removed the duplicate followed by the removal of high energy ionization/tautomeric. After all filtration, the resultant hits were subjected to structure-based virtual screening to identify molecules having more binding affinity towards SARS-CoV-2 main protease (6LU7). The VS workflow contains multi-stage molecular docking options like high-throughput virtual screening (HTVS), standard precision (SP), and extra precision (XP). In each stage, molecules with a higher binding affinity are forwarded to the next level. In the VS workflow HTVS, a total 10% of the output molecules were subjected to Glide SP and then 20% of total output molecules of SP were subjected for Glide XP. Direct docking was also performed without QikProp filtration and selected one top hit. The top hits resulted from VS were further subjected to rigorous molecular docking with the other four crystal structures of the main protease of SARS-CoV-2 to eliminate false positives.

2.5. Molecular Dynamic (MD) simulations

The 50 ns atomistic molecular dynamics simulation of the selected hits with 6LU7 protein of main protease (6LU7) was performed using Desmond to analyze the protein-ligand complex stability. The Desmond v3.3 was used to execute the molecular dynamics simulation for evaluating the stability of 6LU7-ligand docking complexes.[57] The molecular dynamics simulations were performed for seven selected ligand-6LU7 complexes, and also for the 6LU7-N3 and 6LU7-X77 complex systems. The complexes were prepared for MD simulation through solvation within a water-filled 3D cube of 1 Å spacing

using the TIP3P water model, a three-point model of water. The grid box was neutralized by adding appropriate cations (Na^+) and anions (Cl^-) ions using a salt concentration of 0.15 mol/L. The energy of the prepared system was minimized for a convergence threshold of 1.0 kcal/ mol/Å by using the steepest descent method. The NPT (Number of particles, Pressure, and Temperature) ensemble is applied for minimization and relaxation of the system.[58] During the simulation, the temperature was set at 300 K and maintained by Nose-Hoover thermostat, and the pressure was maintained at 1 atm by Martyna-Tobias Klein pressure bath. Finally, the default relaxation protocol integrated in the Desmond was applied for the relaxation of the system. After 2 ns equilibration, the docked complexes were subjected to production MD run for 50 ns in NPT ensemble. The root mean square deviation (RMSD) and root mean square fluctuation (RMSF) of ligand-protein complex backbone were plotted and the stability of the ligand-protein complex was determined. The analysis of RMSD, RMSF, and hydrogen bonding interactions were performed using the trajectory analysis incorporated in the Desmond.[59]

2.6. Binding energy calculation

The binding free energy (ΔG_{bind}) of each ligand towards the SARS-CoV-2 main protease was calculated from 50 ns MD trajectory using MM-GBSA approach through script program[60] to investigate the relative binding affinity. Snapshots from the 50 ns molecular dynamics trajectory with 10 ns interval were collected and used as an input for the g_mmpbsa program. The lower the value of binding energy (higher negative value) of a ligand indicates the higher binding affinity of that ligand towards the target protein.

3. Results and discussion

3.1. Protein selection and docking validation

Validation of docking protocol is a crucial step ahead of performing docking-based virtual screening of the DrugBank database. For this purpose, the co-crystallized ligands of 6LU7, 6 W79, 7 C6S, 7JQ4, and 7JPZ were split from receptor-coligand complexes and the co-ligands were again subjected to Extra Precision (XP) Glide docking with the corresponding receptor grids and calculated the Root Mean Square Deviation (RMSD) of each coligand. The RMSD between the original crystallographic conformation of coligand and the best-docked conformation of 6LU7, 6W79, 7C6S, 7JQ and 7JPZ were 1.47 Å,[52] 0.3956 Å, 0.6252 Å, 2.1393 Å, and 0.6252 Å, respectively (Figure 1). These lower RMSD values indicate that the Glide docking protocol can reproduce a similar docking pose of the ligand with respect to its original crystallographic bound conformation.

3.2. Structure-based VS and molecular docking analysis

The structure-based virtual screening (VS) of the DrugBank database with 6LU7 of SARS-CoV-2 was performed using Glide.
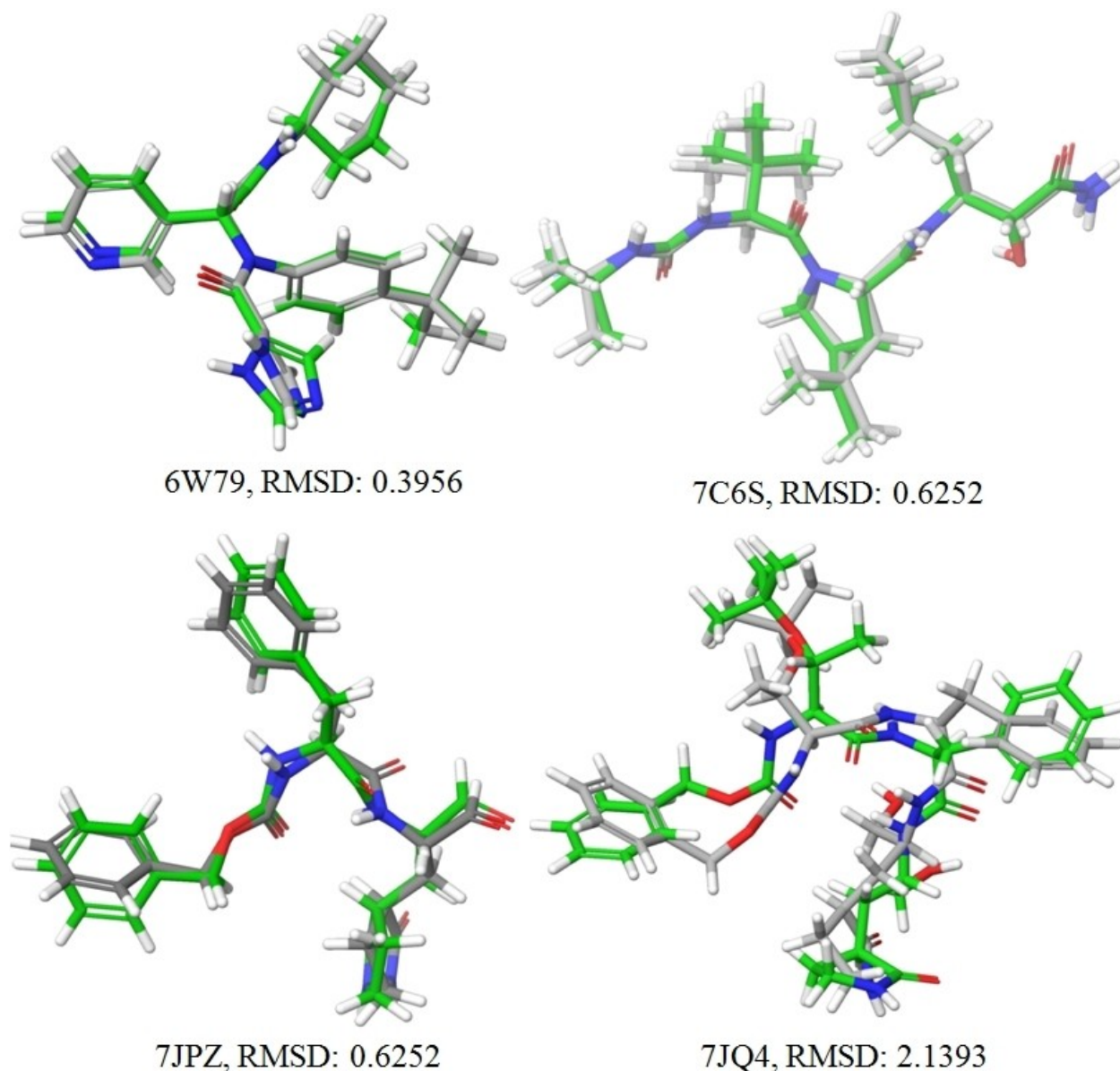
ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

**Chemistry
Europe**
European Chemical
Societies Publishing

Figure 1. Figure illustrates superposition of docked co-ligand (grey) on its originally bound X-ray crystallographic ligand (green) of PDB ID: 6W79, 7C6S, 7JQ4, and 7JPZ.

In each stage of VS (HTVS to SP to XP) molecules with higher binding affinity are forwarded to the next level. Finally, seven best hits having XP glide score ranges from -15.071 to −8.704 kcal/mol were selected. The details of selected hits and co-ligand such as- DrugBank database ID, XP Glide score, and interacting amino acid residues along with distance are shown in Table 1.

The ranges of XP glide score of selected hits are −15.071 to −8.704 kcal/mol. The lower value of glide score i.e. the higher negative value indicates the higher binding affinity of the ligand towards receptor. The 2D and 3D interactions of the seven hits with different active site amino acid residues of 6LU7 are shown in Figure 2 and Figure 3, respectively. The result showed that selected hits have good to excellent glide scores

and exhibit good binding affinity towards the active site of the protein. In the SARS-CoV-2 main protease 6LU7, the co-ligand N3 was covalently bonded with CYS-145 residue of the protease active site. The XP glide score of co-ligand N3 of 6LU7 (PDB ID: 6LU7) is −7.4 kcal/mol and the active site interacting amino acid residues are PHE-140 (π-π), GLY-143 (2.26 Å), CYS-145 (1.98 Å), HIS-163 (2.04 Å), GLU-166 (1.89 Å, 2.02 Å), GLN-189 (1.90 Å, 2.13 Å). The 2D, 3D interactions and binding pose of co-ligand (N3) with the active site of amino acid residues of 6LU7 are shown in Figure 4a. The XP glide score of non-covalently bonded co-ligand X77 of 6 W79 (PDB ID: 6 W79) with 6LU7 was −7.343 kcal/mol. The active site interacting amino acid residue was GLU-166 (Figure 4b).

ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry Europe
European Chemical
Societies Publishing

Table 1. The DrugBank database ID, XP Glide score, and interacting amino acid residues along with distance (Å) of six hits in the active site.

Hit Reg. ID	XP-GS (kcal/mol)	Interacting active site amino acid residues and their interacting distance (Å)	Common interacting amino acid residues with Co-ligand
DB03208	−15.071	GLY-143 (2.61 Å), HIS-163 (2.01 Å), HIS-164 (1.90 Å), GLU-166 (1.89 Å) and THR-190 (1.89 Å)	GLY-143, HIS-163 and GLU-166
DB03949	−10.915	HIE-41 (two, π-π), HIS-163 (1.62 Å), HIS-164 (2.05 Å), GLU-166 (1.86 Å), GLN-189 (2.14 Å, 2.40 Å)	HIS-163, GLU-166 and GLN-189
DB08001	−10.158	HIE-41 (π-π), CYS-145 (2.10 Å), GLU-166 (1.71 Å)	CYS-145 and GLU-166
DB08526	−9.303	LEU-141 (2.17 Å), SER-144 (2.14 Å), CYS-145 (2.63 Å),) HIS-164 (2.06 Å), GLU-166 (2.09 Å), GLN-189 (2.03 Å).	CYS-145, GLU-166 and GLN-189
DB02558	−8.941	HIE-41 (π-π), GLY-143 (2.67 Å), HIS-164 (2.28, 2.60 Å), GLU-166 (1.86 Å, 2.02 Å), GLN-189 (2.06 Å).	GLY-143, GLU-166 and GLN-189
DB12332	−8.787	HIE-41 (two, π-π), LEU-141 (2.29), CYS-145 (2.13), GLU-166 (1.98)	CYS-145 and GLU-166
DB02651	−8.704	HIE-41 (two π-π), THR-25 (2.18 Å), SER-46 (1.54 Å), ASN-142 (2.60 Å) and CYS-145 (2.68 Å)	CYS-145
Co-ligand N3	−7.4	PHE-140 (π-π), GLY-143 (2.26 Å), CYS-145 (1.98 Å), HIS-163 (2.04 Å), GLU-166 (1.89 Å, 2.02 Å), GLN-189 (1.90 Å, 2.13 Å)	–
Co-ligand X77	−7.343	GLU-166	–

The compound DB03208 was the best docked with the target protein (6LU7) with XP glide score of −15.071 kcal/mol. It interacts with target protein through five hydrogen bonds with GLY-143 (2.61 Å), HIS-163 (2.01 Å), HIS-164 (1.90 Å), GLU-166 (1.89 Å), and THR-190 (1.89 Å). This hit exhibits three common amino acid interactions (GLY-143, HIS-163, and GLU-166) with co-ligand N3 in the active site and the interacting distances are comparable with co-ligand N3. The binding energy of hit DB03949 was −10.915 kcal/mol. The interacting amino acid residues of DB03949 were HIE-41 (two π-π), HIS-163 (1.62 Å), HIS-164 (2.05 Å), GLU-166 (1.86 Å), and GLN-189 (2.14 Å, 2.40 Å). The common interacting active site amino acid residues of this hit with co-ligand are HIS-163, GLU-166, and GLN-189, and the interactions occur at a shorter distance as compared to the co-ligand. The binding energy (BE) and K_i values of this hit predicted by Autodock were −7.07 kcal/mole and 6.61 μM, respectively. The binding energy of hit DB08001 was −10.158 kcal/mol. It interacts with target protein through HIE-41 (π-π), CYS-145 (2.10 Å), GLU-166 (1.71 Å), and has two common interacting amino acid residues namely CYS-145 and GLU-166 with co-ligand N3. The interactions are also comparable with co-ligand. The BE and K_i values predicted by Autodock were −8.00 kcal/mole and 1.36 μM, respectively. The XP Glide score of compound DB08526 was −9.303 kcal/mol. This hit interacts with target protein through LEU (2.17 Å), SER-144 (2.14 Å), CYS-145 (2.63 Å), HIS-164 (2.06 Å), GLU-166 (2.09 Å), GLN-189 (2.03 Å) and shows three common interacting amino acid residues (CYS-145, GLU-166, GLN-189) with co-ligand. The predicted BE and K_i values of DB08526 were −7.26 kcal/mole and 2.89 μM, respectively. The XP Glide score of compound

DB02558 was −8.941 kcal/mol and interacting amino acid residues were HIE-41 (π-π), GLY-143 (2.67 Å), HIS-164 (2.28, 2.60 Å), GLU-166 (1.86 Å, 2.02 Å), and GLN-189 (2.06 Å). The common interacting active site amino acid residues of this hit with co-ligand are GLY-143, GLU-166 and GLN-189. The BE and K_i values of DB02558 predicted by Autodock 4.2 were −7.01 and 7.22 μM, respectively. The XP Glide score of DB12332 was −8.787 kcal/mol and the interacting amino acid residues were HIE-41 (π-π), LEU-141 (2.29 Å), CYS-145 (2.13 Å), GLU-166 (1.98 Å). This hit has two common interacting amino acid residues with co-ligand (CYS-145 and GLU-166). The types of interactions of these hits in the active sites are H-bond donor, H-bond acceptor, and π-π staking. The BE and K_i values of DB12332 predicted by Autodock 4.2 were −8.13 kcal/mol and 1.10 μM, respectively. The XP Glide score of hit DB02651 was −8.704 kcal/mol. The interacting amino acid residues of this hit were HIE-41 (two π-π), THR-25 (2.18 Å), SER-46 (1.54 Å), ASN-142 (2.60 Å), and CYS-145 (2.68 Å). This hit has one common interacting active site amino acid residues (CYS-145) with co-ligand N3. The binding energy and K_i values of this hit predicted by Autodock 4.2 were −8.40 kcal/mol and 0.7 μM, respectively. The interacting distance between all the selected hits and active site amino acid residues is shown in the parenthesis.

From the above findings, it is observed that the selected hits exhibited better binding energy and displayed stronger interactions with target protein than co-ligand N3 and co-ligand X77. This indicates that the selected hits have more binding affinity towards the target proteins than their original X-ray crystallographic co-ligand. The hits DB03208, DB03949,
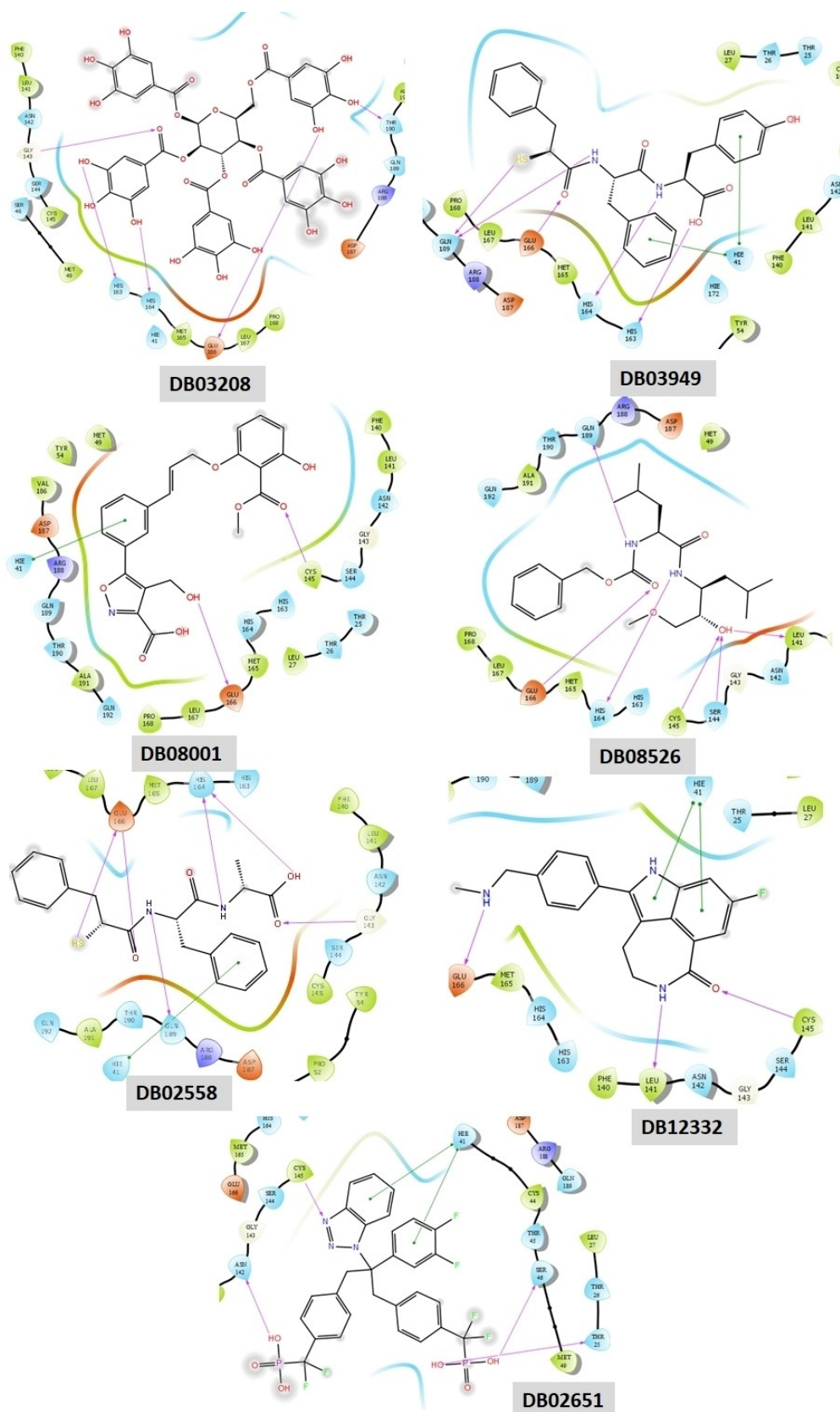
ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry
Europe
European Chemical
Societies Publishing

Figure 2. The 2D docking poses of compounds DB03208, DB03949, DB08001, DB08526, DB02558, DB12332, and DB02651 with 6LU7 of SARS-CoV-2. The hydrogen bond (donor and acceptor) interactions are depicted by the purple colors line, and π-π stacking is presented by the green colors line.
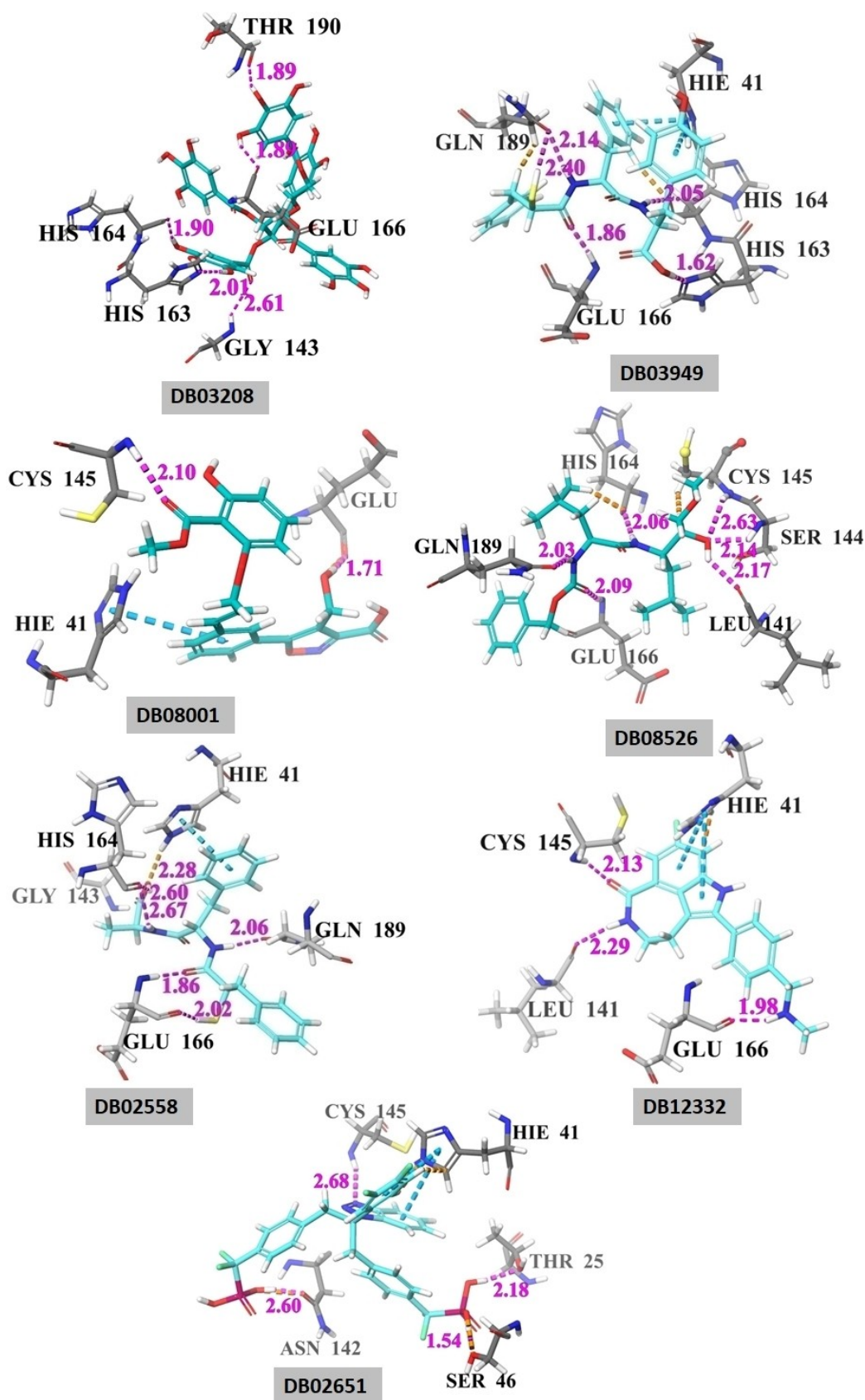
ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry
Europe
European Chemical
Societies Publishing

Figure 3. The 3D docking poses of compounds DB03208, DB03949, DB08001, DB08526, DB02558, DB12332 and DB02651 with 6LU7 of SARS-CoV-2. The hydrogen bond (donor and acceptor) interactions are shown by the dotted purple line, and π-π stacking is presented by doted sky color line.
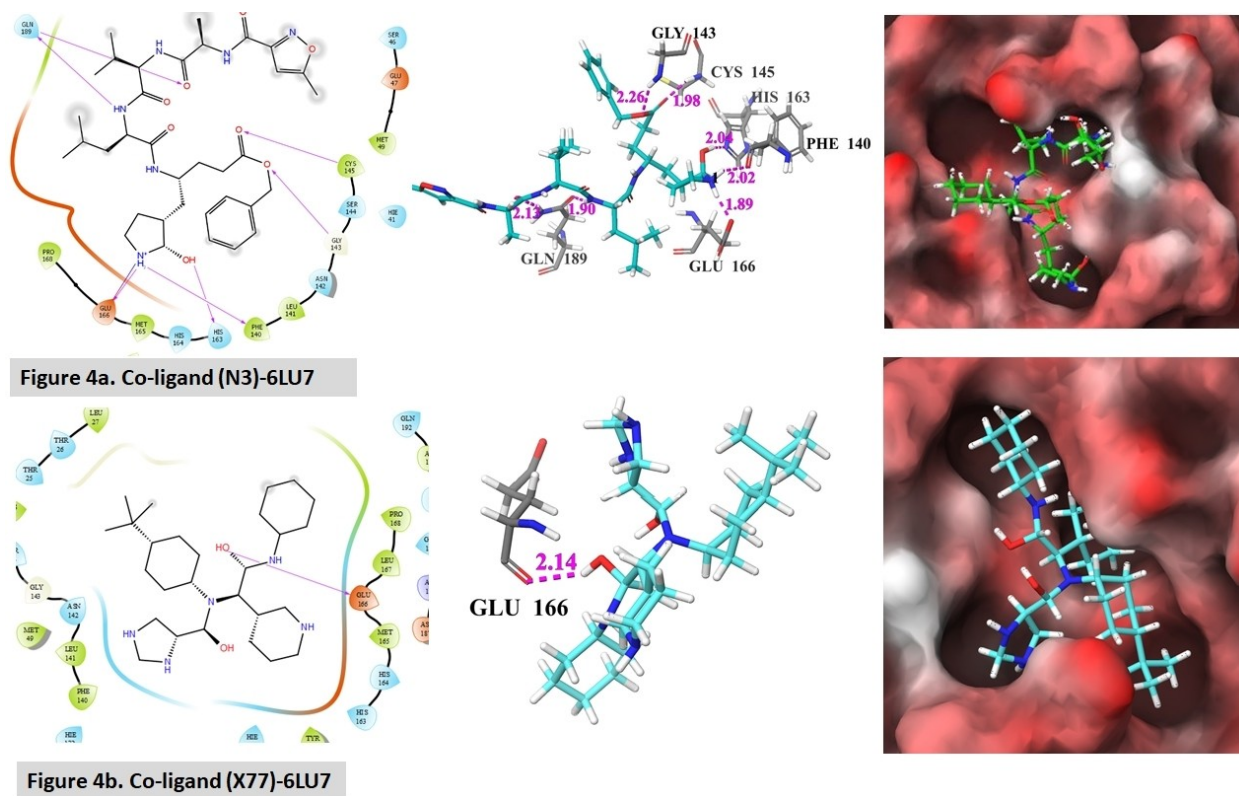
ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry
Europe
European Chemical
Societies Publishing

Figure 4a. Co-ligand (N3)-6LU7

Figure 4b. Co-ligand (X77)-6LU7

Figure 4. The 2D, 3D interactions and binding pose of co-ligand N3 (Figure 4a) and co-ligand X77 (Figure 4b) with active site amino acid residues of 6LU7. The hydrogen bond (donor and acceptor) interactions with distance in Å are shown by dotted purple line.

and DB08001 showed XP glide score lower than −10.00 kcal/mol and interact with the active site amino acids from a close distance. Most of the hits have favorable interactions with crucial amino acid residues HIE-41 and CYS-145. Yang *et al.* also identified HIS-41 and CYS-145 as two crucial amino acids that act as catalytic dyad in the active site of SARS-CoV M^{pro}.[61] A cysteine residue (CYS-145) acts as nucleophile and histidine residue (HIS-41) acts as acid-base. All the hits including co-ligand N3 and co-ligand X77 have a common interaction with 6LU7 through amino acid residue GLU-166. More importantly, the hit DB08526 showed six hydrogen bonding interactions, and each of the hits DB03208, DB03949, and DB02558 showed five hydrogen bonding interactions with the target protein. Out of these interactions, maximum interactions were took place from a close distance of less than 2.0 Å. In a previous study, Chen *et al.* showed that theaflavin-3,3-digallate (TF3), a gallic ester of flavones acts as an inhibitor of the main protease (3CLpro) of SARS-CoV.[62] We also observed that the inhibitor DB03208 is a gallic acid ester, shows strong binding interaction with the main protease of SARS-CoV-2. Although the XP glide score of hits DB12332 and DB02651 are low compared to other selected hits but they interact with two important amino acid residues namely, HIE-41 and CYS-145 with which the co-ligand N3 shows covalent interaction in the original X-ray crystallographic protein-ligand complex. Further, the best-docked pose of the selected hits shows that all the hits are deeply inserted into the active site cavities and the bonded pose is similar to

co-ligand. The individual binding poses of the seven hits are shown in Figure 5.

To check the reproducibility of docking results, the best seven hits were further subjected to molecular docking with other four crystal structures (6W79, 7C6S, 7JQ4, and 7JPZ) of the main protease of SARS-CoV-2. It was observed that all the hits also exhibited comparable binding energy with these proteins. The XP Glide scores of hits and co-ligands with five proteins are shown in Table S1. The docking analysis showed that the ranges of average XP glide score of selected hits are −14.2302 to −7.3518 kcal/mol and the hit DB03208 showed the top average XP glide score of −14.2302 kcal/mol. The ranges of average XP glide score of co-ligands are −9.5992 to −6.9232 kcal/mol. The co-ligand 7C6S has exhibited the highest negative XP glide score of −9.5992 kcal/mol. The average XP glide scores of three hits (DB03208, DB02651, and DB03949) are better than the average XP glide score of co-ligand of 7C6S. Thus, these three ligands are to be considered as the best ligands based on the glide score. The 2D docking poses and interacting amino acids of the hits DB03208, DB03949, DB08001, DB08526, DB02558, DB12332, and DB02651 with different M^{pro} of SARS-CoV-2 are shown in Figures S1–S7. The 2D docking poses of co-ligands with five different proteins (6LU7, 6W79, 7C6S, 7JQ4, and 7JPZ) of SARS-CoV-2 M^{pro} are shown in figures S8-S12.
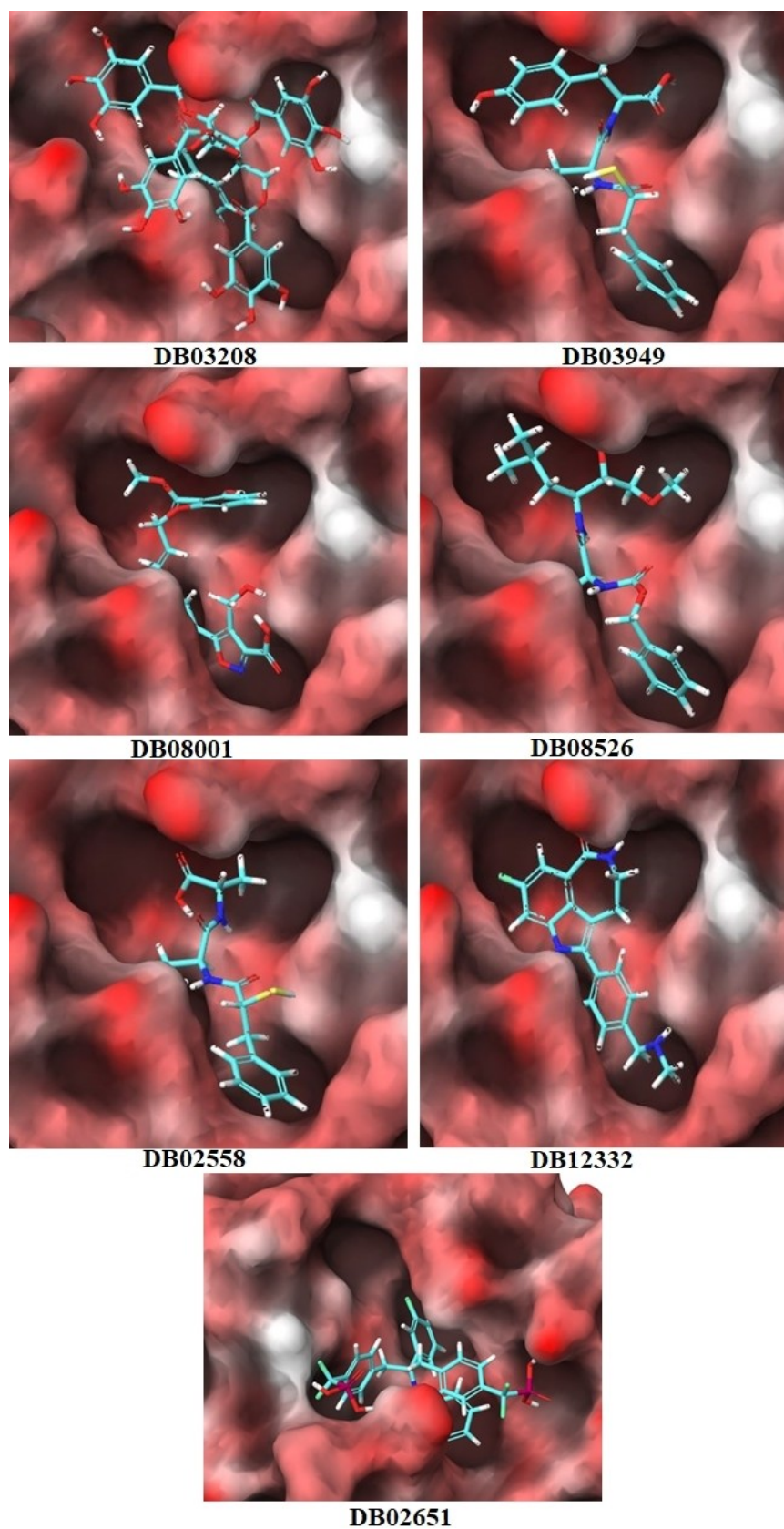
ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry
Europe
European Chemical
Societies Publishing

Figure 5. The individual binding pose of selected seven hits in the active site of 6LU7 of SARS-CoV-2.

ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

**Chemistry
Europe**
European Chemical
Societies Publishing

3.3. Molecular dynamics (MD) simulations study

The best seven hit–main protease complexes obtained from the structure-based virtual screening study were subjected to 50 ns molecular dynamic simulations. The important MD parameters like RMSD, RMSF, 2D-ligand-protein interaction diagram, the fraction of interaction of the ligand with different active site amino acid residues were calculated from molecular dynamics trajectory. These RMSD parameters provide a brief insight into the structural conformations of apo-proteins and protein-ligand complexes. The overall information about the stability of the protein backbone after the formation of the protein-ligand complex can be analyzed by the RMSD parameter. The lower value of RMSD throughout the MD simulation indicates that the high stability of the protein-ligand complex which is ideally acceptable for the protein-ligand system. Large changes in RMSD might indicate that the protein-ligand system during MD simulation time probably undergoes some major deviation or some shorts of conformational changes. Protein backbone RMSD (Å) of each ligand-protein complex and apo-protein were recorded and plotted against time in nanosecond (ns) (Figure 6).

It was observed that the average RMSD of all protein-ligand complexes was less than the RMSD of apo-protein. This indicates that protein-ligand complex systems are more stable than the apo-protein system. Although MD simulation of apo-protein is stable up to 50 ns, it shows a higher average RMSD value than other protein-ligand-complexes. The average RMSD value of apo-protein complex is 2.5 Å whereas protein-ligand complexes show an average RMSD value in the range 1.2 Å to 2.0 Å, which is lower than the average RMSD of apo-protein complex. Most importantly, the protein-ligand complex of compounds DB03949, DB08001, DB02651 and DB08526 exhibit average RMSD value less than 1.5 Å. This observation indicated that these hits are more stable in their complex with target protein as compared to apo-protein. The average RMSD of DB02651, DB12332, and DB02558 bound system exhibited a lower RMSD value than the average RMSD of the co-ligand X77-protein system. However, the protein-ligand complex of the DB02651 system showed the lowest RMSD value amongst the selected hits. The DB08526 system after ∼40 ns showed a slightly higher (∼0.03 Å) RMSD value than the X77-bound system. However, all the protein-ligand complex systems are oscillating within 3.0 Å.
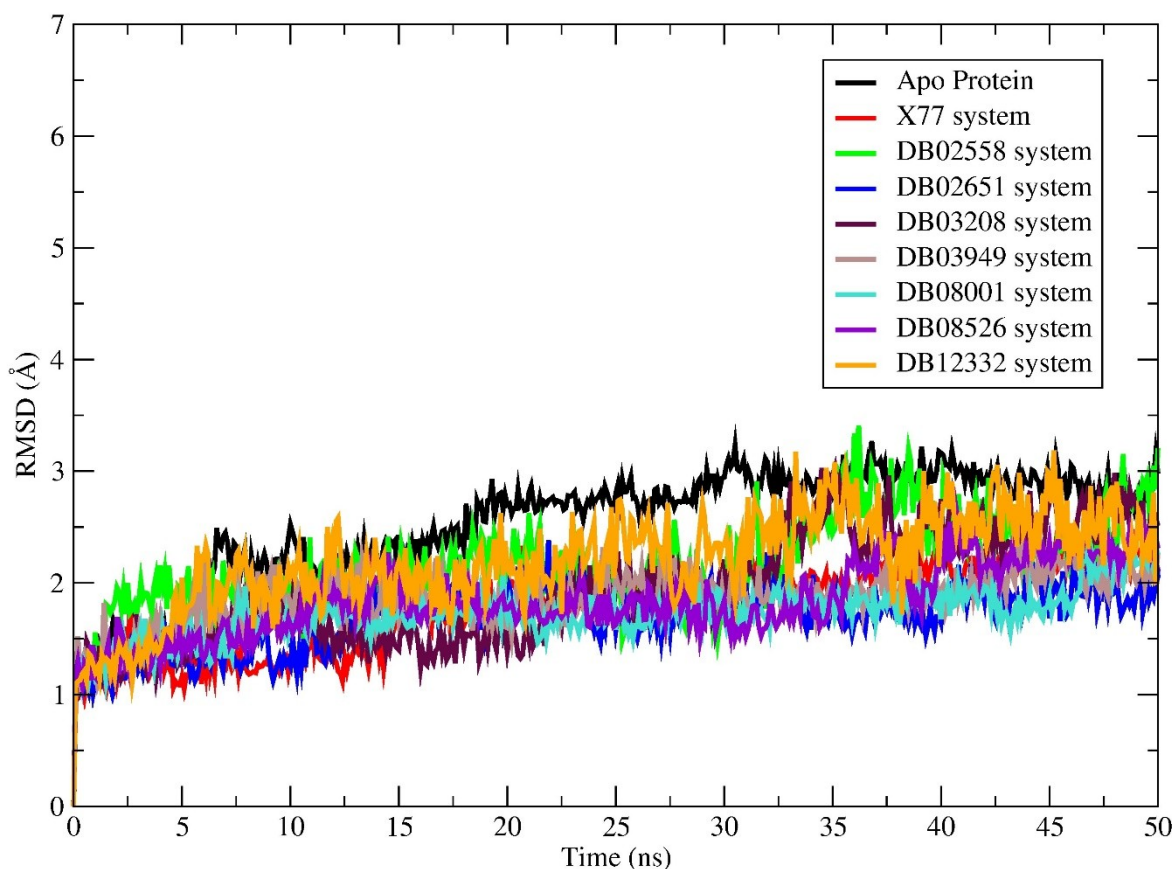

Figure 6. RMSD plot of apo-protein, co-ligand X77, and seven ligand-protein complexes systems.

ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry Europe
European Chemical
Societies Publishing

The average RMSD of ligand DB02558 was within 2.0 Å and shows stable RMSD up to 32 ns (green color). After 32 ns, the RMSD of this ligand-protein system was gradually increased and maximum deviation was observed at 36 ns. Similarly, the hit DB03208 also shows an average RMSD value of 2.0 Å (brown). This complex system shows the highest stability from 11 ns to 22 ns and almost stable up to 32 ns, after that there is a deviation in RMSD up to 50 ns and maximum deviation occurs at 35 ns. The 50 ns MD simulations of DB02651-protein complex (blue), DB03949-protein complex (ochre), and DB08001-protein complex (sky) showed the average RMSD value 1.2 Å and these protein-ligand complexes were nearly stable throughout the 50 ns simulation time. Initially, the RMSD value of the DB03949- M^{pro} system increases and shows the highest RMSD value of 1.8 at 10 ns, after that, it oscillates at ~1.18 Å with consistent deviation in RMSD profile up to 50 ns. The DB08001- M^{pro} system also oscillates at ~1.18 Å with consistent deviation in the RMSD profile. In agreement with the above point, it can be stated that compound DB02651, DB03949, and DB08001 after binding with SARS-CoV-2 main protease does not cause any significant structural changes. The average RMSD of protein-DB08526 complex (violet) was 1.5 Å and showed stable movement in terms of RMSD values throughout 50 ns simulation time. But there is a deviation from 34 ns to 43 ns and the maximum deviation (2.0 Å) was observed at 38 ns. The average RMSD of the protein-DB12332 complex was 2.0 Å and showed stable movement in terms of RMSD values throughout 50 ns simulation time (yellow). From

the above findings, it can be concluded here that these hits form stable bindings with the target protein.

The root mean square fluctuation (RMSF) plot was constructed to understand the deviation of each protein amino acid residues over time from the reference protein amino acid residue position. The RMSF parameter explores the individual fraction of the protein structure that is fluctuating from its mean structure. It is a calculation of individual residue flexibility, or how much a particular residue moves (fluctuates) during a simulation. RMSF per residue is typically plotted against residue number and can indicate structurally which amino acids have more contribution to a molecular motion. The plot was constructed from the 50 ns atomistic molecular dynamics trajectory. Lower the RMSF values indicate the reduced random motions and minimal fluctuations of backbone during the MD simulations run. Therefore, the lower RMSD and RMSF values of the system indicated smaller structural rearrangement, lesser conformational changes, and lesser internal motions around the binding site residues, respectively. On the other hand, the higher degrees of RMSF values infer that the protein structure will attain greater flexibility and higher conformational changes will occur in the protein-ligand complex. The RMSF plot of protein-ligand complexes are shown in Figure 7 and Figure 8, and RMSF plot of apo-protein is shown in Figure 9 to compare the residue-wise fluctuation between the apo-protein and ligand-protein complexes. It was observed that all the seven protein-ligand complexes showed an average RMSF value less than 1.0 Å, whereas X77-protein complex and apo-protein shows average
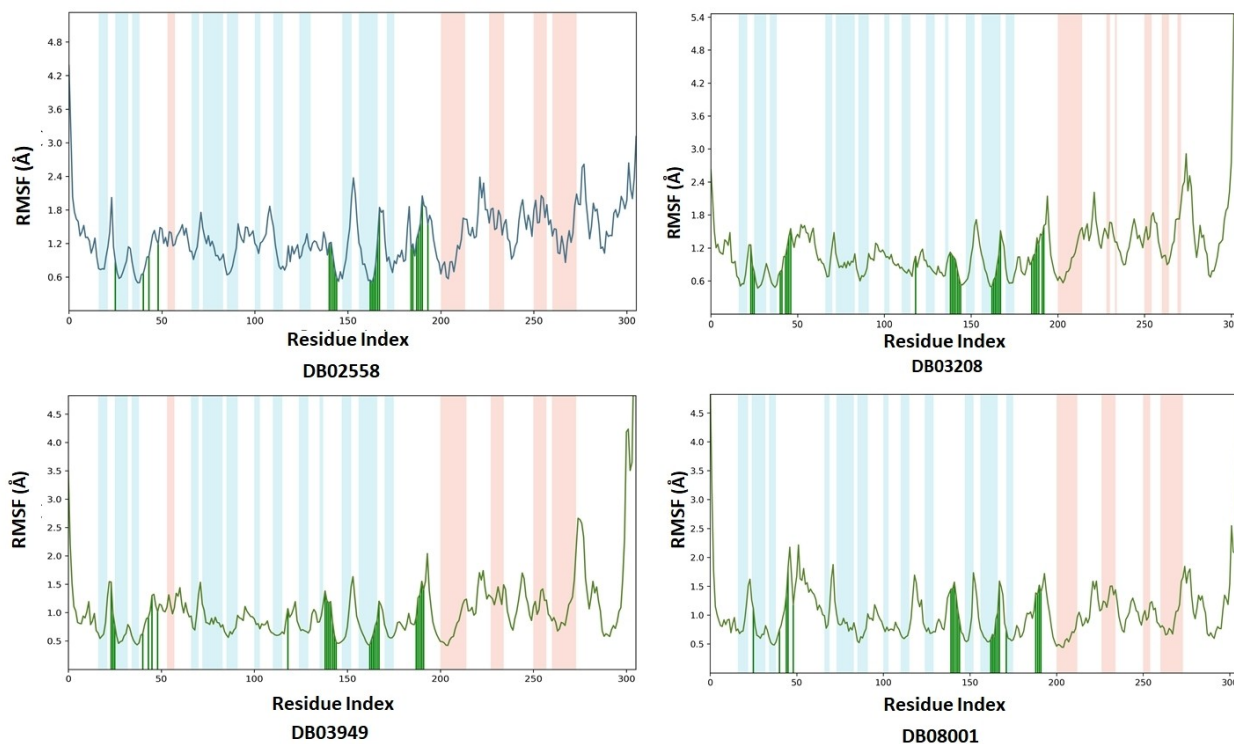

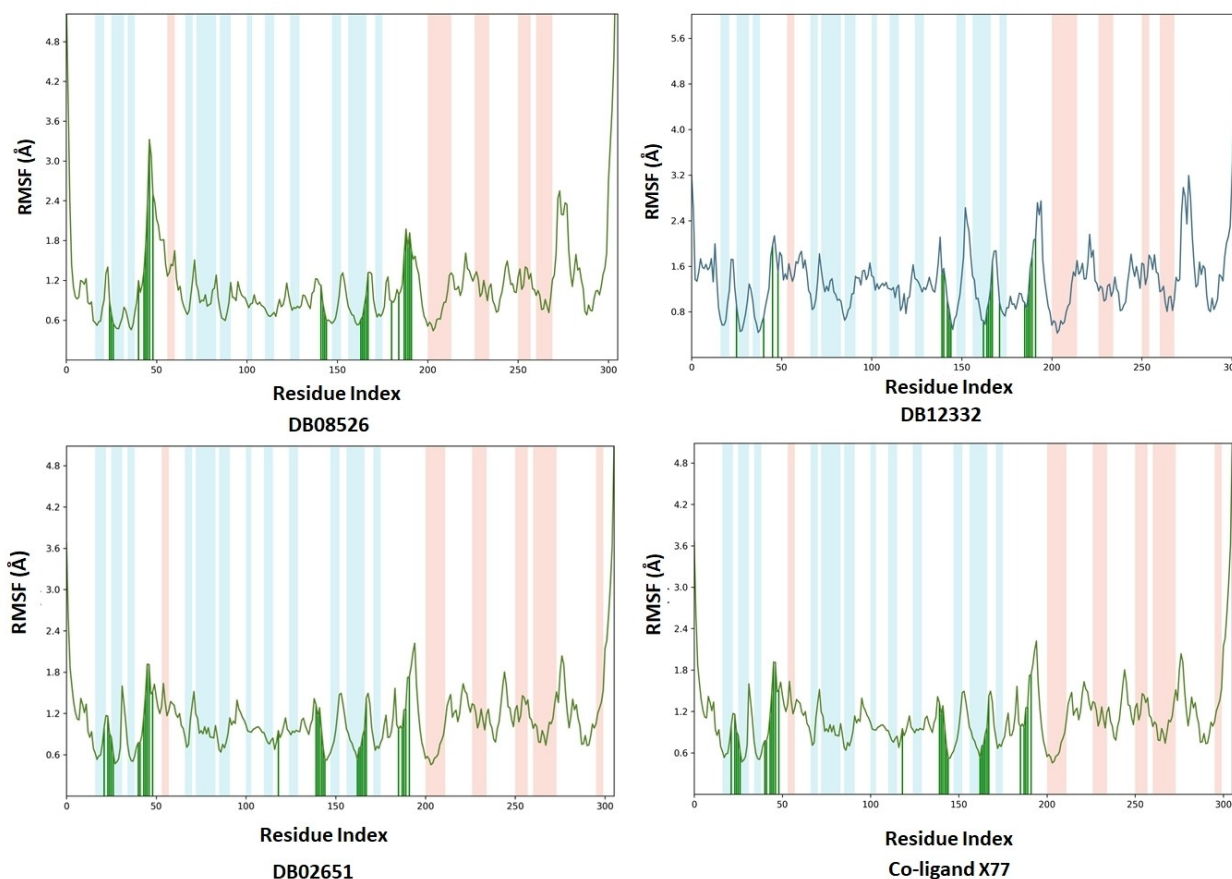**Figure 7.** RMSF plot of ligand-protein complexes.

ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry Europe
European Chemical
Societies Publishing

Figure 8. RMSF plot of ligand-protein complexes and co-ligand X77-6LU7 complex.
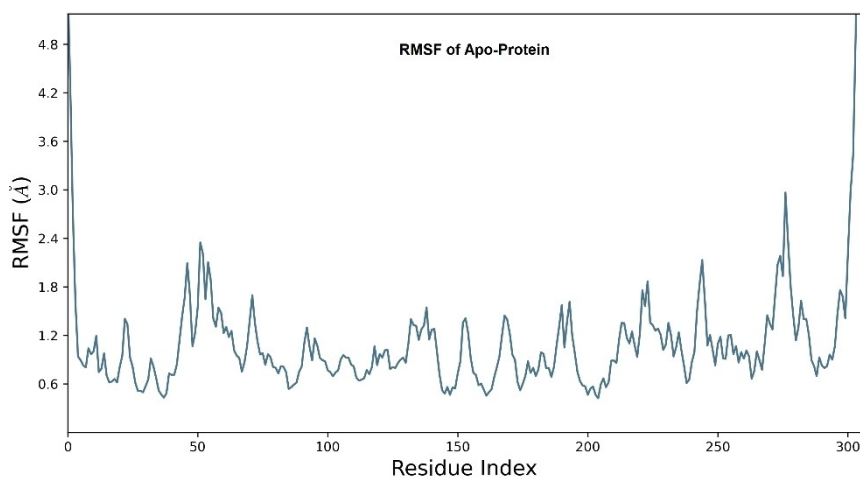

Figure 9. RMSF plot of apo-protein complex.

RMSF values of 1.2 Å and 1.3 Å, respectively. This indicates that protein-ligand complexes are more stable than the apo-protein complex during the molecular dynamics simulation.

The average RMSF of hit DB02558-protein system was 1.0 Å which is lower than the average RMSF of apo-protein and X77-protein complex. This ligand exhibits superior hydrogen

bonding interaction with amino acids HIS-41, GLU-166, HIS-164 and ARG-188. The ligand-protein contacts of this hit through H-bonding were HIS-41 (31% with OH of -COOH gr), GLU-166 (44% with carbonyl of amide gr), HIS-164 (51% with NH of -CONH gr), and ARG-188 (50% with CO of amide gr and 43% with –SH group). The residue GLY-143 interacts with –COOH

ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry Europe
European Chemical
Societies Publishing

group through the water-bridge H-bonding almost 31% of simulation time. This molecule also showed π-π interactions with HIS-164 (56%) and hydrophobic interaction with MET-49. The protein-ligand contacts analysis of hit DB02558 with M^{pro} is shown in Figure 10.

The ligand DB03208 exhibits seven hydrogen bonding interactions with amino acid residues namely- HIS-41, CYS-44, ASN-142, GLY-143, GLU-166, GLN-189 and THR-190. The most of the H-bonding interactions of this hit occur through it's the phenolic group. The ligand-protein contacts of this hit
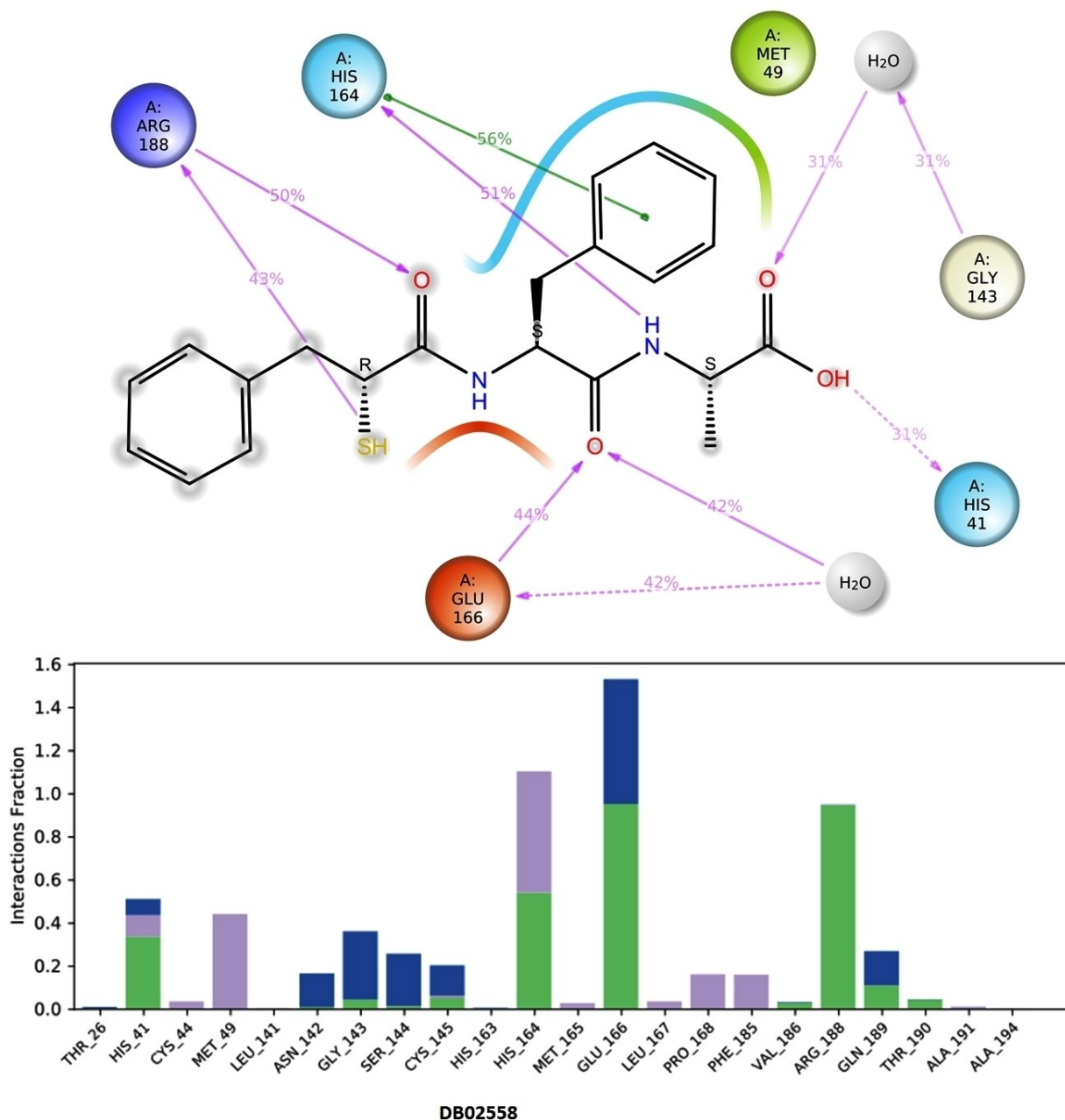

Figure 10. The ligand-protein contacts analysis of hit DB02558 during 50 ns simulation time. The green color indicates the hydrogen bonding interactions, the pink color represented hydrophobic interactions, and the blue color represented water bridge interactions.

ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry Europe
European Chemical
Societies Publishing

(DB03208) were GLY-143 (52%), GLU-166 (56%), ARG-188 (56%), GLN-189 (57%) and THR-190 (81%). All these interaction occurred with phenolic-OH groups of the ligand. . Moreover, the amino acid residue ASN-142, GLU-166 and GLN-189 also interacted with ester carbonyl almost 58%, 89% and 77% of simulation time, respectively. This hit also shows a water-bridge H-bonding interaction with THR-26 almost 52% of simulation time. The protein-ligand contacts analysis of hit DB03208 with M^pro are shown in Figure 11.

The ligand DB03949 showed two H-bonding interactions with amino acid residue GLU-166 and GLN-189 almost 72% and 53% of simulation time, respectively through the amide hydrogens. The hit DB08001 showed three H-bonding inter-actions with GLY-143 (83%), SER-144 (37%), and GLU-166
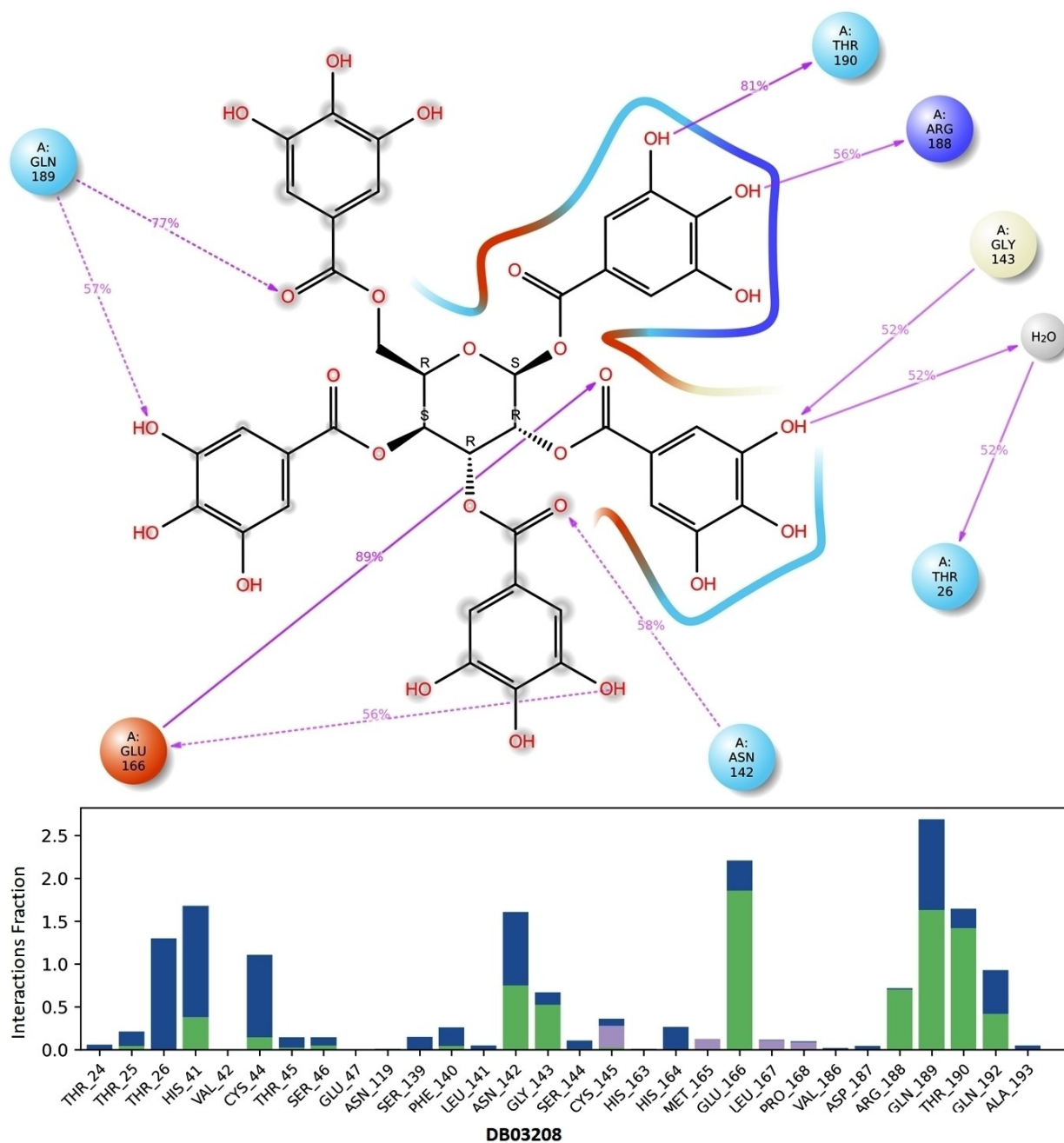

Figure 11. The ligand-protein contacts of hit DB03208 during 50 ns simulation time. The green color indicates the hydrogen bonding interactions, the pink color represented hydrophobic interactions, and the blue color represented water bridge interactions.

ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry
Europe
European Chemical
Societies Publishing

(34%) through the phenolic-OH group of 2,6-dihydroxybenzen-zoate ester moiety. Both the ligands showed hydrophobic interactions with MET-49 and MET-165. The protein-ligand contacts analysis of hits DB03949 and DB08001 with M^pro are shown in Figure 12 and Figure 13, respectively.

The ligand DB08526 shows protein-ligand interactions through two H-bondings and the interacting amino acid residues were HIS-164 (96% through NH of amide group) and GLU-166 (92% with the carbonyl of carbamate). Similarly, the DB12332 shows four superior H-bonding interactions with amino acid residues of M^pro. The ligand-protein contacts of this hit (DB12332) through H-bonding were CYS-145 (82% through carbonyl oxygen), GLY-143 (93% through carbonyl group of azepine moiety), SER-144 (42% with the carbonyl group of azepine and 41% with —NH gr of azepine moiety), GLU-166 (65% with amine gr). This ligand also exhibited a water-bridge H-bonding interaction with GLN-189 (30% with NH of indole). Moreover, this hit showed a π-π interaction with HIS-41(50% with indole ring) and hydrophobic interaction with MET-49. The
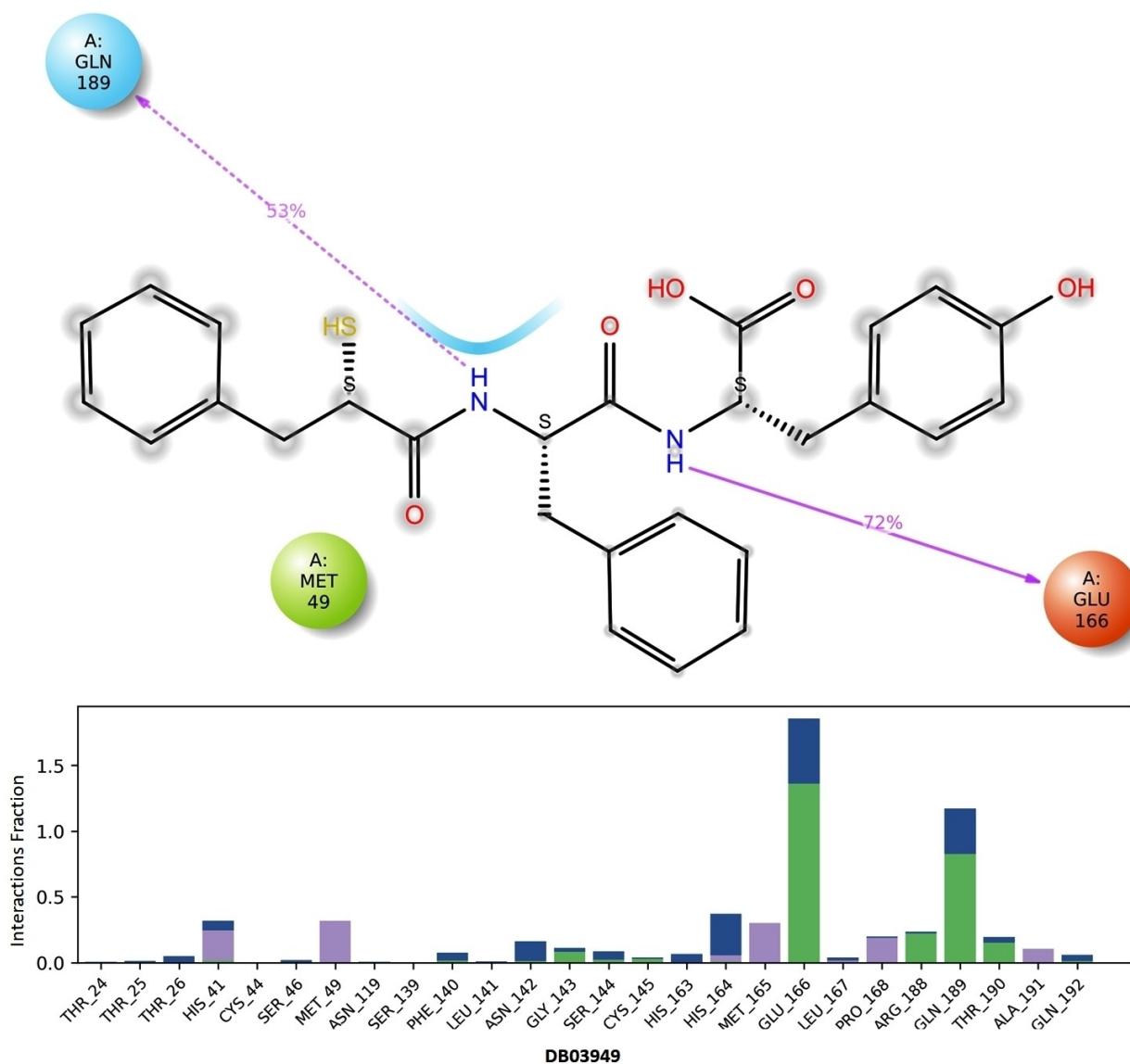

Figure 12. The ligand-protein contacts analysis of hit DB03949 during 50 ns simulation time. The green color indicates the hydrogen bonding interactions, the pink color represented hydrophobic interactions, and the blue color represented water bridge interactions.

ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry Europe
European Chemical
Societies Publishing

Figure 13. The ligand-protein contacts analysis of hit DB08001 during 50 ns simulation time. The green color indicates the hydrogen bonding interactions, the pink color represented hydrophobic interactions, and the blue color represented water bridge interactions.
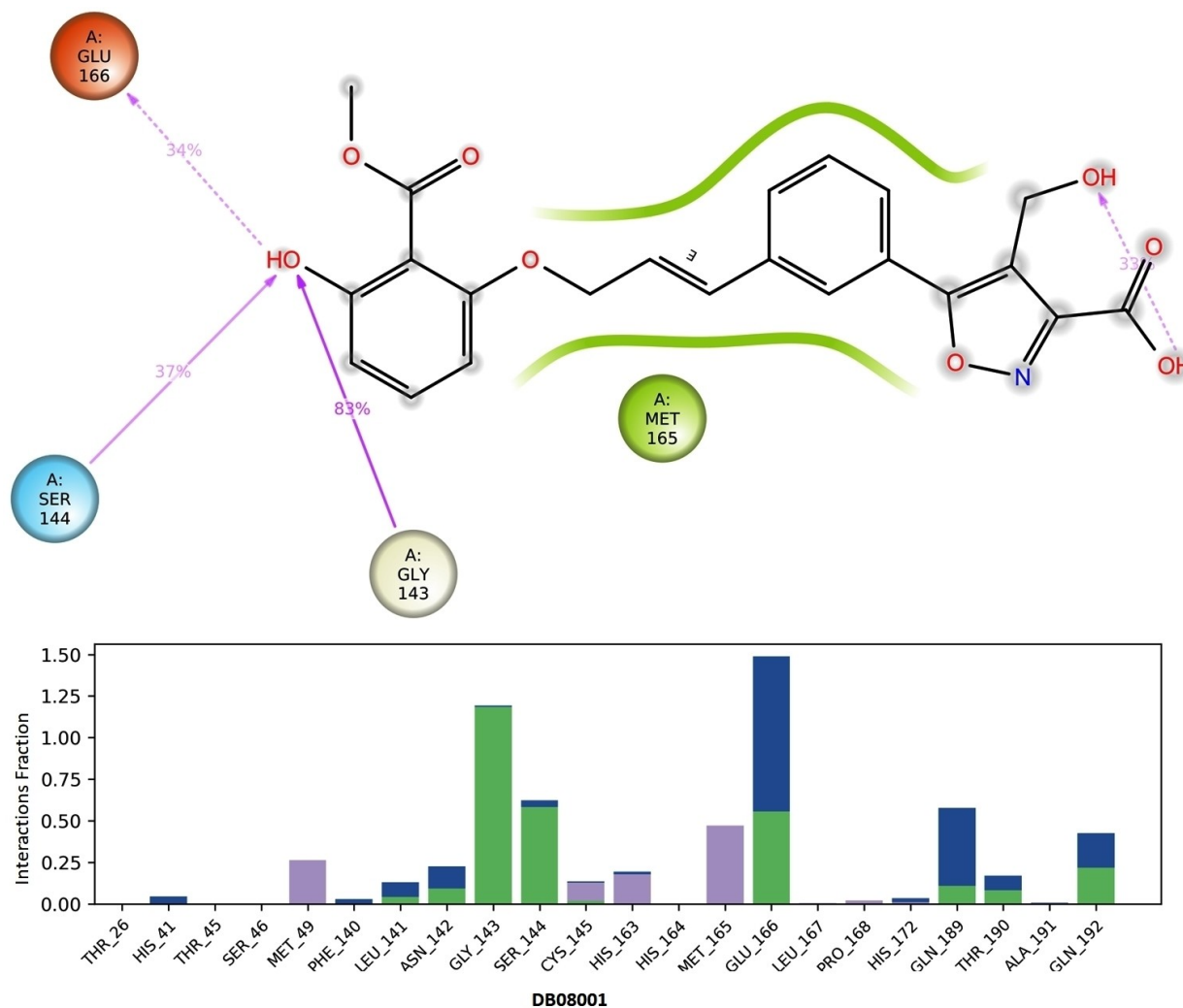
protein-ligand contacts analysis of hits DB08526 and DB12332 with M^{pro} are shown in Figure 14 and Figure 15, respectively.

The ligand DB02651 exhibited major protein-ligand interaction through two H-bonding and one π-π stacking interaction. The interacting amino acid residues were HIS-41 (30% π-π stacking of benzene ring of benzotriazole), CYS-145 (hydrophobic and water bridge interactions) GLU-166 (52% with the hydroxyl group of phosphonic acid moiety), and GLN-192 (57% with phosphonic moiety). The non-covalent inhibitor X77 showed major protein-ligand interactions with HIS-41 (38% of hydrophobic interactions with imidazole moiety), GLY-143

(99% interaction through H-bonding with substituted amide carbonyl moiety), HIS-163 (95% interaction through H-bonding with the nitrogen atom of pyridine ring), GLU-166 (59% interaction through H-bonding with the carbonyl group of substituted amide). The protein-ligand contacts analysis of hits DB02651 and non-covalent inhibitor X77 with M^{pro} are shown in Figure 16 and Figure 17, respectively. The interactions of selected hits are comparable with interactions of non-covalent inhibitor X77. The amino acid residues THR-25 to GLU-47 of the DB08526 system exhibited slightly higher (∼3.1 Å) RMSF. These residues are present in the loop region of this protein, may
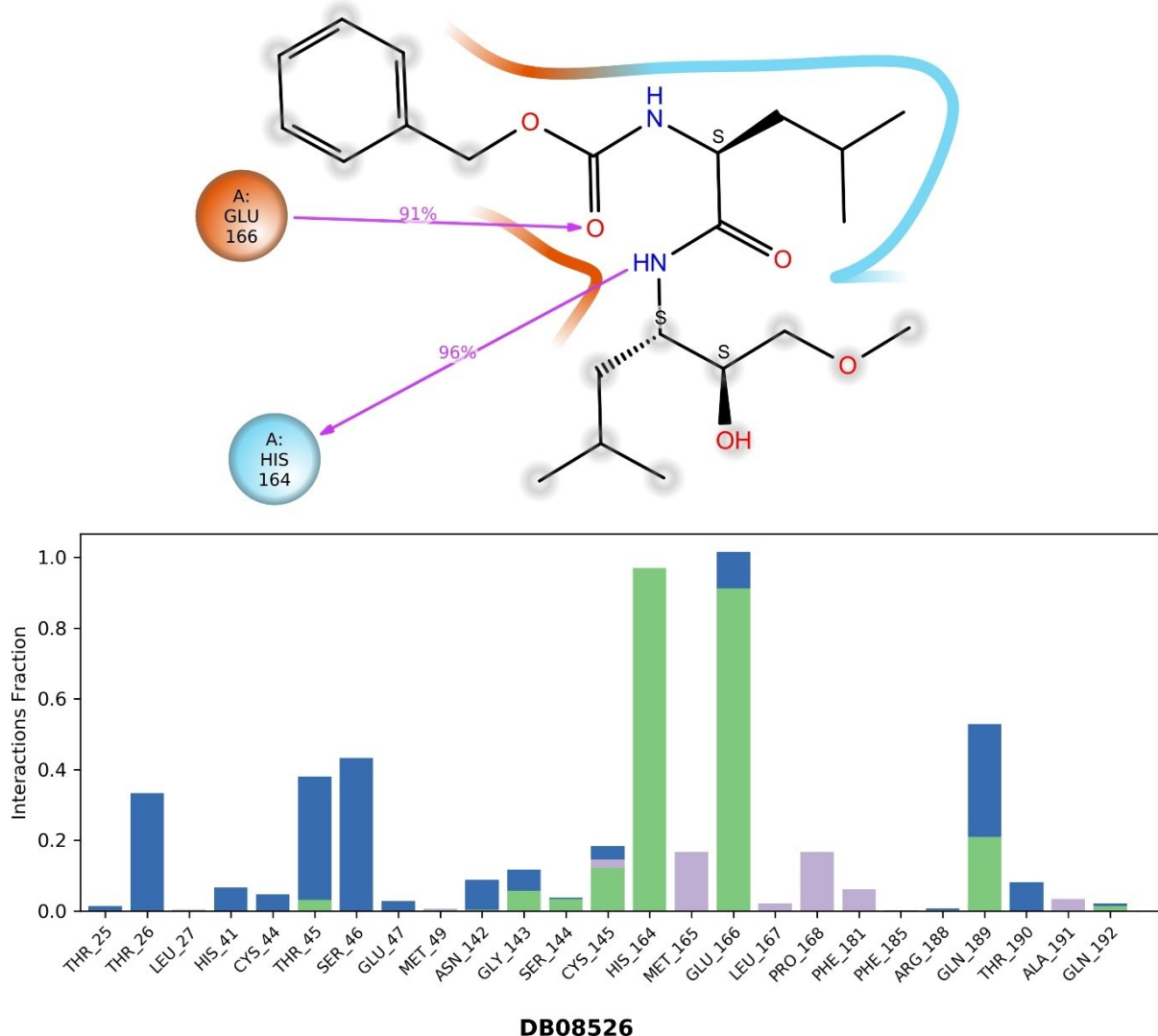
ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

**Chemistry
Europe**
European Chemical
Societies Publishing

Figure 14. The ligand-protein contacts analysis of hit DB08526 during 50 ns simulation time. The green color indicates the hydrogen bonding interactions, pink color represented hydrophobic interactions and blue colour represented water bridges interactions.

cause higher fluctuation. However, no significant changes in the RMSF profile were observed for the identified hit-protein system when compared with the RMSF profile of the control ligand X77 bound system. In-depth analysis of trajectory infers this high fluctuation did not produce any significant effect upon ligand binding. Considering molecular dynamics is a stochastic process each binding site residues of the identified ligand-protein system oscillated within a considerable range. Moreover, during simulation, these ligands exhibited interaction with binding site residues. The green vertical bars present in the RMSF plot (Figure 7 and 8) represents the above phenomenon. The RMSD, RMSF, and protein-ligand contact analysis showed the acceptable range of values that indicate the stable protein-ligand interaction profile. Moreover, the molecular dynamic studies revealed that the selected hits were bound with biologically critical regions of M^{pro} from stable

protein-ligand complexes, indicating their potential to inhibit the functionality of this component.

The binding free energy (ΔG_{bind}) of each ligand with the SARS-CoV-2 main protease was calculated from 50 ns MD trajectory using MM-GBSA approach. For this purpose, snapshots of each 10 ns interval were extracted from the energy equilibrated 50 ns molecular dynamics trajectory and used as an input for the g_mmpbsa program. The energetic behavior of the systems along with MD simulations (binding energy vs time plot) is shown in Figure S13.

The MM-GBSA based binding free energy of compounds DB03949, DB08001, DB02558, DB08526, DB12332, DB02651 and DB03208-main protease complex predicted by prime are −71.8234, −69.4447, −69.8786, −65.0922, −58.6618, −72.1280, and −83.2718 kcal/mol, respectively. The binding energy of co-ligand N3- and X77-main protease complex are −68.927 and −67.86 kcal/mol, respectively. These binding

ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry
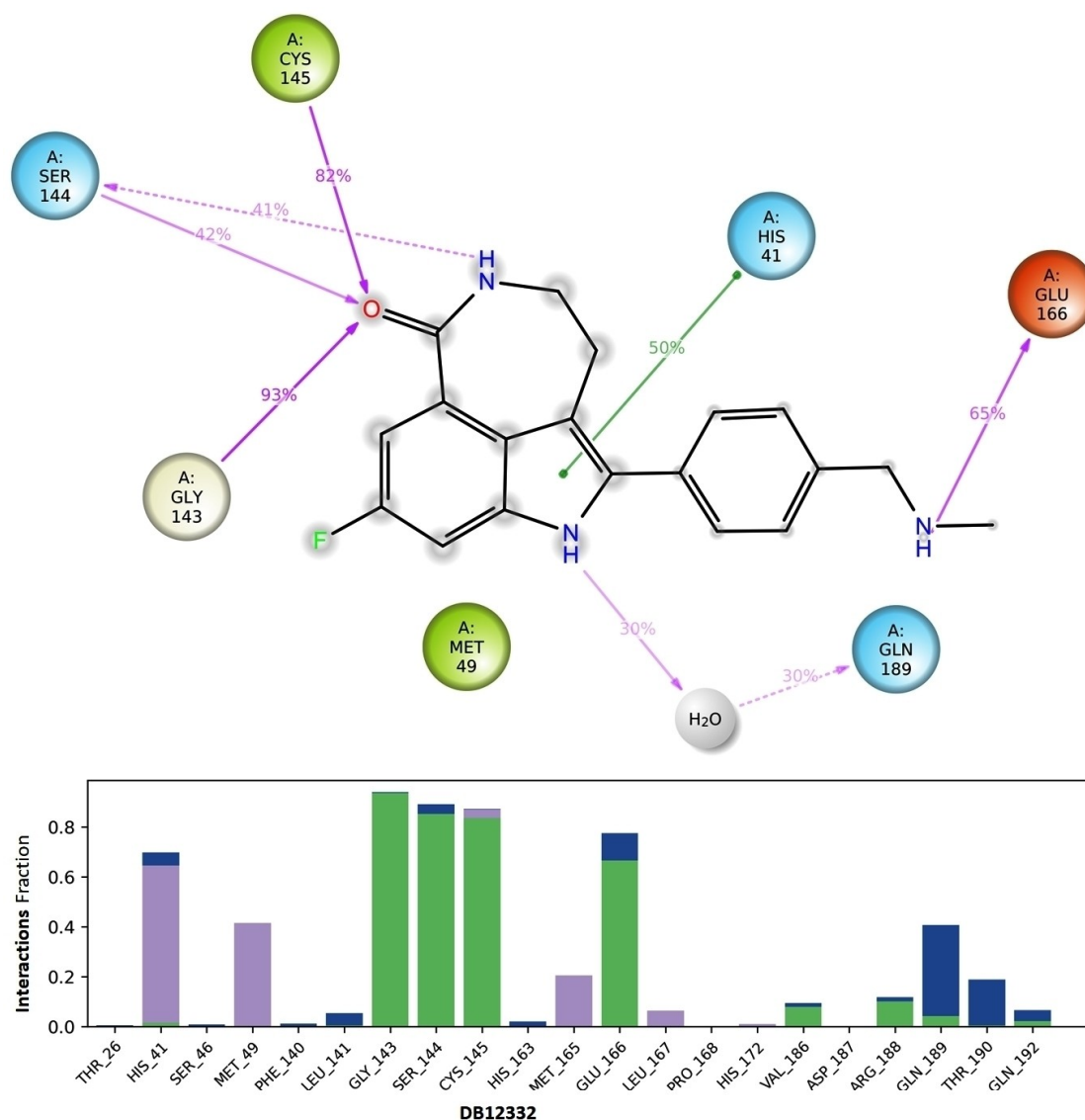**Europe**
European Chemical
Societies Publishing

Figure 15. The ligand-protein contacts analysis of hit DB12332 during 50 ns simulation time. The green color indicates the hydrogen bonding interactions, pink color represented hydrophobic interactions and blue colour represented water bridges interactions.

energy values also support the superiority of compounds DB03208, DB02651, DB03949, DB08001, and DB02558 as SARS-CoV-2 main protease inhibitors compared to the covalent inhibitor N3 and a non-covalent inhibitor X77. The hits DB03949, DB08001, DB02558, DB08526, and DB12332 showed no violation of Lipinski's rule of five. Generally, if any hit shows three or more than three violations in Lipinski's rule of five then the hit does not follow the rule of drug likeliness and does not consider as a drug molecule. The compounds DB02651 and DB03208 violated two and three rules, respectively amongst the five rules of Lipnskis (Table S2) and there-

fore, DB03208 is not considered for further development. The structure and IUPAC name of the selected seven hits are shown in Figure 18.

4. Conclusions

We have carried out structure-based virtual screening (VS) of the DrugBank database containing 11875 numbers of molecules using Glide. The molecular docking resulted in seven potential compounds having a good binding affinity towards active site amino acid residues of the main protease. The
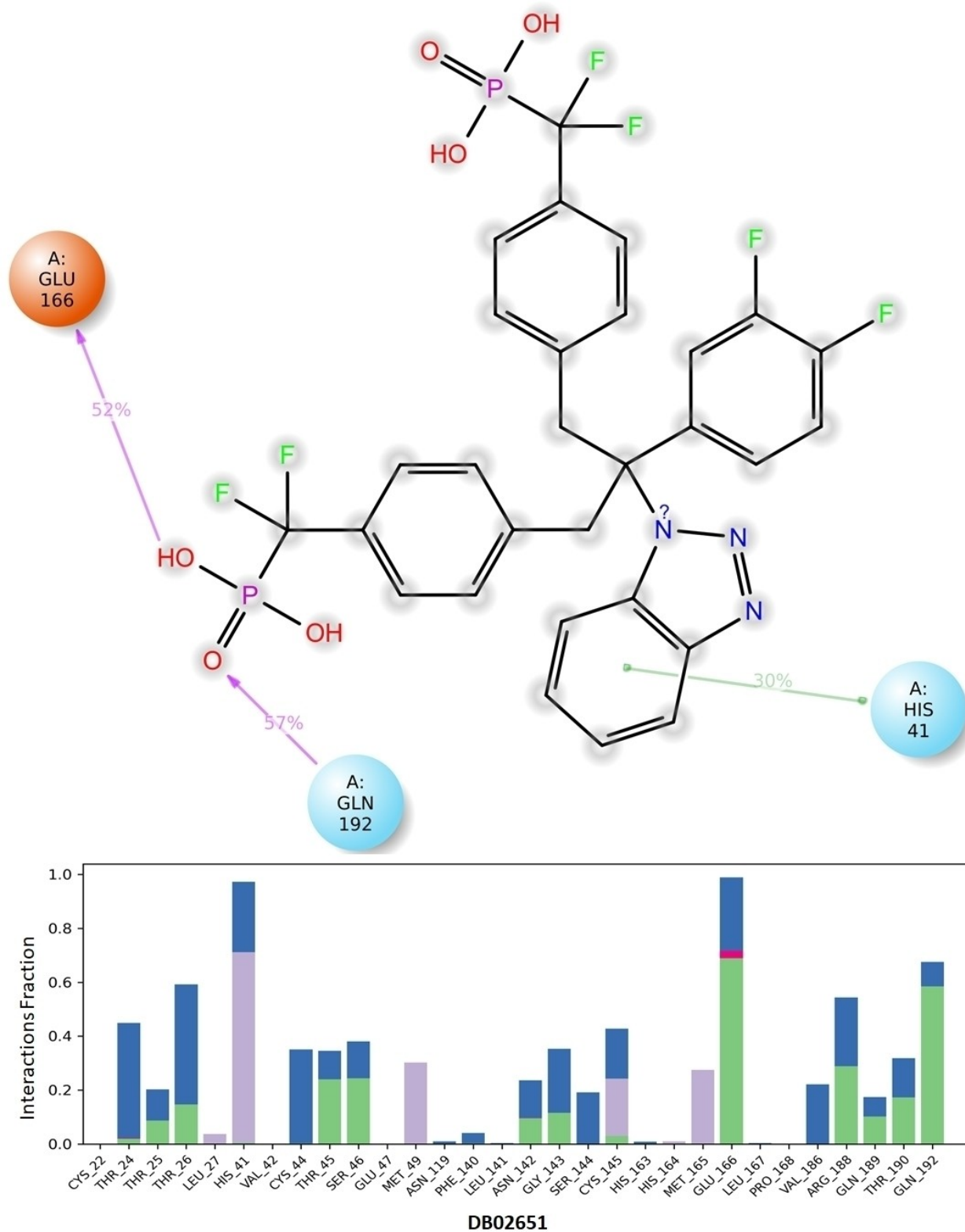
ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry
Europe
European Chemical
Societies Publishing

Figure 16. The ligand-protein contacts analysis of hit DB02651 during 50 ns simulation time. The green color indicates the hydrogen bonding interactions, the pink colour represented hydrophobic interactions, and the blue colour represented water bridge interactions.

ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry Europe
European Chemical
Societies Publishing

Figure 17. The ligand-protein contacts analysis of co-ligand X77 during 50 ns simulation time. The green color indicates the hydrogen bonding interactions, the pink colour represented hydrophobic interactions, and the blue colour represented water bridge interactions.
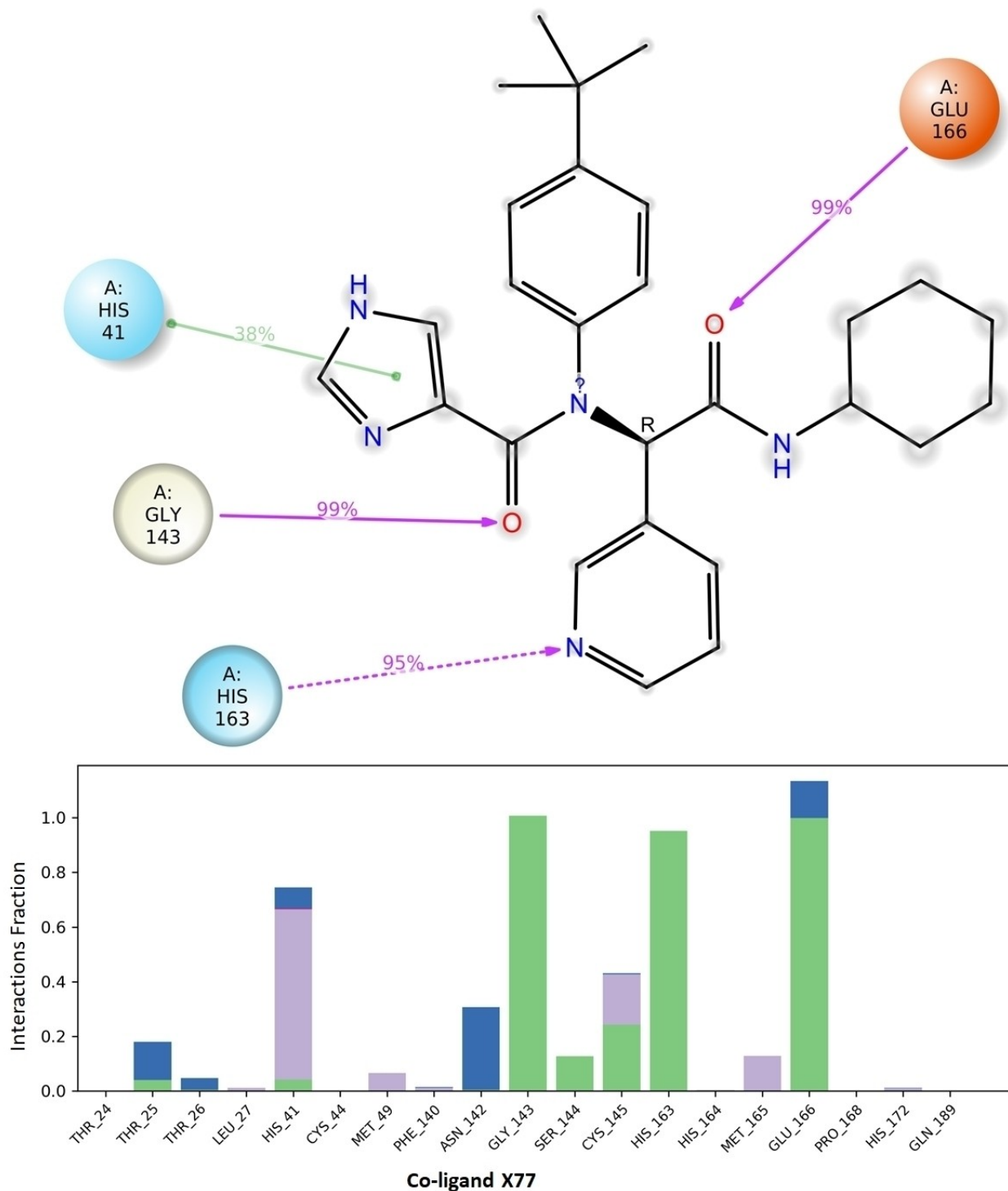
selected inhibitors show admirable XP glide score ranges from −15.071 to −8.704 kcal/mol with 6LU7 and average glide score ranges from −14.2302 to −7.3518 kcal/mol with five crystal structures of M^{pro}. Molecular dynamic simulations showed that all the selected hits formed stable protein-ligand complexes with M^{pro} during the whole simulation time and comparable with the apo-ligand system and non-covalently bonded coligand-protein system. Based on the XP Glide score, RMSD, RMSF, and MM-GBSA binding energy the following five hits DB02651, DB03949, DB08001, DB08526, and DB02558 showed
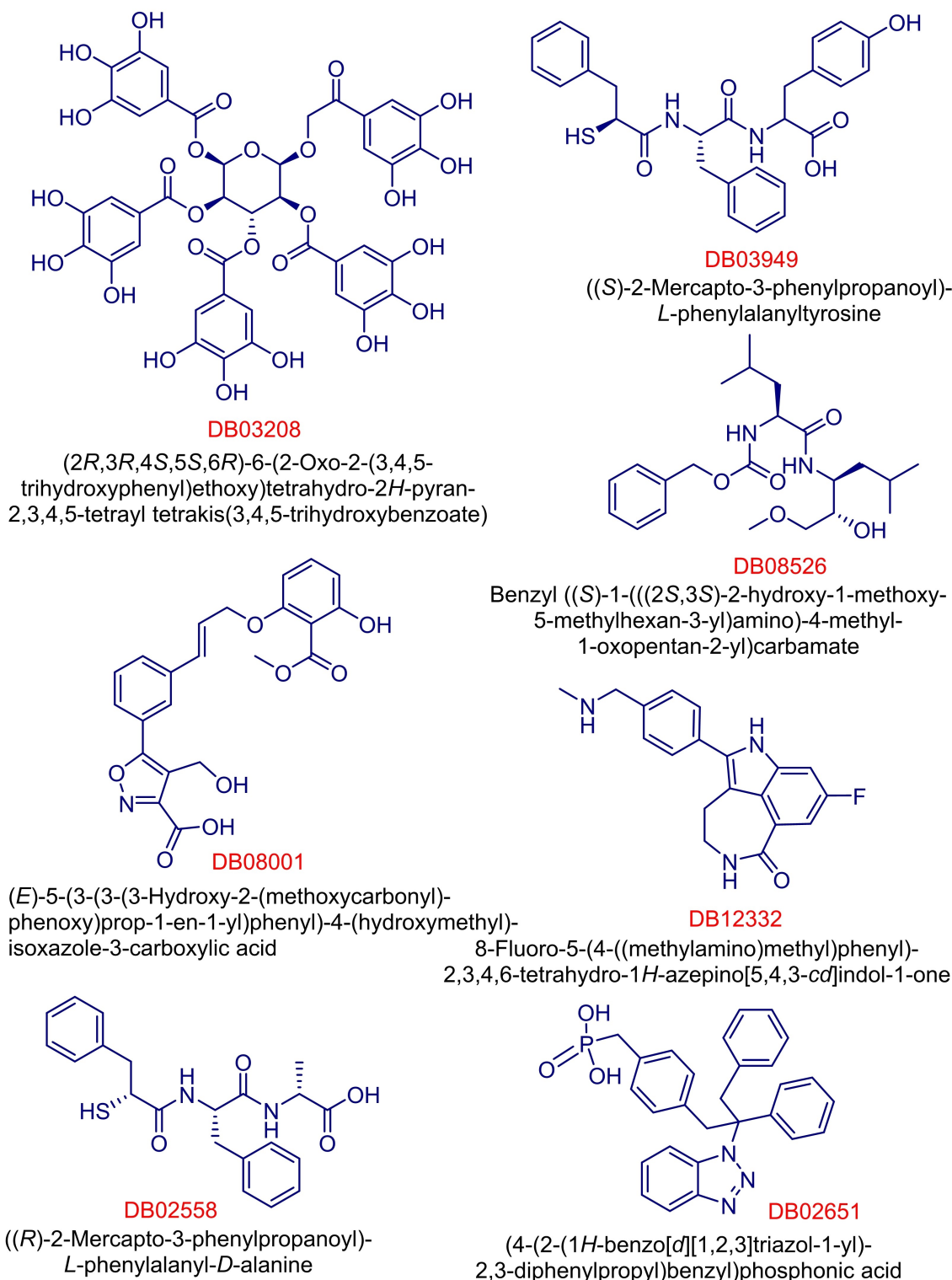
ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry Europe
European Chemical
Societies Publishing

DB03208
(2*R*,3*R*,4*S*,5*S*,6*R*)-6-(2-Oxo-2-(3,4,5-trihydroxyphenyl)ethoxy)tetrahydro-2*H*-pyran-2,3,4,5-tetrayl tetrakis(3,4,5-trihydroxybenzoate)

DB03949
((*S*)-2-Mercapto-3-phenylpropanoyl)-*L*-phenylalanyltyrosine

DB08526
Benzyl ((*S*)-1-(((2*S*,3*S*)-2-hydroxy-1-methoxy-5-methylhexan-3-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate

DB08001
(*E*)-5-(3-(3-(3-Hydroxy-2-(methoxycarbonyl)-phenoxy)prop-1-en-1-yl)phenyl)-4-(hydroxymethyl)-isoxazole-3-carboxylic acid

DB12332
8-Fluoro-5-(4-((methylamino)methyl)phenyl)-2,3,4,6-tetrahydro-1*H*-azepino[5,4,3-*cd*]indol-1-one

DB02558
((*R*)-2-Mercapto-3-phenylpropanoyl)-*L*-phenylalanyl-*D*-alanine

DB02651
(4-(2-(1*H*-benzo[*d*][1,2,3]triazol-1-yl)-2,3-diphenylpropyl)benzyl)phosphonic acid

Figure 18. Structure and name of the selected hits.

excellent binding affinity towards the 6LU7 active site. The hit DB12332 showed excellent interaction with crucial amino acid residue HIS-41, GLY-143, CYS-145. All the inhibitors reported here are commercially available for further characterization and development by the scientific community.

Supporting Information Summary

The supporting information contains the XP glide scores of selected hits, ADME properties, the 2D docking poses and the protein-ligand interactions of all hits with five different crystal

ChemistrySelect

Full Papers
doi.org/10.1002/slct.202100854

Chemistry
Europe
European Chemical
Societies Publishing

structures of SARS-CoV-2 M^{pro} and the energetic behavior of the systems along their MD simulations (total energy vs. time plot).

Conflict of Interest

The authors declare no conflict of interest.

[1] T. Ahmad, H. Haroon, M. Baig, J. Hui, *Pak. J. Med. Sci.* **2020**, *36*, (COVID19-S4):COVID19-S73-S78.

[2] P. Verma, A. Dumka, A. Bhardwaj, M. C. Kestwal, P. Kumar, *S. N. Comput. Sci.* **2021**, *2*, 27.

[3] N. Chen, M. Zhou, X. Dong, J. Qu, F. Gong, Y. Han, Y. Qiu, J. Wang, Y. Liu, Y. Wei, J. Xia, T. Yu, X. Zhang, L. Zhang, *Lancet* **2020**, *395*, 507–513.

[4] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang, Z. Gao, Q. Jin, J. Wang, B. Cao, *Lancet* **2020**, *395*, 497–506.

[5] E. Israeli, *Novel coronavirus that recently emerged in china, Harefuah* **2020**, *159*, 70–71.

[6] F. Wu, S. Zhao, B. Yu, Y. M. Chen, W. Wang, Z. G. Song, Y. Hu, Z. W. Tao, J. H. Tian, Y. Y. Pei, M. L. Yuan, Y. L. Zhang, F. H. Dai, Y. Liu, Q. M. Wang, J. J. Zheng, L. Xu, E. C. Holmes, Y. Z. Zhang, *Nature* **2020**, *579*, 265–269.

[7] P. Zhou, X. L. Yang, X. G. Wang, B. Hu, L. Zhang, W. Zhang, H. R. Si, Y. Zhu, B. Li, C. L. Huang, H. D. Chen, J. Chen, Y. Luo, H. Guo, R. D. Jiang, M. Q. Liu, Y. Chen, X. R. Shen, X. Wang, X. S. Zheng, K. Zhao, Q. J. Chen, F. Deng, L. L. Liu, B. Yan, F. X. Zhan, Y. Y. Wang, G. F. Xiao, Z. L. Shi, *Nature* **2020**, *579*, 270–273.

[8] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F. Gao, W. Tan, *N. Engl. J. Med.* **2020**, *382*, 727–733.

[9] World Health Organization, International Health Regulations Emergency Committee on novel coronavirus in China. https://www.who.int/news-room/events/detail/2020/01/30/defaultcalendar/international-health-regulations-emergency-committee-on-novel-coronavirus-in-china.

[10] Z. Xu, C. Peng, Y. Shi, Z. Zhu, K. Mu, X. Wang, W. Zhu, *bioRxiv* **2020**, https://doi.org/10.1101/2020.01.27.92167.

[11] A. Zumla, D. S. Hui, E. I. Azhar, Z. A. Memish, M. Maeurer, *Lancet* **2020**, *395*, e35–e36.

[12] H. Zhang, K. M. Saravanan, Y. Yang, M. T. Hossain, J. Li, X. Ren, Y. Pan, Y. Wei, *Interdiscip. Sci.* **2020**, 1–9.

[13] C. Wu, Y. Liu, Y. Yang, P. Zhang, W. Zhong, Y. Wang, Q. Wang, Y. Xu, M. Li, X. Li, M. Zheng, L. Chen, H. Li, *Acta Pharm. Sin. B* **2020**, *10*, 766–788.

[14] Y. Zhou, Y. Hu, J. Shen, Y. Huang, W. Martin, F. Cheng, *Cell Discov.* **2020**, *6*, 14.

[15] Y. Li, J. Zhang, N. Wang, H. Li, Y. Shi, G. Guo, K. Liu, H. Zeng, Q. Zou, *bioRxiv.* **2020**, https://doi.org/10.1101/2020.01.28.922922.

[16] A. Fischer, M. Sellner, S. Neranjan, M. A. Lill, M. Smieško, *ChemRxiv.* **2020**, Preprint, https://doi.org/10.26434/chemrxiv.11923239.v1.

[17] Z. Zhu, X. Wang, Y. Yang, Z. Zhang, K. Mu, Y. Shi, C. Peng, Z. Xu, W. Zhu, *ChemRxiv.* **2020**, https://doi.org/10.26434/chemrxiv.11959323.v1.

[18] F. Cheng, R. J. Desai, D. E. Handy, R. Wang, S. Schneeweiss, A. L. Barabási, J. Loscalzo, *Nat. Commun.* **2018**, *9*, 2691.

[19] T. Pillaiyar, S. Meenakshisundaram, M. Manickam, *Drug Discovery Today* **2020**, *25(4)*, 668–688.

[20] A. Zumla, J. F. Chan, E. I. Azhar, D. S. Hui, K. Y. Yuen, *Nat. Rev. Drug Discovery* **2016**, *15*, 327–347.

[21] CDC. Severe Acute Respiratory Syndrome. Available online: https://www.cdc.gov/sars/about/fs-sars.html (accessed on 20 December **2018**).

[22] WHO. Middle East Respiratory Syndrome Coronavirus and severe acute respiratory syndrome coronavirus (SARS-CoV) Available online: http://www.who.int/emergencies/mers-cov/en/ (accessed on 10 November **2018**).

[23] a) D. Wrapp, N. Wang, K. S. Corbett, J. A. Goldsmith, C.-L. Hsieh, O. Abiona, B. S. Graham, J. S. McLellan, *Science* **2020**, *367*, 1260–1263; b) B. J. Bosch, R. van der Zee, C. A. de Haan, P. J. Rottier, *J. Virol.* **2003**, *77*, 8801–8811.

[24] M. Stoermer, *ChemRxiv.* **2020**, Preprint. doi.org/10.26434/chemrxiv.11637294.v3.

[25] Y. M. Baez-Santos, S. E. St. John, A. D. Mesecar, *Antiviral Res.* **2014**, *115*, 21–38.

[26] A. R. Fehr, G. Jankevicius, I. Ahel, S. Perlman, *Trends Microbiol.* **2018**, *26*, 598–610.

[27] a) X. Liu, B. Zhang, Z. Jin, H. Yang, Z. Rao, **2020**, pdb id: 6LU7. https://doi:10.2210/pdb6LU7/pdb; b) A. D. Mesecar, S. St John. **2020**, pdb id: 6 W79. doi: 10.2210/pdb6 W79/pdb; c) L. Fu, L. Y. Feng, **2020**, pdb id: 7 C6S. doi: 10.2210/pdb7 C6S/pdb; d) K. Yang, W. Liu, **2020**, pdb id: 7JQ4. doi: 10.2210/pdb7JQ4/pdb; e) K. Yang. W. Liu, **2020**, pdb id: 7JPZ. doi: 10.2210/pdb7JPZ/pdb.

[28] J. Shang, G. Ye, K. Shi, Y. S. Wan, H. Aihara, F. Li, **2020**, pdb id: 6VW1. https://doi:10.2210/pdb6VW1/pdb.

[29] Y. Kim, R. Jedrzejczak, N. Maltseva, M. Endres, A. Godzik, K. Michalska, A. Joachimiak, **2020**, pdb id: 6VWW. doi:10.2210/pdb6VWW/pdb.

[30] L. Zhang, X. Sun, R. Hilgenfeld, **2020**, pdb id: 6Y2E. doi:10.2210/pdb6Y2E/pdb.

[31] D. L. McKee, A. Sternberg, U. Stange, S. Laufer, C. Naujokat, *Pharmacol. Res.* **2020**, *157*, 104859.

[32] L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox, R. Hilgenfeld, *Science* **2020**, *368*, 409–412.

[33] S. A. Khan, K. Zia, S. Ashraf, R. Uddin, Z. Ul-Haq, *J. Biomol. Struct. Dyn.* **2020**, 1–10.

[34] N. Muralidharan, R. Sakthivel, D. Velmurugan, M. M. Gromiha, *J. Biomol. Struct. Dyn.* **2020**, 1–6.

[35] S. Beura, P. Chetti, *J. Biomol. Struct. Dyn.* **2020**, doi.org/10.1080/07391102.2020.1772111.

[36] M. Wang, R. Cao, L. Zhang, X. Yang, J. Liu, M. Xu, Z. Shi, Z. Hu, W. Zhong, G. Xiao, *Cell Res.* **2020**, 269–271.

[37] L. Guangdi, E. D. Clercq, *Nat. Rev. Drug Discovery* **2020**, *19*, 149–150.

[38] A. Singha, A. Mishra, *J. Biomol. Struct. Dyn.* **2020**, doi.org/10.1080/07391102.2020.1777903.

[39] J. Gao, Z. Tian, X. Yang, *Biosci. Trends* **2020**, *14*, 72–73.

[40] P. Colson, J. M. Rolain, J. C. Lagier, P. Brouqui, D. Raoult, *Int. J. Antimicrob. Agents* **2020**, *55*, 105932.

[41] Schrödinger Release 2020–2: Glide, Schrödinger, LLC, New York, NY, **2020**.

[42] W. C. Wright, J. Chenge, J. Wang, H. M. Girvan, L. Yang, S. C. Chai, A. D. Huber, J. Wu, P. O. Oladimeji, A. W. Munro, T. Chen, *J. Med. Chem.* **2020**, *63*, 1415–1433.

[43] R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin, D. T. Mainz, *J. Med. Chem.* **2006**, *49*, 6177–6196.

[44] Schrödinger Release 2020–2: LigPrep, Schrödinger, LLC, New York, NY, **2020**.

[45] W. C. Wright, J. Chenge, J. Wang, H. M. Girvan, L. Yang, S. C. Chai, A. D. Huber, J. Wu, P. O. Oladimeji, A. W. Munro, T. Chen, *J. Med. Chem.* **2020**, *63*, 1415–1433.

[46] G. Minuesa, S. K. Albanese, W. Xie, Y. Kazansky, D. Worroll, A. Chow, A. Schurer, S.-M. Park, C. Z. Rotsides, J. Taggart, A. Rizzi, L. N. Naden, T. Chou, S. Gourkanti, D. Cappel, M. C. Passarelli, L. Fairchild, C. Adura, J. F. Glickman, J. Schulman, C. Famulare, M. Patel, J. K. Eibl, G. M. Ross, S. Bhattacharya, D. S. Tan, C. S. Leslie, T. Beuming, D. J. Patel, Y. Goldgur, J. D. Chodera, M. G. Kharas, *Nature Commun.* **2019**, *10*, 2691–2705.

[47] Schrödinger Suite **2020**. Protein Preparation Wizard, Epik: Schrödinger, LLC, New York, NY, **2020**.

[48] J. R. Greenwood, D. Calkins, A. P. Sullivan, J. C. Shelley, *J. Comput.-Aided Mol. Des.* **2010**, *24*, 591–604.

[49] a) S. F. Giardina, D. S. Werner, M. Pingle, P. B. Feinberg, K. W. Foreman, D. E. Bergstrom, L. D. Arnold, F. Barany, *J. Med. Chem.* **2020**, *63*, 3004–

3027; b) J. C. Shelley, A. Cholleti, L. Frye, J. R. Greenwood, M. R. Timlin, M. Uchimaya, *J. Comp. Aided Mol. Design* **2007**, *21*, 681–691.

[50] Impact: Schrödinger, LLC, New York, NY, 2020; Prime: Schrödinger, LLC, New York, NY, **2020**.

[51] A. J. Clark, C. Negron, K. Hauser, M. Sun, L. Wang, R. Abel, R. A. Friesner, *J. Mol. Biol.* **2019**, *431*, 1481–1493.

[52] D. Sen, P. Debnath, B. Debnath, S. Bhaumik, S. Debnath, *J. Biomol. Struct. Dyn.* **2020**, doi: 10.1080/07391102.2020.1819883.

[53] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. *J. Comput. Chem.* **2009**, *30*, 2785–2791.

[54] QikProp, Schrödinger, LLC, New York, NY, **2020**.

[55] L. Pérez-Benito, H. Keränen, H. van Vlijmen, G. Tresadern, *Sci. Rep.* **2018**, *8*, 4883.

[56] H. S. Yu, K. Modugula, O. Ichihara, K. Kramschuster, S. Keng, R. Abel, L. Wang, *J. Chem. Theory Comput.* **2021**, *17*, 450–462.

[57] K. J. Bowers, E. Chow, H. Xu, R. O. Dror, M. P. Eastwood, B. A. Gregersen, J. L. Klepeis, I. Kolossvary, M. A. Moraes, F. D. Sacerdoti, J. K. Salmon, Y. Shan, D. E. Shaw, Proceedings of the ACM/IEEE Conference on Super-computing (SC06), Tampa, Florida, **2006**, November 11–17.

[58] C. Selvaraj, R. B. Priya, J. K. Lee, S. K. Singh, *RSC Adv.* **2015**, *5*, 100498–100510.

[59] J. Y. Lee, J. Kim, H. Kim, M. C. Min Suh, *ACS Appl. Mater. Interfaces* **2020**, *12*, 23244–23251.

[60] R. Kumari, R. Kumar, A. Lynn, *J. Chem. Inf. Model.* **2014**, *54*, 1951–1962.

[61] H. Yang, M. Yang, Y. Ding, Y. Liu, Z. Lou, Z. Zhou, L. Sun, L. Mo, S. Ye, H. Pang, G. F. Gao, K. Anand, M. Bartlam, R. Hilgenfeld, Z. Rao, *PNAS USA* **2003**, *100*, 13190–13195.

[62] C. N. Chen, C. P. Lin, K. K. Huang, W. C. Chen, H. P. Hsieh, P. H. Liang, J. T. Hsu, *Ecam* **2005**, *2*, 209–215.