Identification of potential anti-COVID-19 drug leads from Medicinal Plants through Virtual High-Throughput Screening

Rohoullah Firouzi[1*] and Mitra Ashouri[2*]

[1] Department of Physical Chemistry, Chemistry and Chemical Engineering Research Center of Iran, Tehran, Iran.

[2] Department of Physical Chemistry, School of Chemistry, College of Science, University of Tehran, Tehran, Iran.

Corresponding Authors:

* rfirouzi@ccerci.ac.ir and firouzi.chemist@yahoo.com

* m.ashouri@ut.ac.ir

Abstract:

Natural compounds are widely used as attractive and valuable starting points for drug lead discovery. The present study aims to identify phytochemical compounds found in medicinal plants as potential COVID-19 inhibitors, using ensemble docking simulations. To this end, a phytochemical library from the PHCD database – a database of natural chemical compositions of Persian medicinal herbs (https://persianherb.com) – have been virtually screened against four key protein targets in the SARS-CoV-2 life cycle – the M^{pro} and PL^{pro} proteases and the Spike and human ACE2 proteins. Several potential antiviral lead candidates have been identified based on the "Computational Multitarget Screening" approach, in which favourite candidates interact simultaneously with all four targets. Four of the bioactive phytochemicals identified – Chelidimerine, Gallagyldilacton, Hinokiflavone, and Physalin Z – show the highest binding affinities to all the targets and are suggested to be the best choices for drug design research. Also, several important medicinal plants, including *Chelidonium majus* L., *Punica granatum*, *Rhus coriaria*, *Capparis spinose*, *Cichorium intybus*, and *Cynara scolymus*, with the most phytochemicals interacting with all the host and viral proteins, have been identified that can be considered as the most important herbal resources for drug development with the medicinal plant formulations against COVID-19.

Keywords: COVID-19, Multitarget Screening, Ensemble Docking Simulations, Medicinal Plants, Drug Discovery

1. Introduction

Despite previous warnings of scientists about the possible emergence of a global viral pandemic, in 2020, the COVID-19 pandemic caused by the SARS-CoV-2 virus brought all countries to their knees and showed that the world with relatively empty-handed is poorly prepared to combat the pandemic and viral diseases. To tackle the highly complex challenge, many academic research laboratories and pharmaceutical companies from all over the world are involved in massive efforts to find effective vaccines and antiviral drugs against COVID-19. Fortunately, thanks to the admirable and round-the-clock efforts of the scientific community, the arrival of several successful vaccines (especially mRNA-based vaccines) at the beginning of 2021 and the approval of a couple of oral antiviral drugs (Molnupiravir and Paxlovid) at the end of the year have raised hopes for handling the pandemic in the near future (1-6).

Vaccines (and antibodies) target mainly Spike protein on the viral surface which is necessary for recognizing and binding to the human angiotensin-converting enzyme 2 (ACE2) receptor during viral entry (5-12). The antiviral drugs target some key proteins involved in the viral replication machinery, such as the RNA-dependent RNA polymerase (RdRp) which synthesizes viral RNA and two conserved viral cysteine proteases, main protease (M^{pro}) and papain-like protease (PL^{pro}) which process the polyprotein chains translated from the viral RNA and cleave the chains into functional proteins that are required for viral replication, transcription, and assembly (1,4,10-21). It is important to note that since there are many mutation-prone residues in different targets – especially in the viral Spike protein – thus a successful vaccine or drug should be less sensitive to the variants and be effective against these mutations. Therefore, researchers should examine regularly the efficiency of approved antivirals and vaccines against new mutations in the target proteins and continuously design new inhibitors to counter the threat of resistance-causing mutations (1-3,22).

The rational design of new drugs requires precise molecular-level structural information on the target proteins involved in the viral life cycle (14,23). The structural analysis of the proteins in complex with different types of inhibitors is critical for understanding the molecular mechanism of the inhibition of protein function, characterizing the complexity of protein–inhibitor interactions, and identifying active and potential binding/interacting sites. Currently, many available structural and functional data on the target proteins – obtained from experimental and computational methods – have facilitated the structure-based drug design against COVID-19 and have enabled active researchers in the field to analyze and uncover many details of SARS-CoV-2 targets for developing new inhibitors. For more detailed discussion and explanation, the reader is directed to Refs. 9, 19-21, 24-37.

In structure-guided drug discovery campaigns, such structural studies are combined with the computational screening of ligand libraries to rapidly identify potent lead inhibitors against SARS-CoV-2 targets and subsequently, complemented by *in vitro*, *in vivo*, and trial studies to confirm their antiviral activities (9,15,17,20-21,38-44). Various new sophisticated computational

technologies and modeling approaches, such as quantum computing, massively parallel processing, Graphical Processing Units (GPU)-based algorithms, and artificial intelligence (AI) models are now being applied to accelerate computational drug design, especially for high-throughput virtual screening of very large databases and compound libraries (15,20,44-53).

In recent years, natural compounds from terrestrial and marine sources have been widely used as promising starting points for drug discovery projects, due to their high chemical and structural diversity (54-63). Among the natural compounds, phytochemicals from medicinal plants and herbs have attracted the most attention and have been extensively studied and tested for their diverse biological activities and drug-like properties (21,64-72). During the COVID-19 pandemic, several herbal compound databases derived from the medicinal plants of various geographical regions, such as China (65,69,73-76), India (68,77), Vietnam (78), Korea (65), Jordan (79), Africa (80), and Brazil (81) have been evaluated, both computationally and experimentally, for their potential antiviral activities against the SARS-CoV-2. Accordingly, it has been reported that multiple phytochemicals from various compound classes, including polyphenols (55,60,62,63,68,78,82-84), alkaloids (57,63,66-69,85), terpenes (55,63,66,67,69,84,86) and flavonoids (9,55,60,63,66-69,71,76,82-84,87) show a striking antiviral activity against the virus.

It is important to note that the phytochemical compounds present in the plants are highly dependent on climate conditions which means that the phytoconstituents of the same plant may vary in different regions of Earth. Therefore, the screening of new phytochemical libraries from different regions of the world provides new potential opportunities for finding novel drug candidates (89,89). Recently, we have designed and developed a searchable database (called PHCD) containing useful information about 312 famous Persian medicinal herbs and their phytochemical constituents (is freely available at https://persianherb.com) (90). This database contains structural and chemical information about more than 5,500 chemical compounds, about 10% of which are not included in any other database. In this research, the antiviral activity of the phytochemical library from the PHCD database against four key protein targets in the SARS-CoV-2 life cycle – the M^{pro} and PL^{pro} proteases and the Spike and human ACE2 proteins – have been studied and considered. Several potential antiviral drug candidates have been identified based on the "Computational Multitarget Screening" approach (91-93), in which favourite candidates interact simultaneously with multiple targets.

2. Methods

2.1. Protein Structures Preparation

The starting point of this research is the preparation and characterization of three-dimensional structure files in the RCSB PDB database (94) for the target proteins. The number of PDB

structures downloaded and analyzed for the M^{pro} and PL^{pro} proteases and the Spike and ACE2 proteins are 251, 45, 336, and 34, respectively. It should be noted that the structures studied for these two proteases are all X-ray crystal structures with a resolution of less than 3 Å, but for the other two proteins due to the lack of structures obtained by X-ray crystallography, structures resolved via cryo-EM have also been selected for the following analyses (for a complete list of the PDB IDs, see Table **S1** in the Supporting Information).

The PDB files were processed and analyzed according to the following procedure. In the first step, the crucial structural information available for each chain in the PDB entries was extracted. For this purpose, mutated/modified residues, missing atoms/residues, ligands information (names/chains, atomic coordinates, and molecular sizes), and some essential crystallographic data (like the resolution values and alternate locations) were identified and documented for all PDB chains. The next step is to split the PDB files into individual chains for performing the structural alignment of the monomeric chains, which is necessary for clustering and identifying representative structures for docking and virtual screening calculations. Different considerations have been made for each target protein, which are discussed in the following.

For the M^{pro}, 291 monomeric chains were obtained from the original 251 PDB files. To build reliable structural binding-site alignment, the intersection of residues involved in the formation of two antiparallel β-barrels between domains I and II for all monomers of M^{pro} which contains 66 residues were selected based on "SHEET" records in the PDB files (See Table **S2** in the Supporting Information for a list of residues contributing to the β-barrel structures). The 87 monomeric chains were identified from the 45 PDB files of the PL^{pro}, which 36 of the chains were discarded due to having the missing functionally important segments in the catalytic pocket of the PL^{pro} (which includes a Cys111-His272-Asp286 catalytic triad). The intersection of residues involved in the formation of β-sheets in the thumb and palm domains which contain 41 residues was used to drive the structural alignment process (Table **S2**). All M^{pro} and PL^{pro} monomeric chains were aligned on the Cα atoms of the selected residues of their corresponding reference structures, 6LU7 chain A and 6WRH chain A, respectively. The former reference is the first released protein crystal structure for SARS-CoV-2 M^{pro} with a resolution of 2.16 Å and the latter is a high-quality crystal structure of the PL^{pro} at 1.6 Å resolution. The pairwise RMSD values for all above Cα atoms between every pair of the aligned protein structures of the M^{pro} and PL^{pro} did not exceed 0.43 and 0.48 Å, respectively.

The 801 monomeric chains were recognized for the original 336 PDB files of the Spike protein. Due to the inherent flexibility of this protein, most of the chains are of low quality (resolutions up to 12 Å). Therefore, only the chains with a resolution of less than 3 Å (209 chains) were selected for the next steps. Some of these chains contain missing residues at the receptor-binding domain (RBD) (sequences 305 to 530) (19,28,29,38,95-100). The 41 Spike chains with more than 10 missing residues at the RBD domain (more than 5% of the total sequence) were also removed from these 209 selected chains. The remaining 168 protein chains share the same amino acid sequences of 11 residues with β-sheet contents (Table **S2**). The structural alignment was

performed on the Cα atoms of these 11 residues. For the ACE2 protein, the 34 PDBs were analyzed – all of them in complex with the RBD domain of the Spike protein – and in total 58 monomeric chains of ACE2 were extracted. The 52 residues with the helical contents in the α-helix binding domain of the ACE2 protein which directly interacts with the RBD were used to perform the structural alignment of the monomeric chains (Table **S2**). The crystal structure 6M0J (at 2.45 Å resolution) of human ACE2 (chain A) in complex with the Spike RBD (chain E) were used as the reference structures for the structural alignment of all monomeric chains of the ACE2 and Spike proteins. The pairwise RMSD values for all above Cα atoms between every pair of the aligned protein structures of the ACE2 and Spike did not exceed 1.00 and 0.70 Å, respectively.

2.2. Residual Binding Spot Detection

In the structural alignment step, all solvent molecules (water and organic solvents like diglycol, dimethyl sulfoxide, glycerol, and 1,2-ethanediol), metals, and small ions (such as chloride, nitrate, and acetate) were removed from all monomeric chains, except the co-crystallized ligands in the M^{pro} and PL^{pro} complexes. The location of the substrate-binding sites of the two proteases has been defined based on residues around the co-crystal ligands (3.5 Å) in different protein-ligand complexes. To this end, first, the center of mass of the crystallographic ligands in the aligned protein-ligand complexes has been calculated. The principal component analysis (PCA) on the center of mass of the ligands was used to define the binding pocket of both the proteases (for a detailed discussion see, for example, section **2.2** of Ref. 32). According to the analysis, out of a total of 245 ligands present in the PDBs of the M^{pro} complexes, 194 ligands have been localized on the same location in the crystal structure of the complexes - the region between domains I and II, which is known as the substrate-binding site of the M^{pro} (7,9,12,14,101-103). For the PL^{pro} complexes, out of a total of 41 identified ligands, 33 ligands have been located in the active site of the protease - in a cleft between the thumb and palm domains (9,14,20,104-107). The remaining ligands (51 out of 245 for the M^{pro} and 8 out of 41 for the PL^{pro}) have been sparsely distributed in other regions of the surface of the protein and therefore were excluded from further analyses. In the next step, all important amino acid residues interacting with the ligands/inhibitors inside the binding site were identified. For this purpose, the distances between the heavy atoms of the ligands/inhibitors and the residues of the proteases were calculated and then, a list of binding residues for each complex (194 and 33 complexes selected for the M^{pro} and PL^{pro} complexes, respectively) was generated for which at least one heavy atom of the residues falls within a cutoff distance (less than 3.5 Å in this study) of any ligand heavy atom. The union of all obtained list of residues which contains 30 and 21 residues for the M^{pro} and PL^{pro}, respectively, were considered as the binding sites of the proteases. The selected residues for defining the binding site have been tabulated in Table **S3** in the Supporting Information. It is important to note that the catalytic residues Cys145 and His41 in the M^{pro} and Cys111 and His272 in the PL^{pro} which play a key role in the proteolytic cleavage of the viral polyproteins, present in the above lists of the selected residues (14,20,76,102-104).

For the other two proteins, the binding regions were identified based on the contact interface between the residues of the Spike and ACE2 chains (within the cutoff distance of 4.5 Å) in the Spike RBD-ACE2 complex PDBs. From the structural analyses of 58 Spike RBD-ACE2 complexes, 17 hot spot residues of the Spike RBD were detected which interact with the ACE2 chain in more than 80% of the analyzed complexes, while there are 15 interacting residues of the ACE2 which make a network of inter-residue contacts with the Spike RBD chain in more than 80% of the complexes. The identified interacting residues were considered as potential binding sites of the Spike and ACE2 chains (Table **S3**). The found binding spot residues for the two proteins are very consistent with recent computational findings and experimental observations (19,28,29,38,95-98).

2.3. Protein Clustering

In this step, the pairwise RMSD matrices between the binding site residues of all aligned monomeric chains for each target protein were calculated and employed to identify the representative protein structures using the single-linkage hierarchical agglomerative clustering method based on the Ward variance minimization algorithm (108,109). It is interesting to mention that the number of matrix elements for each of the M^{pro}, PL^{pro}, Spike, and ACE2 proteins is (291 × 291 × 30), (51 × 51 × 21), (168 × 168 × 17), and (58 × 58 × 15), respectively, in which the first two numbers indicate the number of aligned monomeric chains and the third numbers denote the number of the binding site residues, as detected in the previous section (see Table **S3**).

Membership in clusters depended on the simultaneous fulfillment of two conditions: first that the number of pairwise RMSD values more than 2.0 Å for each pair of residues between members of one cluster should not be more than a predefined number of residues and, second, that average pairwise RMSD values of each residue between all pairs of members in one cluster should be less than a chosen cutoff value. The success of this strategy in protein clustering and the selection of the representative protein structures with the maximum conformational diversity of the binding site has recently been shown in different molecular docking studies (32,110). The 291 monomeric M^{pro} chains were classified into eight clusters, when the values of the two thresholds were set to seven numbers for the first criterion (the number of residues with RMSD > 2.0 Å between members of each cluster) and 1.3 Å for the second criterion (average RMSD value over the residues of each member in a cluster). Similarly, eight clusters have been identified for the 51 monomeric PL^{pro} chains when the values of the two thresholds were set to three numbers and 1.2 Å for the first and second criteria, respectively.

All 168 and 58 monomeric chains of the Spike and ACE2 proteins were clustered into ten and five groups, respectively, by setting the first and second criteria to seven numbers and 2.5 Å for the Spike and three numbers and 1.6 Å for the ACE2. It should be noted that the selection of more clusters for the Spike protein is due to its high flexibility. The results of clustering have been depicted as a dendrogram representation in **Figure 1**. The structure with the lowest average

RMSD value with respect to all the other members has been selected as the representative structure of a given cluster. The high-quality structures with better resolution have been introduced as the representative structure for the two-membered clusters. For each cluster, PDB IDs of the representative structures have been listed in **Figure 1**.

2.4. Representative Structures Preparation

The representative protein structures were protonated by the REDUCE program (111). The hydrogenated structures were partially relaxed for optimizing the hydrogen atom positions and removing steric clashes or close contacts in the crystal structures using the NAMD.2.13 program (112) with the CHARMM36m force field and generalized Born implicit solvent (GBIS). All the structures were minimized using 20000 steps of conjugate gradient minimization, involving 10000 steps minimization with the protein heavy atoms restrained harmonically using a force constant of 200 (kcal/mol)/$Å^2$ followed by 10000 steps with a harmonic positional restraint of 100 (kcal/mol)/$Å^2$ on all heavy atoms. All the optimized representative structures of the target proteins (31 protein chains in total) were converted into PDBQT format for performing molecular docking simulations.

2.5. Ligands Preparation

All 5546 natural compounds from the PHCD Database (extracted from Persian medicinal herbs) (90) were geometrically optimized in the gas phase using the PM7 (113) semi-empirical quantum mechanics (SQM) method with MOPAC2016 (114). It should be added here that the PM7 is a fast and successful SQM method which reliably describes various types of noncovalent interactions and some important chemical observations, such as the planarity of conjugated rings or molecular fragments (110). The gradient norm was set to 10 kcal $mol^{-1}Å^{-1}$. To ensure that no chemical bond breaking/formation processes take place during the optimization calculations, InChIKey identifiers - generated with InChI software - were generated and compared for the initial structures and final optimized structures. Furthermore, the structures with more than 20 rotatable bonds were excluded from subsequent calculations, due to the well-known fact that the docking success rates decrease with increasing the number of active rotatable bonds (115,116). The remaining optimized compounds (4892 chemical compounds) were prepared in PDBQT format for the virtual screening process.

2.6. Docking Setup and Protocol

In this research, all the 4892 phytochemicals were docked individually to each of the selected representative structures of the protein targets (8, 8, 10, and 5 representative structures for the M^{pro}, PL^{pro}, Spike, and ACE2 proteins, respectively, as described in sections **2.3**). It is noteworthy to point out that the use of an ensemble of multiple rigid receptor conformations for one target protein in docking simulations, often referred to as ensemble docking, is the most common strategy to incorporate the receptor flexibility in the docking that achieves better enrichment than rigid receptor docking to any of the individual members of the ensemble

(17,20,32,117,118). However, in the absence of a common strategy to choose the representative docking poses from ensemble docking results, we have very recently presented a new strategy to pick up the most appropriate docking poses from the ensemble docking results (32). In our proposed protocol, all predicted poses of a given ligand against an ensemble of multiple different conformations of a receptor are collected in a pool of ligand conformations and clustered to identify representative poses. The top-ranked poses (with the lowest-energy poses) from *the first* and *the most populated clusters* are chosen as representative poses of the ligand. In addition, it has been shown that the top-ranked poses of *the most populated clusters* obtained by AutoDock Vina show a very good performance in estimating binding poses and affinity ranking for the available experimental data for M^{pro}-ligand complexes (32).

Accordingly, AutoDock Vina software (version 1.1.2) (119) was used for the molecular docking simulations. The exhaustiveness parameter was set to 200. It seems useful to recall that the default exhaustiveness value is 8, and increasing this to higher values enhances the exploration of the conformational space of the ligand during the docking procedure and increases the probability of finding the proper ligand conformations (120-122). Each Vina run generates 20 poses. Next, all predicted docking poses for each phytochemical – 160 (=8×20) poses for the M^{pro} and 160 (=8×20), 200 (=10×20), and 100 (=5×20) poses for the PL^{pro}, Spike, and ACE2 proteins, respectively – were collected separately for each protein targets and re-clustered based on the *symmetry-corrected* heavy-atom root mean square deviation (RMSD) algorithm implemented in AutoDock4 with an RMSD cutoff of 2.0 Å (32,110,123,124). Then, the top-ranked poses of *the first* and *the most populated clusters* were selected as the representative docking poses of each phytochemical in complex with each of four protein targets for further analyses.

The docking search space for the M^{pro} and PL^{pro} proteases was constructed based on the Cartesian coordinates of the co-crystal ligands found inside the catalytic binding site of the aligned protein-ligand complexes – as described in section **2.2**. The initial docking box covers all the bound ligands with a wide range of sizes (from small molecules to large peptidomimetic inhibitors) at different locations of the catalytic binding site. Then, the box size was extended by 5 Å in each of the three dimensions to ensure that the docking search space is large enough for exploring possible binding conformations of new ligands (119,125). The dimensions and center coordinates of the final docking box have been tabulated in Table **S4** in the Supporting Information.

The docking search space for the other two proteins (Spike and ACE2) were defined based on the binding spot residues identified in the Spike RBD-ACE2 complexes – as described in section **2.2** (Table **S3**). First, an initial docking box containing all of the binding spot residues was constructed for each of these proteins and then, the box size was increased by 5 Å in each of the three dimensions. Due to the large surface of the interaction site of these two proteins in their complexes with each other, and consequently estimating an unwillingly narrow docking search space, the center of the docking box has been displaced in the direction of the Spike RBD-ACE2

interface to create a larger search space for the conformational poses generated during the docking process, without increasing the docking box dimensions (for the docking box information, see Table **S4** in the Supporting Information). The final docking boxes on a superposition of optimized representative protein structures have been displayed graphically in Figure **S5** in the Supporting Information.

3. Results and discussion

As mentioned above, the computational pipeline employed in this research (including the structural clustering strategy to construct the protein ensemble for performing docking calculations and the proposed manner to choose the representative docking poses from the ensemble docking results) has already been designed for correctly predicting experimental binding poses and affinity ranking of M^{pro}-ligand complexes (32) and thus can be utilized to properly produce a rank-ordered list of the phytochemicals of the PHCD Database, according to the Autodock Vina scoring function of the representative docking poses. Consequently, two rank-ordered lists of the docked phytochemicals have been produced for each protein target, one results from the representative poses of *the first clusters* and the other from the representative poses of *the most populated clusters*. The top 10 phytochemicals in the rank-ordered lists of the representative poses of *the first* and *the most populated clusters* for the investigated targets, along with their plant sources are shown in **Table 1** (the top 100 compounds of ranked lists are given in Tables **S6-S13** in the Supporting Information).

The analysis and comparison of the data in the tables show that the same phytochemicals and medicinal plants appear among the rank-ordered lists of the top 100 phytochemicals of the protein targets. In total, 12 common phytochemicals were found in all tables obtained from the representative poses of *the first clusters* (Tables **S6, S8, S10, and S12**) and 7 common phytochemicals were identified in all tables related to *the most populated clusters* (Tables **S7, S9, S11, and S13**). From a computational viewpoint, it means that these phytochemicals simultaneously target the four protein targets and can be considered as potential antiviral drug candidates against various key protein targets in the SARS-CoV-2 life cycle. In addition, 21 and 24 common medicinal plants were identified in the tables related to *the first* and *the most populated clusters*, respectively. The common phytochemicals and medicinal plants for the two rank-ordered lists related to *the first* and *the most populated clusters* are summarized in **Table 2** and the two-dimensional (2D) chemical structures of the multi-target phytochemicals identified are depicted in **Figure 2**.

The Vina scores of the common phytochemicals and their positions in the ranked lists of the top 100 phytochemicals are shown in **Table 3**. The first observation from **Table 3** is that most of the phytochemicals exhibit higher binding affinities for both proteases (especially the M^{pro}) than the two other targets, the Spike and ACE2. A reasonable explanation for this observation is that the

docked ligands should be placed inside the substrate-binding pocket of the proteases and are able to interact with a large number of residues of different regions of the binding pocket surface, while the docked ligands to the Spike and ACE2 targets experience a relatively large flat surface with few or no binding (sub)pockets and consequently, the ligands interact with a few residues of limited regions of the protein surface. For comparison purposes, the Vina scores of the top-ranked poses (with the lowest-energy) belonging to *the first* and *the most populated clusters* for all docked phytochemicals have been summarized in Figure **S15** in the Supporting Information. The highest negative Vina score, which indicates the maximum predicted binding affinity, for the M^{pro}, PL^{pro}, Spike, and ACE2 proteins are -12.1, -10.2, -9.9, and -8.9 $kcal.mol^{-1}$, respectively. It should be added here that the magnitude of the Vina score values calculated for the top-ranked phytochemicals (in **Table 3** and Tables **S6-S13** in the Supporting Information) may be compared with the Vina score values calculated from virtual screening of several sets of FDA-approved drugs and natural compounds against the two proteins M^{pro} and Spike which have recently been reported in some researches (80,126-132). The information about ligand names, the Vina score values, and literature sources are documented from the original papers and given in Table **S14** in the Supporting Information. By comparing the Vina score values reported in **Tables 3** and **S14**, it can be seen obviously that the common phytochemicals identified in this study (**Table 3**) reveal higher binding affinities to the two protein targets than previous literature data on known drugs (Table **S14**). Therefore, it seems that the introduced multi-target phytochemicals in this research with a high tendency to simultaneously target the host and viral proteins create better drug lead candidates against COVID-19.

Another important observation from **Table 3** is that 4 out of 13 phytochemicals – Chelidimerine, Gallagyldilacton, Hinokiflavone, and Physalin Z – show higher binding affinities to all the four protein targets and occupy better ranking positions in their ranked lists (numbers in parentheses) compared to the other phytochemicals. Thus, they are suggested to be the best choices for drug design research. It is very interesting to note that for *Chelidonium majus* L., the only herbal source of Chelidimerine, significant *in vitro* inhibitory activity against cysteine proteinases has already been reported (133). In addition, some Physalin derivatives have very recently been introduced as potent M^{pro} inhibitors (78,134,135) and Hinokiflavone which belongs to biflavonoid compounds, has already been identified as the antiviral potential of H1N1 influenza inhibitor (55). Even more interesting is the fact that, 5 out of 13 phytochemicals – two flavonoid glycosides (6 and 7 in **Figure 2**), two biflavonoid compounds (8 and 13), and Pongamoside A (11) – contain the flavonoid scaffold which is well-known for its striking antiviral potential against diverse coronaviruses (9,55,60,63,66-69,71,76,82-84,87). Two compounds 1 (a steroidal lactone) and 10 have also been proposed as dual inhibitors targeting both the Spike RBD and M^{pro} proteins (68,72,78,136,137). However, to the best of our knowledge, none of the five phytochemicals 3, 4, 9, 11, and 12 have been studied or reported as a potent antiviral agent to date, and then their observed antiviral activity against COVID-19 are suggested for the first time in the present virtual screening study.

As an important note, it should be pointed out that identifying *Capparispine 26-O-beta-D-glucoside* from the ranked lists related to *the most populated clusters* (in **Table 3**) as a potential antiviral agent against all the four protein targets (especially against ACE2 and M^{pro}) and not being seen in the *first clusters* indicates the importance of the selection of the lowest-energy pose in *the most populated cluster* as an important representative pose of the ensemble docking results. Therefore, just taking the top-ranked poses with the lowest-energy (the same representative pose of the *first cluster*) as the best solution of docking calculations, the possibility of identifying some important phytochemicals is lost.

Finally, to determine the importance of common medicinal plants introduced in **Table 2**, the numbers of observations of the medicinal plants in each ranked list (Tables **S6-S13**) have been summarized in **Table 4**. The medicinal plants observed only in one of the two lists related to *the first* and *the most populated clusters* have been discarded from **Table 4**. Assuming that the more phytochemicals belonging to a medicinal plant in the ranked list(s), the more attractive the medicinal plant for the development of potential antiviral drug candidates, the medicinal plants were arranged based on the sum of the number of its phytochemicals observed in the four corresponding lists. Of course, it should also be added that the possible synergistic effects of multiple bioactive phytochemicals belonging to a medicinal plant that can act simultaneously on different key protein targets in the viral life cycle enhance the importance of such a medicinal plant. According to the hypothesis, *Chelidonium majus* L. with the maximum number of its phytochemicals over all the ranked lists of the top 100 phytochemicals (in total, 25 and 23 times in the lists belonging to *the first* and *the most populated clusters*, respectively) can be considered as the most important herbal resource for drug design targeting the key viral proteins, the M^{pro}, PL^{pro}, and Spike RBD (only a couple of phytochemicals obtained from the plant present in ACE2 lists). The next three important medicinal plants, including *Punica granatum*, *Rhus coriaria*, and *Capparis spinose*, with the largest number of their phytochemicals interact simultaneously with all the host and viral proteins and thus, they are suggested to possess antiviral effects against SARS-CoV-2.

4. Conclusion:

Currently, the use of computational techniques for antiviral drug discovery from compound libraries and databases is one of the most powerful tools for combating the COVID-19 pandemic. In this work, the antiviral activity of the phytochemical library from the PHCD database (https://persianherb.com) (90) against four key protein targets in the SARS-CoV-2 life cycle – the M^{pro} and PL^{pro} proteases and the Spike and human ACE2 proteins – have been studied and considered using a new successful computational pipeline in the framework of the ensemble docking strategy.

The computational pipeline employed in this research, including the structural clustering strategy to construct the protein ensemble for performing docking calculations and the proposed manner to choose the representative docking poses from the ensemble docking results, has recently been designed for correctly predicting experimental binding poses and affinity ranking of M^{pro}-ligand complexes (32) and has now been utilized to properly rank about 5,000 phytochemical compounds of the PHCD database during their screening against different protein targets.

Several potential antiviral lead candidates have been identified based on the "Computational Multitarget Screening" approach, in which favourite candidates interact simultaneously with all protein targets. Four of the bioactive phytochemicals identified – Chelidimerine, Gallagyldilacton, Hinokiflavone, and Physalin Z – with different chemical scaffolds show the highest binding affinities to all the targets and are suggested to be the best choices for drug design research. Also, some important medicinal plants have been identified based on the numbers of their phytochemicals in the ranked lists of the top 100 phytochemicals for each protein target, assuming that the more phytochemicals belonging to a medicinal plant in the ranked lists, the more attractive the medicinal plant for drug development with the medicinal plant formulations. These important medicinal plants, including *Chelidonium majus* L., *Punica granatum*, *Rhus coriaria*, *Capparis spinose*, *Cichorium intybus*, and *Cynara scolymus*, with the most phytochemicals interacting with all the host and viral proteins, can be considered as promising potential herbal resources for drug discovery against COVID-19.

Clearly, the plant-based antiviral lead candidates identified in this research should be further evaluated by comprehensive molecular dynamics (MD) simulations, experimental assays, and clinical trials to confirm their actual activity against COVID-19. At present, we are collaborating with another academic research laboratory at Western University for performing large-scale MD simulations of the bioactive phytochemicals identified. We hope that these findings may contribute to the rational drug design against COVID-19.

Acknowledgments

RF thanks National Institute of Genetic Engineering and Biotechnology (NIGEB) for financial support. RF is also grateful to Dr Mohammad Hossein Karimi-Jafari and Dr Saleh Bagheri (both at University of Tehran) and Dr Mehrdad Karimi (Tehran University of Medical Science) for helpful comments.

Conflicts of Interest Statement

There are no conflicts of interest to declare.

Data and Software Availability

MOPAC package and AutoDock Vina (version 1.1.2) were used under a free academic license for ligands preparation and docking simulations. Produced and analyzed data are available upon request.

Author Information
Rohoullah Firouzi
Department of Physical Chemistry, Chemistry and Chemical Engineering Research Center of Iran, Tehran, Iran
Email: rfirouzi@ccerci.ac.ir
orcid.org/0000-0003-2385-831X

Mitra Ashouri
Department of Physical Chemistry, School of Chemistry, College of Science, University of Tehran, Tehran, Iran
Email: m.ashouri@ut.ac.ir
orcid.org/0000-0002-3263-2067

References:

(1) Max Kozlov, Why Scientists Are Racing To Develop More COVID Antivirals, *Nature* **2022**, 601, 496. https://doi.org/10.1038/d41586-022-00112-8

(2) Megan Cully, A tale of two antiviral targets — and the COVID-19 drugs that bind them, *Nat. Rev. Drug Discov.* **2022**, 21, 3−5. https://doi.org/10.1038/d41573-021-00202-8

(3) Heidi Ledford, COVID antiviral pills: what scientists still want to know, *Nature* **2021**, 599, 358−359. https://doi.org/10.1038/d41586-021-03074-5

(4) Elie Dolgin, Stocking the shelves for the next pandemic, *Nature* **2021**, 592, 340−343. https://doi.org/10.1038/d41586-021-00958-4

(5) Tung Thanh Le, Zacharias Andreadakis, Arun Kumar, Raúl Gómez Román, Stig Tollefsen, Melanie Saville and Stephen Mayhew, The COVID-19 vaccine development landscape, *Nat. Rev. Drug Discov.* **2020**, 21, 305−306. https://doi.org/10.1038/d41573-020-00073-5

(6) Yingzhu Li, Rumiana Tenchov, Jeffrey Smoot, Cynthia Liu, Steven Watkins and Qiongqiong Zhou, A Comprehensive Review of the Global Efforts on COVID-19 Vaccine Development, *ACS Cent. Sci.* **2021**, 7, 512−533. https://doi.org/10.1021/acscentsci.1c00120

(7) Pritam Kumar Panda, Murugan Natarajan Arul, Paritosh Patel, Suresh K. Verma, Wei Luo, Horst-Günter Rubahn, Yogendra Kumar Mishra, Mrutyunjay Suar, and Rajeev Ahuja, Structure-

based drug designing and immunoinformatics approach for SARS-CoV-2, *Sci. Adv.* **2020**, 6, eabb8097. https://doi.org/10.1126/sciadv.abb8097

(8) Ariel H. Thames, Kristy L. Wolniak, Samuel I. Stupp, and Michael C. Jewett, Principles Learned from the International Race to Develop a Safe and Effective COVID-19 Vaccine, *ACS Cent. Sci.* **2020**, 6, 1341−1347. https://doi.org/10.1021/acscentsci.0c00644

(9) Antonio Francés-Monerris, Cécilia Hognon, Tom Miclot, Cristina García-Iriepa, Isabel Iriepa,Alessio Terenzi, Stéphanie Grandemange, Giampaolo Barone, Marco Marazzi, and Antonio Monari, Molecular Basis of SARS-CoV-2 Infection and Rational Design of Potential Antiviral Agents: Modeling and Simulation Approaches, *J. Proteome Res.* **2020**, 19, 4291−4315. https://dx.doi.org/10.1021/acs.jproteome.0c00779

(10) StephenPelly and Dennis Liotta, Potent SARS-CoV-2 Direct-Acting Antivirals Provide an Important Complement to COVID-19 Vaccines, *ACS Cent. Sci.* **2021**, 7, 396−399. https://dx.doi.org/10.1021/acscentsci.1c00258

(11) Guangdi Li and Erik De Clercq, Therapeutic options for the 2019 novel coronavirus (2019-nCoV), *Nat. Rev. Drug Discov.* **2020**, 19, 149−150. https://doi.org/10.1038/d41573-020-00016-0

(12) Jared S. Morse, Tyler Lalonde, Shiqing Xu, and Wenshe Ray Liu, Learning from the Past: Possible Urgent Prevention and Treatment Options for Severe Acute Respiratory Infections Caused by 2019-nCoV, *ChemBioChem* **2020**, 21,730 –738. https://doi.org/10.1002/cbic.202000047

(13) Krishna Kumar and Tania J. Lupoli, Exploiting Existing Molecular Scaffolds for Long-Term COVID Treatment, *ACS Med. Chem. Lett.* **2020**, 11, 1357−1360. https://doi.org/10.1021/acsmedchemlett.0c00254

(14) Hailei Su, Feng Zhou, Ziru Huang, Xiaohua Ma, Kathiresan Natarajan, Minchuan Zhang,Yong Huang, and Haibin Su, Molecular Insights into Small-Molecule Drug Discovery for SARS-CoV-2, *Angew. Chem. Int. Ed.* **2021**, 60, 9789 –9802. https://doi.org/10.1002/anie.202008835

(15) Eugene N. Muratov, Rommie Amaro, Carolina H. Andrade, Nathan Brown, Sean Ekins, Denis Fourches, Olexandr Isayev, Dima Kozakov, José L. Medina-Franco, Kenneth M. Merz, Tudor I. Oprea, Vladimir Poroikov, Gisbert Schneider, Matthew H. Todd, Alexandre Varnek, David A. Winkler, Alexey V. Zakharov, Artem Cherkasov, and Alexander Tropsha, A critical overview of computational approaches employed for COVID-19 drug discovery, *Chem. Soc. Rev.*, **2021**, 50, 9121–9151. https://doi.org/10.1039/d0cs01065k

(16) María-JesusPérez-Pérez, Juan-Carlos Saiz, Eva-María Priego, and Miguel A. Martín-Acebes, Antivirals against (Re)emerging Flaviviruses: Should We Target the Virus or the Host? *ACS Med. Chem. Lett.* **2022**, 13, 5−10. https://doi.org/10.1021/acsmedchemlett.1c00617

(17) How to Discover Antiviral Drugs Quickly, Jerry M. Parks and Jeremy C. Smith, *N. Engl. J. Med.* **2020**, 382, 2261−2264. https://doi.org/10.1056/NEJMcibr2007042

(18) Binquan Luan, Tien Huynh, Xuemei Cheng, Ganhui Lan, and Hao-Ran Wang, Targeting Proteases for Treating COVID-19, *J. Proteome Res.* **2020**, 19, 4316−4326. https://doi.org/10.1021/acs.jproteome.0c00430

(19) Siyu Xiu, Alexej Dick, Han Ju, Sako Mirzaie, Fatemeh Abdi, Simon Cocklin, Peng Zhan, and Xinyong Liu, Inhibitors of SARS-CoV-2 Entry: Current and Future Opportunities, *J. Med. Chem.* **2020**, 63, 12256−12274. https://doi.org/10.1021/acs.jmedchem.0c00502

(20) A. Acharya, R. Agarwal, M. B. Baker, J. Baudry, D. Bhowmik, S. Boehm, K. G. Byler, S. Y. Chen, L. Coates, C. J. Cooper, O. Demerdash, I. Daidone, J. D. Eblen, S. Ellingson, S. Forli, J. Glaser, J. C. Gumbart, J. Gunnels, O. Hernandez, S. Irle, D. W. Kneller, A. Kovalevsky, J. Larkin, T. J. Lawrence, S. LeGrand, S.-H. Liu, J.C. Mitchell, G. Park, J.M. Parks, A. Pavlova, L. Petridis, D. Poole, L. Pouchard, A. Ramanathan, D. M. Rogers, D. Santos-Martins, A. Scheinberg, A. Sedova, Y. Shen, J. C. Smith,M. D. Smith, C. Soto, A. Tsaris, M. Thavappiragasam, A. F. Tillack, J. V. Vermaas, V. Q. Vuong, J. Yin, S. Yoo, M. Zahran, and L. Zanetti-Polzi, Supercomputer-Based Ensemble Docking Drug Discovery Pipeline with Application to Covid-19, *J. Chem. Inf. Model.* **2020**, 60, 5832−5852. https://doi.org/10.1021/acs.jcim.0c01010

(21) Catherine S. Adamson, Kelly Chibale, Rebecca J. M. Goss, Marcel Jaspars, David J. Newman, and Rosemary A. Dorrington, Antiviral drug discovery: preparing for the next pandemic, *Chem. Soc. Rev.*, **2021**, 50, 3647–3655. https://doi.org/10.1039/d0cs01118e

(22) Rui Wang, Jiahui Chen, and Guo-Wei Wei, Mechanisms of SARS-CoV-2 Evolution Revealing Vaccine-Resistant Mutations in Europe and America, *J. Phys. Chem. Lett.* **2021**, 12, 11850−11857. https://doi.org/10.1021/acs.jpclett.1c03380

(23) Montserrat Bárcena, Christopher O. Barnes, Martin Beck, Pamela J. Bjorkman, Bruno Canard, George F. Gao, Yunyun Gao, Rolf Hilgenfeld, Gerhard Hummer, Ardan Patwardhan, Gianluca Santoni, Erica Ollmann Saphire, Christiane Schaftzel, Sharon L. Schendel, Janet L. Smith, Andrea Thorn, David Veesler, Peijun Zhang and Qiang Zhou, Structural biology in the fight against COVID-19, *Nat. Struct. Mol. Biol.* **2021**, 28, 2-7. https://doi.org/10.1038/s41594-020-00544-8

(22) Sean Ekins, Melina Mottin, Paulo R.P.S. Ramos, Bruna K.P. Sousa, Bruno Junior Neves, Daniel H. Foil, Kimberley M. Zorn, Rodolpho C. Braga, Megan Coffee, Christopher Southan, Ana C. Puhl, and Carolina Horta Andrade, Déjà vu: Stimulating open drug discovery for SARS-CoV-2, *Drug Discov. Today* **2020**, 25, 928−941. https://doi.org/10.1016/j.drudis.2020.03.019

(25) Danuta Witkowska, Mass Spectrometry and Structural Biology Techniques in the Studies on the Coronavirus-Receptor Interaction, *Molecules* **2020**, 25, 4133. https://doi.org/10.3390/molecules25184133

(26) Cong Xu, Yanxing Wang, Caixuan Liu, Chao Zhang, Wenyu Han, Xiaoyu Hong, Yifan Wang, Qin Hong, Shutian Wang, Qiaoyu Zhao, Yalei Wang, Yong Yang, Kaijian Chen, Wei Zheng, Liangliang Kong, Fangfang Wang, Qinyu Zuo, Zhong Huang, Yao Cong, Conformational dynamics of SARS-CoV-2 trimeric spike glycoprotein in complex with receptor ACE2 revealed by cryo-EM, *Sci. Adv.* **2021**, 7, eabe5575. https://doi.org/10.1126/sciadv.abe5575

(27) Yongfei Cai, Jun Zhang, Tianshu Xiao, Hanqin Peng, Sarah M. Sterling, Richard M. Walsh Jr., Shaun Rawson, Sophia Rits-Volloch, Bing Chen, Distinct conformational states of SARS-CoV-2 spike protein, *Science* **2020**, 369, 1586–1592. https://doi.org/10.1126/science.abd4251

(28) Senbiao Fang, Ruoqian Zheng, Chuqi Lei, Jianxin Wang, Ruiqing Zheng, and Min Li, Key residues influencing binding affinities of 2019-nCoV with ACE2 in different species, *Brief. Bioinformatics*, **2021**, 22, 963–975. https://doi.org/10.1093/bib/bbaa329

(29) Wenpeng Cao, Chuqiao Dong, Seonghan Kim, Decheng Hou, Wanbo Tai, Lanying Du, Wonpil Im, and X. Frank Zhang, Biomechanical characterization of SARS-CoV-2 spike RBD and human ACE2 protein-protein interaction, *Biophysical J.* **2021**, 120, 1011–1019. https://doi.org/10.1016/j.bpj.2021.02.007

(30) Long-Can Mei, Yin Jin, Zheng Wang, Ge-Fei Hao, and Guang-Fu Yang, Web resources facilitate drug discovery in treatment of COVID-19, *Drug Discov. Today* **2021**, 26, 2358-2366. https://doi.org/10.1016/j.drudis.2021.04.018

(31) Wanwisa Dejnirattisai, Daming Zhou, Helen M. Ginn, Helen M.E. Duyvesteyn, Piyada Supasa, James Brett Case, Yuguang Zhao, Thomas S. Walter, Alexander J. Mentzer, Chang Liu, Beibei Wang, Guido C. Paesen, Jose Slon-Campos, César López-Camacho, Natasha M. Kafai, Adam L. Bailey, Rita E. Chen, Baoling Ying, Craig Thompson, Jai Bolton, Alex Fyfe, Sunetra Gupta, Tiong Kit Tan, Javier Gilbert-Jaramillo, William James, Michael Knight, Miles W. Carroll, Donal Skelly, Christina Dold, Yanchun Peng, Robert Levin, Tao Dong, Andrew J. Pollard, Julian C. Knight, Paul Klenerman, and Nigel Temperton, The antigenic anatomy of SARS-CoV-2 receptor binding domain, *Cell* **2021**, 184, 2183–2200. https://doi.org/10.1016/j.cell.2021.02.032

(32) Rohoullah Firouzi, Mitra Ashouri, and Mohammad Hossein Karimi-Jafari, Structural insights into the substrate-binding site of main protease for the structure-based COVID-19 drug discovery, *Proteins*. **2022**, https://doi.org/10.1002/prot.26318

(33) Kaifu Gao, Rui Wang, Jiahui Chen, Jetze J. Tepe, Faqing Huang, and Guo-Wei Wei, Perspectives on SARS-CoV-2 Main Protease Inhibitors, *J. Med. Chem*. **2021**, 64, 16922−16955. https://doi.org/10.1021/acs.jmedchem.1c00409

(34) Chamandi S. Dampalla, Athri D. Rathnayake, Krishani Dinali Perera, Abdul-Rahman M. Jesri, Harry Nhat Nguyen, Matthew J. Miller, Hayden A. Thurman, Jian Zheng, Maithri M. Kashipathy, Kevin P. Battaile, Scott Lovell, Stanley Perlman, Yunjeong Kim, William C. Groutas, and Kyeong-Ok Chang, Structure-Guided Design of Potent Inhibitors of SARS-CoV-2 3CL Protease: Structural, Biochemical, and Cell-Based Studies, *J. Med. Chem.* **2021**, 64, 17846−17865. https://doi.org/10.1021/acs.jmedchem.1c01037

(35) Zheng Zhao and Philip E. Bourne, Structural Insights into the Binding Modes of Viral RNA-Dependent RNA Polymerases Using a Function-Site Interaction Fingerprint Method for RNA Virus Drug Discovery, *J. Proteome Res.* **2020**, 19, 4698−4705. https://doi.org/10.1021/acs.jproteome.0c00623

(36) Fabian Bylehn, Cintia A. Menendez, Gustavo R. Perez-Lemus, Walter Alvarado, and Juan J. de Pablo, Modeling the Binding Mechanism of Remdesivir, Favilavir, and Ribavirin to SARS-CoV-2 RNA-Dependent RNA Polymerase, *ACS Cent. Sci.* **2021**, 7, 164−174. https://doi.org/10.1021/acscentsci.0c01242

(37) Wanchao Yin, Xiaodong Luan, Zhihai Li, Ziwei Zhou, Qingxing Wang, Minqi Gao, Xiaoxi Wang, Fulai Zhou, Jingjing Shi, Erli You, Mingliang Liu, Qingxia Wang, Yi Jiang, Hualiang Jiang, Gengfu Xiao, Leike Zhang, Xuekui Yu, Shuyang Zhang, and H. Eric Xu, Structural basis for inhibition of the SARS-CoV-2 RNA polymerase by suramin, *Nat. Struct. Mol. Biol.* **2021**, 28, 319–325. https://doi.org/10.1038/s41594-021-00570-0

(38) Mahdi Ghorbani, Bernard R. Brooks, and Jeffery B. Klauda, Critical Sequence Hotspots for Binding of Novel Coronavirus to Angiotensin Converter Enzyme as Evaluated by Molecular Simulations, *J. Phys. Chem. B* **2020**, 124, 10034−10047. https://doi.org/10.1021/acs.jpcb.0c05994

(39) Pablo R. Arantes, Aakash Saha and Giulia Palermo, Fighting COVID-19 Using Molecular Dynamics Simulations, *ACS Cent. Sci*. **2020**, *6*, 1654−1656. https://doi.org/10.1021/acscentsci.0c01236

(40) Gennaro Ciliberto and Luca Cardone, Boosting the arsenal against COVID-19 through computational drug repurposing, *Drug Discov. Today*, **2020**, *25*, 946−948. https://doi.org/10.1016/j.drudis.2020.04.005

(41) Rommie E. Amaro and Adrian J. Mulholland, A Community Letter Regarding Sharing Biomolecular Simulation Data for COVID-19, *J. Chem. Inf. Model.* **2020**, *60*, 2653−2656. https://doi.org/10.1021/acs.jcim.0c00319

(42) Rommie E. Amaro and Adrian J. Mulholland, Biomolecular Simulations in the Time of COVID-19, and After, *Comput. Sci. Eng.* **2020**, *22*, 30−36. https://doi.org/10.1109/MCSE.2020.3024155

(43) Aditya K. Padhi, Soumya Lipsa Rath, and Timir Tripathi, Accelerating COVID-19 Research Using Molecular Dynamics Simulation, *J. Phys. Chem. B* **2021**, 125, 9078−9091. https://doi.org/10.1021/acs.jpcb.1c04556

(44) Kaifu Gao, Rui Wang, Jiahui Chen, Limei Cheng, Jaclyn Frishcosy, Yuta Huzumi, Yuchi Qiu, Tom Schluckbier, and Guo-Wei Wei, Methodology-centered review of molecular modeling, simulation, and prediction of SARS-CoV-2, *arXiv*, **2021**, 2102.00971v1. https://doi.org/10.48550/arXiv.2102.00971

(45) Ren Kong, Guangbo Yang, Rui Xue, Ming Liu, FengWang, Jianping Hu, Xiaoqiang Guo and Shan Chang, COVID-19 Docking Server: a meta server for docking small molecules, peptides and antibodies against potential targets of COVID-19, *Bioinformatics*, **2020**, 36, 5109–5111. https://doi.org/10.1093/bioinformatics/btaa645

(46) Garrett A. Stevenson, Derek Jones, Hyojin Kim, W. F. Drew Bennett, Brian J. Bennion, Monica Borucki, Feliza Bourguet, Aidan Epstein, Magdalena Franco, Brooke Harmon, Stewart He, Max P. Katz, Daniel Kirshner, Victoria Lao, Edmond Y. Lau, Jacky Lo, Kevin McLoughlin, Richard Mosesso, Deepa K. Murugesh, Oscar A. Negrete, Edwin A. Saada, Brent Segelke, Maxwell Stefan, Marisa W. Torres, Dina Weilhammer, Sergio Wong, Yue Yang, Adam Zemla, Xiaohua Zhang, Fangqiang Zhu, Felice C. Lightstone and Jonathan E. Allen, High-Throughput Virtual Screening of Small Molecule Inhibitors for SARS-CoV-2 Protein Targets with Deep Fusion Models, *arXiv* **2021**, 2104.04547v3. https://doi.org/10.48550/arXiv.2104.04547

(47) Christopher Monroe, Remote quantum computing is the future, *Nature*, **2020**, 583, 10. https://doi.org/10.1038/d41586-020-01937-x

(48) Xueyan Mei, Hao-Chih Lee, Kai-yue Diao, Mingqian Huang, Bin Lin, Chenyu Liu, Zongyu Xie, Yixuan Ma, Philip M. Robson, Michael Chung, Adam Bernheim, Venkatesh Mani, Claudia Calcagno, Kunwei Li, Shaolin Li, Hong Shan, Jian Lv, Tongtong Zhao, Junli Xia, Qihua Long, Sharon Steinberger, Adam Jacobi, Timothy Deyer, Marta Luksza, Fang Liu, Brent P. Little, Zahi A. Fayad and Yang Yang, Artificial intelligence–enabled rapid diagnosis of patients with COVID-19, *Nat. Med.* **2020**, 26, 1224−1228. https://doi.org/10.1038/s41591-020-0931-3

(49) Xiangxiang Zeng, Xiang Song, Tengfei Ma, Xiaoqin Pan, Yadi Zhou, Yuan Hou, Zheng Zhang, Kenli Li, George Karypis, and Feixiong Cheng, Repurpose Open Data to Discover Therapeutics for COVID-19 Using Deep Learning, *J. Proteome Res.* **2020**, 19, 4624−4636. https://doi.org/10.1021/acs.jproteome.0c00316

(50) Govinda B. KC, Giovanni Bocci, Srijan Verma, Md Mahmudulla Hassan, Jayme Holmes, Jeremy J. Yang, Suman Sirimulla, and Tudor I. Oprea, A machine learning platform to estimate anti-SARS-CoV-2 activities, *Nat. Mach. Intell.* **2021**, 3, 527–535. https://doi.org/10.1038/s42256-021-00335-w

(51) Agastya P. Bhati, Shunzhou Wan, Dario Alfè, Austin R. Clyde, Mathis Bode, Li Tan, Mikhail Titov, Andre Merzky, Matteo Turilli, Shantenu Jha, Roger R. Highfield, Walter Rocchia, Nicola Scafuri, Sauro Succi, Dieter Kranzlmüller, Gerald Mathias, David Wifling, Yann Donon, Alberto Di Meglio, Sofia Vallecorsa, Heng Ma, Anda Trifan, Arvind Ramanathan, Tom Brettin, Alexander Partin, Fangfang Xia, Xiaotan Duan, Rick Stevens, and Peter V. Coveney, Pandemic Drugs at Pandemic Speed: Accelerating COVID-19 Drug Discovery with Hybrid Machine Learning- and Physics-based Simulations on High Performance Computers, *arXiv*, **2021**, 2103.02843v2. https://doi.org/10.1098/rsfs.2021.0018

(52) Scott LeGrand, Aaron Scheinberg, Andreas F. Tillack, Mathialakan Thavappiragasam, Josh V. Vermaas, Rupesh Agarwal, Jeff Larkin, Duncan Poole, Diogo Santos-Martins, Leonardo Solis-Vasquez, Andreas Koch, Stefano Forli, Oscar Hernandez, Jeremy C. Smith, and Ada Sedova, GPU-Accelerated Drug Discovery with Docking on the Summit Supercomputer: Porting, Optimization, and Application to COVID-19 Research, *arXiv*, **2020**, 2007.03678v1. https://doi.org/10.1145/3388440.3412472

(53) Mohit Pandey, Michael Fernandez, Francesco Gentile, Olexandr Isayev, Alexander Tropsha, Abraham C. Stern, and Artem Cherkasov, The transformational role of GPU computing and deep learning in drug discovery, *Nat. Mach. Intell.* **2022**, 4, 211–221. https://doi.org/10.1038/s42256-022-00463-x

(54) Alan L. Harvey, Natural products in drug discovery. *Drug Discov. Today*, **2008**, 13, 894–901. https://doi.org/10.1016/j.drudis.2008.07.004

(55) Ulrike Grienke, Michaela Schmidtke, Susanne von Grafenstein, Johannes Kirchmair, Klaus R. Liedl and Judith M. Rollinger, Influenza neuraminidase: A druggable target for natural products, *Nat. Prod. Rep.* **2012**, 29, 11–36. https://doi.org/10.1039/c1np00053e

(56) Alexey A. Lagunin, Rajesh K. Goel, Dinesh Y. Gawande, Priynka Pahwa, Tatyana A. Gloriozova, Alexander V. Dmitriev, Sergey M. Ivanov, Anastassia V. Rudik, Varvara I. Konova, Pavel V. Pogodin, Dmitry S. Druzhilovsky, and Vladimir V. Poroikov, Chemo- and bioinformatics resources for *in silico* drug discovery from medicinal plants beyond their traditional use: a critical review. *Nat. Prod. Rep.* **2014**, 13, 1585−1611. https://doi.org/10.1039/c4np00068d

(57) J. P. Martinez, F. Sasse, M. Brönstrup, J. Diezc, and A. Meyerhans, Antiviral drug discovery: broad-spectrum drugs from nature, *Nat. Prod. Rep.*, **2015**, 32, 29−48. https://doi.org/10.1039/c4np00085d

(58) Carlos Jiménez, Marine Natural Products in Medicinal Chemistry, *ACS Med. Chem. Lett.* **2018**, 9, 959−961. https://doi.org/10.1021/acsmedchemlett.8b00368

(59) Thomas E. Smith, Christopher D. Pond, Elizabeth Pierce, Zachary P. Harmer, Jason Kwan, Malcolm M. Zachariah, Mary Kay Harper, Thomas P. Wyche, Teatulohi K. Matainaho, Tim S. Bugni, Louis R Barrows, Chris M. Ireland, and Eric W. Schmidt, Accessing chemical diversity from the uncultivated symbionts of small marine animals, *Nat. Chem. Biol.* **2018**, 14, 179–185. https://doi.org/10.1038/nCHeMBIO.2537

(60) Mitchell P. Christy, Yoshinori Uekusa, Lena Gerwick, and William H. Gerwick, Natural Products with Potential to Treat RNA Virus Pathogens Including SARS-CoV-2, *J. Nat. Prod.* **2021**, 84, 161−182. https://doi.org/10.1021/acs.jnatprod.0c00968

(61) Avinash Mishra, Wajihul Hasan Khan, and Anurag S. Rathore, Synergistic Effects of Natural Compounds Toward Inhibition of SARS-CoV-2 3CL Protease, *J. Chem. Inf. Model.* **2021**, 61, 5708−5718. https://doi.org/10.1021/acs.jcim.1c00994

(62) Jaime Rubio-Martínez, Ana Jiménez-Alesanco, Laura Ceballos-Laita, David Ortega-Alarcón, Sonia Vega, Cristina Calvo, Cristina Benítez, Olga Abian, Adrián Velázquez-Campoy, Timothy M. Thomson, José Manuel Granadino-Roldán, Patricia Gómez-Gutiérrez, and Juan J. Pérez, Discovery of Diverse Natural Products as Inhibitors of SARS-CoV-2 Mpro Protease through Virtual Screening, *J. Chem. Inf. Model.* **2021**, 61, 6094−6106. https://doi.org/10.1021/acs.jcim.1c00951

(63) Rudra Chakravarti, Rajveer Singh, Arijit Ghosh, Dhritiman Dey, Priyanka Sharma, Ravichandiran Velayutham, Syamal Roy and Dipanjan Ghosh, A review on potential of natural products in the management of COVID-19, *RSC Adv.*, **2021**, 11, 16711–16735. https://doi.org/10.1039/d1ra00644d

(64) Muhammad T. Islam, Chandan Sarka, Dina M. El-Kersh, Sarmin Jamadda, Shaikh J. Uddin, Jamil A. Shilpi, and Mohammad S. Mubarak, Natural products and their derivatives against coronavirus: A review of the non-clinical and pre-clinical data. *Phytother. Res.* **2020**, 34, 2471–2492. https://doi.org/10.1002/ptr.6700

(65) Lin Ang, Hye Won Lee, Jun Yong Choi, Junhua Zhang, and Myeong Soo Lee, Herbal medicine and pattern identification for treating COVID-19: a rapid review of guidelines. *Integr. Med. Res.* **2020**, 9, 100407. https://doi.org/10.1016/j.imr.2020.100407

(66) Ahmed M. Sayed, Amira R. Khattab, Asmaa M. AboulMagd, Hossam M. Hassan, Mostafa E. Rateb, Hala Zaidg, and Usama Ramadan Abdelmohsen, Nature as a treasure trove of potential anti-SARS-CoV drug leads: a structural/mechanistic rationale, *RSC Adv.*, **2020**, 10, 19790–19802. https://doi.org/10.1039/D0RA04199H

(67) Salar Hafez Ghoran, Mohamed El-Shazly, Nazim Sekeroglu, and Anake Kijjoa, Natural Products from Medicinal Plants with Anti-Human Coronavirus Activities, *Molecules* **2021**, 26, 1754. https://doi.org/10.3390/molecules26061754

(68) Arumugam Vijaya Anand, Balasubramanian Balamuralikrishnan, Mohandass Kaviya, Kathirvel Bharathi, Aluru Parithathvi, Meyyazhagan Arun, Nachiappan Senthilkumar, Shanmugam Velayuthaprabhu, Muthukrishnan Saradhadevi, Naif Abdullah Al-Dhabi, Mariadhas Valan Arasu, Mohammad Iqbal Yatoo, Ruchi Tiwari, and Kuldeep Dhama, Medicinal Plants, Phytochemicals, and Herbs to Combat Viral Pathogens Including SARS-CoV-2, *Molecules* **2021**, 26, 1775. https://doi.org/10.3390/molecules26061775

(69) Jia-Tsrong Jan, Ting-Jen Rachel Cheng, Yu-Pu Juang, Hsiu-Hua Ma, Ying-Ta Wu, Wen-Bin Yang, Cheng-Wei Cheng, Xiaorui Chen, Ting-Hung Chou, Jiun-Jie Shie, Wei-Chieh Cheng, Rong-Jie Chein, Shi-Shan Mao, Pi-Hui Liang, Che Ma, Shang-Cheng Hung, and Chi-Huey Wong, Identification of existing pharmaceuticals and herbal medicines as inhibitors of SARS-CoV-2 infection, *PNAS* **2021**, 118, e2021579118. https://doi.org/10.1073/pnas.2021579118

(70) Eyana Thomas, Laura E. Stewart, Brien A. Darley, Ashley M. Pham, Isabella Esteban, and Siva S. Panda, Plant-Based Natural Products and Extracts: Potential Source to Develop New Antiviral Drug Candidates, *Molecules* **2021**, 26, 6197. https://doi.org/10.3390/molecules26206197

(71) Myriam Merarchi, Namrata Dudha, Bhudev C Das, and Manoj Garg, Natural products and phytochemicals as potential anti-SARS-CoV-2 drugs, *Phytother. Res.* **2021**, 35, 5384–5396. https://doi.org/10.1002/ptr.7151

(72) Rambod Abiri, Hazandy Abdul-Hamid, Oksana Sytar, Ramin Abiri, Eduardo Bezerra de Almeida, Jr., Surender K. Sharma, Victor P. Bulgakov, Randolph R. J. Arroo, and Sonia Malik, A Brief Overview of Potential Treatments for Viral Diseases Using Natural Plant Compounds: The Case of SARS-Cov, *Molecules* **2021**, 26, 3868. https://doi.org/10.3390/molecules26133868

(73) Canrong Wu, Yang Liu, Yueying Yang, Peng Zhang, Wu Zhong, Yali Wang, Qiqi Wang, Yang Xu, Mingxue Li, Xingzhou Li, Mengzhu Zheng, Lixia Chen, Hua Li, Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods, *Acta Pharm. Sin. B* **2020**, 10, 766−788. https://doi.org/10.1016/j.apsb.2020.02.008

(74) Deng-hai Zhang, Kun-lun Wu, Xue Zhang, Sheng-qiong Deng, and Bin Peng, *In silico* screening of Chinese herbal medicines with the potential to directly inhibit 2019 novel coronavirus, *J. Integr. Med.* **2020**, 18, 152–158. https://doi.org/10.1016/j.joim.2020.02.005

(75) De-an Guo, Traditional Chinese medicine played a crucial role in battling COVID-19. *Chin. Herb. Med.* **2020**, 12, 205–206. https://doi.org/10.1016/j.chmed.2020.07.001

(76) Anuj Gahlawat, Navneet Kumar, Rajender Kumar, Hardeep Sandhu, Inder Pal Singh, Saranjit Singh, Anders Sjöstedt, and Prabha Garg, Structure-Based Virtual Screening to Discover Potential Lead Molecules for the SARS-CoV-2 Main Protease, *J. Chem. Inf. Model.* **2020**, 60, 5781–5793. https://doi.org/10.1021/acs.jcim.0c00546

(77) R.P. Vivek-Ananth, Abhijit Ran, Nithin Rajan, Himansu S. Biswal, and Areejit Samal, In Silico Identification of Potential Natural Product Inhibitors of Human Proteases Key to SARS-CoV-2 Infection, *Molecules* **2020**, 25, 3822. https://doi.org/10.3390/molecules25173822

(78) Son Tung Ngo, Ngoc Quynh Anh Pham, Ly Thi Le, Duc-Hung Pham, and Van V. Vu, Computational Determination of Potential Inhibitors of SARS-CoV-2 Main Protease, *J. Chem. Inf. Model.* **2020**, 60, 5771–5780. https://doi.org/10.1021/acs.jcim.0c00491

(79) Sherif S. Ebada, Nariman A. Al-Jawabri, Fadia S. Youssef, Dina H. El-Kashef, Tim-Oliver Knedel, Amgad Albohy, Michal Korinek, Tsong-Long Hwang, Bing-Hung Chen, Guan-Hua Lin, Chia-Yi Lin, Sa'ed M. Aldalaien, Ahmad M. Disi, Christoph Janiak, and Peter Proksch, Anti-inflammatory, antiallergic and COVID-19 protease inhibitory activities of phytochemicals from the Jordanian hawksbeard: identification, structure–activity relationships, molecular modeling and impact on its folk medicinal uses, *RSC Adv.*, **2020**, 10, 38128–38141. https://doi.org/10.1039/d0ra04876c

(80) Samuel K. Kwofie, Emmanuel Broni, Seth O. Asiedu, Gabriel B. Kwarko, Bismark Dankwa, Kweku S. Enninful, Elvis K. Tiburu, and Michael D. Wilson, Cheminformatics-Based Identification of Potential Novel Anti-SARS-CoV-2 Natural Compounds of African Origin, *Molecules* **2021**, 26, 406. https://doi.org/10.3390/molecules26020406

(81) Mohd Saeed, Amir Saeed, Md Jahoor Alam, and Mousa Alreshidi, Identification of Persuasive Antiviral Natural Compounds for COVID-19 by Targeting Endoribonuclease NSP15: A Structural-Bioinformatics Approach, *Molecules* **2020**, 25, 5657. https://doi.org/10.3390/molecules25235657

(82) Giovanna Giovinazzo, Carmela Gerardi, Caterina Uberti-Foppa, and Lucia Lopalco, Can Natural Polyphenols Help in Reducing Cytokine Storm in COVID-19 Patients? *Molecules* **2020**, 25, 5888. https://doi.org/10.3390/molecules25245888

(83) Simona Piccolella, Giuseppina Crescente, Shadab Faramarzi, Marialuisa Formato, Maria Tommasina Pecoraro, and Severina Pacifico, Polyphenols vs. Coronaviruses: How Far Has Research Moved Forward? *Molecules* **2020**, 25, 4103. https://doi.org/10.3390/molecules25184103

(84) Roberta Bernini and Francesca Velotti, Natural Polyphenols as Immunomodulators to Rescue Immune Response Homeostasis: Quercetin as a Research Model against Severe COVID-19, *Molecules* **2021**, 26, 5803. https://doi.org/10.3390/molecules26195803

(85) Amr El-Demerdash, Afnan Hassan, Tarek Mohamed Abd El-Aziz, James D. Stockand, and Reem K. Arafa, Marine Brominated Tyrosine Alkaloids as Promising Inhibitors of SARS-CoV-2, *Molecules* **2021**, 26, 6171. https://doi.org/10.3390/molecules26206171

(86) Mahmoud A. A. Ibrahim, Alaa H. M. Abdelrahman, Tarik A. Mohamed, Mohamed A. M. Atia, Montaser A. M. Al-Hammady, Khlood A. A. Abdeljawaad, Eman M. Elkady, Mahmoud F. Moustafa, Faris Alrumaihi, Khaled S. Allemailem, Hesham R. El-Seedi, Paul W. Paré, Thomas Efferth, and Mohamed-Elamir F. Hegazy, In Silico Mining of Terpenes from Red-Sea Invertebrates for SARS-CoV-2 Main Protease (Mpro) Inhibitors, *Molecules* **2021**, 26, 2082. https://doi.org/10.3390/molecules26072082

(87) Muchtaridi Muchtaridi, M. Fauzi, Nur Kusaira Khairul Ikram, Amirah Mohd Gazzali, and Habibah A. Wahab, Natural Flavonoids as Potential Angiotensin-Converting Enzyme 2 Inhibitors for Anti-SARS-CoV-2, *Molecules* **2020**, 25, 3980. https://doi.org/10.3390/molecules25173980

(88) George Nicola, Tiqing Liu, and Michael K. Gilson, Public Domain Databases for Medicinal Chemistry, *J. Med. Chem.* **2012**, 55, 6987−7002. https://doi.org/10.1021/jm300501t

(89) Maria Sorokina and Christoph Steinbeck, Review on natural products databases: where to find data in 2020, *J. Cheminform.* **2020**, 12, 20. https://doi.org/10.1186/s13321-020-00424-9

(90) Mohammad Hossein Karimi-Jafari, Rohoullah Firouzi, Mitra Ashouri, Atefeh Poursoleiman, PHCD: A database of chemical compositions of Persian medicinal herbs, *ChemRxiv*. **2022**, https://doi.org/10.26434/chemrxiv-2022-8rrwp

(91) Ekachai Jenwitheesuk, Jeremy A. Horst, Kasey L Rivas, Wesley C. Van Voorhis, and Ram Samudrala, Novel paradigms for drug discovery: computational multitarget screening, *Trends Pharmacol. Sci.* **2008**, 29, 62−71. https://doi.org/10.1016/j.tips.2007.11.007

(92) Johannes Kirchmair, Patrick Markt, Simona Distinto, Daniela Schuster, Gudrun M. Spitzer, Klaus R. Liedl, Thierry Langer, and Gerhard Wolber, The Protein Data Bank (PDB), Its Related Services and Software Tools as Key Components for In Silico Guided Drug Discovery, *J. Med. Chem.* **2008**, 51, 22, 7021–7040. https://doi.org/10.1021/jm8005977

(93) Hongguang Ma, Boshi Huang, and Yan Zhang, Recent advances in multitarget-directed ligands targeting G-protein-coupled receptors, *Drug Discov. Today* **2020**, 25, 1682–1692. https://doi.org/10.1016/j.drudis.2020.07.004

(94) Helen M. Berman, John Westbrook, Zukang Feng, Gary Gilliland, T. N. Bhat, Helge Weissig, Ilya N. Shindyalov, and Philip E. Bourne, The protein data bank, *Nucleic Acids Res.*, **2000**, 28, 235–242. https://doi.org/10.1093/nar/28.1.235

(95) Renhong Yan, Yuanyuan Zhang, Yaning Li, Lu Xia, Yingying Guo, Qiang Zhou, Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2, *Science* **2020**, 367, 1444–1448. https://doi.org/10.1126/science.abb2762

(96) Yi Zhang and Tatiana G. Kutateladze, Molecular structure analyses suggest strategies to therapeutically target SARS-CoV-2, *Nat. Commun.*, **2020**, 11, 2920. https://doi.org/10.1038/s41467-020-16779-4

(97) Chiduru Watanabe, Yoshio Okiyama, Shigenori Tanaka, Kaori Fukuzawaef, and Teruki Honma, Molecular recognition of SARS-CoV-2 spike glycoprotein: quantum chemical hot spot and epitope analyses, *Chem. Sci*., **2021**, 12, 4722−4739. https://doi.org/10.1039/d0sc06528e

(98) Jun Lan, Jiwan Ge, Jinfang Yu, Sisi Shan, Huan Zhou, Shilong Fan, Qi Zhang, Xuanling Shi, Qisheng Wang, Linqi Zhang, and Xinquan Wang, Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor, *Nature*, **2020**, 581, 215−220. https://doi.org/10.1038/s41586-020-2180-5

(99) Thassanai Sitthiyotha and Surasak Chunsrivirot, Computational Design of 25-mer Peptide Binders of SARS-CoV-2, *J. Phys. Chem. B* **2020**, 124, 10930−10942. https://doi.org/10.1021/acs.jpcb.0c07890

(100) Xiaoqiang Huang, Chengxin Zhang, Robin Pearce, Gilbert S. Omenn, and Yang Zhang, Identifying the Zoonotic Origin of SARS-CoV-2 by Modeling the Binding Affinity between the Spike Receptor-Binding Domain and Host ACE2, *J. Proteome Res.* **2020**, 19, 4844−4856. https://doi.org/10.1021/acs.jproteome.0c00717

(101) Zhenming Jin, Xiaoyu Du, Yechun Xu, Yongqiang Deng, Meiqin Liu, Yao Zhao, Bing Zhang, Xiaofeng Li, Leike Zhang, Chao Peng, Yinkai Duan, Jing Yu, Lin Wang, Kailin Yang, Fengjiang Liu, Rendi Jiang, Xinglou Yang, Tian You, Xiaoce Liu, Xiuna Yang, Fang Bai, Hong Liu, Xiang Liu, Luke W. Guddat, Wenqing Xu, Gengfu Xiao, Chengfeng Qin, Zhengli Shi, Hualiang Jiang, Zihe Rao, and Haitao Yang, Structure of M^{pro} from SARS-CoV-2 and discovery of its inhibitors, *Nature* **2020**, 582, 289–293. https://doi.org/10.1038/s41586-020-2223-y

(102) Daniel W. Kneller, Gwyndalyn Phillips, Hugh M. O'Neill, Robert Jedrzejczak, Lucy Stols, Paul Langan, Andrzej Joachimiak, Leighton Coates, and Andrey Kovalevsky, Structural plasticity of SARS-CoV-2 3CL M^{pro} active site cavity revealed by room temperature X-ray crystallography, *Nat. Commun.* **2020**, 11, 3202. https://doi.org/10.1038/s41467-020-16954-7

(103) Alice Douangamath, Daren Fearon, Paul Gehrtz, Tobias Krojer, Petra Lukacik, C. David Owen, Efrat Resnick, Claire Strain-Damerell, Anthony Aimon, Péter Ábrányi-Balogh, José Brandão-Neto, Anna Carbery, Gemma Davison, Alexandre Dias, Thomas D. Downes, Louise Dunnett, Michael Fairhead, James D. Firth, S. Paul Jones, Aaron Keeley, György M. Keserü, Hanna F. Klein, Mathew P. Martin, Martin E. M. Noble, Peter O'Brien, Ailsa Powell, Rambabu

N. Reddi, Rachael Skyner, Matthew Snee, Michael J. Waring, Conor Wild, Nir London, Frank von Delft, and Martin A. Walsh, Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease, *Nat. Commun.* **2020**, 11, 5047. https://doi.org/10.1038/s41467-020-18709-w

(104) Sk. Abdul Amin, Suvankar Banerjee, Kalyan Ghosh, Shovanlal Gayen, and Tarun Jha, Protease targeted COVID-19 drug discovery and its challenges: Insight into viral main protease (Mpro) and papain-like protease (PLpro) inhibitors, *Bioorg. Med. Chem.* **2021**, 29, 115860. https://doi.org/10.1016/j.bmc.2020.115860

(105) Xiaopan Gao, Bo Qin, Pu Chen, Kaixiang Zhu, Pengjiao Hou, Justyna Aleksandra Wojdyla, Meitian Wang, Sheng Cui, Crystal structure of SARS-CoV-2 papain-like protease, *Acta Pharm. Sin. B* **2021**, 11, 237–245. https://doi.org/10.1016/j.apsb.2020.08.014

(106) Mostafa Jamalan, Ebrahim Barzegari, and Fathollah Gholami-Borujeni, Structure-Based Screening to Discover New Inhibitors for Papain-like Proteinase of SARS-CoV-2: An *In Silico* Study, *J. Proteome Res.* **2021**, 20, 1015−1026. https://doi.org/10.1021/acs.jproteome.0c00836

(107) Theresa Klemm, Gregor Ebert, Dale J. Calleja, Cody C. Allison, Lachlan W. Richardson, Jonathan P. Bernardini, Bernadine G. C. Lu, Nathan W. Kuchel, Christoph Grohmann, Yuri Shibata, Zhong Yan Gan, James P. Cooney, Marcel Doerflinger, Amanda E. Au, Timothy R. Blackmore, Gerbrand J. van der Heden van Noort, Paul P Geurink, Huib Ovaa, Janet Newman, Alan Riboldi-Tunnicliffe, Peter E. Czabotar, Jeffrey P. Mitchell, Rebecca Feltham, Bernhard C. Lechtenberg, Kym N. Lowes, Grant Dewson, Marc Pellegrini, Guillaume Lessene, and David Komander, Mechanism and inhibition of the papain-like protease, PLpro, of SARS-CoV-2,

EMBO J. **2020**, 39, e106275. https://doi.org/10.15252/embj.2020106275

(108) Fionn Murtagh and Pedro Contreras, Algorithms for hierarchical clustering: an overview, II, *WIREs Data Mining Knowl. Discov*. **2017**, 7, e1219. http://doi.org/10.1002/widm.1219

(109) Naomi Altman and Martin Krzywinski, Clustering, *Nat. Methods*, **2017**, 14, 545−546. https://doi.org/10.1038/nmeth.4299

(110) Saleh Bagheri, Hassan Behnejad, Rohoullah Firouzi, and Mohammad Hossein Karimi-Jafari, Using the Semiempirical Quantum Mechanics in Improving the Molecular Docking: A Case Study with CDK2, *Mol. Inf.* **2020**, 39, 2000036. https://doi.org/10.1002/minf.202000036

(111) J. Michael Word, Simon C. Lovell, Jane S. Richardson, and David C. Richardson, Asparagine and Glutamine: Using Hydrogen Atom Contacts in the Choice of Side-Chain Amide Orientation. *J. Mol. Biol.* **1999**, 285, 1735−1747. https://doi.org/10.1006/jmbi.1998.2401

(112) James C Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad ajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D Skeel, Laxmikant Kalé, Klaus Schulten, Scalable

molecular dynamics with NAMD, *J. Comput. Chem.* **2005**, 26, 1781–1802. https://doi.org/10.1002/jcc.20289

(113) James J. P. Stewart, Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters, *J. Mol. Model.* **2013**, 19, 1–32. https://doi.org/10.1007/s00894-012-1667-x

(114) MOPAC2016, James J. P. Stewart, Stewart Computational Chemistry, Colorado Springs, CO, USA, HTTP://OpenMOPAC.net (**2016**).

(115) Douglas B. Kitchen, Hélène Decornez, John R. Furr, and Jürgen Bajorath, Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.* **2004**, 3, 935–949. https://doi.org/10.1038/nrd1549

(116) Emanuele Perola, W. Patrick Walters, Paul S. Charifson, A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins*. **2004**, 56, 235–249. https://doi.org/10.1002/prot.20088

(117) Ian R. Craig, Jonathan W. Essex, and Katrin Spiegel, Ensemble Docking into Multiple Crystallographically Derived Protein Structures: An Evaluation Based on the Statistical Analysis of Enrichments, *J. Chem. Inf. Model.* **2010**, 50, 511–524. https://doi.org/10.1021/ci900407c

(118) Shashidhar Rao, Paul C. Sanschagrin, Jeremy R. Greenwood, Matthew P. Repasky, Woody Sherman, and Ramy Farid, Improving database enrichment through ensemble docking, *J. Comput. Aided. Mol. Des.* **2008**, 22, 621–627.https://doi.org/10.1007/s10822-008-9182-y

(119) Oleg Trott and Arthur J. Olson, AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading, *J. Comput. Chem.* **2010**, 31, 455–461. https://doi.org/10.1002/jcc.21334

(120) Stefano Forli, Ruth Huey, Michael E. Pique, Michel F. Sanner, David S. Goodsell, and Arthur J. Olson, Computational protein–ligand docking and virtual drug screening with the AutoDock suite, *Nat. Protoc.* **2016**, 11, 905–919. https://doi.org/10.1038/nprot.2016.051

(121) Mohammad Mahdi Jaghoori, Boris Bleijlevens, and Silvia D. Olabarriaga, 1001 Ways to run AutoDock Vina for virtual screening, *J. Comput. Aided Mol. Des.* **2016**, 30, 237–249. https://doi.org/10.1007/s10822-016-9900-9

(122) Christoph Gorgulla, Andras Boeszoermenyi, Zi-Fu Wang, Patrick D. Fischer, Paul W. Coote, Krishna M. Padmanabha Das, Yehor S. Malets, Dmytro S. Radchenko, Yurii S. Moroz, David A. Scott, Konstantin Fackeldey, Moritz Hoffmann, Iryna Iavniuk, Gerhard Wagner, and Haribabu Arthanari, An open-source drug discovery platform enables ultra-large virtual screens, *Nature*, **2020**, 580, 663–668. https://doi.org/10.1038/s41586-020-2117-z

(123) William J. Allen and Robert C. Rizzo, Implementation of the Hungarian Algorithm to Account for Ligand Symmetry and Similarity in Structure-Based Design, *J. Chem. Inf. Model.* **2014**, 54, 2, 518–529. https://doi.org/10.1021/ci400534h

(124) Flavio Ballante and Garland R. Marshall, An Automated Strategy for Binding-Pose Selection and Docking Assessment in Structure-Based Drug Design, *J. Chem. Inf. Model.* **2016**, 56, 1, 54–72. https://doi.org/10.1021/acs.jcim.5b00603

(125) Garrett M. Morris, Ruth Huey, William Lindstrom, Michel F. Sanner, Richard K. Belew, David S. Goodsell, Arthur J. Olson, AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility, *J. Comput. Chem.* **2009**, 30, 2785–2791. https://doi.org/10.1002/jcc.21256

(126) Romulo O. Barros, Fabio L. C. C. Junior, Wildrimak S. Pereira, Neiva M. N. Oliveira, and Ricardo M. Ramos, Interaction of Drug Candidates with Various SARS-CoV-2 Receptors: An in Silico Study to Combat COVID-19, *J. Proteome Res.* **2020**, 19, 4567–4575. https://doi.org/10.1021/acs.jproteome.0c00327

(127) Rajib Islam, Md. Rimon Parves, Archi Sundar Paul, Nizam Uddin, Md. Sajjadur Rahman, Abdulla Al Mamun, Md. Nayeem Hossain, Md. Ackas Ali and Mohammad A. Halim, A molecular modeling approach to identify effective antiviral phytochemicals against the main protease of SARS-CoV-2, *J. Biomol. Struct. Dyn.* **2021**, 39, 3213–3224. https://doi.org/10.1080/07391102.2020.1761883

(128) Simona De Vita, Maria Giovanna Chini, Gianluigi Lauro and Giuseppe Bifulco, Accelerating the repurposing of FDA-approved drugs against coronavirus disease-19 (COVID-19), *RSC Adv.*, **2020**, 10, 40867–40875. https://doi.org/10.1039/d0ra09010g

(129) Sohini Chakraborti, Sneha Bheemireddy and Narayanaswamy Srinivasan, Repurposing drugs against the main protease of SARS-CoV-2: mechanism-based insights supported by available laboratory and clinical data, *Mol. Omics*, **2020**, 16, 474–491. https://doi.org/10.1039/d0mo00057d

(130) Liya Thurakkal, Satyam Singh, Rajarshi Roy, Parimal Kar, Sushabhan Sadhukhan and Mintu Porel, An *in-silico* study on selected organosulfur compounds as potential drugs for SARS-CoV-2 infection via binding multiple drug targets, *Chem. Phys. Lett.* **2021**, 763, 138193. https://doi.org/10.1016/j.cplett.2020.138193

(131) Gema Lizbeth Ramírez-Salinas, Marlet Martínez-Archundia, José Correa-Basurto and Jazmín García-Machorro, Repositioning of Ligands That Target the Spike Glycoprotein as Potential Drugs for SARS-CoV-2 in an In Silico Study, *Molecules* **2020**, 25, 5615. https://doi.org/10.3390/molecules25235615

(132) Preeti Pandey, Jitendra Subhash Rane, Aroni Chatterjee, Abhijeet Kumar, Rajni Khan , Amresh Prakash and Shashikant Ray, Targeting SARS-CoV-2 spike protein of COVID-19 with naturally occurring phytochemicals: an in silico study for drug development, *J. Biomol. Struct. Dyn.* **2021**, 39, 6306–6316. https://doi.org/10.1080/07391102.2020.1796811

(133) Boris Rogelj, Tatjana Popovič, Anka Ritonja, Borut Štrukelj, Jože Brzin, Chelidocystatin, a novel phytocystatin from Chelidonium majus, *Phytochemistry* **1998**, 94, 1645–1649. https://doi.org/10.1016/S0031-9422(98)00281-7

(134) Anna Maria Sardanelli, Camilla Isgrò, and Luigi Leonardo Palese, SARS-CoV-2 Main Protease Active Site Ligands in the Human Metabolome, *Molecules* **2021**, 26, 1409. https://doi.org/10.3390/molecules26051409

(135) Minh Quan Pham, Khanh B. Vu, T. Ngoc Han Pham, Le Thi Thuy Huong, Linh Hoang Tran, Nguyen Thanh Tung, Van V. Vu, Trung Hai Nguyen, and Son Tung Ngo, Rapid prediction of possible inhibitors for SARS-CoV-2 main protease using docking and FPL simulations, RSC Adv.,2020, 10,31991–31996. https://doi.org/10.1039/d0ra06212j

(136) Riddhidev Banerjee, Lalith Perera and L.M. Viranga Tillekeratne, Potential SARS-CoV-2 main protease inhibitors, *Drug Discov. Today* 26, 2021, 804–816. https://doi.org/10.1016/j.drudis.2020.12.005

(137) Suresh Gangadevi, Vishnu Nayak Badavath, Abhishek Thakur, Na Yin, Steven De Jonghe, Orlando Acevedo, Dirk Jochmans, Pieter Leyssen, Ke Wang, Johan Neyts, Tao Yujie, and Galia Blum, Kobophenol A Inhibits Binding of Host ACE2 Receptor with Spike RBD Domain of SARS-CoV-2, a Lead Compound for Blocking COVID-19. *J. Phys. Chem. Lett.* **2021**, 12, 1793−1802. https://doi.org/10.1021/acs.jpclett.0c03119

Table 1. The top 10 compounds from the PHCD database (obtained from the top-ranked poses of *the first* and *the most populated clusters*) against the protein targets, along with their plant sources.

Protein targets	Top 10 phytochemicals in the rank-ordered list belonging to *the first cluster*			Top 10 phytochemicals in the rank-ordered list belonging to *the most populated cluster*		
	Rank	**Chemical Name**	**Herb Name**	**Rank**	**Chemical Name**	**Herb Name**
M^{pro}	**1**	Chelidimerine	*Chelidonium majus* L.	**1**	Chelidimerine	*Chelidonium majus* L.
	2	Gallagyldilacton	*Punica granatum*	**2**	Gallagyldilacton	*Punica granatum*
	3	mulberrofuran G	*Morus nigra*	**3**	Pedunculagin	*Juglans regia*
	4	Pedunculagin	*Juglans regia*	**4**	Bisindigotin	*Isatis tinctoria*
	5	Granatin A	*Punica granatum*	**5**	Akyrogenin	*Polygonatum orientale*
	6	Bisindigotin	*Isatis tinctoria*	**6**	Capparispine	*Capparis spinosa*
	7	Dehydrated tergallic C-glucoside	*Ipomea purpura*	**7**	Sennidin A	*Cassia Italica*
	8	Cyclomorusin	*Morus alba*	**8**	cimiracemoside I	*Cimicifuga racemosa*
	9	Amentoflavon	*Ginkgo biloba*	**9**	Hinokiflavone	*Rhus coriaria*
	10	Scaiadopitysin	*Ginkgo biloba*	**10**	Procyanidin B2	*Ziziphus jujuba*
PL^{pro}	**1**	Amentoflavon	*Ginkgo biloba*	**1**	Amentoflavon	*Ginkgo biloba*
	2	Bisindigotin	*Isatis tinctoria*	**2**	Bisindigotin	*Isatis tinctoria*
	3	Hinokiflavone	*Rhus coriaria*	**3**	perylene	*Melissa officinalis*
	4	perylene	*Melissa officinalis*	**4**	Isocodonocarpine	*Capparis decidua*
	5	Isocodonocarpine	*Capparis decidua*	**5**	Pongamoside A	*Hyoscyamus niger*
	6	Pongamoside A	*Hyoscyamus niger*	**6**	temazepame	*Artemisia dracunculus*
	7	temazepame	*Artemisia dracunculus*	**7**	Hinokiflavone	*Rhus coriaria*
	8	physalin Z	*Physalis alkekengi*	**8**	Norsaguinarine	*Papaver bracteatum*
	9	Silybin B	*Silybum marianum*	**9**	Hispaglabridin B	*Glycyrrhiza glabra*
	10	Licoisoflavanone	*Glycyrrhiza glabra*	**10**	valoneic acid dilactone	*Oenothera biennis*
ACE2	**1**	Capparispine 26-O-beta-D-glucoside	*Capparis spinosa*	**1**	Capparispine 26-O-beta-D-glucoside	*Capparis spinosa*
	2	Hinokiflavone	*Rhus coriaria*	**2**	Hinokiflavone	*Rhus coriaria*
	3	Gallagyldilacton	*Punica granatum*	**3**	Gallagyldilacton	*Punica granatum*
	4	physalin Z	*Physalis alkekengi*	**4**	physalin Z	*Physalis alkekengi*
	5	Granatin A	*Punica granatum*	**5**	Granatin A	*Punica granatum*
	6	Neolutein b	*Pistacia vera*	**6**	Neolutein b	*Pistacia vera*
	7	cimiracemoside P	*Cimicifuga racemosa*	**7**	cimiracemoside P	*Cimicifuga racemosa*
	8	(all-E)-Lutein	*Isatis tinctoria*	**8**	(all-E)-Lutein	*Isatis tinctoria*
	9	Beta-carotene	*Pistacia vera*	**9**	Beta-carotene	*Pistacia vera*
	10	Lawsowaseem	*Lawsonia inermis*	**10**	Lawsowaseem	*Lawsonia inermis*
Spike	**1**	Chelidimerine	*Chelidonium majus* L.	**1**	Chelidimerine	*Chelidonium majus* L.
	2	15-N-acetyl capparisine	*Capparis decidua*	**2**	Hinokiflavone	*Rhus coriaria*
	3	Isoglaucanone	*Adiantum capillus-veneris*	**3**	9-cis-beta-Carotin	*Spinacia oleracea*
	4	Hinokiflavone	*Rhus coriaria*	**4**	Apigenin-7-O-rutinoside	*Theobroma cacao*
	5	Gallagyldilacton	*Punica granatum*	**5**	Gallagyldilacton	*Punica granatum*
	6	10-Hydroxy phaeophorbide	*Isatis tinctoria*	**6**	cimiracemoside I	*Cimicifuga racemosa*
	7	9-cis-beta-Carotin	*Spinacia oleracea*	**7**	Dihydrocoptisine	*Chelidonium majus* L.
	8	farnesiferone A	*Ferula persica*	**8**	Chelidonine	*Fumaria parviflora*
	9	Apigenin-7-O-rutinoside	*Theobroma cacao*	**9**	Atroposide A	*Hyoscyamus niger*
	10	Apigenin glucuronide	*Rhus coriaria*	**10**	Amentoflavon	*Ginkgo biloba*

Table 2. The list of common phytochemicals and medicinal plants for all tables related to *the first* and *the most populated clusters*.

Common phytochemicals in all tables related to *the first clusters*	Hinokiflavone, Gallagyldilacton, Chelidimerine, physalin Z, Bisindigotin, mulberrofuran G, Amentoflavo, Apigenin-7-O-rutinoside, Pedunculagin, Withanolide Luteolin-3'-O-di-rhamnoside-7-O-rhamnoside, Pongamoside A
Common phytochemicals in all tables related to *the most populated clusters*	Capparispine 26-O-beta-D-glucoside, Gallagyldilacton, mulberrofuran G, Chelidimerine, Luteolin-3'-O-di-rhamnoside-7-O-rhamnoside, Hinokiflavone, Bisindigotin
Common medicinal plants in all tables related to *the first clusters*	*Adiantum capillus-veneris, Capparis decidua, Capparis spinosa, Chelidonium majus* L., *Cichorium intybus, Cynara scolymus, Ginkgo biloba, Hyoscyamus niger, Ipomea purpura, Isatis tinctoria, Juglans regia, Morus nigra, Onopordon acanthium, Passiflora caerulea, Physalis alkekengi, Punica granatum, Rhus coriaria, Withania somnifera, Securigera securidaca* (L.) Degen & Dorfl, *Silybum marianum, Theobroma cacao*
Common medicinal plants in all tables related to *the most populated clusters*	*Adiantum capillus-veneris, Amygdalus communis, Capparis decidua, Capparis spinosa, Cassia Italica, Chelidonium majus* L., *Cichorium intybus, Crataegus microphylla, Cynara scolymus, Ginkgo biloba, Hyoscyamus niger, Ipomea purpura, Isatis tinctoria, Juglans regia, Morus alba, Morus nigra, Oenothera biennis, Onopordon acanthium, Passiflora caerulea, Punica granatum, Rhus coriaria, Securigera securidaca* (L.) Degen & Dorfl, *Silybum marianum, Theobroma cacao*

Table 3. Vina scores (in kcal.mol^{-1}) of the common phytochemicals and their positions (in parentheses) in the rank-ordered subsets of the top 100 phytochemicals.

Common phytochemicals in the top 100 compounds of ranked lists	Vina score and the position in the ranked lists (related to *the first* clusters)				Vina score and the position in the ranked lists (related to *the most populated* clusters)			
	M^{pro}	PL^{pro}	ACE2	Spike RBD	M^{pro}	PL^{pro}	ACE2	Spike RBD
Amentoflavon	-9.9 (9)	-10.2 (1)	-7.5 (100)	-8.2 (23)	–	–	–	–
Apigenin-7-O-rutinoside	-9.3 (56)	-8.8 (61)	-8.5 (6)	-8.5 (9)	–	–	–	–
Bisindigotin	-10.0 (6)	-10.2 (2)	-7.6 (87)	-8.4 (14)	-10.0 (4)	-10.2 (2)	-7.5 (61)	-8.0 (17)
Capparispine 26-O-beta-D-glucoside	–	–	–	–	-9.5 (11)	-8.5 (90)	-8.6 (1)	-7.7 (64)
Chelidimerine	-12.1 (1)	-9.4 (11)	-8.2 (13)	-9.9 (1)	-12.1 (1)	-8.5 (93)	-7.9 (18)	-9.9 (1)
Gallagyldilacton	-10.4 (2)	-9.1 (27)	-8.9 (1)	-8.7 (5)	-10.4 (2)	-9.1 (15)	-8.4 (3)	-8.5 (5)
Hinokiflavone	-9.5 (30)	-10.1 (3)	-8.5 (5)	-8.7 (4)	-9.5 (9)	-9.4 (7)	-8.5 (2)	-8.7 (2)
Luteolin-3'-O-di-rhamnoside-7-O-rhamnoside	-9.5 (38)	-9.2 (22)	-7.9 (38)	-7.9 (83)	-9.0 (51)	-8.7 (49)	-7.9 (17)	-7.9 (32)
Mulberrofuran G	-10.1 (3)	-9.2 (19)	-7.8 (47)	-8.1 (36)	-9.4 (14)	-8.8 (34)	-7.8 (22)	-7.8 (38)
Pedunculagin	-10.1 (4)	-9.0 (37)	-7.8 (54)	-7.9 (82)	–	–	–	–
Physalin Z	-9.6 (22)	-9.5 (8)	-8.3 (9)	-8.4 (11)	–	–	–	–
Pongamoside A	-9.1 (81)	-9.5 (6)	-7.5 (96)	-7.9 (66)	–	–	–	–
Withanolide	-9.2 (77)	-8.7 (83)	-7.7 (74)	-8.2 (32)	–	–	–	–

Table 4. The numbers of phytochemicals belonging to the medicinal plants in the ranked lists of the top 100 phytochemicals for each protein target (obtained from Tables **S6-S13**).

Common medicinal plants in all the top 100 compounds of ranked lists	**Number of observations of common medicinal plants in each ranked list (related to *the first* clusters)**					**Number of observations of common medicinal plants in each ranked list (related to *the most populated* clusters)**				
Medicinal Plant Names	**M^{pro}**	**PL^{pro}**	**ACE2**	**Spike**	**Total count**	**M^{pro}**	**PL^{pro}**	**ACE2**	**Spike**	**Total count**
Chelidonium majus L.	10	6	1	8	25	9	5	2	7	23
Punica granatum	7	4	6	6	23	6	5	5	6	22
Rhus coriaria	7	7	5	3	22	8	4	3	6	21
Capparis spinosa	4	2	5	10	21	4	3	7	6	20
Cichorium intybus	4	5	4	6	19	3	4	2	4	13
Cynara scolymus	3	1	4	11	19	2	1	3	6	12
Ipomea purpura	4	5	4	3	16	6	4	1	3	14
Theobroma cacao	1	4	3	7	15	3	3	2	6	14
Securigera securidaca (L.) Degen & Dorfl	4	3	4	4	15	7	3	1	3	14
Ginkgo biloba	6	6	2	1	15	3	3	1	4	11
Hyoscyamus niger	3	3	6	2	14	3	2	2	3	10
Capparis decidua	1	2	4	6	13	1	3	2	4	10
Adiantum capillus-veneris	2	2	4	4	12	4	1	2	2	9
Onopordon acanthium	2	1	2	6	11	2	1	1	5	9
Isatis tinctoria	2	4	3	1	10	1	3	2	2	8
Juglans regia	3	1	3	2	9	2	1	2	4	9
Silybum marianum	3	2	2	2	9	1	2	1	1	5
Morus nigra	3	1	1	3	8	2	1	1	3	7
Passiflora caerulea	3	1	1	2	7	2	1	1	1	5

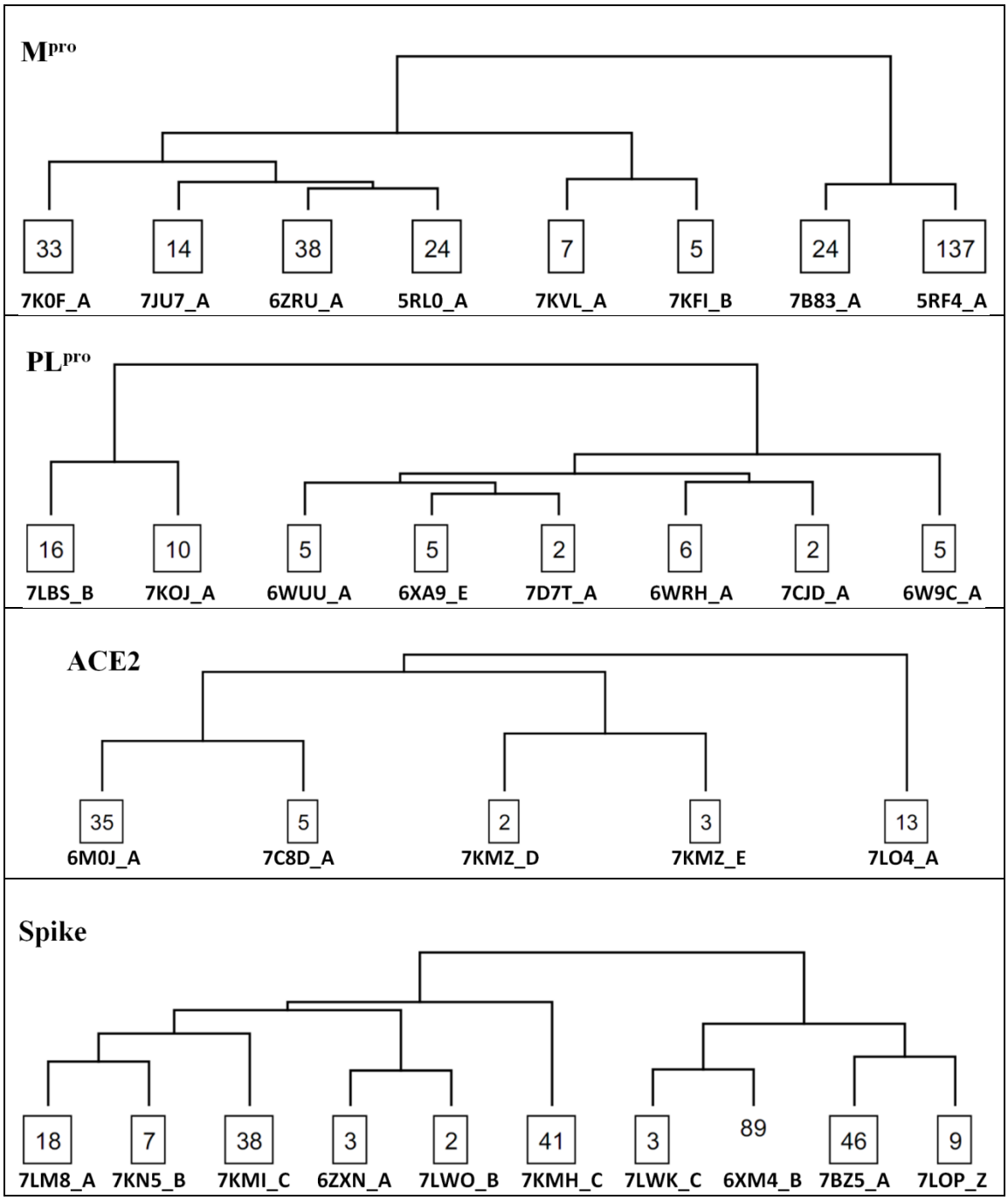

Figure 1. Bottom-up tree dendrogram of the clusters obtained using the Ward's hierarchical method. The population of each cluster is given in each box and the PDB IDs of the representative structures for each cluster are also displayed below the boxes.

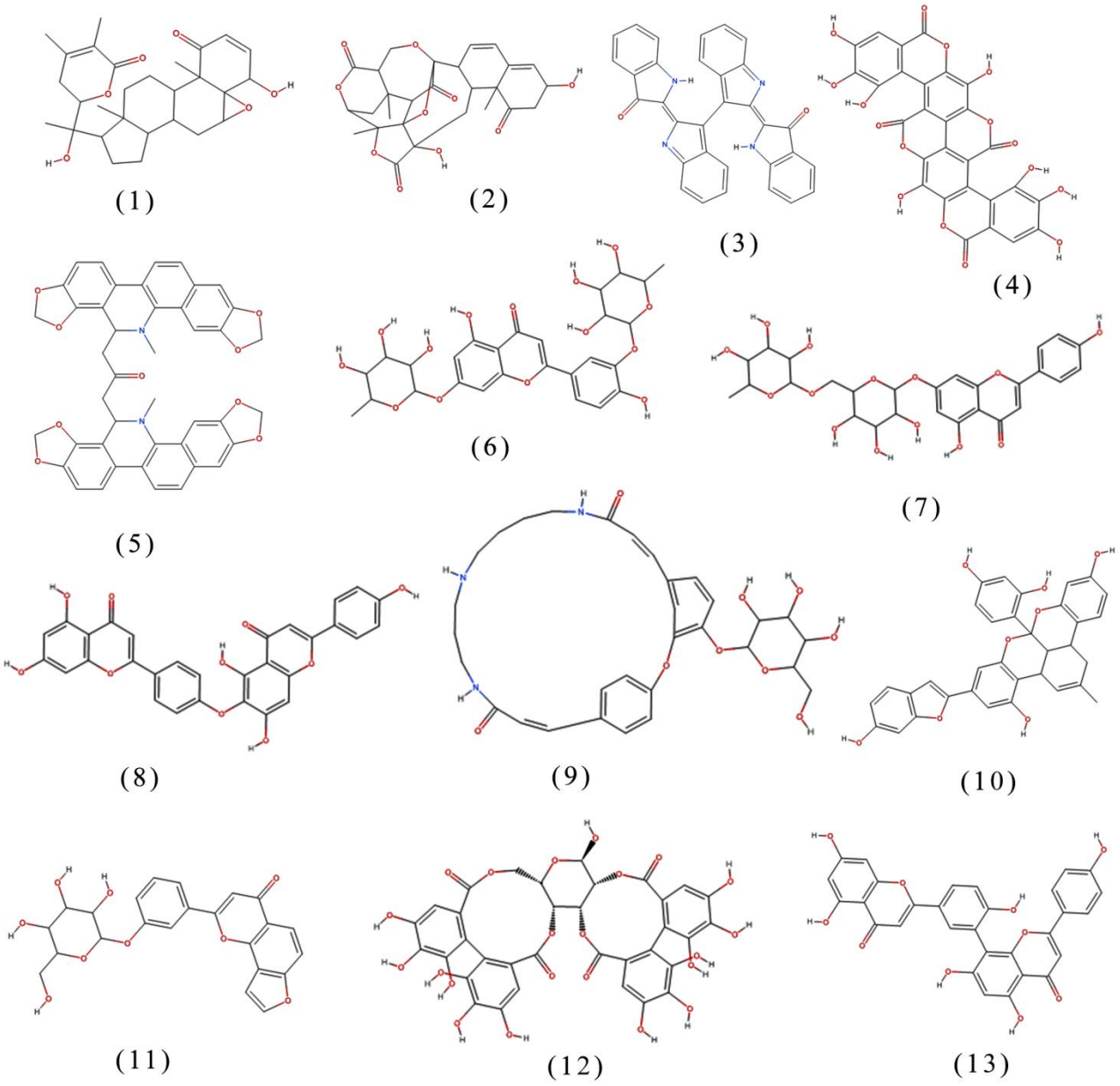

Figure 2. 2D chemical structures of the common multi-target phytochemicals identified from the virtual screening process. Withanolide (1), Physalin Z (2), Bisindigotin (3), Gallagyldilacton (4), Chelidimerine (5), Luteolin-3'-O-di-rhamnoside-7-O-rhamnoside (6), Apigenin-7-O-rutinoside (7), Hinokiflavone (8), Capparispine 26-O-beta-D-glucoside (9), Mulberrofuran G (10), Pongamoside A (11), Pedunculagin (12), and Amentoflavon (13).

For Table of Contents Only