Molecular docking analysis of selected natural products from plants for inhibition of SARS-CoV-2 main protease

Megha Hastantram Sampangi-Ramaiah[1], Ram Vishwakarma[2] and R. Uma Shaanker[1,3,*]

[1]School of Ecology and Conservation, University of Agricultural Sciences, GKVK, Bengaluru 560 065, India
[2]Indian Institute of Integrative Medicine, Council of Scientific and Industrial Research, Jammu 180 001, India
[3]Department of Crop Physiology, University of Agricultural Sciences, GKVK, Bengaluru 560 065, India

In this article, we report results of a molecular docking analysis of commonly occurring natural product compounds against COVID-19 6LU7 and 6Y2E proteases. Our results show that several of these compounds have binding affinity against both the COVID-19 proteases, and compare well with a known anti-HIV drug, Saquinavir. Many of the compounds form an integral component of many cuisines, both Indian as well as others. We propose that some of these compounds could be easily and quickly positioned to hold fort against the COVID-19 virus, until of course newer therapies are discovered and detailed studies are taken to empirically validate some of the compounds for their ability to inhibit the virus.

Keywords: Affinity/binding energy, COVID-19 protease, drug discovery, ligands, natural products.

As the world awaits eagerly for a therapy for COVID-19 virus (SARS-CoV-2 (2019-nCoV), there have been several attempts in the recent weeks to ask if existing solutions and compounds could be repositioned to address the spectre of COVID-19, if not as a curative, at least as a preventive (for complete reference of such efforts please see refs 1, 2). However, many of these initiatives await further laboratory and clinical validation. In the meanwhile, efforts have been initiated[1] to examine the potential of several plant secondary metabolites in inhibiting the COVID-19 main protease (M^{pro})/chymotrypsin-like protease ($3CL^{pro}$) using molecular docking analysis to arrive at binding affinity. For example and interestingly, Khaerunnisa *et al.*[1] show that several plant secondary metabolites such as kaempferol, quercetin, luteolin-7-glucoside, demethoxycurcumin, naringenin, apigenin-7-glucoside, oleuropein, curcumin, catechin, and epicatechin-gallate have the potential to inhibit the COVID-19 M^{Pro} protease. In a more recent paper, first released on 20 March 2020, Zhang *et al.*[2] report a potent α-ketoamide inhibitor of COVID-19 main protease (M^{pro}, $3CL^{pro}$) and suggest that subject to further investigation, it could form an attractive drug candidate.

In this study, we examine the binding affinity of 27 ligands, most of which form an integral component of many cuisines, both Indian as well as others, to both COVID-19 6LU7 and 6Y2E proteases. The express purpose of this study is to ask if one or more of these ligands offer potential for further studies in developing natural product based solutions to the challenge of the COVID-19 pandemic. If yes, the use of some of the identified natural products can be deployed to fortify defences against COVID-19 by developing such natural products based drugs. Considering the fact that most of the natural product ligands chosen for the study occur in commonly used spices, vegetables and fruits, it is likely that their eventual deployment as drug candidates could hasten regulatory approvals. Finally our studies, as evident from the results presented below, show that indeed, few of the natural products ligand have acceptable binding affinity with both COVID-19 6LU7 protease and 6Y2E protease and compare well with a known anti-HIV drug, Saquinavir[4].

Methods

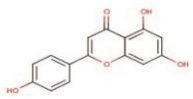
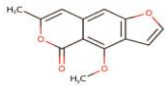
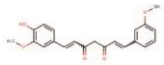
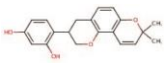
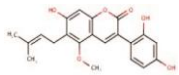
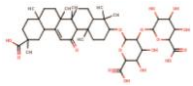
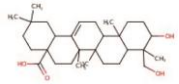
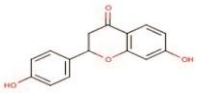
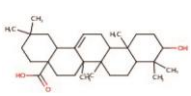
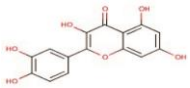
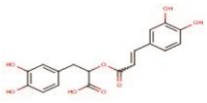
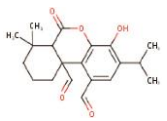
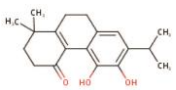
Data sets

The COVID-19 $3CL^{pro}$/M^{pro} (PDB ID: 6LU7)[1,5] and free enzyme of the SARS-CoV-2 (2019-nCoV) main protease (PBD ID: 6Y2E)[3] structures were obtained from the https://www.rcsb.org/ website in .PDB format. The active site of the proteases were determined using the Computed Atlas for Surface Topography of Proteins (CASTp) (http://sts.bioe.uic.edu/castp/index.html?3igg)[6].

The 3D structures of the selected ligands were obtained from the https://pubchem.ncbi.nlm.nih.gov/ website in the .SDF format. The structural formula of the ligands as represented in Table 1 was retrieved from the SWISSADME website. Further, the ligands were optimized using the Avogadro version 1.2, with Force Field

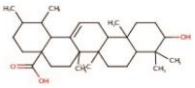
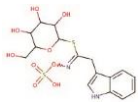
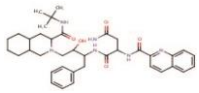
*For correspondence. (e-mail: umashaanker@gmail.com)

Table 1. Natural product compounds along with their binding affinity to COVID-19 proteases, 6LU7 and 6Y2E. Also given is the molecular formula, structural formula and major plant sources of these compounds

Compound name	Molecular formula	Structural formula	Binding affinity (kcal/mol) 6LU7 protease	Binding affinity (kcal/mol) 6Y2E protease	Ligand source
Apigenin	$C_{15}H_{10}O_5$		–7.8	–7.0	Apple (*Malus domestica*) Thyme (*Thymus vulgaris*) Chamomile (*Matricaria chamomilla*) Red pepper (*Capsicum annuum*) Garlic (*Allium sativum var. sativum*) Celery (*Apium graveolens*)
Coriandrin	$C_{13}H_{10}O_4$		–6.4	–6.3	Coriander (*Coriandrum sativum*)
Curcumin	$C_{21}H_{20}O_6$		–7.0	–6.4	Turmeric (*Curcuma longa*)
Glabridin	$C_{20}H_{20}O_4$		–8.0	–7.1	Licorice (*Glycyrrhiza glabra*)
Glycoumarin	$C_{21}H_{20}O_6$		–7.5	–7.1	Licorice-root (*Glycyrrhiza glabra*)
Glycyrrhizin	$C_{42}H_{62}O_{16}$		–7.2	–8.4	Licorice (*Glycyrrhiza glabra*)
Hederagenin	$C_{30}H_{48}O_4$		–7.6	–7.7	Black cumin (*Nigella sativa*) Spiny gourd (*Momordica dioica*) Sponge gourd (*Luffa cylindrica*)
Liquiritigenin	$C_{15}H_{12}O_4$		–7.7	–6.9	Chinese Licorice (*Glycyrrhiza uralensis*) Licorice (*Glycyrrhiza glabra*)
Oleanolic acid	$C_{30}H_{48}O$		–7.8	–8.0	Thyme (*Thymus vulgaris*) Rosemary (*Rosmarinus officinalis*) Olive (*Olea europaea*) Java Apple (*Syzygium samarangense*)
Quercetin	$C_{15}H_{10}O_7$		–7.3	–7.4	Onion (*Allium cepa*) Green tea (*Camellia sinensis*) Apple (*Malus domestica*) Buckwheat (*Fagopyrum esculentum*)
Rosmarinic acid	$C_{18}H_{16}O_8$		–7.1	–7.3	Basil (*Ocimum basilicum*) Holy basil (*Ocimum tenuiflorum*) Lemon balm (*Melissa officinalis*) Rosemary (*Rosmarinus officinalis*) Sage (*Salvia officinalis*) Thyme (*Thymus vulgaris*) Peppermint (*Mentha piperita*)
Safficinolide	$C_{20}H_{24}O_5$		–6.8	–6.9	Sage (*Salvia officinalis*)
Sageone	$C_{19}H_{24}O_3$		–7.1	–6.6	Sage (*Salvia officinalis*)

(*Contd*)

Table 1. (*Contd*)

Compound name	Molecular formula	Structural formula	Binding affinity (kcal/mol) 6LU7 protease	Binding affinity (kcal/mol) 6Y2E protease	Ligand source
Ursolic acid	$C_{30}H_{48}O_3$		–7.6	–8.2	Apple peel (*Malus domestica*) Basil (*Ocimum basilicum*) Bilberries (*Vaccinium myrtillus*) Cranberries (*Vaccinium oxycoccos*) Rosemary (*Rosmarinus officinalis*) Thyme (*Thymus vulgaris*) Peppermint (*Mentha piperita*) Prunes (*Prunus domestica*)
Glucobrassicin	$C_{16}H_{20}N_2O_9S_2$		–8.1	–7.4	Cabbage (*Brassica oleracea*) Mustard *(Brassica juncea*) Broccoli (*Brassica oleracea* var. *italica*)
Saquinavir	$C_{38}H_{50}N_6O_5$		–9.2	–7.9	Synthetic (HIV drug)

type MMFF94, and saved in .mol2 format[7]. The ligands chosen for the study are listed in Table 1. We used the anti-HIV drug, Saquinavir, as a positive control.

Molecular docking

The protease files were separately prepared for both, COVID-19 6LU7 protease and 6Y2E protease using Autodock 4.2 (ref. 8). The macromolecule was prepared using 'A' chain of the protease, the water molecule was deleted and polar hydrogen atoms were added. The file was saved in the .PDBQT format for further analysis. The amino acids in the active site of the macromolecule was selected and the grid box was used to obtain the X, Y and Z coordinates. The ligand files in the .mol2 format were converted to the .PDBQT format after detecting the torsion root.

Using the protease .PDBQT file, ligand .PDBQT file and the X, Y and Z coordinates, binding affinity was calculated using AutoDock–Vina[9]. The 3D structure of the protease-ligand interaction was visualized using PyMOL[10] and the 2D structure of the molecular interaction was visualized using the Biovia Discovery Studio 20.1.0.

Results and discussion

Of the 27 compounds evaluated (Supplementary Table 1), for 15 natural products the binding energies were lesser than the upper threshold (–6 kcal/mol), generally regarded as a cut-off in ligand-binding studies[11] (Table 1). The binding affinity for COVID-19 6LU7 protease ranged from –6.4 kcal/mol (for Coriandrin) to –8.0 kcal/mol (for Glabridin) and –8.1 kcal/mol (for Glucobrassicin). Coriandrin is a component of the essential oil of *Coriandrum sativum* and is present in leaf, stem and seeds[12]. This plant is used widely in Indian cuisines across both north and south India. Glabridin is present in *Glycyrrhiza glabra*, often referred to a licorice or liquorice[13]. The plant is a herbaceous perennial legume plant and occurs naturally in the Middle East, Europe and parts of Asia. It is an important source of antioxidants and is well known as an important medicinal plant[13–15]. Glucobrassicin is a glucosinolate found in many cruciferous vegetables (such as the common cabbage and broccoli) and mustard (Brassica). Besides these compounds, several others are also noteworthy based on their binding affinity to the protease. For example, the affinity energies of Ursolic acid (–7.6 kcal/mol), Hederagenin (–7.6 kcal/mol), Apigenin (–7.8 kcal/mol), Oleanolic acid (–7.8 kcal/mol) and Rosemarinic acid (–7.1 kcal/mol) make them useful candidates for further empirical validation. Ursolic acid also known as urson is a pentacyclic triterpenoid and is found abundantly in peel of fruits such as apple, besides also in several other plants used commonly such as rosemary and thyme[16]. Oleanolic acid[17] is an integral component of clove oil and again used extensively in culinary preparations. Finally, Hederagenin is a constituent of *Nigella sativa* and several Cucurbitaceae vegetables including *Luffa cylindrica* and *Momordica dioica*, making it again a potentially useful candidate[18,19]. The binding affinities of these compounds compare reasonably well with that of the positive control, Saquinavir (–9.2 kcal/mol).

The binding affinity of the natural products to the COVID-19 protease 6Y2E was only slightly different from that to 6LU7. There was largely a positive correlation between the binding affinities ($R^2 = 0.236$). The binding affinities to COVID-19 protease 6Y2E ranged from –6.3 kcal/mol (for Coriandrin) to –8.2 kcal/mol (for

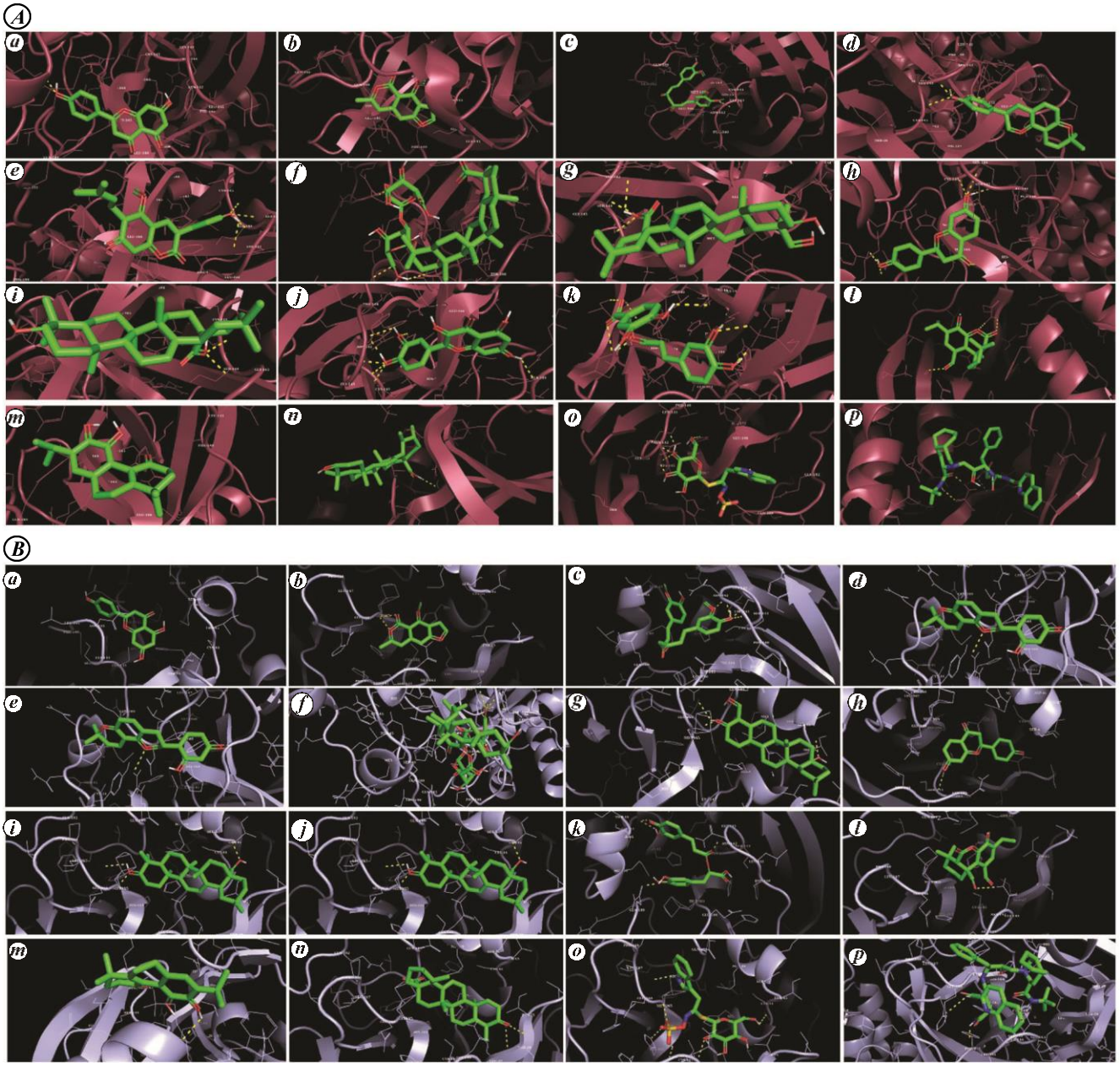


Figure 1. ***A***, 3D visualization of docking analysis of 6LU7 protease binding with Apigenin (***a***), Coriandrin (***b***), Curcumin (***c***), Glabridin (***d***), Glycoumarin (***e***), Glycyrrhizin (***f***), Hederagenin (***g***), Liquiritigenin (***h***), Oleanolic acid (***i***), Quercitin (***j***), Rosmarinic acid (***k***), Saffinicinolide (***l***), Sageone (***m***), Ursolic acid (***n***), Glucobrassicin (***o***), Saquinavir (***p***; positive control). ***B***, 3D visualization of docking analysis of 6Y2E protease binding with Apigenin (***a***), Coriandrin (***b***), Curcumin (***c***), Glabridin (***d***), Glycoumarin (***e***), Glycyrrhizin (***f***), Hederagenin (***g***), Liquiritigenin (***h***), Oleanolic acid (***i***), Quercitin (***j***), Rosmarinic acid (***k***), Saffinicinolide (***l***), Sageone (***m***), Ursolic acid (***n***), Glucobrassicin (***o***), Saquinavir (***p***; positive control).

Ursolic acid). Oleanolic acid (–8.0 kcal/mol) and Glycyrrhizin (–8.4 kcal/mol) were other potential candidates. It is interesting to note that a diammonium salt of glycyrrhizin is recommended for the cure of COVID-19 in China[20]. The binding affinity of the positive control, Saquinavir was –7.9 kcal/mol.

The 3D and 2D visualization of binding of the various compounds to the active sites of the COVID-19 protease, 6LU7 and 6Y2E are presented in Figures 1 and 2 respectively. These results clearly show that each of the ligands bind to the active sites of the proteases and thus could be expected to inhibit the activity of the enzyme and hence restrain viral replication.

In summary, our study shows that several of the natural product compounds that are commonly used in Indian and other cuisines may have the potential to provide the first line of defence against the COVID-19 virus and in time to come, after detailed studies are taken up, might also offer exciting candidates for repurposing them as effective therapeutic drugs.

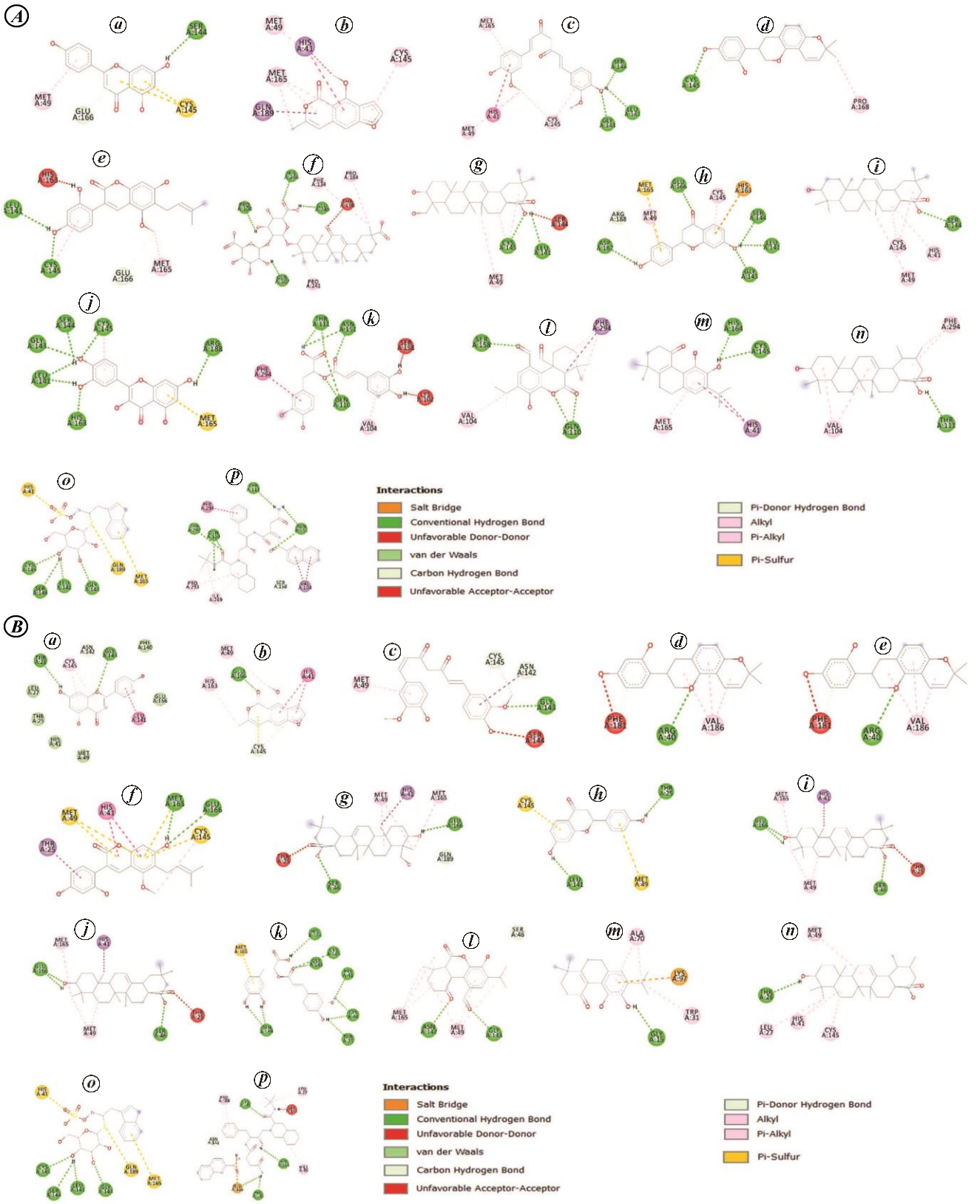


Figure 2. ***A***, 2D visualization of molecular interaction of 6LU7 protease with Apigenin (***a***), Coriandrin (***b***), Curcumin (***c***), Glabridin (***d***), Glycoumarin (***e***), Glycyrrhizin (***f***), Hederagenin (***g***), Liquiritigenin (***h***), Oleanolic acid (***i***), Quercitin (***j***), Rosmarinic acid (***k***), Saffinicinolide (***l***), Sageone (***m***), Ursolic acid (***n***), Glucobrassicin (***o***), Saquinavir (***p***; positive control). ***B***, 2D visualization of molecular interaction of 6Y2E protease with Apigenin (***a***), Coriandrin (***b***), Curcumin (***c***), Glabridin (***d***), Glycoumarin (***e***), Glycyrrhizin (***f***), Hederagenin (***g***), Liquiritigenin (***h***), Oleanolic acid (***i***), Quercitin (***j***), Rosmarinic acid (***k***), Saffinicinolide (***l***), Sageone (***m***), Ursolic acid (***n***), Glucobrassicin (***o***), Saquinavir (***p***; positive control).

Conclusions

At times of adversity, such as what the world is passing through today due to the COVID-19 virus pandemic, there is an urgent need to develop a rapid yet fairly reliable strategy to counter the disease. One such approach that we have employed and presented here is to explore the repositioning or repurposing of some of the most common natural products that are used in our kitchens, especially Indian and other cuisines. We used 27 natural products, many of them very familiar, used as spices and condiments or present in common vegetables and explored if they bind to the active sites of COVID-19 proteases, 6LU7 and 6Y2E, critical for viral replication. Our results showed that 15 are effective in binding the viral protease and hence are likely to hinder viral replication. Besides the very common compounds, curcumin and coriandrin, that is used almost every day in the Indian cuisine, compounds found in apple peels (ursolic acid), in cucurbit vegetables (hederagenin), in olive oil (oleanolic acid), rosemary and thyme or mint family plants (sageone), in red pepper (apigenin), in *Glycyrrhiza glabra* (glabridin) to name a few are very promising and could serve as potential candidates for further research. Admittedly the results presented here are preliminary and the optimism needs to be tempered with more detailed dynamic simulation studies for getting a better insight into the mode of inhibition. However as mentioned elsewhere, the purpose of this article is to draw urgent attention to the promise of some of these compounds, and galvanize attention of the Indian scientific community to these possibilities and hopefully lead to the development of some drugs against COVID-19 virus in the near future.

Author contributions: M.H.S. developed the idea and performed the molecular docking analysis and created the 2D and 3D visualization models. M.H.S., R.U.S. and R.V. prepared the manuscript. R.V. contributed to the medicinal chemistry issues highlighted in the manuscript.

1. Khaerunnisa, S., Kurniawan, H., Awaluddin, R., Suhartati, S. and Soetjipto, S., Potential inhibitor of COVID-19 main protease (Mpro) from several medicinal plant compounds by molecular docking study. 2020. Preprints 2020, 2020030226 (doi: 10.20944/preprints202003.0226.v1).
2. Wu, C. *et al*., Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharmaceut. Sin. B*, 2020.
3. Zhang, L. *et al*., Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science*, 2020, eabb3405.
4. Noble, S. and Faulds, D., Saquinavir. *Drugs*, 1996, **52**, 93–112.
5. Liu, X., Zhang, B., Jin, Z., Yang, H. and Rao, Z., The crystal structure of 2019-NCoV main protease in complex with an inhibitor N3. *RCSB Protein Data Bank*, 2020.
6. Tian, W., Chen, C., Lei, X., Zhao, J. and Liang, J., CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Res.*, 2018, **46**, W363–W367.
7. Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., and Hutchison, G. R., Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.*, 2012, **4**, 17.
8. Morris, G. M., Goodsell, D. S., Huey, R. and Olson, A. J., Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. *J. Comp.-Aided Mol. Design*, 1996, **10**, 293–304.
9. Trott, O. and Olson, A. J., 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.
10. DeLano, W. L., PyMOL: An open-source molecular graphics tool. *CCP4 Newslett. Protein Crystallogr.*, 2002, **40**(1), 82–92.
11. Shityakov, S. and Förster, C., *In silico* predictive model to determine vector-mediated transport properties for the blood–brain barrier choline transporter. *Adv. Appl. Bioinform. Chem.*, 2014, **7**, 23.
12. Ceska, O., Chaudhary, S. K., Warrington, P., Ashwood-Smith, M. J., Bushnell, G. W. and Poultont, G. A., Coriandrin, a novel highly photoactive compound isolated from *Coriandrum sativum*. *Phytochemistry*, 1988, **27**, 2083–2087.
13. Tian, M., Yan, H. and Row, K. H., Simultaneous extraction and separation of liquiritin, glycyrrhizic acid, and glabridin from licorice root with analytical and preparative chromatography. *Biotechnol. Bioproc. Eng.*, 2008, **13**, 671–676.
14. Ishtiyaq, A., Alam, A., Siddiqui, J. I. and Kazmi, M. H., Therapeutic potential of widely used unani drug Asl-Us-Soos (*Glycyrrhiza glabra* Linn.): a systematic review. *J. Drug Deliv. Therap.*, 2019, **9**, 765–773.
15. Kumari, A. and Kumar, J., Phyto-chemical screening of root extracts of *Glycyrrhiza glabra* by spectroscopic methods (UV-VIS spectrophotometer, FTIR and HPLC). *Int. J. Pharmaceut. Sci. Drug Res.*, 2019, **11**, 376–381.
16. Lin, J.-H., Chen, S.-Y., Lu, C.-C., Lin, J.-A. and Yen, G.-C., Ursolic acid promotes apoptosis, autophagy, and chemosensitivity in gemcitabine-resistant human pancreatic cancer cells. *Phytother. Res.*, 2020, 1–14, doi:10.1002/ptr.6669.
17. Khwaza, V., Oyedeji, O. O. and Aderibigbe, B. A., Antiviral activities of oleanolic acid and its analogues. *Molecules*, 2018, **23**, 2300.
18. Wu, A.-G. *et al*., Hederagenin and α-hederin promote degradation of proteins in neurodegenerative diseases and improve motor deficits in MPTP-mice. *Pharmacol. Res.*, 2017, **115**, 25–44.
19. Zeng, J. *et al*., Current knowledge and development of hederagenin as a promising medicinal agent: a comprehensive review. *RSC Adv.*, 2018, **8**, 24188–24202.
20. Redeploying plant defences. *Nature Plants*, 2020, **6**, 177.

ACKNOWLEDGEMENTS. M.H.S. is supported by a DBT grant on 'Chemical Ecology of the North East Region (NER) of India: A collaborative programme Linking NER and Bangalore Researchers' (DBT-NER/Agri/24/2013). R.U.S. is supported by an ICAR Emeritus Scientist Grant. R.V. is supported by the Council for Scientific and Industrial Research, Government of India.

Received and accepted 25 March 2020

doi: 10.18520/cs/v118/i7/1087-1092