

# Potential drugs against COVID-19 revealed by gene expression profile, molecular docking and molecular dynamic simulation

Claudia Cava\*<sup>1</sup>, Gloria Bertoli<sup>1</sup> & Isabella Castiglioni<sup>2</sup>

<sup>1</sup>Institute of Molecular Bioimaging & Physiology, National Research Council (IBFM-CNR), Via F. Cervi 93, Segrate-Milan, Milan, 20090, Italy

<sup>2</sup>Department of Physics "Giuseppe Occhialini", University of Milan-Bicocca Piazza dell'Ateneo Nuovo, Milan, 20126, Italy

\*Author for correspondence: [claudia.cava@ibfm.cnr.it](mailto:claudia.cava@ibfm.cnr.it)

**Aim:** SARS-CoV-2, an emerging betacoronavirus, is the causative agent of COVID-19. Currently, there are few specific and selective antiviral drugs for the treatment and vaccines to prevent contagion. However, their long-term effects can be revealed after several years, and new drugs for COVID-19 should continue to be investigated. **Materials & methods:** In the first step of our study we identified, through a gene expression analysis, several drugs that could act on the biological pathways altered in COVID-19. In the second step, we performed a docking simulation to test the properties of the identified drugs to target SARS-CoV-2. **Results:** The drugs that showed a higher binding affinity are bardoxolone (-8.78 kcal/mol), irinotecan (-8.40 kcal/mol) and pyrotinib (-8.40 kcal/mol). **Conclusion:** We suggested some drugs that could be efficient in treating COVID-19.

First draft submitted: 27 November 2020; Accepted for publication: 30 June 2021; Published online: 20 July 2021

**Keywords:** 3CL main protease • bioinformatics • COVID-19 • drug • dynamic simulation • gene expression • interactions • molecular docking • SARS-CoV • therapy

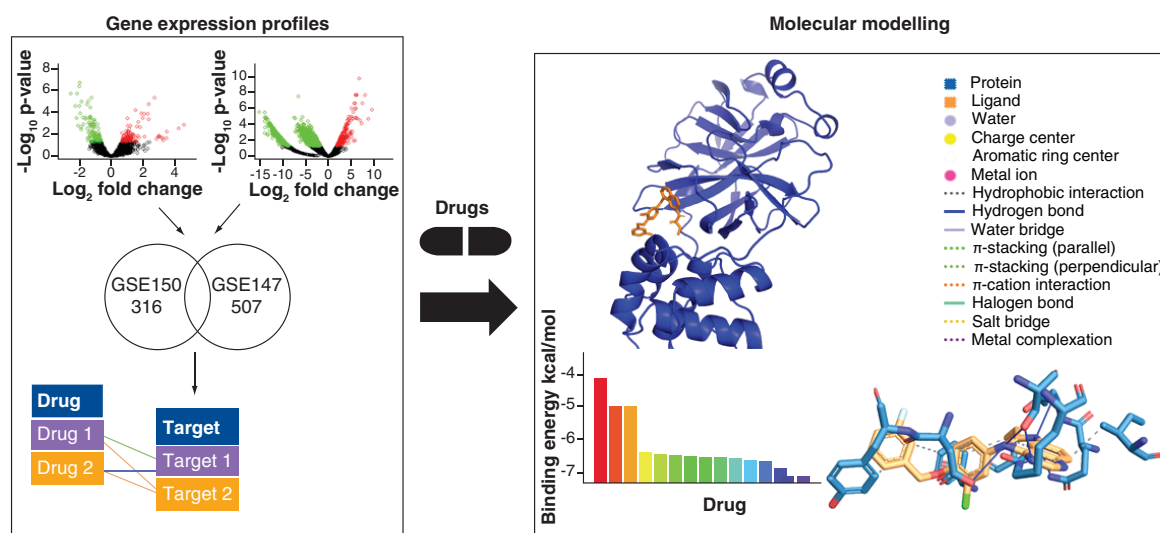
A state of global health emergency has been declared by WHO for COVID-19. The causative agent is SARS-CoV-2, which like SARS-CoV and MERS-CoV, belongs to the betacoronavirus genus [1].

The most functionally relevant proteins in SARS-CoV-2 are the spike glycoprotein and 3CL main protease. The spike protein binds to a host cell membrane and allows the entry of the virus into the host cell. It has been described that SARS-CoV entrance is mediated by angiotensin-converting enzyme 2 (ACE2) that is particularly overexpressed by epithelial cells in oral mucosa and intestine [1].

Previous studies have shown that the interaction between the spike glycoprotein and ACE2 has a key role in SARS pathogenesis. Thus, initially, these two proteins have been proposed as possible and promising drug targets. However, anti-viral agents targeting ACE2 and spike glycoprotein did not proceed clinically due to significant side effects [2,3].

Several studies also demonstrated the potential of the 3CL main protease as a promising drug target [4,5]. 3CL main protease is responsible for the regulation of the polyproteins that are translated from the viral RNA [6]. The inhibition of this enzyme could block the replication processes of the virus, which makes it a potential target for drug discovery. Recently x-ray crystal structure of the 3CL main protease was determined providing an excellent basis for structure-based drug discovery [7]. To date, no effective SARS-CoV-3CL inhibitor has been proposed yet to treat COVID-19.

In order to identify new potential targets to block COVID-19 disease, *in silico* drug discovery approaches have been demonstrated to be effective, efficient and cost saving and can follow two main strategies. The first strategy is based on the analysis of gene expression profiles of patients affected by COVID-19, to identify new potential drug targets acting on molecular mechanisms altered by SARS-CoV-2. Several studies based on this approach allowed researchers to discovery candidate drugs for different viruses, such as SARS-CoV-2, MERS-CoV, Ebola and Zika virus [8–11]. The second strategy is to test existing drugs with a virtual screening approach based on the study of the



**Figure 1. Workflow of the proposed approach.** A computational pipeline based on gene expression data analysis, molecular docking and molecular dynamic simulation was adopted for *in silico* drug discovery.

3D structures of the target protein. This approach, called molecular docking, can predict how drug candidates bind to a receptor of known 3D structure [12], generating a score (binding energy value) that defines the binding affinities between ligand and target protein. There have been successful applications that demonstrated how molecular docking is becoming a powerful tool in the discovery of drug candidates [13,14]. However, most of these studies tested the drugs without considering the effect of drugs on genes deregulated as effect of diseases [1,15]. Indeed, the novelty of the present study is the combination of gene expression analyses with molecular docking and dynamic simulation. In our study, the selection of a gene as therapeutic target was performed via a gene expression analysis and, since the optimal target was identified, the known interactions between drugs and targets were investigated through the study of their binding energies using molecular docking and simulation studies. To our knowledge there are no studies in the literature that integrate these two techniques in COVID-19. Currently, published works screened a library of drugs or molecules against SARS-CoV-2, but without a prior selection based on drugs that could regulate genes altered in COVID-19 as obtained from a gene expression analysis [16–18]. Other studies utilized quantitative structure–activity relationship (such as [19]) or binding energies [20] without considering COVID-19 affected cascades of genes. The gene expression profiles altered by COVID-19 supported the docking results and offered a new perspective for subsequent studies.

Although docking programs are fast and easy to use, they are affected by too low accuracy and stability. Indeed, several studies reported that binding free energies based on molecular dynamic simulations are much more reliable [21].

In this study, we proposed a bioinformatic approach based on molecular docking, molecular dynamic simulations and gene expression profiles of COVID-19 to be used *in silico* to test existing drugs as therapies against COVID-19.

## Materials & methods

In the first step of our study we identified, through a gene expression analysis, several drugs that could act on the biological pathways altered by COVID-19. In particular, we selected differentially expressed genes between COVID-19 and healthy samples, obtained from two published datasets. In the second step, we performed a docking simulation in order to test the properties of these drugs to target the 3CL main protease of SARS-CoV-2. Figure 1 shows the workflow of the proposed approach.

## Gene expression data

We analyzed the gene expression levels of COVID-19 positive and negative patients from two Gene Expression Omnibus (GEO) datasets: GSE150316 and GSE147507.

Specifically, we selected five negative-control samples of lung tissues and 16 COVID-19-positive lung samples from GSE150316; and two lung biopsies from COVID-19-negative controls and two lung biopsies from post-mortem COVID-19-positive patients from GSE147507.

We defined differentially expressed genes between COVID-19-negative and -positive samples if the genes have an absolute value of log fold change  $>0.5$  and adjusted p-value  $< 0.05$ . Pre-processing and differential expression analyses were performed with the Bioconductor package TCGAbiolinks [22].

### Drug interactions & overrepresentation analysis

Matador and DGIdb databases were used to identify the interactions between the drug and protein [23,24].

We performed a Fisher's test to verify if there is an overrepresentation of drug targets in the list of differentially expressed genes. We considered drugs acting on the differentially expressed genes if FDR  $<0.05$ . P-values were adjusted with Benjamini–Hochberg procedure for multiple testing correction [25].

### Molecular modeling

We performed molecular docking studies in order to identify potential agents with antiviral properties against COVID-19. Molecular modeling was performed with AutodockTools 1.5.6 [26].

Drugs considered in the overrepresentation analysis were selected and PubChem database was used to extract 3D structure of the selected drugs [27]. Discovery studio was used to prepare ligand (drugs) and generate Protein Data Bank (PDB) format [28]. PDB format is a standard representation for molecular structure data originated from x-ray diffraction.

The crystal structure of COVID-19 main protease in complex with an inhibitor N3 was extracted from PDB (IDs: 6LU7). PDB ID 6LU7 consists of two chains: chain A and chain C. The chain C, representing the complex-bound inhibitor to the protein receptor molecule, was removed. During the pre-processing step of protein polar hydrogen atoms were added and water molecules were removed.

Molecular docking identifies the amino acids that interact between the selected protein and drugs. Minimum energy binding of the ligand with the receptor was considered. Binding energy is described as a decreasing of the overall energy of the complex when a drug is associated with a target protein. Binding energy also defines the ligand binding affinity.

Protein–ligand interaction profiler was used to visualize the binding interactions of the selected drugs with 3D model of protease of COVID-19 [29].

### Molecular dynamics

The investigated main protease of SARS-CoV-2 and screened molecules resulting from molecular docking were subjected to molecular dynamic simulation. The results obtained through the molecular docking were extended, as molecular dynamics are considered a more solid analysis for molecular study of ligand recognition.

Molecular dynamic simulations of the main protease of SARS-CoV-2 with screened molecules were performed using BIOVIA Discovery Studio Client (Dassault Systèmes, Vélizy-Villacoublay, France) [30]. Molecular dynamics simulations were carried out using the standard dynamics cascade protocol in four steps: minimization, heating, equilibration run, and production run. We set simulation time (ps) parameters to 20 and 200 in equilibration and production step, respectively. For the other parameters, we used the default values.

Minimization of the 3D structures to get the most stable confirmations of the complexes was performed using the default algorithms: the steepest descent algorithm and the adopted basis NR algorithm.

After the minimization step, all the complexes were subjected to gradual heating using the default values.

In the equilibration phase, the system was stabilized at a target temperature, as the energy has to be distributed appropriately among all systems.

Furthermore, complexes were subjected to the production run at constant temperature, using constant-volume ensemble (NVT). The results of this step are stored in the simulation trajectory.

The average binding energy was calculated for equilibrated molecular dynamic trajectory. Trajectory file contains 100 conformations. The binding energy was calculated by:

$$\Delta G_{\text{bind}} = \Delta G_{\text{complex}} - \Delta G_{\text{protein}} - \Delta G_{\text{ligand}} \quad (\text{Eq. 1})$$

where  $\Delta G_{\text{complex}}$ ,  $\Delta G_{\text{protein}}$  and  $\Delta G_{\text{ligand}}$  are the total free energy of protein–ligand complex, protein and ligand in solvent, respectively.

### Pharmacokinetics parameters

Lipinski's rule is an algorithm consisting of a set of rules to be considered for the design and development of a drug [30]. It is based on four criteria which must be respected: molecular mass should be less than 500 daltons; the drug should not have more than five hydrogen bond donors in its skeleton; the drug should not have more than five hydrogen bond acceptors in its skeleton; and the fat solubility of the molecule expressed by the partition coefficient must be less than 5. In summary, Lipinski's rule considers two main concepts: absorption (drug candidate should be better absorbed if small in size) and permeation (drug candidate should better pass through the cell membranes if it is not too hydrophilic). We used the SwissADME tool to verify if the selected drugs respect the above-reported criteria [30]. We used ChEBI database to download mol file for the selected drugs [31]. Mol file, a commonly used chemical structure file format is used as input for SwissADME tool.

## Results

### Gene expression analysis

We obtained 198 differentially expressed genes from GSE150316 and 2670 genes from GSE147507. We found 47 genes that are differentially expressed in both datasets.

We then identified 38 drugs (FDR <0.0211) that have an overrepresentation of gene targets in the list of 47 differentially expressed genes. Some of these drugs showed an antiviral property against different viral diseases such as bardoxolone, irinotecan, olaparib, inosine and anthraquinone. The mechanism of action of these drugs are different such as inhibition of protein kinase, inhibition of epidermal growth factor receptor and induction of apoptosis.

27 out of 38 drugs were tested for docking analysis and are depicted in the Table 1. The 3D structure of Hemay-022, lapuleucel-T, T-DM1, ertumaxomab, seribantumab, onartuzumab, depatuxizumab mafodotin, MDX-210, MM-302, trodusquemine and margetuximab are not available from PubCHEM.

### Docking studies

In order to identify potential drugs for treating COVID-19 patients, a molecular docking analysis was performed on the 27 drugs revealed by the differential expression analysis.

The results from the SARS-CoV-2 3CL main protease are reported in Table 2. 12 out of 27 drugs obtained a binding energy less than -7 kcal/mol (AZD-8055, olaparib, irinotecan, tyrphostin AG 879, topotecan hydrochloride, MP-412, S-222611, allitinib, 7-ethyl-10-hydroxy-camptothecin, faldamamol, pyrotinib and bardoxolone). The study revealed no violation of Lipinski's rule for these drugs. Tyrphostin AG 879, MP-412, S-222611, allitinib, pyrotinib, bardoxolone are not available in ChEBI.

The best binding energy values were obtained for bardoxolone (-8.78 kcal/mol), irinotecan (-8.40 kcal/mol) and pyrotinib (-8.40 kcal/mol). Figure 2 shows the protein-ligand complex for bardoxolone, irinotecan and pyrotinib.

Bardoxolone has four types of interactions with protein residues: van der Waals, hydrogen bonds (conventional and carbon hydrogen bonds), alkyl and pi-alkyl interactions. In particular, the residues LEU282 and LYS5 form alkyl interactions with aromatic ring of the molecule. Moreover, the residue GLN127 interacts with a hydrogen bond with the cyano group of the molecule.

Irinotecan interacts with protein structure through van der Waals, hydrogen bonds, alkyl, pi-alkyl interactions, pi-sigma and pi-pi stacked. The residue TYR126 forms pi-pi stacked and alkyl interactions with aromatic ring of the molecule. Hydroxyl group interacts with GLU288.

Pyrotinib interacts with protein structure through van der Waals, hydrogen bonds, halogen, pi-pi stacked and alkyl interactions. In particular, chloro of the ligand interacts with the residue THR111 through halogen interaction and with the residue ASN151 through hydrogen interaction. PHE294 forms pi-pi stacked interactions with the aromatic rings of the molecule.

Supplementary File 1 shows the binding site predictions for each of the 27 drugs against SARS-CoV-2 3CL main protease.

### Molecular dynamic simulation

COVID-19 main protease (PDB IDs: 6LU7) in complex with the top three (bardoxolone, irinotecan and pyrotinib) docked ligands have been selected to compute dynamic properties through molecular dynamic simulations. The calculated binding energies are presented in Table 3. All three ligands obtained negative binding energies. Pyrotinib

Table 1. Drugs enriched with differentially-expressed gene targets.

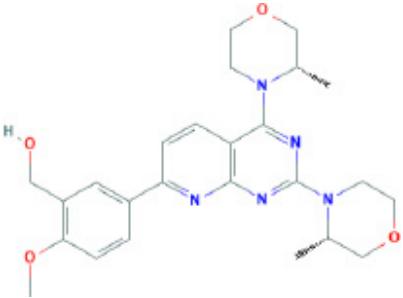
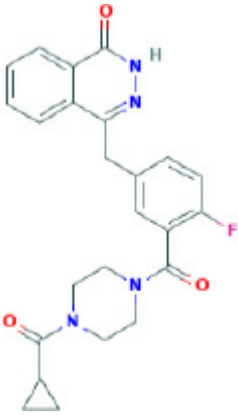
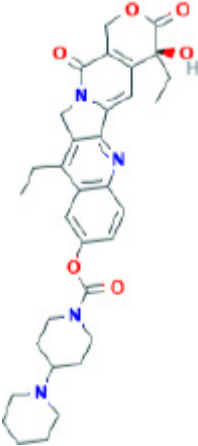
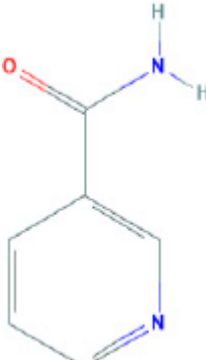
Drug	Structure	Target gene	Mechanism
1 AZD-8055		<i>ERBB2, MAPKAP1</i>	mTOR inhibitor
2 Olaparib		<i>PARP1, ERBB2</i>	Inhibitor of the nuclear enzyme poly(ADP-ribose) polymerase
3 Irinotecan		<i>ERBB2, TOP1MT</i>	Inhibitor of topoisomerase I activity
4 Niacinamide		<i>PARP1</i>	Inhibitor of poly(ADP-ribose) polymerases

Table 1. Drugs enriched with differentially-expressed gene targets (cont.).				
Drug	Structure	Target gene	Mechanism	
5	Heptanoate	<i>PARP1</i>	Plant metabolite	
6	Iniparib	<i>PARP1</i>	Potential cytotoxic activity	
7	Tyrphostin AG 879	<i>ERBB2</i>	Inhibitor of tyrosine kinase	
8	Irbinitinib	<i>ERBB2</i>	Inhibitor of the human epidermal growth factor receptor tyrosine kinase ErbB-2	
9	Inosine	<i>PARP1</i>	Purine nucleoside	

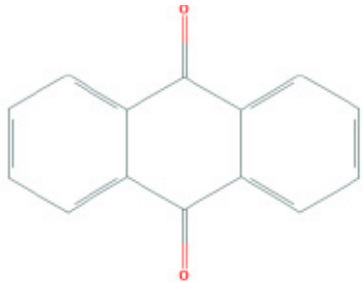
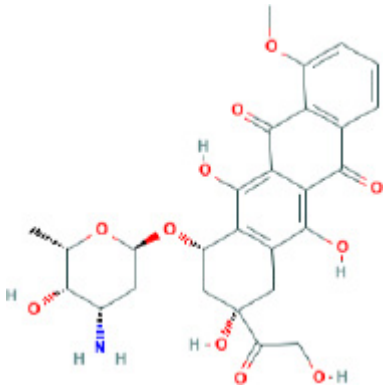
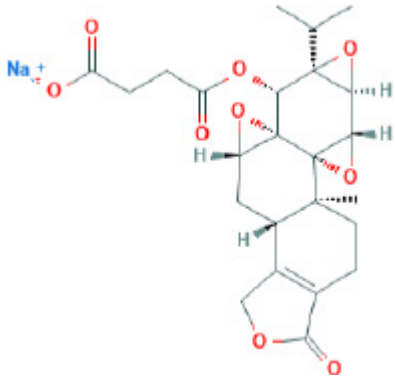
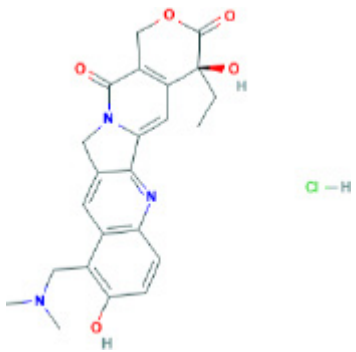
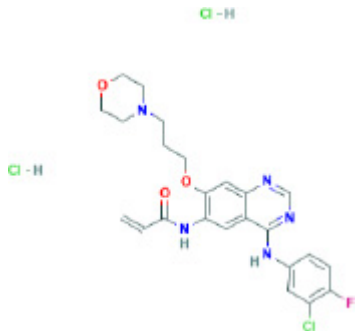
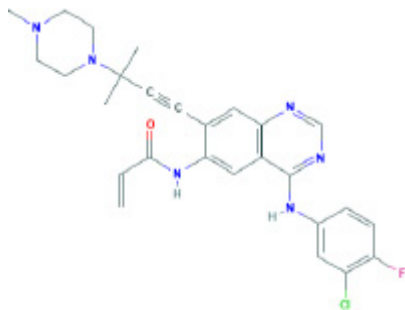
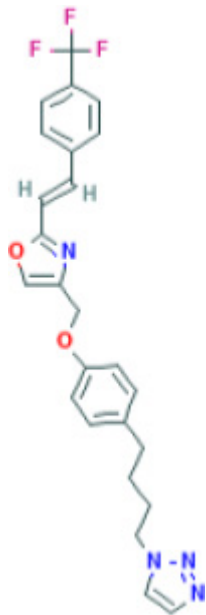
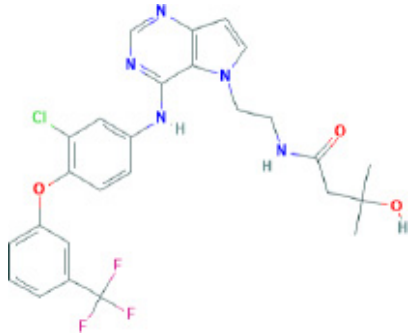
Table 1. Drugs enriched with differentially-expressed gene targets (cont.).			
Drug	Structure	Target gene	Mechanism
10 Anthraquinone		<i>ERBB2</i>	Polycyclic aromatic hydrocarbon derived from anthracene or phthalic anhydride
11 Doxorubicin		<i>GADD45A, ERBB2</i>	Inhibitor of protein synthesis
12 Omtriptolide Sodium		<i>NFKB1</i>	Inhibitor of topoisomerase II
13 Topotecan Hydrochloride		<i>TOP1MT</i>	Inhibitor of topoisomerase

Table 1. Drugs enriched with differentially-expressed gene targets (cont.).			
Drug	Structure	Target gene	Mechanism
14	TAK-285	<i>ERBB2</i>	Inhibitor of epidermal growth factor receptor
15	Mubritinib	<i>ERBB2</i>	Protein kinase inhibitor
16	MP-412	<i>ERBB2</i>	Protein kinase inhibitor
17	Canertinib	<i>ERBB2</i>	Protein kinase inhibitor





**Table 1. Drugs enriched with differentially-expressed gene targets (cont.).**

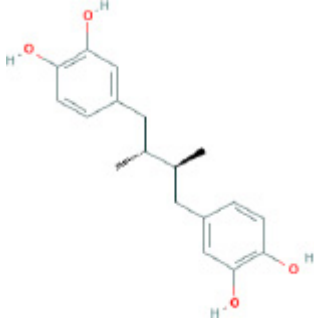
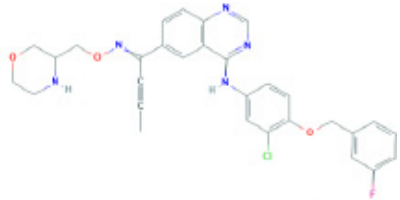
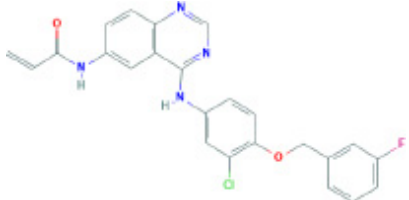
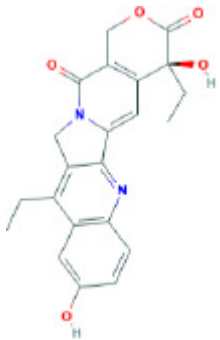
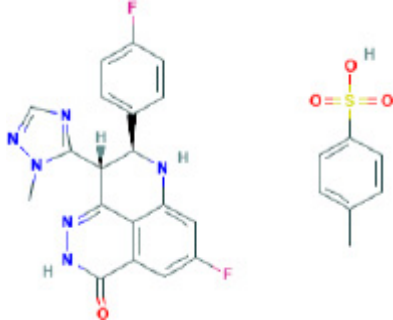
Drug	Structure	Target gene	Mechanism
18 INSM-18		<i>ERBB2</i>	Inhibitor of two receptor tyrosine kinases, the insulin-like growth factor receptor and the c-erbB2/HER2/neu receptor
19 5-222611		<i>ERBB2</i>	Protein kinase inhibitor
20 Allitinib		<i>ERBB2</i>	Inhibitor of the epidermal growth factor receptor 1 and human epidermal receptor 2
21 7-Ethyl-10-Hydroxy-Camptothecin		<i>TOP1MT</i>	Apoptosis inducer
22 Talazoparib Tosylate		<i>PARP1</i>	Inhibitor of the nuclear enzyme poly(ADP-ribose) polymerase

Table 1. Drugs enriched with differentially-expressed gene targets (cont.).			
Drug	Structure	Target gene	Mechanism
23	Varlitinib	<i>ERBB2</i>	Inhibitor of the epidermal growth factor receptor 1 and human epidermal receptor 2
24	Falnidamol	<i>ERBB2</i>	Epidermal growth factor receptor inhibitor
25	Pyrotinib	<i>ERBB2</i>	Inhibitor of the epidermal growth factor receptor 1 and human epidermal receptor 2
26	CP-724714	<i>ERBB2</i>	Inhibitor of human epidermal receptor 2
27	Bardoxolone	<i>NFKB1</i>	Apoptosis inducer

exhibited the lowest binding energy (-291 kcal/mol). Irinotecan achieved a binding energy of -240 kcal/mol and bardoxolone of -160 kcal/mol.

## Discussion

Previous studies demonstrated the emergent ability of mesenchymal stem cells (MSCs), and adipose-derived mesenchymal stem cells (AD-MSCs) as antiviral agents in patients with COVID-19 having properties of immunomodulation, regeneration and repair and antimicrobial [32–35]. MSC-based treatments have been reported in

**Table 2. Docking study results for SARS-CoV-2 3CL main protease.**

Drug	Binding energy (kcal/mol)
AZD-8055	-7.5
Olaparib	-7.39
Irinotecan	-8.4
Niacinamide	-4.43
Heptanoate	-4.19
Iniparib	-6.28
Tyrphostin AG 879	-7.67
Irbinitinib	-6.6
Inosine	-4.58
Anthraquinone	-6.2
Doxorubicin	-6.14
Omtriptolide sodium	-6.62
Topotecan hydrochloride	-7.4
TAK-285	-5.43
Mubritinib	-6.51
MP-412	-7.93
Canertinib	-6.04
INSM-18	-6.05
S-222611	-7.15
Allitinib	-7.1
7-ethyl-10-hydroxy-camptothecin	-7.07
Talazoparib tosylate	-6.43
Varlitinib	-6.2
Falnidamol	-7.64
Pyrotinib	-8.4
CP-724714	-6.43
Bardoxolone	-8.78

**Table 3. Binding energies based on molecular dynamic simulations for three drugs: bardoxolone, irinotecan and pyrotinib.**

Ligands	$\Delta G$ kcal/mol
Bardoxolone	-160,887 $\pm$ 5,877
Irinotecan	-240,473 $\pm$ 12,521
Pyrotinib	-291,571 $\pm$ 15,627

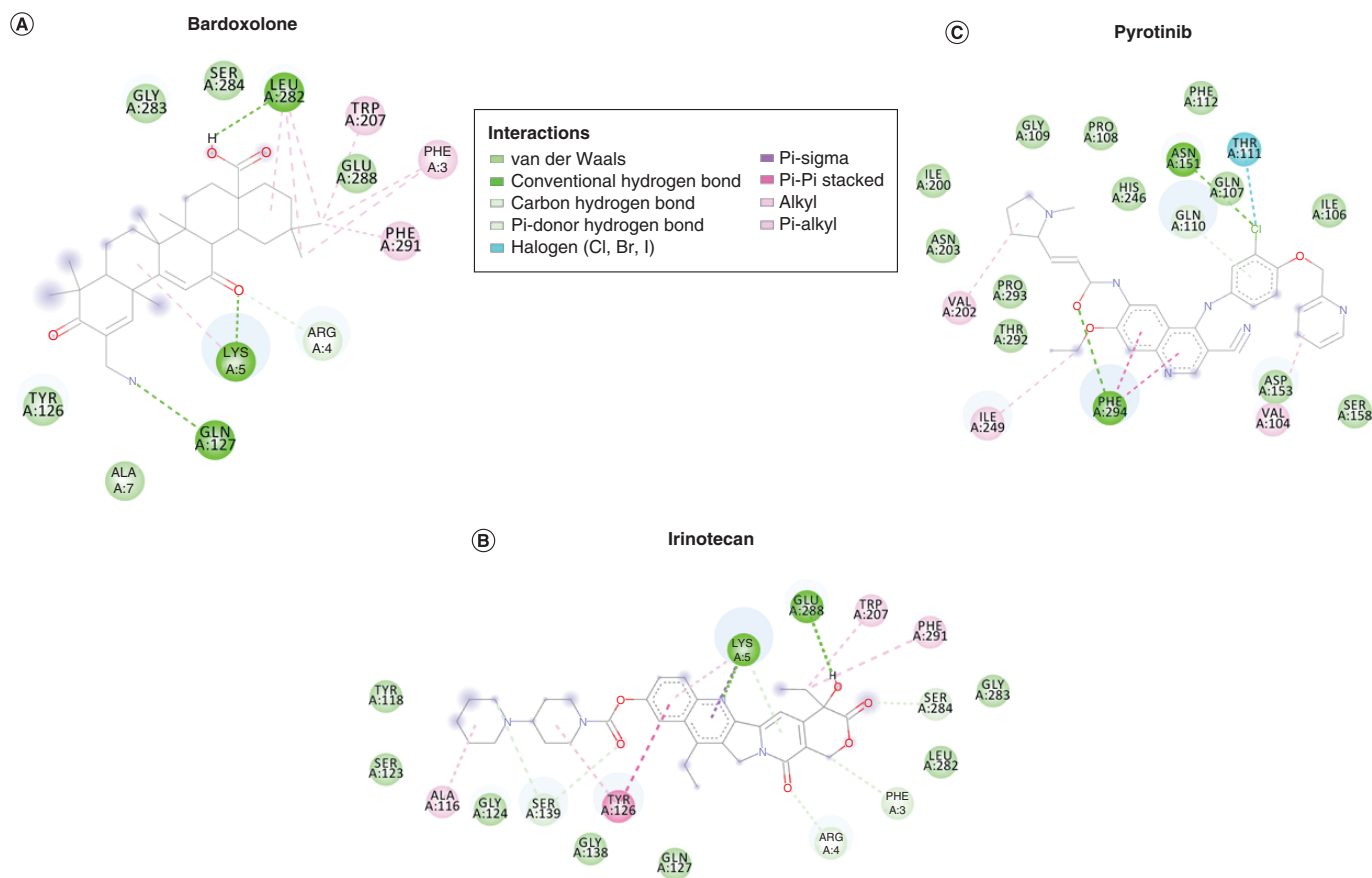
several autoimmune disorders successfully, but they have not still been led to a standard therapy [36]. Indeed, the US FDA did not approve MSCs for the treatment of COVID-19 because of limited data studied [36].

To date, FDA has approved five therapies for COVID-19: remdesivir, sotrovimab, combination of bamlanivimab and etesevimab, combination of casirivimab and imdevimab and COVID-19 convalescent plasma [37].

Remdesivir was one of the first antiviral drugs used in COVID-19 treatment and was authorised in over 50 countries [36]. It is an adenosine nucleotide analogue and inhibits the viral RNA-dependent RNA polymerase blocking the virus replication [38].

Sotrovimab is a recombinant human monoclonal antibody that binds to spike protein of SARS-CoV-2 and acts by avoiding the virus' entry into human cells [39].

Bamlanivimab and etesevimab, monoclonal antibodies, were approved by FDA to be administered together for the treatment of COVID-19. They are specifically directed against different but overlapping sites of the spike protein of SARS-CoV-2 stopping its binding to the human ACE2 receptor [40]. Similarly, the combination of two recombinant human monoclonal antibodies, casirivimab, and imdevimab, was reported for use in the treatment



**Figure 2.** Interaction diagram of three drugs (bardoxolone, irinotecan and pyrotinib) with the main protease of SARS-CoV-2. The figure shows detailed information about protein residues, involved amino acid and type of interactions.

of COVID-19 [41]. Convalescent plasma (CP), obtained from plasma derived from healthy donors recovered from COVID-19, contains virus-specific antibodies that could provide passive immunity in infected patients [42].

Despite the current standard therapy for COVID-19, new drugs should continue to be investigated as, the long-term effects of drugs and vaccines will be revealed in several years. In addition, recent studies demonstrated that a minimal variation in the sequence of SARS-CoV-2 could lead to modifications in the viral proteins making the vaccines or drugs ineffective [43].

To date, no mutations have been identified on 3CL main protease rendering it an appealing target for novel antiviral drugs [44]. Indeed, another innovative aspect of our computational approach is the focus on 3CL main protease that could help in identifying its inhibitors [45].

In this study, we proposed a novel computational approach integrating gene expression analysis and molecular docking to test *in silico* existing drugs against 3CL main protease of SARS-CoV-2.

In the first step of the study, we analyzed gene expression profiles of COVID-19 patients from two published GEO datasets. We identified 47 common differentially-expressed genes in COVID-19 patients compared with healthy subjects, and we verified the overrepresentation of these genes in the list of drug targets. We obtained 38 drugs as potential promising treatment for COVID-19 patients.

As the 3D structure for several drugs was not available, 27 out of 38 drugs were tested for docking analysis. Docking results suggested three potential protease inhibitors namely bardoxolone (binding energy: -8.78 kcal/mol), irinotecan (binding energy: -8.40 kcal/mol) and pyrotinib (binding energy: -8.40 kcal/mol). However, the best binding energy was obtained by bardoxolone. Based on differential expression and overrepresentation analysis, bardoxolone appears to be the best drug for treating COVID-19 patients. However, as 12 out of 27 drugs obtained a binding energy less than -7 kcal/mol (AZD-8055, olaparib, irinotecan, tyrphostin AG 879, topotecan hydrochloro-

ride, MP-412, S-222611, allitinib, 7-ethyl-10-hydroxy-camptothecin, faldinamol, pyrotinib and bardoxolone), these drugs can all be considered potential protease inhibitor ligands.

Bardoxolone was suggested as drug with novel potent antiviral properties against hepatitis B and C viruses in human hepatocyte cell culture systems and herpesvirus [46,47]. Bardoxolone could act as an agonist of activator of NRF2 impairing viral replication [46,47]. NRF2 induces HO-1 that was shown to suppress genome replication in hepatitis B and C. High *NRF2* activity suggests the role of *NRF2* as cellular defence against the progression of the infection. Indeed, several studies reported that the cells with a high level of *NRF2* are less likely to progress the infection. *NRF2* activity was also associated with a decreasing virus production in rotavirus [48].

Irinotecan (CPT-11) is a semisynthetic plant alkaloid derived from camptothecin and acts as an anticancer due to its ability to inhibit topoisomerase I during S phase; it is used to treat various cancers (i.e. colorectal cancer, pancreatic cancer, etc.), but its side effects limits its use. Nevertheless, it has been demonstrated in *in vitro* experiments to be effective in blocking herpes simplex virus-1 replication and lytic oncolysis [49]. Several nanoparticle- or lipid-based formulations have been developed to be used in different viral infectious disease, such as hepatitis A or influenza [50] carriers it is necessary to control the possible reactivation of the virus to avoid the development of impaired liver function or liver cancer. About 20% of HBV carriers developed reactivation [51]. In these subjects, irinotecan is used as a palliative treatment for advanced colorectal cancer, and its combination with lamivudine that inhibits HBV reverse transcription, could be helpful in controlling the virus reactivation [51].

Pyrotinib (Irene) is an anti-HER2 therapeutic target drug. This drug acts by covalently binding to ATP binding sites of the kinase domain: in this way the drug inhibits the formation of homo/heterodimer and auto-phosphorylation of HER family, blocking the signaling of the downstream pathway (RAS/RAF/MEK/MAPK, PI3K/AKT) and stopping the cells in G1. With its mechanism of action, this drug is effective on metastatic breast cancer [52]. Being an irreversible tyrosine kinase inhibitor, it blocks signal transduction through the erythroblastic leukemia viral oncogene homolog (erbB) receptors. Although no publication is present in the literature regarding the possible use of pyrotinib as antiviral agent, it is possible that this molecule interferes with viral entry by inhibiting tyrosine kinase activity of surface receptor. Indeed, it has been already reported that tyrosine kinase activity of Abl protein is necessary for the viral spike S-induced syncytia formation prior to the hemifusion step during infectious bronchitis virus (IBV) infection [53]. Moreover, the authors demonstrated that Abl inhibitor drugs are able to block IBV infection process.

Another promising drug against COVID-19 patients that achieved a binding energy of -7.39 kcal/mol is olaparib. It is utilized in current clinical use as poly adenosine diphosphate-ribose polymerase (PARP) inhibitors. PARP/ARTD enzymes are involved in a wide range of cellular processes such as antiviral response. PARP inhibitors have been shown to be effective in several models of acute respiratory distress syndrome [54]. Furthermore, a consequence of SARS-CoV-2 infection can be inflammatory response flaring out of control induced by IL-1 and IL-6. PARP inhibitors can reduce the expression of IL-1 and IL-6 decreasing the inflammatory response [55].

In conclusion, we proposed several existing drugs with antiviral activity that, alone or in combination, could be considered as drugs against COVID-19. However, it is vital to better understand the clinicopathological features of COVID-19 to design treatment plans leading to favorable outcomes in SARS-CoV-2 infected patients [56,57].

Although our work is the first study on the integration of gene expression, molecular docking and dynamic simulation in COVID-19 to our knowledge, it presents some limitations. Indeed, further docking and dynamic simulation studies are needed to validate the inhibitory effects of these drugs against COVID-19 using reference molecules as gold standard.

## Conclusion

In this study we have demonstrated that an *in silico* approach, based on integration of gene expression analysis of COVID-19-specific genes and molecular docking of therapeutic drugs on target proteins, revealed a new antiviral use of existing drugs on the basis of the interaction between each drug and specific domains on SARS-CoV-2 viral proteins. Dynamic simulation studies confirmed the promising role of three drugs: bardoxolone, irinotecan and pyrotinib.

### Summary points

- SARS-CoV-2 3CL main protease as a promising drug target.
- Gene expression profiles investigated the differentially-expressed genes in COVID-19 patients.
- The best drugs were chosen for their known interactions with genes altered by COVID-19.
- Molecular docking and dynamic simulation analyses were applied on the best drugs.

### Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: [www.futuremedicine.com/doi/suppl/10.2217/fvl-2020-0392](http://www.futuremedicine.com/doi/suppl/10.2217/fvl-2020-0392)

### Author contributions

C Cava has made substantial contribution to the acquisition, analysis and interpretation of data. G Bertoli has made substantial contribution to the interpretation of data. I Castiglioni revised the manuscript.

### Acknowledgments

IBFM-CNR thanks the generous support from Dassault Systèmes BIOVIA for the use of BIOVIA Discovery Studio Academic Research Suite in this work.

### Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

### References

1. Hall DC Jr, Ji HF. A search for medications to treat COVID-19 via *in silico* molecular docking models of the SARS-CoV-2 spike glycoprotein and 3CL protease. *Travel. Med. Infect. Dis.* 35, 101646 (2020).
2. Han DP, Penn-Nicholson A, Cho MW. Identification of critical determinants on ACE2 for SARS-CoV entry and development of a potent entry inhibitor. *Virology* 350(1), 15–25 (2006).
3. Brown JD. Antihypertensive drugs and risk of COVID-19? *Lancet Respir. Med.* 8(5), e28 (2020).
4. Lu IL, Mahindroo N, Liang PH *et al.* Structure-based drug design and structural biology study of novel nonpeptide inhibitors of severe acute respiratory syndrome coronavirus main protease. *J. Med. Chem.* 49(17), 5154–5161 (2006).
5. Blanchard JE, Elowe NH, Huitema C *et al.* High-throughput screening identifies inhibitors of the SARS coronavirus main proteinase. *Chem. Biol.* 11(10), 1445–1453 (2004).
6. Zhang L, Lin D, Sun X *et al.* Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved  $\alpha$ -ketoamide inhibitors. *Science* 368(6489), 409–412 (2020).
7. Jin Z, Du X, Xu Y *et al.* Structure of M<sup>pro</sup> from SARS-CoV-2 and discovery of its inhibitors. *Nature* 582(7811), 289–293 (2020).
8. Cava C, Bertoli G, Castiglioni I. *In silico* discovery of candidate drugs against COVID-19. *Viruses* 12(4), 404 (2020).
9. Cheng F, Murray JL, Rubin DH. Drug repurposing: new treatments for Zika virus infection? *Trends Mol. Med.* 22, 919–921 (2016).
10. Cheng F, Murray JL, Zhao J, Sheng J, Zhao Z, Rubin DH. Systems biology-based investigation of cellular antiviral drug targets identified by gene-trap insertional mutagenesis. *PLoS Comput. Biol.* 12, e1005074 (2016).
11. Mustafa S, Balkhy H, Gabere M. Peptide–protein interaction studies of antimicrobial peptides targeting middle east respiratory syndrome coronavirus spike protein: an *in silico* approach. *Adv. Bioinform.* 2019, doi: 10.1155/2019/6815105 (2019).
12. Lionta E, Spyrou G, Vassilatis DK, Cournia Z. Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr. Top. Med. Chem.* 14, 1923–1938 (2014).
13. Cheng LS, Amaro RE, Xu D, Li WW, Arzberger PW, McCammon JA. Ensemble-based virtual screening reveals potential novel antiviral compounds for avian influenza neuraminidase. *J. Med. Chem.* 51, 3878–3894 (2008).
14. Zhang L, Ai HX, Li SM *et al.* Virtual screening approach to identifying influenza virus neuraminidase inhibitors using molecular docking combined with machine-learning-based scoring function. *Oncotarget* 8(47), 83142–83154 (2017).
15. Shah B, Modi P, Sagar SR. *In silico* studies on therapeutic agents for COVID-19: drug repurposing approach. *Life Sci.* 252, 117652 (2020).

16. Aatif M, Muteeb G, Alsultan A, Alshoabi A, Khelif BY. Dieckol and its derivatives as potential inhibitors of SARS-CoV-2 spike protein (UK strain: VUI 202012/01): a computational study. *Mar. Drugs* 19(5), 242 (2021).
17. Prabhu D, Rajamanikandan S, Sureshan M, Jeyakanthan J, Saraboji K. modeling studies reveal the importance of the C-terminal inter motif loop of NSP1 as a promising target site for drug discovery and screening of potential phytochemicals to combat SARS-CoV-2. *J. Mol. Graph. Model.* 106, 107920 (2021).
18. Vivek-Ananth RP, Rana A, Rajan N, Biswal HS, Samal A. *In silico* identification of potential natural product inhibitors of human proteases key to SARS-CoV-2 infection. *Molecules* 25(17), 3822 (2020).
19. Gentile F, Agrawal V, Hsing M *et al.* Deep docking: a deep learning platform for augmentation of structure based drug discovery. *ACS Cent. Sci.* 6(6), 939–949 (2020).
20. Nguyen DD, Gao K, Chen J, Wang R, Wei GW. Potentially highly potent drugs for 2019-nCoV. *Biorxiv: the Preprint Server for Biology* doi: <https://doi.org/10.1101/2020.02.05.936013> (2020).
21. Salmaso V, Moro S. Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: an overview. *Front. Pharmacol.* 9, 923 (2018).
22. Colaprico A, Silva TC, Olsen C *et al.* TCGAAbiLinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 44(8), e71 (2016).
23. Günther S, Kuhn M, Dunkel M *et al.* SuperTarget and Matador: resources for exploring drug-target relationships. *Nucleic Acids Res.* 36(Database issue), D919–922 (2008).
24. Cotto KC, Wagner AH, Feng YY *et al.* DGIdb 3.0: a redesign and expansion of the drug-gene interaction database. *Nucleic Acids Res.* 46(D1), D1068–D1073 (2018).
25. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* 57, 289–300 (1995).
26. Morris GM, Huey R, Lindstrom W *et al.* AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* 30(16), 2785–2791 (2009).
27. Kim S, Chen J, Cheng T *et al.* PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res.* 47(D1), D1102–D1109 (2019).
28. Dassault Systèmes. BIOVIA, Discovery Studio Modeling Environment, release. (2020). <https://www.3ds.com/products-services/biovia/products/molecular-modeling-simulation/biovia-discovery-studio/>
29. Salentin S, Schreiber S, Haupt VJ, Adasme MF, Schroeder M. PLIP: fully automated protein–ligand interaction profiler. *Nucleic Acids Res.* 43(W1), W443–447 (2015).
30. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* 7, 42717 (2017).
31. Hastings J, Owen G, Dekker A *et al.* ChEBI in 2016: improved services and an expanding collection of metabolites. *Nucleic Acids Res.* 44(D1), D1214–D1219 (2016).
32. Gentile P, Sterodimas A, Pizzicannella J, Calabrese C, Garcovich S. Research progress on mesenchymal stem cells (MSCs), adipose-derived mesenchymal stem cells (AD-MSCs), drugs, and vaccines in inhibiting COVID-19 disease. *Aging Dis.* 11(5), 1191–1201 (2020).
33. Gentile P, Sterodimas A. Adipose stem cells (ASCs) and stromal vascular fraction (SVF) as a potential therapy in combating (COVID-19)-disease. *Aging Dis.* 11(3), 465–469 (2020).
34. Gentile P, Sterodimas A. Adipose-derived stromal stem cells (ASCs) as a new regenerative immediate therapy combating coronavirus (COVID-19)-induced pneumonia. *Expert Opin. Biol. Ther.* 20(7), 711–716 (2020).
35. Durand N, Mallea J, Zubair AC. Insights into the use of mesenchymal stem cells in COVID-19 mediated acute respiratory failure. *NPJ Regen. Med.* 5(1), 17 (2020).
36. Rad F, Ghorbani M, Mohammadi Roushandeh A, Habibi Roudkenar M. Mesenchymal stem cell-based therapy for autoimmune diseases: emerging roles of extracellular vesicles. *Mol. Biol. Rep.* 46(1), 1533–1549 (2019).
37. Food and Drug Administration. Emergency Use Authorization. <https://www.fda.gov/emergency-preparedness-and-response/mcm-legal-regulatory-and-policy-framework/emergency-use-authorization#coviddrugs>
38. Gordon CJ, Tchesnokov EP, Feng JY, Porter DP, Götte M. The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus. *J. Biol. Chem.* 295(15), 4773–4779 (2020).
39. Ledford H. COVID antibody treatments show promise for preventing severe disease. *Nature* 591(7851), 513–514 (2021).
40. Huang DT, McCreary EK, Bariola JR *et al.* The UPMC OPTIMISE-C19 (optimizing treatment and impact of monoclonal antibodies through evaluation for COVID-19) trial: a structured summary of a study protocol for an open-label, pragmatic, comparative effectiveness platform trial with response-adaptive randomization. *Trials* 22(1), 363 (2021).
41. Hansen J, Baum A, Pascal KE *et al.* Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science* 369(6506), 1010–1014 (2020).

42. Devarasetti PK, Rajasekhar L, Baisya R, Sreejitha KS, Vardhan YK. A review of COVID-19 convalescent plasma use in COVID-19 with focus on proof of efficacy. *Immunol. Res.* 69(1), 18–25 (2021).
43. Naqvi AAT, Fatima K, Mohammad T *et al.* Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: structural genomics approach. *Biochim. Biophys. Acta. Mol. Basis Dis.* 1866(10), 165878 (2020).
44. Klemm T, Ebert G, Calleja DJ *et al.* Mechanism and inhibition of the papain-like protease, PL<sup>pro</sup>, of SARS-CoV-2. *EMBO J.* 39(18), e106275 (2020).
45. Bepari AK, Reza HM. Identification of a novel inhibitor of SARS-CoV-2 3CL-PRO through virtual screening and molecular dynamics simulation. *Peer J.* 9, e11261 (2021).
46. Nio Y, Sasai M, Akahori Y *et al.* Bardoxolone methyl as a novel potent antiviral agent against hepatitis B and C viruses in human hepatocyte cell culture systems. *Antiviral Res.* 169, 104537 (2019).
47. Wyler E, Franke V, Menegatti J *et al.* Single-cell RNA-sequencing of herpes simplex virus 1-infected cells connects NRF2 activation to an antiviral program. *Nat. Commun.* 10(1), 4878 (2019).
48. Patra U, Mukhopadhyay U, Sarkar R, Mukherjee A, Chawla-Sarkar M. RA-839, a selective agonist of Nrf2/ARE pathway, exerts potent anti-rotaviral efficacy *in vitro*. *Antiviral Res.* 161, 53–62 (2019).
49. Kulu Y, Kawasaki H, Donahue JM *et al.* Concurrent chemotherapy inhibits herpes simplex virus-1 replication and oncolysis. *Cancer Gene Ther.* 20(2), 133–140 (2013).
50. *Lipid Nanocarriers for Drug Targeting* Grumezescu AM (Ed.). William Andrew (2018).
51. Lui LCY, Chan RTT, Kwok CCH *et al.* Reactivation of Hepatitis B after Irinotecan. *J. Hong Kong College of Radiologists* 7(2), 72 (2004).
52. Gourd E. Pyrotinib shows activity in metastatic breast cancer. *Lancet Oncol.* 18(11), e643 (2017).
53. Sisk JM, Frieman MB, Machamer CE. Coronavirus S protein-induced fusion is blocked prior to hemifusion by Abl kinase inhibitors. *J. Gen. Virol.* 99(5), 619–630 (2018).
54. Curtin N, Bányai K, Thaventhiran J, Le Quesne J, Helyes Z, Bai P. Repositioning PARP inhibitors for SARS-CoV-2 infection(COVID-19); a new multi-pronged therapy for acute respiratory distress syndrome? *Br. J. Pharmacol.* 177(16), 3635–3645 (2020).
55. Morrow DA, Brickman CM, Murphy SA *et al.* A randomized, placebo-controlled trial to evaluate the tolerability, safety, pharmacokinetics, and pharmacodynamics of a potent inhibitor of poly(ADP-ribose) polymerase (INO-1001) in patients with ST-elevation myocardial infarction undergoing primary percutaneous coronary intervention: results of the TIMI 37 trial. *J. Thromb. Thrombolysis.* 27(4), 359–364 (2009).
56. Sabetian S, Castiglioni I, Jahromi BN, Mousavi P, Cava C. *In silico* identification of miRNA-lncRNA interactions in male reproductive disorder associated with COVID-19 infection. *Cells.* 10(6), 1480 (2021).
57. Cava C, Bertoli G, Castiglioni I. A protein interaction map identifies existing drugs targeting SARS-CoV-2. *BMC Pharmacol Toxicol.* 21(1), 65 (2020).