Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Docking-based virtual screening studies aiming at the covalent inhibition of SARS-CoV-2 M\textsuperscript{Pro} by targeting the cysteine 145

Laurent Soulère *, Thibaut Barbier, Yves Queneau

Univ Lyon, Université Claude Bernard Lyon 1, INSA Lyon, CPE Lyon, UM3 5246, CNRS, ICBMS, Institut de Chimie et de Biochimie Moléculaires et Supramoléculaires, Chimie Organique et Bioorganique, Bat. E. Lederer, 1 rue Victor Grignard, F-69622, Villeurbanne, France

1. Introduction

COVID-19 is an infectious disease caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Lai et al., 2020). In November 2020, this disease has infected more than 55 300 000 people worldwide, including more than 1 300 000 deaths (https://covid19.who.int). Consequently, there is an urgent need to identify new antiviral drugs targeting this virus. Several strategies for identifying coronavirus anti-viral drugs have been described in the literature (Kitchen et al., 2004), docking studies for covalent protein inhibition are less common (Kumalo et al., 2015; Sotriffer, 2018), in particular with SARS-CoV-2 main protease (Liu et al., 2020; Paul et al., 2020). Despite covalent inhibition approaches are less studied because the requirement of a nucleophilic residue is a structural limitation and they can be considered as harmful, the resurgence of covalent drugs encourage to also consider covalent inhibition (Dalton and Campos, 2020; Ghosh et al., 2019; Singh et al., 2011). Irreversible specific protein inhibitors are now reported, such as in the case of the Ras protein possessing a G12C mutation, a promising example of potential anticancer strategy (Goody et al., 2019). Recent studies describe the inactivation of the SARS-CoV-2 M\textsuperscript{Pro} with either the N3 inhibitor (Jin et al., 2020), originally discovered for SARS-CoV (Yang et al., 2005), or the alpha-ketoamide inhibitor 1 (Fig. 1) (L. Zhang et al., 2020). These studies show the importance of the catalytic Cys145 residue to design covalent inhibitors. Taking advantage of the X-ray crystal structure of the complex M\textsuperscript{Pro}-compound 1 (Zhang et al., 2020), we report herein flexible docking studies to identify potential irreversible inhibitors using...
2. Materials and methods

2.1. Proteins used in the present study

The main protein of SARS-CoV-2 M\textsuperscript{Pro} was obtained from the protein data bank (PDB code 6Y2F, 5RHF, 5REN, 5REK).

Docking studies of known covalent inhibitors 2-4 with crystallographic data

Compounds 2-4 were generated as 3D mol file using Arguslab (Thompson, 2004) and they were then docked within the M\textsuperscript{Pro} active site (PDB codes 5RHF, 5REN, 5REK respectively) using Arguslab software with the Argusdock engine with default parameters. The obtained binding modes were compared to crystallographic data to validate the method. The distance between the sulfur and the methylene group of the chloroacetamide group was monitored to establish the structural bases of the M\textsuperscript{Pro} covalent inhibition (distance \(< 4 \text{ Å})

2.2. Docking-based virtual screening of potential covalent inhibitors

Compounds libraries were obtained from PubChem using structure search of acrylamide, halogenoacetamide, cyanoacrylamide and vinylsulfonamide. Asinex and Enamine soft electrophilic compounds libraries were obtained from the corresponding website and hydrogen atoms were added using OpenBabel 2.4.1 (O’Boyle et al., 2011). Each

![Fig. 1. Structure, representation of the alpha-ketoamide inhibitor 1 within the M\textsuperscript{Pro} active site with surface and hydrogen bonds network. (L. Zhang, et al., 2020).](image1)

![Fig. 2. Docking studies for covalent inhibition of SARS-CoV-2 M\textsuperscript{Pro} by targeting the Cys145 residue.](image2)
compound from the different libraries (SDF file) were then docked within the MPro active site centered on the β-keto amide 1 using Arguslab software with the Argusdock engine with default parameters. Based on their docking score, binding modes were analyzed for the 20 best compounds of each library. The distance between the sulfur and the electrophilic center was monitored and compounds were selected when a distance < 4 Å was measured. Binding modes were examined with PyMOL and hydrogen bonds networks were generated using LigPlot+ (Laskowski and Swindells, 2011). For all selected compounds, further flexible docking experiments with a genetic algorithm engine implemented in Arguslab were achieved to corroborate the docking results (see supplementary information).

2.3. Docking studies of the approved covalent drugs library

The library was constructed based on the list of covalent drugs established by Vasudevan and co-workers (Vasudevan et al., 2019), from which β-lactam derivatives, drugs withdrawn from market and mechanism based covalent drugs were excluded. For five of the 29 remaining compounds, their metabolites which are the active covalent binders (Shin and Kim, 2013) were used in the library. Overall, 32 compounds were used for the docking studies. When available, the SDF file for each compound was obtained from PubChem and was converted to PDB file using Accelrys Visualizer 2.0. For omeprazole, lansoprazole, pantoprazole and rabeprazole active metabolites, they were drawn using Vega ZZ (Pedretti et al., 2002, 2004) and were saved as PDB files. The 32 compounds were then docked within the MPro active site centered on the β-keto amide 1 using Arguslab software (Thompson, 2004) with the Argusdock engine with default parameters. The distance between the sulfur and the electrophilic center was monitored and compounds were selected when a distance < 4 Å was measured. Binding modes were examined with PyMOL and hydrogen bonds networks were generated using LigPlot+ (Laskowski and Swindells, 2011). Further flexible docking experiments with a genetic algorithm engine implemented in Arguslab were achieved to corroborate the docking results (see supplementary information).

3. Results and discussion

When applicable, the search for covalent inhibitors by docking can be carried out according to the scheme depicted in Fig. 2 by targeting a nucleophilic residue such as a cysteine residue. It is then necessary that the electrophilic center is placed at the vicinity of the thiol function of this residue. This approach can be investigated with the MPro protein that contains a catalytic cysteine residue (Cys145). In order to explore

![Fig. 3. Docking results (colored in cyan) obtained from docking experiments conducted with compounds 2-4 within the active site of MPro (co-crystallized compounds are indicated in green for comparison). Distance (Å) between the thiol function of Cys145 and the electrophilic center is shown in red color.](image)

### Table 1

| Compounds libraries | Structures          | Compounds number/potential inhibitors | Adducts               |
|---------------------|---------------------|---------------------------------------|-----------------------|
| PubChem acrylamide  | ![Acrylamide](image) | 3905/1                                | Michael adduct        |
| PubChem halogenoacetamide | ![Halogenoacetamide](image) | 4281/2                               | Nucleophilic substitution |
| PubChem cyanoacrylamide | ![Cyanoacrylamide](image) | 3356/1                               | Reversible covalent adduct |
| PubChem vinylsulfonamide | ![Vinylsulfonamide](image) | 1070/2                               | Michael adduct        |
| Asinex Soft electrophile | ![Soft electrophile](image) | 8098/3                               | Structure dependent   |
| Enamine Covalent Screening Library | ![Covalent Screening Library](image) | 15684/2                             | Structure dependent   |
| Enamine Halogenoacetamide | ![Halogenoacetamide](image) | 2210/4                               | Nucleophilic substitution |
| Enamine Acrylamide | ![Acrylamide](image) | 3153/2                               | Michael adduct        |
the potentialities to covalently inhibit $M^{\text{Pro}}$, a flexible docking of compounds was performed to establish the structural bases for binding recognition with high affinity to improve the selectivity to the molecular target to form the complex $M^{\text{Pro}} + I$. In a second time, examination of the potentialities of the covalent inhibition of this protein due to the formation of a covalent adduct $M^{\text{Pro}}-I$ can be examined. For this, the location of the electrophilic moiety of compounds, in particular the distance between the electrophilic center and the thiol of Cys145 has to be adequate (Zhang et al., 2016) (Fig. 2). As a consequence, we used standard docking experiments targeting Cys145 with a cutoff of 4Å as a theoretical study (Choi et al., 2016; Lin et al., 2011).

3.1. Docking studies of known covalent inhibitors of SARS-CoV-2 $M^{\text{Pro}}$

In order to explore the potentialities to covalently inhibit SARS-CoV-2 $M^{\text{Pro}}$, we first studied docking of the covalent inhibitors 2-4 co-crystallized within the protease to establish the structural bases for the covalent inhibition of this protein (Douangamath et al., 2020). In particular, the location of the electrophilic moiety of compounds i.e. the distance between the electrophilic center and the thiol of Cys145 was investigated. Docking experiments were thus conducted on these three compounds; the results are depicted in Fig. 3. These experiments show

![Fig. 4. Structures of potential covalent inhibitors of $M^{\text{Pro}}$, with PubChem CID number.](image_url)

Table 2
Potential covalent inhibitors of $M^{\text{Pro}}$: docking score, hydrogen bonds numbers and $S_n$ indicates the binding pockets interacting with the ligands (Zhang et al., 2020), sulfur – electrophilic group distances.

| Compounds | Docking score (kcal. mol$^{-1}$) | H-bonds numbers, $S_n$ | Sulfur – electrophilic center distances (Å) |
|-----------|---------------------------------|------------------------|------------------------------------------|
| 1658938   | −9.19                           | 1, S1 S3 S4            | 3.62                                     |
| 1625245   | −9.83                           | 3, S1 S3 S4            | 4.03                                     |
| 4868406   | −9.28                           | 2, S1 S2 S4            | 3.40                                     |
| 68782938  | −9.21                           | 5, S1 S2 S3 S4         | 3.48                                     |
| 2011299   | −10.32                          | 3, S1 S2 S4            | 3.18                                     |
| 134294169 | −9.43                           | 6, S1 S2 S3 S4         | 3.80                                     |
| 3207595   | −10.18                          | 1, S1 S2 S4            | 3.46                                     |
| 54693381  | −9.99                           | 3, S1 S2 S3            | 2.97                                     |
| 1102141   | −10.41                          | 1, S1 S3 S4            | 3.82                                     |
| 132327024 | −7.22                           | 3, S1                  | 3.94                                     |
| 132349371 | −7.35                           | 3, S1 S2 S3            | 3.76                                     |
| 2405938   | −8.19                           | 3, S1 S2 S3            | 3.27                                     |
| 2404847   | −7.94                           | 1, S1 S2 S3            | 3.51                                     |
| 2513345   | −7.96                           | 2, S1 S3               | 3.57                                     |
| 7788967   | −8.49                           | 2, S1 S2 S3            | 3.89                                     |
| 53520431  | −8.81                           | 3, S1 S3               | 2.27                                     |
| 91944335  | −8.39                           | 4, S1 S2 S3            | 3.99                                     |
consistent results with the crystallographic data (PDB codes 5RHF, 5REN, 5REK respectively) and show that the electrophilic center of compounds i.e. the methylene group of the chloroacetamide functional group is located at the vicinity of Cys145 with distance values between the carbon and the thiol atoms ranging from 2.83 Å to 3.19 Å.

Based on these docking experiments of compounds 2-4, we defined the structural basis for the covalent inhibition of Mpro as 1) binding the protein with a good affinity and 2) having an adequate orientation of the electrophilic moiety towards Cys145 with a distance inferior to 4 Å with the sulfur (Fig. 2).

3.2. Docking-based virtual screening studies for covalent inhibition of M\textsuperscript{Pro}

We then investigated docking-based virtual screening studies for covalent inhibition of M\textsuperscript{Pro} using libraries of electrophilic inhibitors (Table 1). A total of 41,757 compounds in diverse libraries constructed from PubChem or available from Enamine or Asinex were employed for docking simulations. Docking poses for compounds ranked in the top 20 in each library were then examined to evaluate the distance between the electrophilic center and the sulfur atom.

Structures of all potential covalent inhibitors of M\textsuperscript{Pro} are shown in Fig. 4. All compounds were found to be aromatic ones with a high diversity of structures. Among the 17 compounds, three compounds are structurally related to acrylamides, two to vinylsulfonamides and six to electrophilic chlorinated compounds. An alkyne derivative and three activated cyano-compounds could be also identified as well as two cyanoacrylamides. These latter compounds are particularly interesting leading to reversible covalent inhibition (Serafimova et al., 2012).

Binding mode of each compound was carefully examined in terms of hydrogen bonds with the protein M\textsuperscript{Pro}, and the distance between the thiol and the electrophilic center was also measured (Table 2). Hydrogen bonds numbers were ranging from one to six for the compound 134294169. The docking score was found to be variable depending of the compounds library and the structure of the compounds. These values were ranging from – 7.22 to – 10.41 kcal mol\textsuperscript{−1} for the compound 1102141.

As example, we chose to depict the binding modes of two compounds with high docking score values. The binding modes of the compound 2011299, which exhibits a low distance between the thiol and the electrophilic center and the one of the compound 134294169 with six hydrogen bonds, are described in Fig. 5. The two compounds fit well in the active site. The compound 2011299 interacts with the Gly143, Ser144 and Cys145 residues via hydrogen bonds and with a distance of 3.18 Å between the sulfur and the electrophilic carbon atom of the cyanoacrylamide functional group. This group is of particular interest by potentially promoting reversible covalent inhibition. The compound 134,294,169 interacts tightly within the active site with the His 41, Ser144 and Cys145 residues via hydrogen bonds and with a distance of 3.80 Å between the sulfur and the electrophilic carbon atom of the vinyl sulfonamide functional group. We also mention the compound 1,658,938 with a valuable docking score and a distance of 3.62 Å. Indeed, this compound shows a low cytotoxicity which has been demonstrated in several bioassays (see PubChem

Fig. 5. Binding modes of compounds 2011299 (A) and 134294169 (B) with the corresponding hydrogen bonds network.
3.3. Flexible docking studies of an approved covalent drugs library targeting the Cys145 residue

Based on the approach described in Fig. 2, we investigated a library of known approved covalent drugs. For this, we used a list of 52 FDA (Food and Drug administration) approved drugs described by Anil Vasudevan and colleagues in 2019 (Vasudevan, et al., 2019). Among these, we excluded some molecules for our study, such as drugs withdrawn from the market and drugs with a mechanism-based inhibition, resulting in a small library of 32 compounds. On this library, the workflow described in Fig. 2 was applied by means of flexible docking experiments and subsequent distance measurement between the Cys145 residue and the electrophilic center. Following this experiment, four compounds showed interesting results and they are presented in Table 3.

Careful analysis of the proposed binding mode of these compounds with the corresponding hydrogen bonds network revealed that they could be able to bind tightly within the $\text{M}^{\text{Pro}}$ binding site. The distance between the sulfur and the electrophilic center is ranging from 2.98 and 3.78 Å showing their potentialities to act as covalent inhibitors. The 2D-representation is described in Fig. 6 for the four compounds, showing the hydrogen bonds networks.

These four compounds are approved drugs for a designated pathology (Table 4) and some literature data can be found about relationships between COVID-19 and these compounds. First, Vittorio Mantero and colleagues reported that the use of dimethylfumarate for patients suffering from multiple sclerosis seems to have a positive impact against COVID-19 (Mantero et al., 2020). For ibrutinib, two studies show that this compound seems to have a protective role against COVID-19, although they hypothesize that it may be due to the anti-inflammatory effect of the Bruton’s tyrosine kinase pathway inhibition (Thibaud et al., 2020; Treon et al., 2020). For saxagliptin, Alicja Krejner-Bienias and colleagues hypothesized that antidiabetic drugs, gliptins, could prevent the virus from binding to dipeptidyl peptidase IV (DPP IV) (Krejner-Bienias et al., 2020). Saxagliptin could then have a double effect on both $\text{M}^{\text{Pro}}$ and DPP IV. For fosfomycin, no relationship between this compound and COVID-19 can be found in the literature to date.

4. Conclusion

To summarize, we described docking studies directed toward the identification of potential covalent inhibitors of SARS-CoV-2 $\text{M}^{\text{Pro}}$ with significant structural diversity by targeting the cysteine 145 residue. A docking-based virtual screening approach of 41,757 compounds belonging to libraries of soft electrophilic small molecules, allowed us to unveil 17 more potential covalent inhibitors. Then, using a library of biological assays 435019, 1825, 504648, 602141.

### Table 3

| Compounds (electrophilic moiety) | Binding modes | H-bonds, Sn | S – electrophilic center distances |
|---------------------------------|---------------|-------------|---------------------------------|
| Dimethylfumarate                |               | 4, S1       | 2.98 Å                          |
| (α,β-unsaturated ester)         |               |            |                                 |
| Fosfomycin                      |               | 3, S1       | 3.73 Å                          |
| (epoxide)                       |               |            |                                 |
| Ibrutinib                       |               | 5, S1 S3 S4 | 3.39 Å                          |
| (acrylamide)                    |               |            |                                 |
| Saxagliptin                     |               | 4, S1 S3    | 3.78 Å                          |
| (nitride)                       |               |            |                                 |

L. Soulere et al.  Computational Biology and Chemistry 92 (2021) 107463
approved covalent drugs, four compounds namely dimethylfumarate, fosfomycin, ibrutinib and saxagliptin were identified for their theoretical ability to first bind to the active site of the protein and second to form a covalent bond with the catalytic cysteine. This study provides structural insights in the covalent inhibition of MPro, which might be useful in the search for therapeutic approaches fighting COVID-19.

CRediT authorship contribution statement

Laurent Soulé: Conceptualization, Investigation, Methodology, Writing - review & editing. Thibaut Barbier: Investigation, Methodology, Writing - review & editing. Yves Queneau: Writing - review & editing.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgment

The authors are grateful to CNRS, France MESRI, France for financial

Table 4

| International nonproprietary name | Biological Target | Therapeutic Domain |
|----------------------------------|-------------------|--------------------|
| Dimethylfumarate                 | NF-κB activation  | Multiple sclerosis |
| Fosfomycin                       | UDP-N-acetylglucosamine-3-enolpyruvyltransferase | Anti-infective |
| Ibrutinib                        | Bruton’s tyrosine kinase (BTK) | Cancer |
| Saxagliptin                      | Dipeptidyl peptidase IV (DPP IV) | Anti-diabetic drug |

Fig. 6. 2D representation of the proposed binding modes of the four potential covalent inhibitors within the binding site of MPro: dimethylfumarate (A), fosfomycin (B), ibrutinib (C), saxagliptin (D).
support and to the MESRI, France for a scholarship to Thibaut Barbier.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.compbiolchem.2020.21.107463

References

Arun, K.G., Sharanya, C.S., Abhithaj, J., Francis, D., Sadasivan, C., 2020. Drug repurposing against SARS-CoV-2 using E-pharmacophore based virtual screening, molecular docking and molecular dynamics with main protease as the target. J. Biomol. Struct. Dyn. 1–12. https://doi.org/10.1080/07391102.2020.1779819.

Barrett, A., 2006. DFP-4 inhibitors and their potential role in the management of type 2 diabetes. Int. J. Clin. Pract. 60, 1454–1470.

Burger, J.A., Tedeschi, A., Barr, P.M., Robak, T., Owen, C., Ghia, P., Bairey, O., 2020. Drug discovery for the treatment of the epidemic and the challenges. Int. J. Antimicrob. Agents 55, 105924.

Douangamath, A., Fearon, D., Gehrtz, P., Krojer, T., Lukacik, P., Owen, C.D., Resnick, E., 2020. Covalent small molecules as enabling platforms for drug discovery: methods and applications. Nat. Rev. Drug Discov. 3, 668–688.

Pillayar, T., Manimack, M., Namastivayam, V., Hayashi, Y., Jung, S.H., 2016. An overview of severe acute respiratory syndrome-coronavirus (SARS-CoV) 3CL protease inhibitors: peptidomimetics and small molecule chemotherapy. J. Med. Chem. 59, 6695–6626.

Pillayar, T., Meenakshisundaram, S., Manimack, M., 2020. Recent discovery and development of inhibitors targeting coronavirus. Drug Disc. Today. 25, 668–688.

Serafinova, I.M., Pufall, M.A., Krishnan, S., Duda, K., Cohen, M.S., Maglathlin, R.L., McFarland, J.M., Miller, R.M., Frodia, M., Tauto, J., 2012. Reversible targeting of noncatastytic cystines with chemically tuned electrophiles. Nat. Chem. Biol. 8, 471–476.

Shin, J.M., Kim, N., 2013. J. Neurogastroenterol. Motil. 19, 25–39.

Singh, J., Petter, R.C., Baillie, T.A., Whitty, A., 2011. The resurgence of covalent drugs. J. Med. Chem. 59, https://doi.org/10.1021/jm203730s.

Thibaud, S., Tremblay, D., Bhalla, S., Ziget, K., Gabrilove, J., 2020. Protective role of Bruton tyrosine kinase inhibitors in patients with chronic lymphocytic leukemia and COVID-19. Br. J. Haematol. 190, e73–e75.

Thompson, M.A., 2004. ArgusLab 4.0.1, Seattle, WA, Planetary Software LLC.

Ton, A.T., Gentle, F., Hsing, M., Ban, F., Cherkasov, A., 2020. Rapid identification of potential inhibitors of SARS-CoV-2 main protease by deep docking of 1.3 billion products. Mol. Inform. 39 e2000028.

Weis, D., Raff, S., Riedl, R., 2013. Disease-modifying treatments for multiple sclerosis - a review of approved medications. Eur. J. Neurol. 23 (Suppl. 1), 18–27.

Treon, S.P., Castillo, J.J., Skarabin, A.P., Soumerai, J.D., Ghobrial, I.M., Guerrera, M.L., Meid, K., Yang, G., 2020. The BTK inhibitor ibrutinib may protect against pulmonary injury in COVID-19-infected patients. Blood 135, 1912–1915.

Vasudevan, A., Argiriadi, M.A., Baranczak, A., Friedman, M.M., Gavrilyuk, J., Hobson, A.D., Hulse, J.J., Osman, S., Wilson, N.S., 2019. Covalent binders in drug discovery. Prog. Med. Chem. 1–62.

Wang, J., 2020. Fast identification of possible drug treatment of coronavirus Disease-19 (COVID-19) through computational drug repurposing study. J. Chem. Inf. Model. 60, 3277–3286.

Yang, H., Xie, W., Xue, X., Yang, K., Mu, J., Liang, W., Zhao, Z., Pei, D., Ziebaur, J., Hilgenfeld, R., Yu, K.Y., Wang, G., Chen, S., Chen, Z., Mu, D., Bartlam, M., Rao, Z., 2005. Design of wide-spectrum inhibitors targeting coronavirus main proteases. PLoS Biol. 3 e324.

Zhang, Y., Zhang, D., Tian, H., Jiao, Y., Shi, Z., Ran, T., Liu, H., Lu, S., Xu, A., Xiao, C., Pan, Y., Yin, L., Zhou, Y., Lin, T., 2016. Identification of covalent binding sites targeting cystines based on computational approaches. Mol. Pharm. 13, 3106–3118.

Zhang, L., Liu, D., Sun, X., Caruth, U., Drosten, C., Sauerhering, L., Becker, S., Rux, K., Hilgenfeld, H., 2020. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved alpha-keotimide inhibitors. Science 368, 409–412.