



HAL
open science

Design and synthesis of naturally-inspired SARS-CoV-2 inhibitors

Haitham Hassan, Jeanne Chiavaralli, Afnan Hassan, Loay Bedda, Tim Krischuns, Kuang-Yu Chen, Alice Shi Ming Li, Adrien Delpal, Etienne Decroly, Masoud Vedadi, et al.

► **To cite this version:**

Haitham Hassan, Jeanne Chiavaralli, Afnan Hassan, Loay Bedda, Tim Krischuns, et al.. Design and synthesis of naturally-inspired SARS-CoV-2 inhibitors. RSC Medicinal Chemistry, 2023, 14 (3), pp.507-519. 10.1039/D2MD00149G . hal-03981817

HAL Id: hal-03981817

<https://hal.science/hal-03981817>

Submitted on 6 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Design and Synthesis of Naturally-Inspired SARS-CoV-2 Inhibitors

Haitham Hassan,^{a*} Jeanne Chiavaralli,^b Afnan Hassan,^{c,d} Loay Bedda,^{c,d} Tim Krischuns,^e Kuang-Yu Chen,^e Alice Shi Ming Li,^f Adrien Delpal,^g Etienne Decroly,^g Masoud Vedadi,^{f, h} Nadia Naffakh,^e Fabrice Agou,^b Sergio Mallart,^a Reem K. Arafa^{c,d*}, and Paola B. Arimondo^{a*}

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

A naturally inspired chemical library of 25 molecules was synthesised guided by 3-D dimensionality and natural product likeness factors to explore a new chemical space. The synthesised chemical library, consisting of fused-bridged dodecahydro-2a,6-epoxyazepino[3,4,5-cd] indole skeletons, followed lead likeness factors in terms of molecular weight, C-sp³ fraction and CLogP. Screening of the 25 compounds against lung cells infected with SARS-CoV-2 led to the identification of 2 hits. Although the chemical library showed cytotoxicity, the two hits (**3b**, **9e**) showed the highest antiviral activity (EC₅₀ values of 3.7 and 1.4 μM, respectively) with acceptable cytotoxicity difference. Computational analysis based on docking and molecular dynamics simulations against main protein targets in SARS-CoV-2 (Main Protease M^{pro}, Nucleocapsid phosphoprotein, Non-structural protein nsp10/nsp16 complex and RBD/ACE2 complex) were performed. The computational analysis proposed the possible binding targets to be either M^{pro} or nsp10/nsp16 complex. Biological assays were performed to confirm this proposition. A cell-based assay for M^{pro} protease activity using a Reverse-Nanoluciferase (Rev-Nluc) reporter confirmed that **3b** is targeting M^{pro}. These results open the way towards further hit-to-lead optimisations.

Introduction

The late 2019 coronavirus disease outbreak turned out into the global COVID-19 pandemic that is still on-going.¹ The causative agent of the COVID-19 disease was identified as the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).²⁻⁴ This highly transmissible airborne virus is a positive-sense single-stranded RNA virus, whose genome encodes 29 proteins including four structural proteins.⁵ Although there are currently different vaccines and treatments available to tackle this pandemic, research efforts should continue to find therapeutic agents against the virus, as well as variants of concern or other zoonotic coronaviruses that might emerge in the future¹⁴⁻¹⁶. There are now several treatments for COVID19. For example, Remdesivir, which has been granted emergency-use authorisation, showed modest benefit to the patients with COVID19.⁶⁻⁹ Paxlovid, the FDA approved oral SARS-CoV-2 main protease M^{pro} inhibitor, showed benefit for the treatment of COVID-19 patients in severe conditions.^{9b} Although the structure of the active site of M^{pro} is highly conserved among coronaviruses^{9c}, the emergence of drug-resistant variants cannot be excluded. Furthermore, the massive vaccination campaign currently on-going does not guarantee the herd immunity and that vaccinated people with SARS-CoV-2 antibodies will be immune to reinfection.¹⁰⁻¹³ Therefore, designing small molecule libraries from new unexplored chemical space could open an avenue towards new antiviral drugs. To get better hit rates in drug discovery programs, a beneficial factor is natural product likeness.¹⁷ In addition, enhancing the 3-D character and increasing the fraction of sp³ carbons in the designed chemical libraries increase the opportunities to discover hits and consequently potential drug candidates.^{18,19} Herein, we designed a naturally-inspired high 3-D fused-bridged dodecahydro-2a,6-epoxyazepino[3,4,5-c,d]indole-based chemical library, with up to three points of diversity and six controlled chiral centres, two of them are all-carbon quaternary, accessible in one step. The synthesised 25 compounds were assayed in lung cell-lines infected with SARS-CoV-2 and screened for cytotoxicity leading to the confirmation of two hits promising for further optimisation. Furthermore, to identify the possible target, high-throughput virtual screening was carried out on four main protein targets in SARS-CoV-2: (i) the main protease (M^{pro}) that is a non-structural cysteine protease and plays a key role in the release of 16 non-structural proteins involved in the virus replication;²⁰ (ii) the nucleocapsid phosphoprotein that packages the viral RNA into a

^a Institut Pasteur, Université Paris Cité, CNRS UMR n°3523 Chem4Life, Epigenetic Chemical Biology, Department of Structural Biology and Chemistry, F-75015 Paris, France.

^b Institut Pasteur, Université Paris Cité, CNRS UMR n°3523 Chem4Life, Chemogenomic and Biological Screening platform, Center for Technological Resources and Research (C2RT), F-75015 Paris, France.

^c Drug Design and Discovery Lab, Zewail City of Science and Technology, 12578 Cairo, Egypt.

^d Biomedical Sciences Program, University of Science and Technology, Zewail City of Science and Technology, 12578 Cairo, Egypt.

^e Institut Pasteur, Université Paris Cité, CNRS UMR3569, Unité Biologie des ARN et Virus Influenza, F-75015 Paris, France.

^f Department of Pharmacology and Toxicology, University of Toronto, Canada.

^g CNRS - UMR7257 - AFMB - Aix-Marseille Université, Marseille, France.

^h QBI COVID-19 Research Group (QCRG), San Francisco, CA, USA.

* to whom correspondence should be addressed:

Haitham Hassan email: ibnel.haitham.hassan@gmail.com

Reem K. Arafa: rkhidr@zewailcity.edu.eg

Paola B. Arimondo: paola.arimondo@cnrs.fr

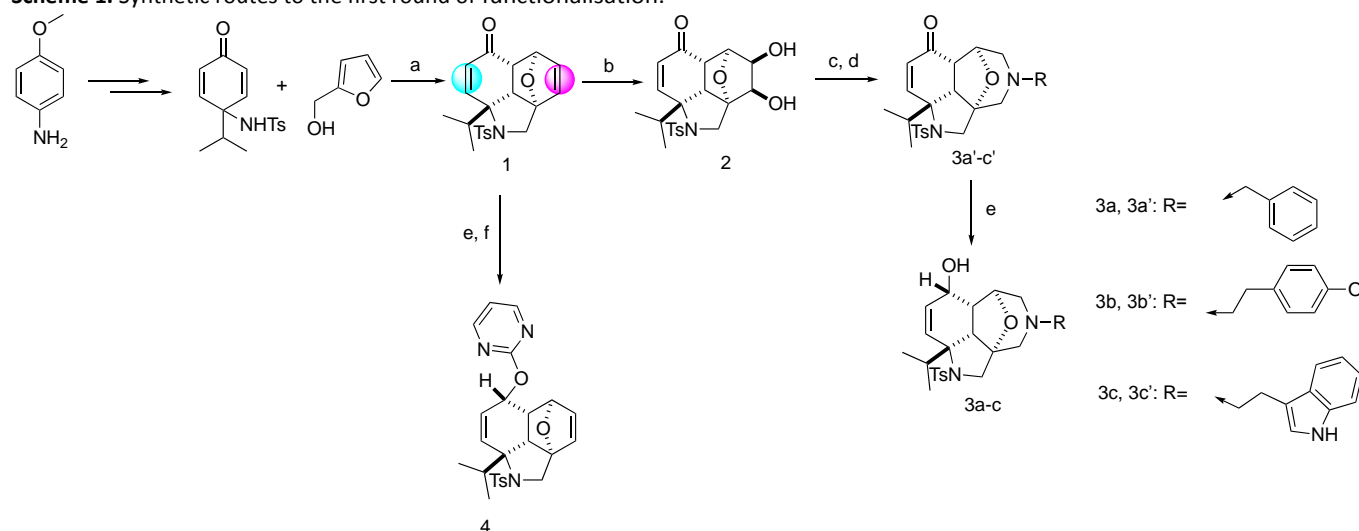
helical ribonucleocapsid (RNP);²¹ (iii) the non-structural protein nsp10/nsp16 complex;^{22, 30} and (iv) the RBD/ACE2 membrane

glycoprotein complex, responsible for the entry in the host cells.²³ The computational data led to the possible binding of the best two hits **3b** and **9e** to either M^{pro} or the nsp10/nsp16 complex. A cellular and a biochemical assay for M^{pro} protease activity confirmed the anti-M^{pro} activity of **3b**, while no activity was validated for **9e**. These findings open a promising starting point for the identification of new antiviral lead candidates.

Results and Discussion

Chemistry. Tuning the right combination between 3-D dimensionality, natural product likeness, C-sp³ fraction, ClogP and number of heteroaromatic rings bearing H-donor/acceptor atoms in one molecule is a challenging task. Finding the right blend leads to a better hit rate. Inspired by the synthesis by Shi and his co-workers, compound **1** can be accessed in three steps in large scale from the commercially available *p*-anisidine (Scheme 1).^{24a} This hydro-epoxybenzo[*c,d*]indole skeleton is a natural product-like and has a high 3-D character and high C-sp³ fraction along with optimal ClogP value that is known to favour bioactivity. Modulation of **1** by introducing relevant groups bearing different H-donor/acceptor atoms can increase its biological activity potential. Compound **1** has two double bonds (highlighted in colours in Scheme 1) and a ketone function allowing further chemical modification. Because of the diverse stereo electronic environment, the two double bonds could be differentiated. The non-conjugated double bond (in magenta) was selectively di-hydroxylated by potassium osmate(VI) to give diol **2** (Scheme 1). This intermediate could be transformed into decahydro epoxy-azepino indoles (**3a'-c'**), which incorporate a constrained morpholine ring present in several bioactive molecules.^{24b} Ring expansion was achieved by oxidation of **2** with sodium periodate to give the intermediate bisaldehyde, that was reacted without isolation with different primary amines under reductive amination conditions to give **3a'-c'**. Reduction of the ketone with sodium borohydride afforded alcohols **3a-c** as single diastereoisomers. These transformations allowed to introduce a new accessible point of diversity at the nitrogen atom.

Scheme 1. Synthetic routes to the first round of functionalisation.^a



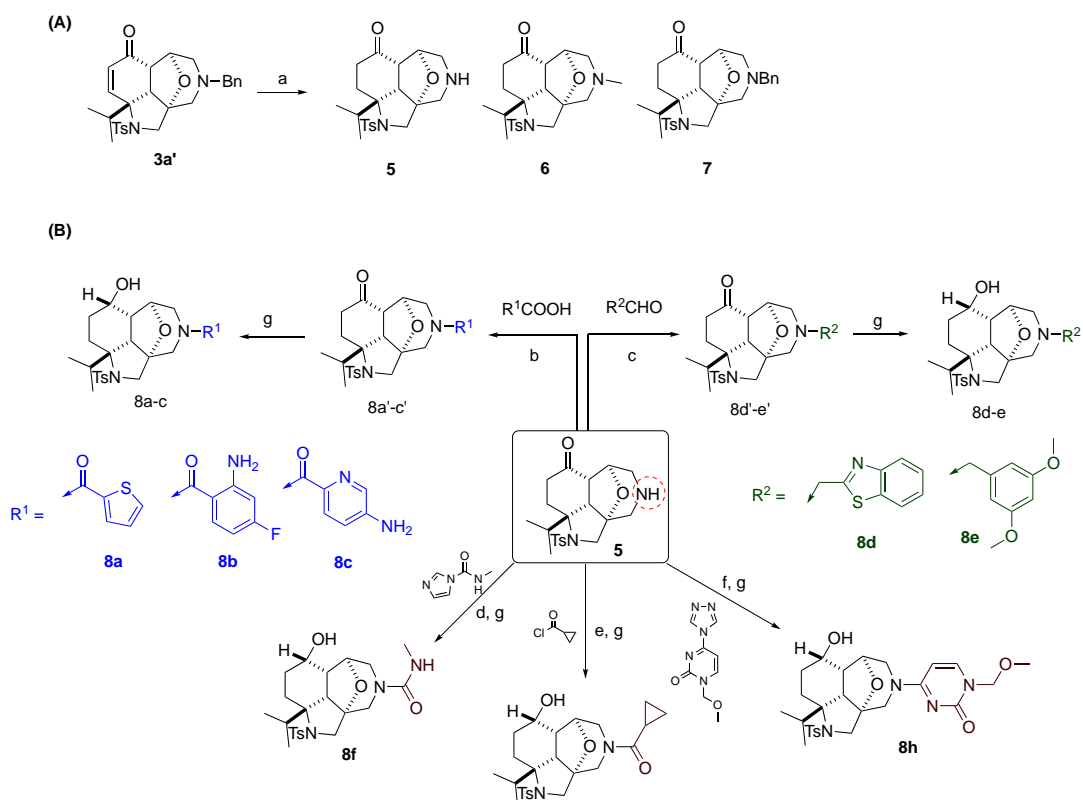
^a **Reagents and conditions:** (a) DIAD (1.1 equiv.), PPh₃ (1.2 equiv.), 4 Å M.S. in THF, 77%. (b) NMO (1.2 equiv.), K₂OsO₄·2H₂O (0.8 mol%) in THF/H₂O 27% over two steps. (c) NaIO₄ in H₂O (1.4 equiv.), silica in DCM. (d) RNH₂ (1.2 equiv.), STAB (3.0 equiv.), AcOH (0.1 equiv.), M.S. 4 Å in THF. (e) NaBH₄ (1.1 equiv.) in MeOH. (f) NaH (60% dispersion in mineral oil) in dry THF, 0°C, 1 h then 2-chloropyrimidine, r.t., overnight, 67%.

The enone moiety in **1** or **2** was resistant to reductive amination conditions. Indeed, in the preparation of **3a'-c'** no reaction at the ketone function was observed and several attempts to transform **1** into the corresponding secondary amines using primary amines under diverse reductive amination conditions failed to produce the desired secondary amines. Thus, further modulation was conducted by reducing the ketone of compound **1** to the corresponding alcohol and then reacted with electrophiles such as 2-chloropyrimidine to afford **4** (Scheme 1).

To introduce other relevant groups bearing different H-donor/acceptor atoms, scaffold **3a'** was hydrogenated in MeOH to give the deprotected secondary amine **5** with concomitant reduction of the double bond (Scheme 2). A by-product **6** was also isolated as the result of the reductive amination of the newly formed secondary amine **5** with adventitious formaldehyde, formed *in-situ* from the reaction of MeOH with Pd/C and hydrogen.²⁵ Changing the solvent to ethyl acetate slowed down the reaction and gave **7** resulting from the reduction of the double bond only without removal of the benzyl group (Scheme 2, panel A). Secondary amine **5** allowed to prepare a diverse set of compounds, showing the reactivity of this position. Reaction with aldehydes under reductive conditions gave tertiary amines and reaction with acyl chloride under basic conditions, or with carboxylic acids under peptide coupling conditions gave the corresponding tertiary amides. As reactive functional groups are not desirable, all the intermediate ketones were directly subjected to reduction with NaBH₄ to give **8a-g**. In parallel, the secondary

amine **5** was reacted with triazolo-pyrimidine, and then the ketone was reduced by NaBH₄ to give compound **8h** (Scheme 2, panel B).

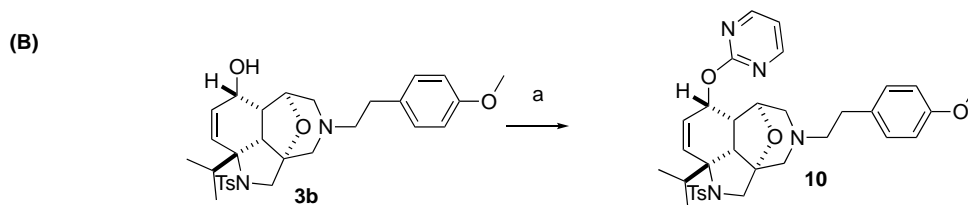
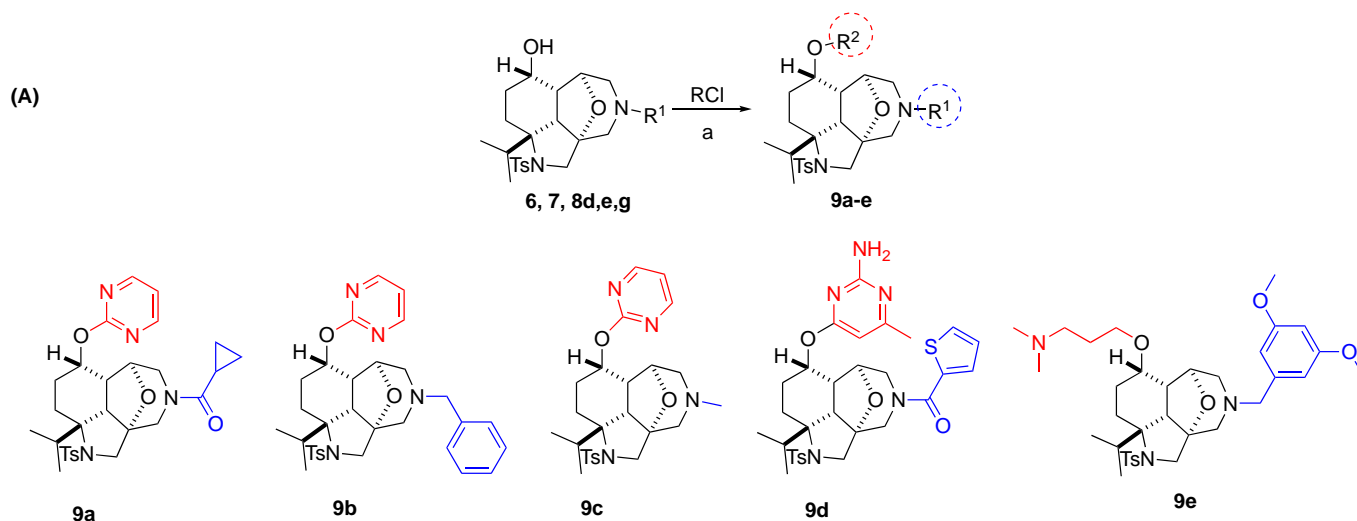
Scheme 2. Synthetic routes to the second round of functionalisation.^a



^a **Reagents and conditions:** (a) H₂ (1.0 bar), Pd/C (10 mol%), MeOH (**5** and **6**) or EtOAc (**7**). (b) RCOOH (1.1 equiv.), DIPEA (1.5 equiv.), TBTU (2.5 equiv.) in DCM. (c) RCHO (1.2 equiv.), STAB (3.0 equiv.), AcOH (0.1 equiv.), M.S. 4 Å in DCM. (d) *N*-methyl-1*H*-imidazole-1-carboxamide (1.1 equiv.), Et₃N (1.1 equiv.) in DCM. (e) RCOCl (1.1 equiv.), Et₃N (1.1 equiv.) in DCM. (f) 1-(methoxymethyl)-4-(1*H*-1,2,4-triazol-1-yl)pyrimidin-2(1*H*)-one (1.2 equiv.), Et₃N (1.2 equiv.) in MeCN, 70 °C. (g) NaBH₄ (1.1 equiv.) in MeOH.

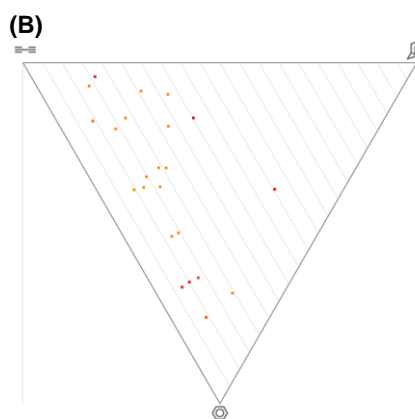
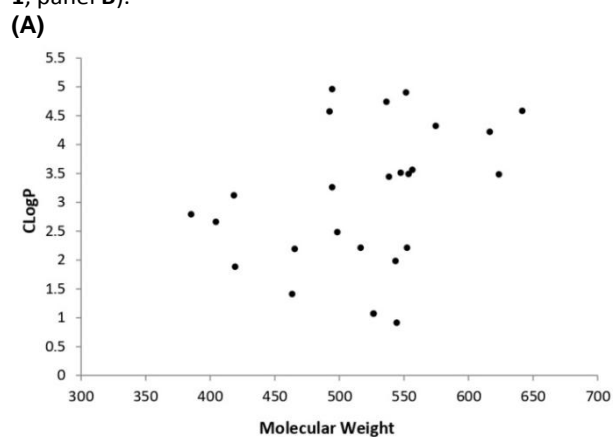
The third round of functionalisation was conducted on **6**, **7**, **8a**, **e**, **g** (Scheme 3). Reaction of the free alcohol with halogenated heterocycles or alkanes gave compounds **9a-e** (Scheme 3, panel A). Similarly, compound **3b** reacted with 2-chloropyrimidine to give compound **10**. Surprisingly, the di-hydroxylation of the left-hand site (LHS) double bond in compound **3b** by potassium osmate(VI) was not successful. The isolated product was ketone **3b'**, resulting from the oxidation of the allylic alcohol catalysed by osmium salts²⁶ (Supplementary information at page S18).

Scheme 3. Synthetic routes to the third round of functionalisation.^a



^aReagents and conditions: (a) NaH (60% dispersion in mineral oil) in dry THF, 0°C, 1h then RCl, r.t., overnight. †

Molecular properties. The 3-D dimensionality and molecular properties of the 25 molecules were analysed using the web-free tool LLAMA (Lead-Likeness And Molecular Analysis; llama.leeds.ac.uk).²⁷ The chemical library showed average C-sp³ fraction of 0.5-0.7, average ClogP of 1.0-5.0 and average molecular weight of 400-600, which are all features consistent with lead-likeness factors (Figure 1, panel A). It also provided a high 3-D character (Figure 1, panel B), a property which was proven to give better hit rates.¹⁹ The natural product likeness was assessed using Natural Product Likeness Score calculator (NaPLoS) web-free tool and it indicated that all the synthesised compounds fall in the natural product likeness space with an average score of 0.2-1.3 (Figure 1, panel B).²⁸



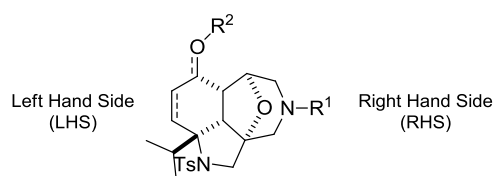
(C)




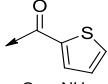
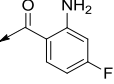
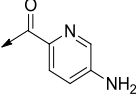
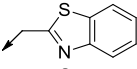
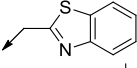
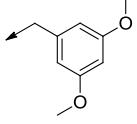
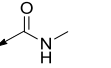
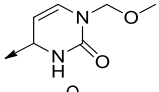





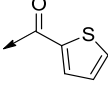

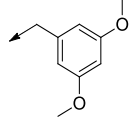
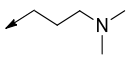
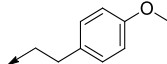

Figure 1. Molecular Properties and diversity. Panel A: Molecular weight vs CLogP (CLogP was calculated by ChemDraw Version 15). Panel B: Principal Moment of Inertia PMI Plot. Panel C: Natural Product Likeness score, the green line indicates All Synthetic Products, the orange line the Natural Products and the red dots the here synthesised chemical library

Biological evaluation. The 25 compounds were tested for inhibition of SARS-CoV-2 replication of the Beta variant in the lung cancer A549 cell line stably transduced with a lentiviral construct bearing the human ACE2 receptor.⁴⁰ The measurement of viral replication was carried out by quantitative RT-PCR as described in the Supplemental Methods section (Table 1). The key scaffolds **1** and its analogue **4** were not active, indicating the importance of the right-hand side (RHS) diversity point. Introducing dihydroxyl groups on compound **1** showed to be beneficial with an antiviral activity for compound **2** of $EC_{50} = 14 \mu\text{M}$. Structure Activity Relationship (SAR) revealed that the presence of large aromatic groups in R^1 position on the RHS is important for the antiviral activity as compounds **5**, **6** and **8f** were inactive. It was also found that the presence of the ketone did not significantly change the viral activity when comparing compounds **3a'** and **8d'** to **3a** and **8d**, respectively. Substitution in R^2 with aromatic groups did not increase the activity; the aliphatic chain gave better activity. In parallel, the cytotoxicity activity of the compounds was assessed as the cytotoxicity could interfere with the antiviral activity of the compound (Supplemental Methods section). Several hits showed a favourable ratio between the antiviral activity and cytotoxicity. The most active compounds with favourable antiviral activity to cytotoxicity ratio and do not have electrophilic groups, which could favour covalent binding are **3b** and **9e** (Figure 2). In addition, **3b** and **9e** remained active against cells infected with the SARS-CoV2 Omicron BA1 variant with an EC_{50} of 3.5 and 4.4 μM respectively, with **3b** showing no cytotoxicity, while **9e** giving cytotoxicity above 20 μM (Table 1 and Supporting information Figure S1).

Table 1. Anti-SARS-CoV-2 Activity on A549-ACE2 cells (RT-qPCR) and cell cytotoxicity of the 22 compounds, expressed as EC_{50} in μM , the concentration at which is observed 50% of inhibition of the maximal signal and its 95% confidential interval.



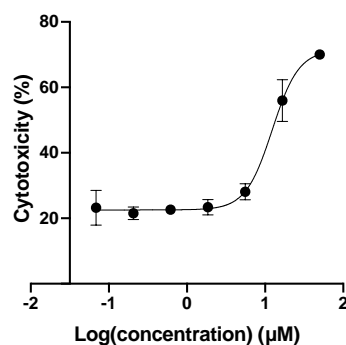
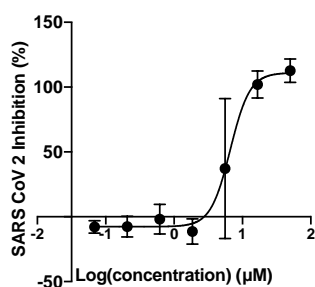
Entry	R^1	R^2	Saturation on the LHS	RT-qPCR Beta EC_{50} (μM) [95% CI]	Cytotoxicity CC_{50} (μM) [95% CI]	Ratio cytotoxicity/RT-qPCR
3a'		-	Double bond	16 (N/A)	14 [2.8 to 73]	0.9
3a		H	Double bond	17 [13 to 23]	> 50	>3
3b'		-	Double bond	13 [7.2 to 24]	6.6 [2.7 to 16]	0.5
3b		H	Double bond	3.7 [2.0 to 6.9] 3.5 [2.3to 5.6]*	12 [10 to 15] > 50*	3.3
3c		H	Double bond	8.3 [6.3 to 11]	13 (N/A)	1.6

5	H	-	Saturated	> 50	> 50	N/A
6	Me	-	Saturated	> 50	> 50	N/A
7		-	Saturated	8.4 [5.5 to 13]	5.8 [3.0 to 11]	0.7
8a		H	Saturated	16 [5.2 to 51]	> 50	>3
8b		H	Saturated	14 [8.5 to 22]	20 [8.2 to 48]	1.4
8c		H	Saturated	37 [23 to 59]	> 50	1.4
8d'		-	Saturated	5.3 [4.0 to 7.1]	37 N/A	6.9
8d		H	Saturated	6.6 [4.3 to 10]	12 [6.7 to 21]	1.6
8e		H	Saturated	5.6 N/A	13 [7.6 to 21]	2.3
8f		H	Saturated	> 50	> 50	N/A
8h		H	Saturated	> 17	> 50	N/A
9a			Saturated	>40 [N/A]	> 50 [N/A]	N/A
9b			Saturated	6.2 [N/A]	6.3 [3.7 to 11]	1.0
9c	Me		Saturated	20 [N/A]	17 [14 to 18]	0.8
9d			Saturated	> 50	> 50	N/A
9e			Saturated	1.4 [0.9 to 2.2] 4.4 [2.0 to 9.0]*	7.7 [5.2 to 12] > 20*	5.7
10			Double bond	> 50	-	N/A

N/A= not applicable

* Assays against cells infected with SARS-CoV-2 variant Omicron BA1.

A) Compound 3b



B) Compound 9e

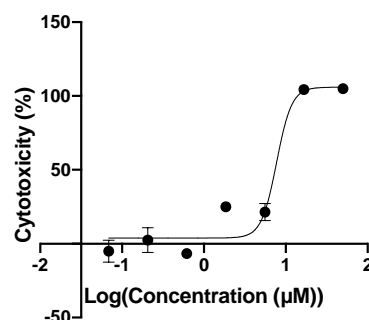
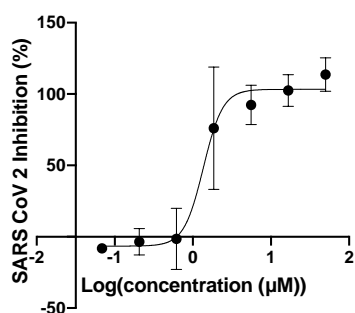


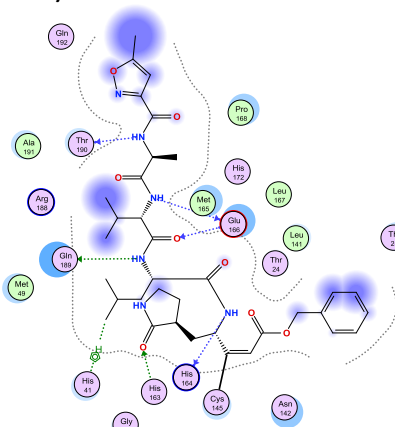
Figure 2. Anti-SARS-CoV-2 dose response curve (left) and cytotoxicity (right) in log of the concentration in μM of the compound. Panel A: Compound **3b**. Panel B: Compound **9e**. Error bars of triplicates are shown.

In Silico studies. To identify potential target(s) behind the observed activity against SARS-CoV-2 of compounds **3b** and **9e**, a docking and virtual screening study followed by a molecular dynamics simulation (MD) was conducted against four SARS-CoV-2 proteins. The first selected potential target was the SARS-CoV-2 main protease (M^{pro}), a non-structural cysteine protease that is one of the key players in the release of 16 non-structural proteins involved in the virus replication²⁰. Second was the nucleocapsid phosphoprotein that packages the viral RNA into a helical ribonucleocapsid²¹. The third selected target was the non-structural nsp10/nsp16 complex that has a 2'-*O*-methyltransferase activity (2'-*O*-MTase) activity involved in the methylation of RNA cap structure^{22,30}. Finally, the receptor-binding domain (RBD) of the viral spike protein responsible for the viral entry in the host cell was also virtually screened as the fourth plausible target²³. 3D structures of the four target proteins were obtained from the protein data bank (PDB) as follows: the SARS-CoV-2 main protease M^{pro} (PDB ID: 6LU7)³¹, the 2'-*O*-methyltransferase nsp10/nsp16 complex (PDB ID: 6W4H)³⁰, the nucleocapsid phosphoprotein RNA binding domain (PDB IDs: 6VYO)²⁹, and the RBD domain/ACE2/B(0)AT1 complex (PDB ID: 6M17)³³. These 3D structures were selected based on their high resolution and lack of mutations.

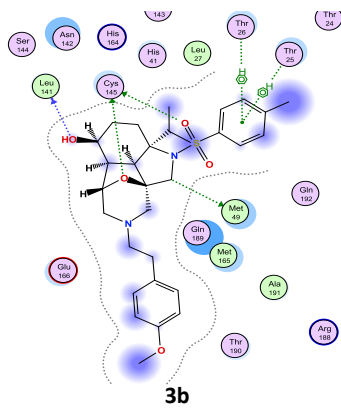
Docking validation and high-throughput virtual screening. First, we validated the use of the method MOE 2019.01 for the virtual screening study by self-docking the co-crystallized ligands in the adopted crystal structures for the nsp10/nsp16 complex (PDB ID: 6W4H)²⁹ and the SARS-CoV-2 main protease M^{pro} (PDB ID: 6LU7)³⁰ with the respective co-crystallized ligands *S*-adenosyl-L-methionine (SAM) and the peptide-like inhibitor PRD_00221 (Supplementary Figure S2). The proteins were initially subjected to structure preparation and protonation state fixation, then London dG and GBVI/WSA dG scoring functions were used to assess the binding poses and binding interactions. The root mean square deviation (RMSD) of the best scoring poses was calculated using the DockRMSD online server (<https://zhanggroup.org/DockRMSD/>)³¹. Since the optimum RMSD for docking validation is conventionally $\leq 2 \text{ \AA}$ and the results obtained by MOE 2019.01 were within this range, the adopted docking protocol was considered reliable.

The binding potential of the 25 synthesised compounds was assessed and their potential binding affinity was reported as *S*-scores (MOE internal scoring function) and compared to those of the co-crystallized ligands, when applicable, for each of the four target proteins (Supplementary Table S1). The 2D interactions of the molecules compared to the co-crystallized ligands showed that compounds **3b** and **9e** were among the best virtual hits, which is in line with the antiviral activity profile of these compounds (table 1). The 2D binding interactions of the two compounds with the four-screened SARS-CoV-2 proteins are shown in Figure 3.

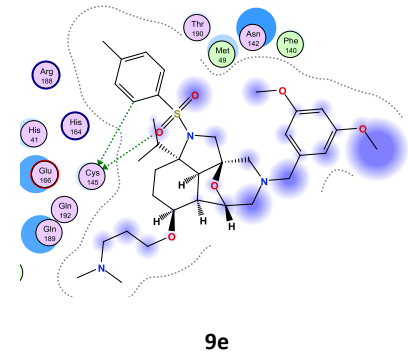
A) M^{Pro}



Peptide like inhibitor PRD_002214

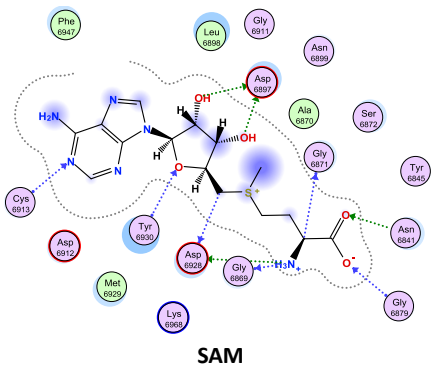


3b

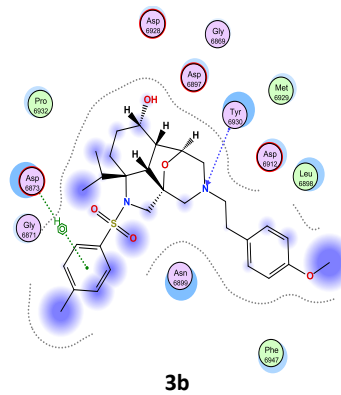


9e

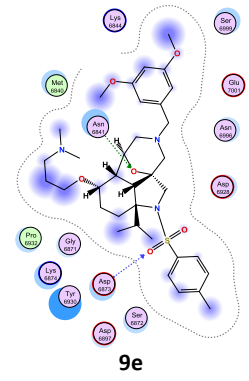
B) Nsp10/nsp16 complex



SAM

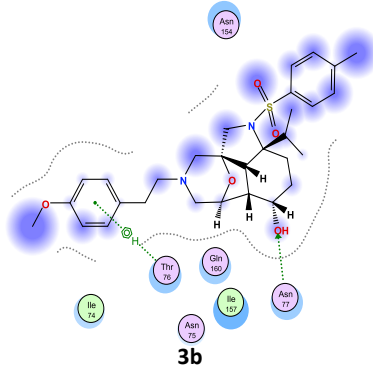


3b

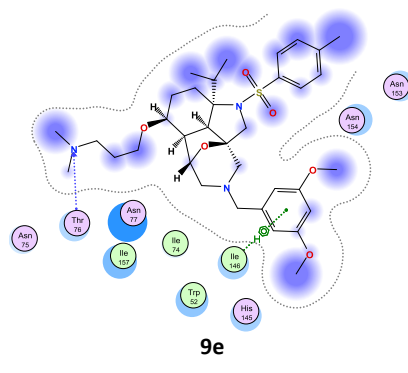


9e

C) Nucleocapsid phosphoprotein

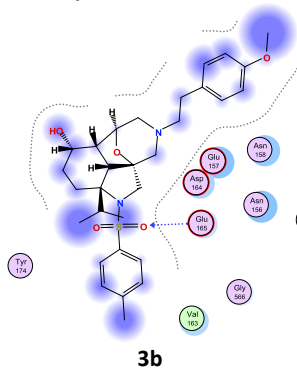


3b

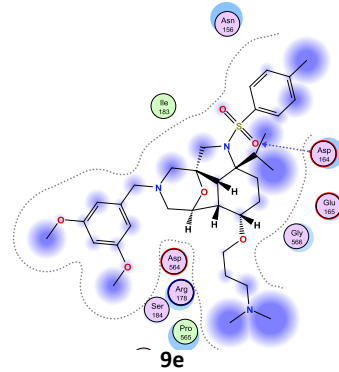


9e

D) RBD/ACE2 complex



3b



9e

Figure 3. 2D interactions of the best 2 hits, compounds **3b** and **9e**. (A) Binding of co-crystallized ligand (PRD_002214), **3b** and **9e** with M^{pro} (PDB ID. 6LU7). (B) Binding of co-crystallized ligand (SAM), **3b** and **9e** with nsp10/nsp16 complex (PDB 6W4H). (C) Binding of **3b** and **9e** with the nucleocapsid phosphoprotein (PDB 6VYO). (D) Binding of **3b** and **9e** with the RBD/ACE2 complex (PDB 6M17).

Docking studies showed that the peptide-like inhibitor PRD_002214 is covalently bound to Cys145 and forms seven other non-covalent interactions with the active pocket of M^{pro} (Figure 3.A and Supplementary Table S2). Compound **3b** forms six interactions: four as a H-bond donor—with distances 3.14 Å, 3.42 Å, 3.61 Å and 3.72 Å, respectively, and with energy scores of -0.6 Kcal/mol, -0.8 Kcal/mol, -0.8 Kcal/mol, and -1.6 Kcal/mol, respectively. It also indicated two pi-H interactions between the ligand's 6-membered ring and Thr25 and Thr26 in the pocket with distances 4.39 Å and 4.22 Å, respectively and with energy scores of -1.1 Kcal/mol, and -2.0 Kcal/mol respectively (Figure 3.A and Supplementary Table S2). This suggests that compound **3b** might interact with this target protein. On contrary, compound **9e** makes only two interactions with M^{pro} (Figure 3.A and Supplementary Table S2). For Nsp10/nsp16 complex, we mapped strong 11 interactions with the co-crystallized ligand SAM (Figure 3.B and Supplementary Table S2), while compounds **3b** and **9e** formed only two interactions, H bonds and pi-H interactions. Concerning the nucleocapsid phosphoprotein, the docking suggests that both compounds **3b** and **9e** form two interactions only (Figure 3.C). Finally, with the spike protein RBD/ACE2 complex (Figure 3.D), only one H-bond was observed between O62 of compound **3b** and Glu 165 and one hydrogen bond between O72 of compound **9e** and Asp164 suggesting a reduced interaction with this target complex (Supplementary Table S1).

Molecular Dynamics Simulation for Compounds 3b and 9e. We then carried out a 100 ns MD simulation of **3b** and **9e** to assess their binding stability with the target M^{pro} and nsp10/nsp16 complex (Supplementary data and Figures S3, S4, S5 and S6). To measure how much the protein and ligand conformations change along the MD simulation trajectory of the target proteins-ligand complex, RMSD values were calculated for the proteins' C α atomic coordinates and the screened ligands using GROMACS 2021.1^{32, 34}. The complex of **3b** with M^{pro} formed a very stable complex reaching stability in the first nanoseconds of the simulation with very low perturbation (less than 1 Å) throughout the whole 100 ns simulation (figure 4.A). Additionally, the stability of the complex was confirmed by the RMSD of **3b** itself, as displayed in figure 4.B, eliciting perturbations of less than 1 Å throughout the whole simulation. On the other hand, the RMSD of nsp10/nsp16 complex C α backbone in its complex with **3b** showed small perturbations, while the RMSD of the **3b** itself gave very high fluctuations after 50 ns.

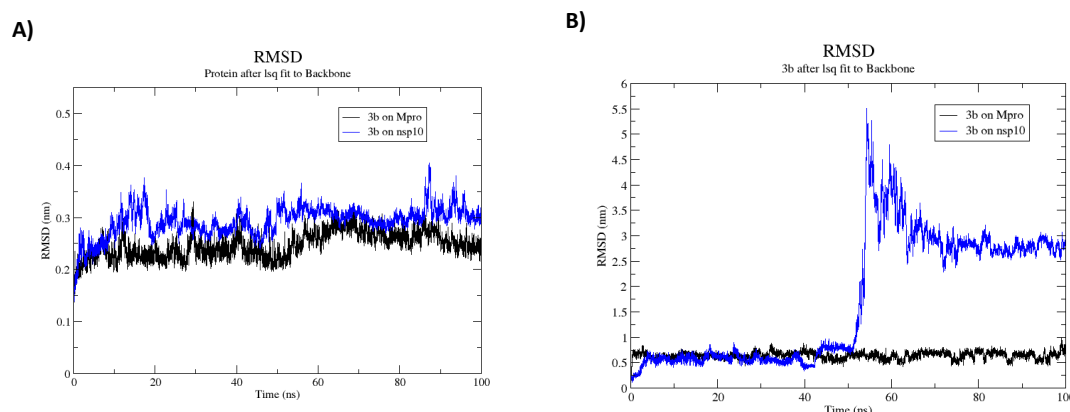


Figure 4. Structural dynamics of compound **3b** bound to M^{pro} (black curve) and nsp10/nsp16 complex (blue curve); RMSD of the C α backbone of the proteins in nm (A) and RMSD of **3b** in nm along the MD trajectory (B).

Then the stability of **9e** in complex with M^{pro} and nsp10/nsp16 complex was assessed over 100 ns simulation (Figure 5). The RMSD of the protein referenced to the backbone was calculated. On one hand, the complex of **9e** bound to M^{pro} achieved stability at almost 5 ns, showing small perturbations (less than 1 Å) in the region from 65 ns to the end of the simulation (Figure 5.A, black curve). This indicates a stability of the protein upon interaction with **9e**. On the other hand, the RMSD of **9e** itself showed significant perturbations at the beginning (from 0 ns to 40 ns), representing the ligand jumping away from the binding pocket, and at the end (from 70 ns to 100 ns) (Figure 5.B, black curve). In parallel, nsp10/nsp16 complex exhibited convergence at 12 ns with small perturbations (~ 1 Å) thus indicating a stabilisation of the protein (Supplementary Figure S5). Besides that, the RMSD of **9e** complexed to nsp10/nsp16 complex exhibited a very low RMSD (less than 6 Å) and RMSD perturbations, which reflect high **9e** complex stability as shown in figure 5.B.

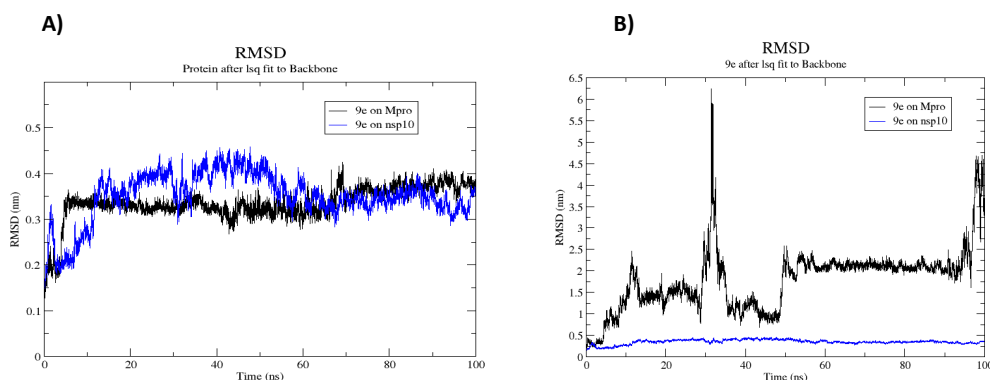


Figure 5. Structural dynamics of compound **9e** bound M^{pro} (black curve) and nsp10/nsp16 complex (blue curve); RMSD of the C_{α} backbone of the proteins in nm (A) and RMSD of **9e** in nm along the MD trajectory (B).

Stable binding of **3b** and **9e** should reflect on the stability of the residues within the binding pocket. The rigidity and flexibility of such residues can be investigated through the root mean square fluctuation (RMSF) of the C_{α} atoms during the simulation (Supplementary Figure S5).³⁵ The RMSF of **3b** and M^{pro} complex showed low fluctuations., especially Cys145 and His41 that constitute the catalytic dyad.³⁵ **9e** exhibited fluctuations (~ 1.3 - 3.2 Å) in some residues, with higher fluctuations at residues (185:192) than **3b**. In parallel, nsp10/nsp16 complex shows residues within 5 Å proximity from **3b** and **9e**, exhibiting lower fluctuations (~ 0.4 - 1.4 Å) (Supplementary Figure S5). The collective results for the *in-silico* studies suggest that the target of **3b** is the main protease M^{pro} , while **9e** preferentially interacts with nsp10/nsp16 complex. To verify the *in-silico* results, compounds **3b** and **9e** were tested in *in-vitro* for inhibition of 2'-O-methyltransferase activity of nsp10-nsp16 complex, and nsp14 guanine-N7-methyltransferase activity, and in cellular assay involving M^{pro} .

Biological targets validation:

Inhibition of nsp10-nsp16 complex. The nsp10-nsp16 complex is responsible for catalysing the final step of the coronaviral mRNA capping. Nsp16 is member of the (2'-O-MTase) family, which catalyse the transfer of methyl group to the RNA substrates from the methyl donor SAM. Nsp16 requires nsp10 for methyltransferase activity and stability. The *in-vitro* activity of SARS-CoV-2 nsp10-nsp16 complex was assessed by monitoring the transfer of 3H -SAM to the biotinylated N7-meGpppACCCCC RNA substrate (Supplementary information). The subsequently methylated RNA was captured using scintillation proximity assay (SPA) beads followed by quantifying the level of incorporated 3H -methyl by measuring the radioactivity level (counts per minute [CPM]).⁴¹ In dose response experiments (Figure S7), commercially available *S*-adenosyl-L-homocystein (SAH) showed IC_{50} values of 3.9 μM and 0.22 μM for nsp10-nsp16 complex and nsp14 methyltransferase activities, respectively (Supplementary Figure S7 and Table 2). However, **3b** and **9e** did not show any inhibitory effect on nsp10-nsp16 2'O MTase activity, which did not support the computational analysis prediction for this target. Although **3b** showed no inhibitory effect on the methyltransferase activity of nsp14, **9e** showed a weak activity against nsp14 at high concentrations (Table 2 and Supplementary Figure S7).

Table 2. Effect of **3b** and **9e** on the methyltransferase activities of nsp10-nsp16 complex and nsp14. *S*-adenosyl-L-homocystein (SAH) was used as control.

Compound Name	$IC_{50}^{nsp10-nsp16}$ (μM)	Hill Slope	IC_{50}^{nsp14} (μM)	Hill Slope
SAH	3.9	1.1	0.22	0.8
3b	NI	NA	NI	NA
9e	NI	NA	>100	NA

The corresponding plots are shown in Figure S7. NI: No Inhibition, NA: Not applicable.

Inhibition of SARS-Cov-2 main protease (M^{pro}). M^{pro} is a validated target for treating COVID19 with the recently approval of the drug Paxlovid.^{9b} Encouraged by the importance of this target and guided by the results obtained from the computational analysis through docking and molecular dynamics simulations, we tested the two hits for their ability to inhibit M^{pro} in a highly sensitive cell-based luciferase assay that we developed to monitor SARS-CoV-2 main protease activity.³⁶ This gain-of-function assay is based on a Reverse-Nanoluciferase (Rev-Nluc) reporter in which two Nanoluciferase domains are permuted and linked together by a cleavage site recognized by M^{pro} . Co-expression with the wild-type SARS-CoV-2 M^{pro} results in cleavage of the reporter and thereby a significant reduction in luciferase activity. The addition of an inhibitor of M^{pro} results in a dose-dependent restoration in luciferase activity. The assay is run in parallel on a wild-type and a catalytically inactive M^{pro} , to assess the specificity of the drug (details are provided in the Supplemental Methods section). The commercially available GC376 compound was used as a positive control. The dose response curve showed that GC376 reduced M^{pro}

activity, as seen by a gradual increase of Nanoluciferase activity measured in the presence of M^{pro} WT (Figure 6.A), while the luciferase signal measured in the presence of catalytically inactive mutant remained unchanged. Full inhibition of M^{pro} was achieved at 10 μ M GC376 and the estimated IC₅₀ was in the micromolar range.³⁷⁻³⁹ Only compound **3b** showed a specific activity against M^{pro} in Rev-Nluc assay from 10 μ M, as indicated by an increase in luciferase signal, which was not observed with the catalytically inactive M^{pro} mutant (Figure 6.B). An *in vitro* enzymatic assay on the purified protein confirmed a weak inhibitory activity of compound **3b** (Supplementary Figure S8), while compound **9e** was inactive. This observed activity of **3b** against SARS-CoV-2 M^{pro} was consistent with the computational study.

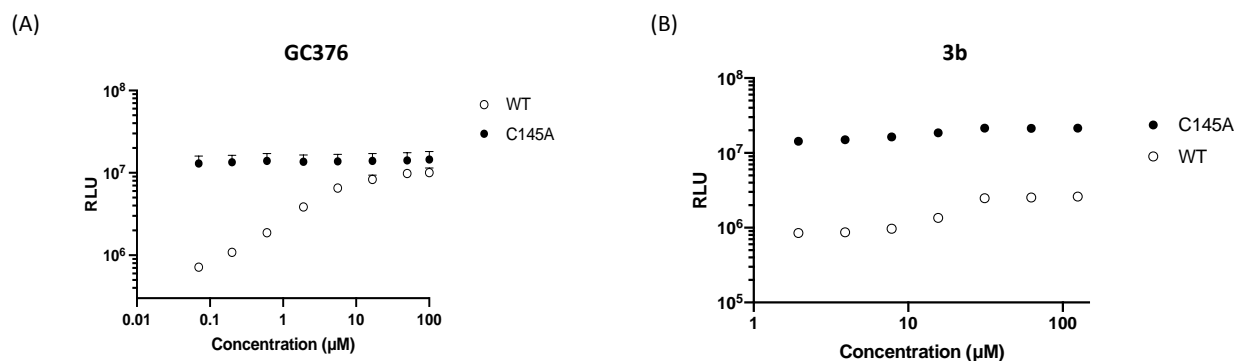


Figure 6. Cell-based assay luciferase assay to monitor SARS-CoV-2 main protease activity of the wild-type (WT, open symbols) and catalytically inactive C145A mutant (black symbols) of M^{pro}: (A) Dose response curve of the positive control GC376. (B) Dose response curve of compound **3b**. The inhibition was assessed using the Rev-Nluc-based assay. The assay was performed as described in the Supplemental Methods section. The inhibitor concentrations correspond to 3-fold serial dilutions from 125 to 0.02 μ M. The RLU are shown as the mean \pm SD of three independent experiments performed in technical triplicates.

Conclusions

In conclusion, to explore a new chemical space in the design of inhibitors of SARS-CoV-2, a chemical library of 25 fused-bridged dodecahydro-2a,6-epoxyazepino[3,4,5-cd]indole compounds guided by natural product likeness, 3-D dimensionality and lead likeness was synthesised. The screening in a cellular system of viral infection led to the identification of 2 hits against the SARS-CoV-2 virus, confirming that following lead likeness guidelines and enhancing 3-D dimensionality increases the hit rate in the design of chemical libraries. Although the compounds also showed cytotoxicity, the two hits were confirmed to have an antiviral activity against SARS-CoV-2 with EC₅₀ between 3.7-1.4 μ M with acceptable cytotoxicity difference. They also showed anti-viral activity against SARS-CoV-2 variant Omicron BA1, which consolidate their inhibition profile. Computational analysis through docking and molecular dynamics simulations was carried out against four main target proteins of SARS-CoV-2: Main Protease M^{pro}, nucleocapsid phosphoprotein, non-structural protein nsp10/nsp16 complex and RBD/ACE2 membrane glycoprotein complex. The analysis identified as possible binding targets either the main protease M^{pro} or nsp10/nsp16 complex. A cellular assay using a reverse-Nanoluciferase (Rev-Nluc) reporter system confirmed that compound **3b** shows some activity against M^{pro} in cells and a weak activity in an *in vitro* enzymatic assay with purified M^{pro}. In contrast, *in-vitro* inhibition assay of nsp10/nsp16 complex and nsp14 did not confirm the computational analysis suggesting the methyltransferases as a potential target. The confirmed hits **3b** and **9e** pave the way to the chemical optimisation of the fused-bridged dodecahydro-2a,6-epoxyazepino[3,4,5-cd] indole skeleton for the inhibition of SARS-CoV-2. This also confirms the interest in inspired natural product scaffolds with improved drug likeness properties for the identification of novel hits with antiviral activity. Finally, the chemical pathway was designed to favour modulation of the substituents and further chemical optimisation is underway to increase antiviral activity and, in particular, to improve the activity against the potential target, the main protease M^{pro}.

Author Contributions

HH designed, synthesised the library and wrote the chemistry, molecular properties and biological activity parts, and revised the manuscript. JC, FA set-up and carried out the biological evaluation in cells and participated to the writing and revision of the manuscript. TK, KC and NN set-up and carried out the cell-based assay for M^{pro} protease activity and NN proofread the manuscript. RKA conceptualized the *in-silico* screening. AH, LB carried out the *in-silico* study and primary writing of the relevant sections, both of them participated equally in the *in-silico* studies part. RKA, AH, and LB revised the results, data analysis and wrote the *in-silico* studies. ED and AD developed and performed M^{pro} enzymatic assay), AL, AD, MV carried out the assay against the methyltransferase activity of nsp10/nsp16 and nsp14. SM synthesised key compounds and revised chemistry and biological activity parts. PBA supervised the project, obtained the grants, wrote and revised the whole manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors thank Dr Benjamin BARDIAUX, Structural Bioinformatics Laboratory, Institut Pasteur, for constructive proofreading of the *in-silico* studies part. The authors thank the Institut Pasteur and Fondation de France grants for the installation and the Drug pipeline against SARS-CoV-2 project. Sylvain Paisant and Marine Ghazarian are acknowledged for providing technical help for the M^{PRO} cell-based assay. TK, KYC and NN were funded by the «URGENCE COVID-19» fundraising campaign of Institut Pasteur and by the Agence Nationale de la Recherche (grants ANR-18-CE18-0026, ANR-18-CE18-0028 and ANR-10-LABX-62-IBEID).

Fundings

Fondation de France grant Programme Fédérateur de Recherche sur SARS-CoV2 & COVID-19 PRF7 SARS-CoV2 drug pipeline project (to PBA). Vedadi's lab is funded by US NIH 1U19AI171110-01.

Notes and references

1. F. Wu, S. Zhao, B. Yu, Y.-M. Chen, W. Wang, Z.-G. Song, Y. Hu, Z.-W. Tao, J.-H. Tian, Y.-Y. Pei, M.-L. Yuan, Y.-L. Zhang, F.-H. Dai, Y. Liu, Q.-M. Wang, J.-J. Zheng, L. Xu, E. C. Holmes and Y.-Z. Zhang, *Nature* 2020, **579**, 265.
2. P. Zhou, X. L. Yang, X. G. Wang, B. Hu, L. Zhang, W. Zhang, H. R. Si, Y. Zhu, B. Li, C. L. Huang, H. D. Chen, J. Chen, Y. Luo, H. Guo, R. Di Jiang, M. Q. Liu, Y. Chen, X. R. Shen, X. Wang, X. S. Zheng, K. Zhao, Q. J. Chen, F. Deng, L. L. Liu, B. Yan, F. X. Zhan, Y. Y. Wang, G. F. Xiao and Z. L. Shi, *Nature*, 2020, **579**, 270.
3. Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. M. Leung, E. H. Y. Lau, J. Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang, S. Zhou, R. Wang, H. Liu, Y. Luo, Y. Liu, G. Shao, H. Li, Z. Tao, Y. Yang, Z. Deng, B. Liu, Z. Ma, Y. Zhang, G. Shi, T. T. Y. Lam, J. T. Wu, G. F. Gao, B. J. Cowling, B. Yang, G. M. Leung and Z. N. Feng, *Engl. J. Med.* 2020, **382**, 1199.
4. H. Mitsuya and N. Kokudo *Glob. Heal. Med.* 2020, **2**, 53.
5. (a)D. Baltimore *Bacteriol. Rev.* 1971, **35**, 235; (b)J. J. Berman Group IV Viruses: Single-Stranded (+)Sense RNA. Taxonomic Guide to Infectious Diseases 2012, 237–246.
6. C. Scavone, S. Brusco, M. Bertini, L. Sportiello, C. Rafaniello, A. Zoccoli, L. Berrino, G. Racagni, F. Rossi and A. Capuano, *Br. J. Pharmacol.*, 2020, **177**, 4813.
7. M. Wang, R. Cao, L. Zhang, X. Yang, J. Liu, M. Xu, Z. Shi, Z. Hu, W. Zhong and G. Xiao, *Cell Res.* 2020, **30**, 269.
8. E. K. McCreary and D. C. Angus, *JAMA* 2020, **324**, 1041.
9. (a) C. D. Spinner, R. L. Gottlieb, G. J. Criner, J. R. Arribas López, A. M. Cattelan, A. Soriano Viladomiu, O. Ogbuagu, P. Malhotra, K. M. Mullane, A. Castagna, L. Y. A. Chai, M. Roestenberg, O. T. Y. Tsang, E. Bernasconi, P. le Turnier, S.-C. Chang, D. SenGupta, R. H. Hyland, A. O. Osinusi, H. Cao, C. Blair, H. Wang, A. Gaggard, D. M. Brainard, M. J. McPhail, S. Bhagani, M. Y. Ahn, A. J. Sanyal, G. Huhn and F. M. Marty, *JAMA* 2020, **324**, 1048. (b) D. R. Owen, C. M. N. Allerton, A. S. Anderson, L. Aschenbrenner, M. Avery, S. Berritt, B. Boras, R. D. Cardin, A. Carlo, K. J. Coffman, A. Dantonio, L. Di, H. Eng, R. Ferre, K. S. Gajiwala, S. A. Gibson, S. E. Greasley, B. L. Hurst, E. P. Kadar, A. S. Kalgutkar, J. C. Lee, J. Lee, W. Liu, S. W. Mason, S. Noell, J. J. Novak, R. S. Obach, K. Ogilvie, N. C. Patel, M. Pettersson, D. K. Rai, M. R. Reese, M. F. Sammons, J. G. Sathish, R. S. P. Singh, C. M. Steppan, A. E. Stewart, J. B. Tuttle, L. Updyke, P. R. Verhoest, L. Wei, Q. Yang and Y. Zhu, *Science*, 2021, **374**, 1586. (c) X. Gao, B. Qin, P. Chen, K. Zhu, P. Hou, J. A. Wojdyla, M. Wang and S. Cui, *Acta Pharmaceutica Sinica B*, 2021, **11**, 237..
10. S. P. Kaur and V. Gupta, *Virus Res.* 2020, **288**, No. 198114.
11. On behalf of the COVID-19 Commission of Accademia Nazionale dei Lincei, Rome, G. Forni and A. Mantovani, *Cell Death Differ.*, 2021, **28**, 626.
12. B. Spellberg, T. B. Nielsen and A. Casadevall, *JAMA Intern. Med.*, 2021, **181**, 460.
13. G. A. Poland, I. G. Ovsyannikova and R. B. Kennedy, *Lancet*, 2020, **396**, 1595.
14. T. Pillaiyar, M. Manickam, V. Namasivayam, Y. Hayashi and S. H. Jung *J. Med. Chem.* 2016, **59**, 6595.
15. C. Drosten, S. Günther, W. Preiser, S. van der Werf, H.-R. Brodt, S. Becker, H. Rabenau, M. Panning, L. Kolesnikova, R. A. M. Fouchier, A. Berger, A.-M. Burguiere, J. Cinatl, M. Eickmann, N. Escriou, K. Grywna, S. Kramme, J.-C. Manuguerra, S. Müller, V. Rickerts, M. Stürmer, S. Vieth, H.-D. Klenk, A. D. M. E. Osterhaus, H. Schmitz and H. W. Doerr *N. Engl. J. Med.* 2003, **348**, 1967.
16. S.K. P. Lau, P. Lee, A.K. L. Tsang, C.C. Y. Yip, H. Tse, R. A. Lee, L.-Y. So, Y.-L. Lau, K.-H. Chan, P. C. Y. Woo and K.-Y. Yuen, *J. Virol.* 2011, **85**, 11325.
17. D.J. Newman and G.M. Cragg, *J. Nat. Prod.* 2016, **79**, 629.
18. A. Hung, A. Ramek, Y. Wang, T. Kaya, J. Wilson, P. Clemons and D. Young, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 6799.
19. W. R. J. D. Galloway, A. Isidro-Llobet and D. R. Spring, *Nat. Commun.*, 2010, **1**, 80.
20. L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox and R. Hilgenfeld, *Science* 2020, **368**, 409.
21. S. Lu, Q. Ye, D. Singh, Y. Cao, J. K. Diedrich, J. R. Yatas III, E. Villa, D.W. Cleaveland and K.D. Corbett, *Nat Commun.*, 2021, **12**, 502.
22. M. Bouvet, A. Lugari, C.C. Posthuma, J.C. Zevenhoven, S. Bernard, S. Betzi, I. Imbert, B. Canard, J.-C. Guillemot and P. J. Lécine, *Biol. Chem.* 2014, **289**, 25783.

23. J. Yang, S.J.L. Petitjean, M. Koehler, Q. Zhang, A.C. Dumitru, W. Chen, S. Derclaye, S.P. Vincent, P. Soumillion and D. Alsteens, *Nat Commun.*, 2020, **11**, 4541.
24. (a) B. Cao, Y. Wei and M. Shi, *Org. Biomol. Chem.*, 2019, **17**, 3737. (b) T. Flagstad, G. Min, K. Bonnet, R. Morgentin, D. Roche, M. H. Clausen and T. E. Nielsen, *Org. Biomol. Chem.*, 2016, **14**, 4943
25. C.-P. Xu, Z.-H. Xiao, B.-Q. Zhuo, Y.-H. Wang and P.-Q. Huang, *Chem. Commun.*, 2010, **46**, 7834.
26. (a) R. A. Fernandes and V. Bethi, *RSC Adv.*, 2014, **4**, 40561; (b) M. Shibuya, M. Tomizawa and Y. J. Iwabuchi, *J. Org. Chem.* 2008, **73**, 4750.
27. I. Colomer, C. J. Empson, P. Craven, Z. Owen, R. G. Doveston, I. Churcher, S. P. Marsden and A. Nelson, *Chem. Commun.*, 2016, **52**, 7209.
28. M. Sorokina and C. J. Steinbeck, *Cheminform.*, 2019, **11**, 55.
29. C. Chang, K. Michalska, R. Jedrzejczak, N. Maltseva, M. Endres, A. Godzik, Y. Kim, and A. Joachimiak. To be published.
30. M. Rosas-Lemus, G. Minasov, L. Shuvalova, N. L. Inniss, O. Kiryukhina, J. Brunzelle and K. J. F. Satchell, *Sci Signal.* 2020, **13**, 1202.
31. Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, J. Yu, L. Wang, K. Yang, F. Liu, R. Jiang, X. Yang, T. You, X. Liu, X. Yang, F. Bai, H. Liu, X. Liu, L. W. Guddat, W. Xu, G. Xiao, C. Qin, Z. Shi, H. Jiang, Z. Rao and H. Yang, *Nature*, 2020, **582**, 289.
32. E. W. Bell and J. Zhang. *Cheminform.* 2019, **11**, 40.
33. R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo and Q. Zhou. *Science.* 2020, **367**, 1444.
34. E. Lindahl, M. J. Abraham, B. Hess and D. van der Spoel. Gromacs 2021.1 manual, March 2021, <https://doi.org/10.5281/zenodo.5053220>.
35. J. C. Ferreira, S. Fadl, A. J. Villanueva and W. M. Rabeh, *Frontiers Chemistry*, 2021, **9**, 491.
36. K. Y. Chen, T. Krischuns, L. O. Varga, E. Harigua-Souiai, S. Paisant, A. Zettor, J. Chiaravalli, A. Delpal, D. Courtney, A. O'Brien, S. C. Baker, E. Decroly, C. Isel, F. Agou, Y. Jacob, A. Blondel and N. Naffakh, *Antiviral Research*, 2022, **201**, 105272.
37. H. M. Froggatt, B. E. Heaton, N. S. Heaton, *J. Viro.*, 2020, **94**, e01265-01220.
38. W. Vuong, M. B. Khan, C. Fischer, E. Arutyunova, T. Lamer, J. Shields, H. A. Saffran, R. T. McKay, M. J. van Belkum, M. A. Joyce, H. S. Young, D. L. Tyrrell, J. C. Vederas and M. J. Lemieux, *Nat. Commun.*, 2020, **11**, 4282.
39. S. Iketani, F. Forouhar, H. Liu, S. J. Hong, F.-Y. Lin, M. S. Nair, A. Zask, Y. Huang, L. Xing, B. R. Stockwell, A. Chavez and D. D. Ho, *Nat. Commun.*, 2021, **12**, 2016.
40. J. Buchrieser, J. Duffloo, M. Hubert, B. Monel, D. Planas, M. M. Rajah, C. Planchais, F. Porrot, F. Guivel-Benhassine, S. Van der Werf, N. Casartelli, H. Mouquet, T. Bruel and O. Schwartz, *The EMBO Journal*, 2020, **39**, e106267.
41. A. K. Yazdi, F. Li, K. Devkota, S. Perveen, P. Ghiabi, T. Hajian, A. Bolotokova, and M. Vedadi, *SLAS Discovery*, 2021, **26**, 757.