RSC Medicinal Chemistry

RESEARCH ARTICLE

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

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

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-2*a*,6-epoxyazepino[3,4,5-*c*,*d*]indole skeletons, followed lead likeness factors in terms of molecular weight, C-sp^3 fraction and Clog*P*. 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 (EC$_{50}$ values of 3.7 and 1.4 μM, respectively) with an 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 the 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** targets 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

[a] Institut Pasteur, Department of Structural Biology and Chemistry, CNRS UMR no 3523 Chem4Life, Epigenetic Chemical Biology, Université Paris Cité, F-75015 Paris, France. E-mail: ibnel.haitham.hassan@gmail.com, paola.arimondo@cnrs.fr

[b] Institut Pasteur, Center for Technological Resources and Research (C2RT), CNRS UMR no 3523 Chem4Life, Chemogenomic and Biological Screening platform, Université Paris Cité, F-75015 Paris, France

[c] Drug Design and Discovery Lab, Zewail City of Science and Technology, 12578 Cairo, Egypt. E-mail: rkhidr@zewailcity.edu.eg

[d] Biomedical Sciences Program, University of Science and Technology, Zewail City of Science and Technology, 12578 Cairo, Egypt

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

[f] Department of Pharmacology and Toxicology, University of Toronto, Canada

[g] CNRS – UMR7257 – AFMB – Aix-Marseille Université, France

[h] QBI COVID-19 Research Group (QCRG), San Francisco, CA, USA

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d2md00149g

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 future.[14–16] There are now several treatments for COVID-19. For example, remdesivir, which has been granted emergency-use authorisation, showed modest benefit to patients with COVID-19.[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 condition.[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 ongoing does not guarantee 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 sp^3 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-2*a*,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 being 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 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 into 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 assay 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-sp^3 fraction, Clog$P$ and the 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 a large scale from the commercially

available *p*-anisidine (Scheme 1).[24a] This hydro-epoxybenzo[*c*, *d*]indole skeleton is a natural product-like compound and has a high 3-D character and high C-sp^3 fraction along with an optimal Clog$P$ 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 stereoelectronic environment, the two double bonds could be differentiated. The non-conjugated double bond (in magenta) was selectively dihydroxylated with potassium osmate(VI) to give diol **2** (Scheme 1). This intermediate could be transformed into decahydroepoxy-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 an intermediate bisaldehyde, which 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 the introduction of a new accessible point of diversity at the nitrogen atom.

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 group of compound **1** to the corresponding alcohol group and then reacted with electrophiles such as 2-chloropyrimidine to afford **4** (Scheme 1).
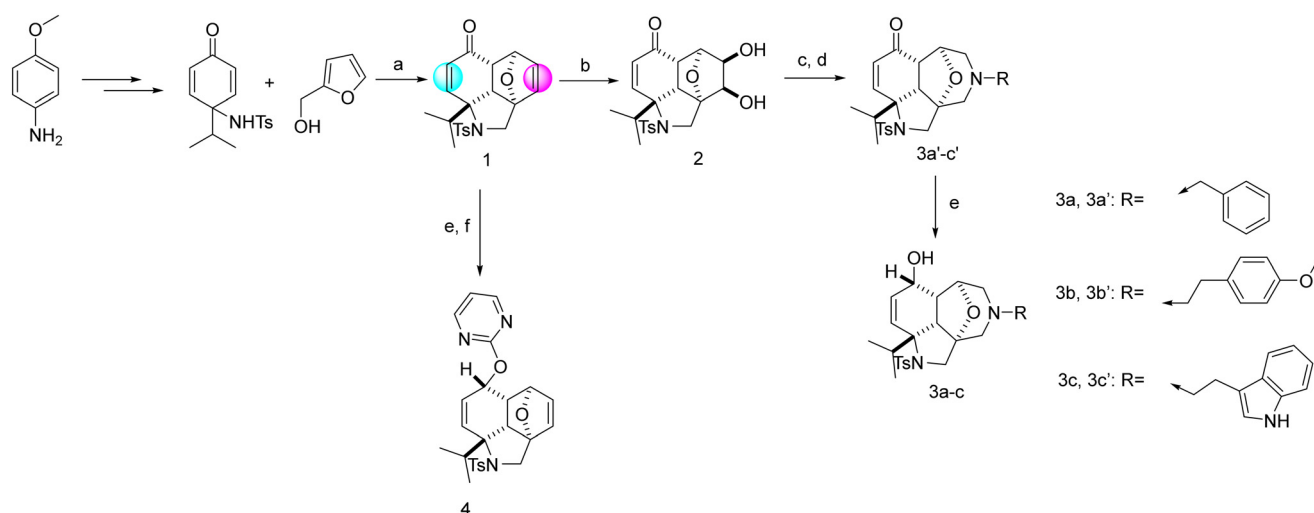
Scheme 1 Synthetic routes to the first round of functionalisation. Reagents and conditions: (a) DIAD (1.1 equiv.), PPh₃ (1.2 equiv.), 4 Å M.S. in THF, 77%. (b) NMO (1.2 equiv.), K₂OsO₄·2H₂O (0.8 mol%) in THF/H₂O 27% over two steps. (c) NaIO₄ in H₂O (1.4 equiv.), silica in DCM. (d) RNH₂ (1.2 equiv.), 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 and then 2-chloropyrimidine, r.t., overnight, 67%.
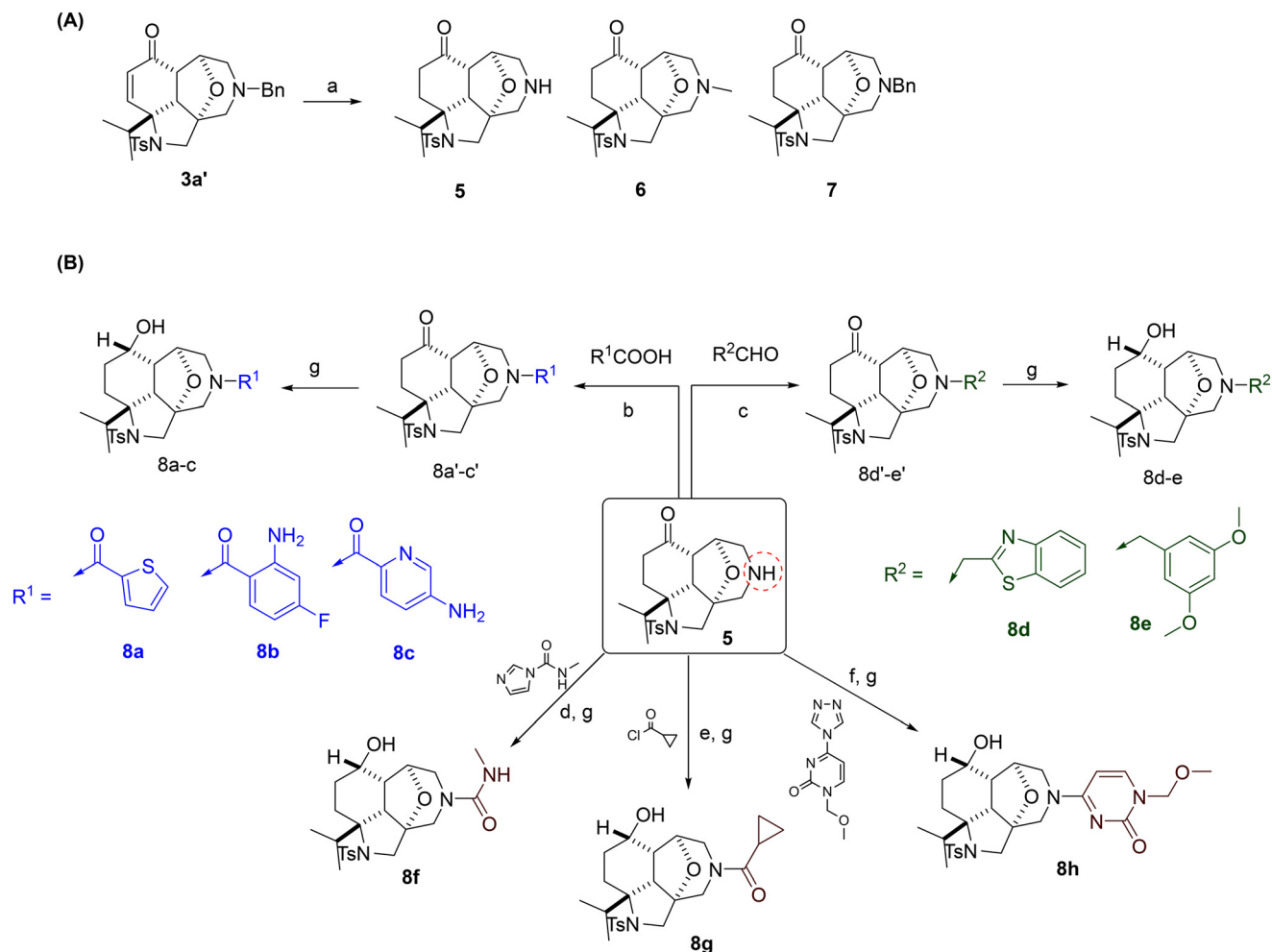
Scheme 2 Synthetic routes to the second round of functionalisation. Reagents and conditions: (a) H₂ (1.0 bar), Pd/C (10 mol%), MeOH (**5** and **6**) or EtOAc (**7**). (b) RCOOH (1.1 equiv.), DIPEA (1.5 equiv.), TBTU (2.5 equiv.) in DCM. (c) RCHO (1.2 equiv.), STAB (3.0 equiv.), AcOH (0.1 equiv.), M.S. 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.

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). 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 the preparation of a diverse set of compounds, showing the reactivity of this position. Reaction with aldehydes under reductive conditions gave tertiary amines and reaction with acyl chlorides 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

triazolopyrimidine, and then the ketone was reduced with NaBH₄ to give compound **8h** (Scheme 2, panel B).

The third round of functionalisation was conducted on **6**, **7**, and **8a**, **e**, and **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 dihydroxylation 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 salts[26] (page S18 in the ESI†).

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; **https://llama.leeds. ac.uk**).[27] The chemical library showed an average C-sp³ fraction of 0.5–0.7, average Clog*P* of 1.0–5.0 and average
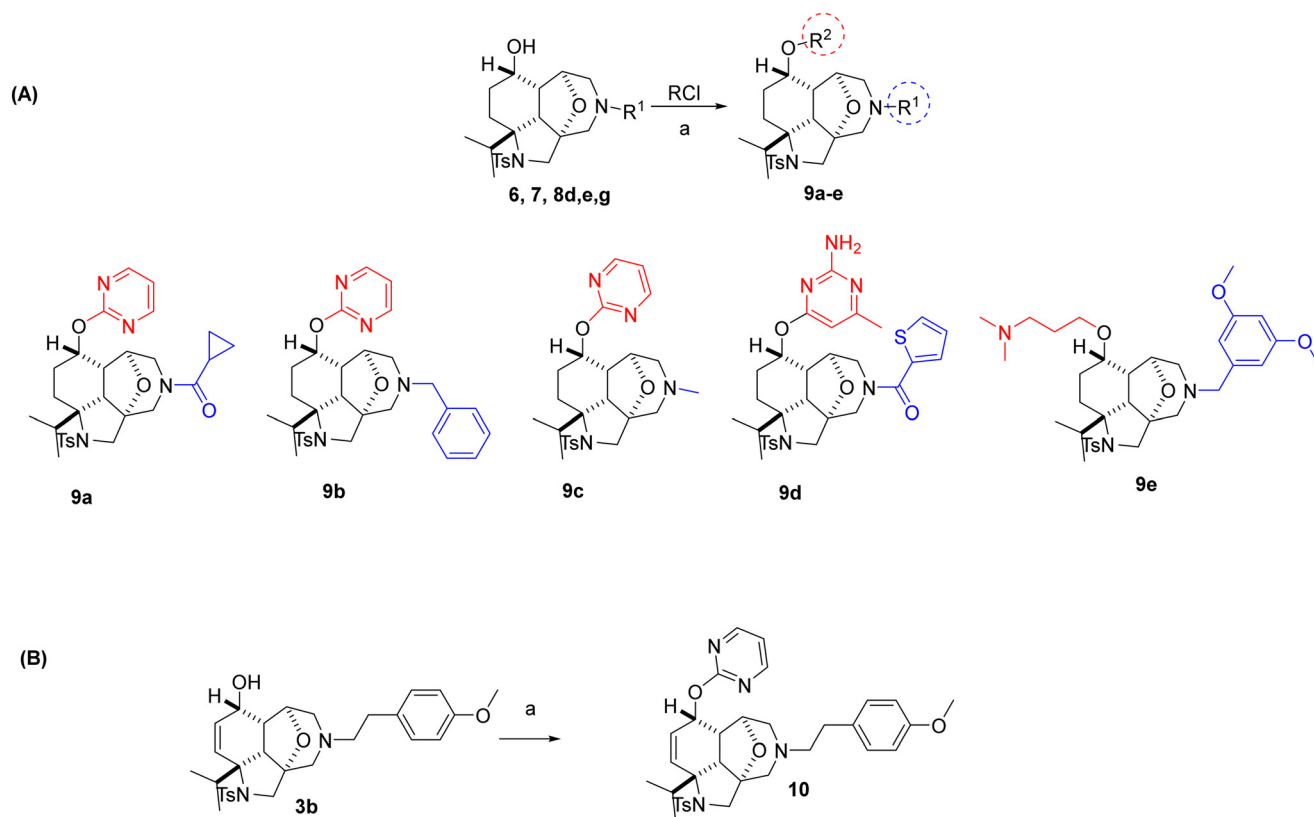
Scheme 3 Synthetic routes to the third round of functionalisation. Reagents and conditions: (a) NaH (60% dispersion in mineral oil) in dry THF, 0 °C, 1 h and then RCl, r.t., overnight.

molecular weight of 400–600, which are all features consistent with lead-likeness factors (Fig. 1, panel A). It also provided a high 3-D character (Fig. 1, panel B), a property that was proven to give better hit rates.[19] The natural product likeness was assessed using the Natural Product Likeness Score calculator (NaPLeS) 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 (Fig. 1, panel B).[28]

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 (adenocarcinomic human alveolar basal epithelial cells, source ATTC, reference CCL-158) 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 ESI† 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** appeared to be beneficial with an antiviral activity for compound **2** of EC_{50} = 14 μM. Structure activity relationship (SAR) revealed that the presence of large aromatic groups in the 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 group 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 with a basic centre gave better activity. In parallel, the cytotoxic activity of the compounds was assessed as the cytotoxicity could interfere with the antiviral activity of the compound (ESI† Methods section). Several hits showed a favourable ratio between the antiviral activity and cytotoxicity. The most active compounds with a favourable antiviral activity to cytotoxicity ratio and do not have electrophilic groups, which could favour covalent binding, are **3b** and **9e** (Fig. 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** exhibiting cytotoxicity above 20 μM (Table 1 and ESI† Fig. S1).

In silico studies

To identify the 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-

(A)

(B)

(C)
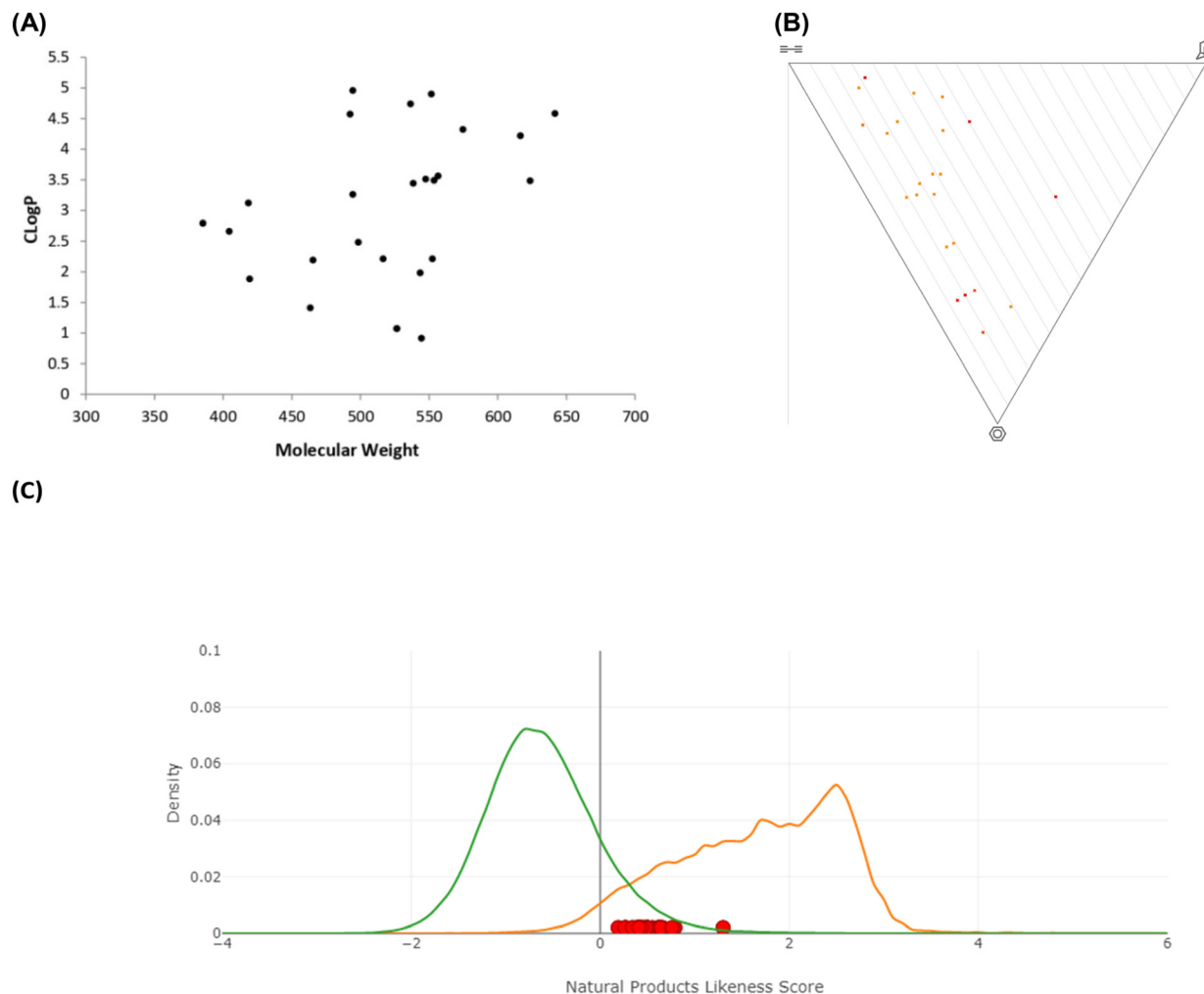
Fig. 1 Molecular properties and diversity. Panel A: Molecular weight *vs.* Clog *P* (Clog *P* was calculated using ChemDraw version 15). Panel B: Principal moment of inertia PMI plot. Panel C: Natural product likeness score, where the green line indicates all synthetic products, the orange line the natural products and the red dots the synthesised chemical library.

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 replication.[20] The second was the nucleocapsid phosphoprotein that packages the viral RNA into a helical ribonucleocapsid.[21] The third selected target was the non-structural nsp10–nsp16 complex that has a 2′-O-methyltransferase (2′-O-MTase) activity involved in the methylation of the RNA cap structure.[22,30] Finally, the receptor-binding domain (RBD) of the viral spike protein responsible for the viral entry into host cells was also virtually screened as the fourth plausible target.[23] 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 virtual screening

First, we validated the use of the method MOE 2019.01 for the virtual screening study by self-docking 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 (ESI† Fig. S2). The proteins were initially subjected to structure preparation and protonation state fixation, and then the 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 ≤2 Å and the results obtained by MOE 2019.01 were within this range, the adopted docking protocol was considered reliable.

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 50% inhibition of the maximal signal is observed at 95% confidence 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′	benzyl	—	Double bond	16 (N/A)	14 [2.8 to 73]	0.9
3a	benzyl	H	Double bond	17 [13 to 23]	>50	>3
3b′	4-methoxyphenethyl	—	Double bond	13 [7.2 to 24]	6.6 [2.7 to 16]	0.5
3b	4-methoxyphenethyl	H	Double bond	3.7 [2.0 to 6.9] 3.5 [2.3 to 5.6]^a	12 [10 to 15] >50^a	3.3
3c	indol-3-ylethyl	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	benzyl	—	Saturated	8.4 [5.5 to 13]	5.8 [3.0 to 11]	0.7
8a	thiophene-2-carbonyl	H	Saturated	16 [5.2 to 51]	>50	>3
8b	2-amino-4-fluorobenzoyl	H	Saturated	14 [8.5 to 22]	20 [8.2 to 48]	1.4
8c	5-aminopyridine-2-carbonyl	H	Saturated	37 [23 to 59]	>50	1.4
8d′	benzothiazol-2-ylmethyl	—	Saturated	5.3 [4.0 to 7.1]	37 N/A	6.9
8d	benzothiazol-2-ylmethyl	H	Saturated	6.6 [4.3 to 10]	12 [6.7 to 21]	1.6
8e	3,5-dimethoxybenzyl	H	Saturated	5.6 N/A	13 [7.6 to 21]	2.3
8f	N-methylcarbamoyl	H	Saturated	>50	>50	N/A
8h	3-(methoxymethyl)-2-oxo-tetrahydropyrimidinyl	H	Saturated	>17	>50	N/A
9a	cyclopropanecarbonyl	pyrimidin-2-yl	Saturated	>40 [N/A]	>50 [N/A]	N/A
9b	benzyl	pyrimidin-2-yl	Saturated	6.2 [N/A]	6.3 [3.7 to 11]	1.0
9c	Me	pyrimidin-2-yl	Saturated	20 [N/A]	17 [14 to 18]	0.8
9d	thiophene-2-carbonyl	4-methylpyrimidin-2-amine	Saturated	>50	>50	N/A

Table 1 *(continued)*

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
9e			Saturated	1.4 [0.9 to 2.2] 4.4 [2.0 to 9.0]^a	7.7 [5.2 to 12] >20^a	5.7
10			Double bond	>50	—	N/A

N/A = not applicable. ^{a} Assays against cells infected with SARS-CoV-2 variant Omicron BA1.

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 (ESI† Table S1). The 2D interactions of the molecules compared to those of 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 Fig. 3.

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} (Fig. 3A and ESI† Table S2). Compound **3b** forms six interactions: four as a H-bond donor with distances of 3.14 Å, 3.42 Å, 3.61 Å and 3.72 Å, respectively, and with energy scores of −0.6 kcal mol^{-1}, −0.8 kcal mol^{-1}, −0.8 kcal mol^{-1}, and −1.6 kcal mol^{-1}, respectively. It also indicated two pi–H interactions between the ligand's 6-membered ring and Thr25 and Thr26 in the pocket with distances of 4.39 Å and 4.22 Å, respectively, and with energy scores of −1.1 kcal
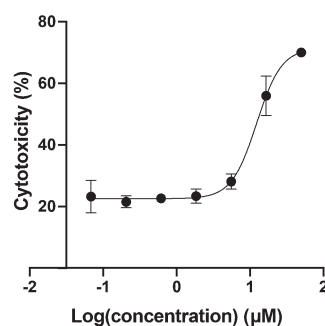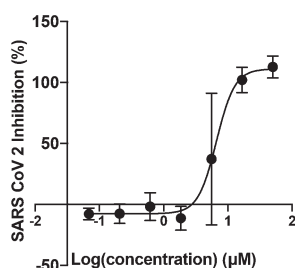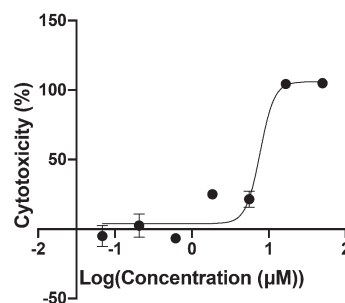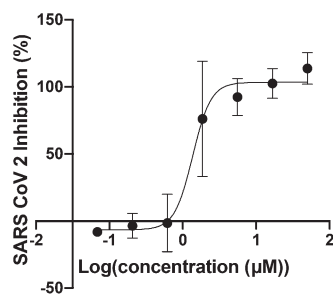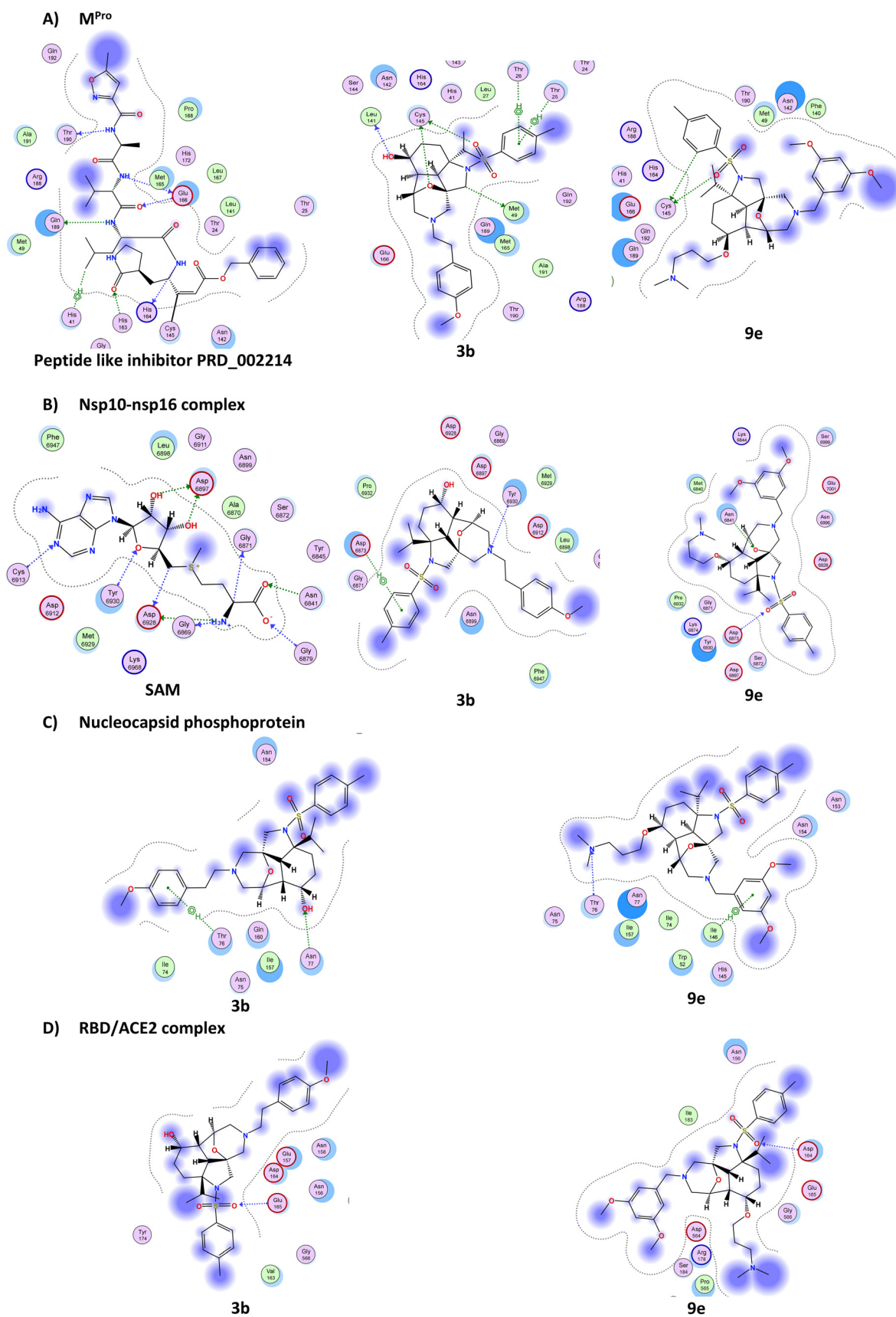

Fig. 2 Anti-SARS-CoV-2 dose response (left) and cytotoxicity (right) curves with respect to the log concentration in μM of the compound. Panel A: Compound **3b**. Panel B: Compound **9e**. Error bars of triplicates are shown.

Fig. 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 the 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).

mol^{-1}, and −2.0 kcal mol^{-1} respectively (Fig. 3A and ESI† Table S2). This suggests that compound **3b** might interact with this target protein. In contrast, compound **9e** makes only two interactions with M^{pro} (Fig. 3A and ESI† Table S2). For the nsp10–nsp16 complex, we mapped strong 11 interactions with the co-crystallized ligand SAM (Fig. 3B and ESI† 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 (Fig. 3C). Finally, with the spike protein RBD/ACE2 complex (Fig. 3D), only one H-bond was observed between O62 of compound **3b** and Glu165 and one hydrogen bond between O72 of compound **9e** and Asp164, suggesting a reduced interaction with this target complex (ESI† 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 (ESI† data and Fig. S3–S6). To measure how much the protein and ligand conformations change along the MD simulation trajectory of the target protein–ligand complex, RMSD values were calculated for the proteins' C$_\alpha$ atomic coordinates and the screened ligands using GROMACS 2021.1.[32,34] **3b** and 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 (Fig. 4A). Additionally, the stability of the complex was confirmed by the RMSD of **3b** itself, as displayed in Fig. 4B, eliciting perturbations of less than 1 Å throughout the whole simulation. On the other hand, the RMSD of the nsp10–nsp16 complex C$_\alpha$ backbone in its complex with **3b** showed small perturbations, while the RMSD of **3b** itself gave very high fluctuations after 50 ns.

Then the stability of **9e** in complex with M^{pro} and the nsp10–nsp16 complex was assessed over a 100 ns simulation (Fig. 5). The RMSD of the protein referenced to the backbone was calculated. On the 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 (Fig. 5A, black curve). This indicates the 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) (Fig. 5B, black curve). In parallel, the nsp10–nsp16 complex exhibited convergence at 12 ns with small perturbations (∼1 Å), thus indicating the stabilisation of the protein (ESI† Fig. S5). Besides that, the RMSD of **9e** complexed to the nsp10–nsp16 complex exhibited a very low RMSD (less than 6 Å) and RMSD perturbations, which reflect the high **9e** complex stability as shown in Fig. 5B.

Stable binding of **3b** and **9e** should reflect 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$_\alpha$ atoms during the simulation (ESI† Fig. S5).[35] The RMSF of the **3b** and M^{pro} complex showed low fluctuations, especially Cys145 and His41 that constitute the catalytic dyad. **9e** exhibited fluctuations (∼1.3–3.2 Å) in some residues, with higher fluctuations at residues (185∶192) than **3b**. In parallel, the nsp10–nsp16 complex shows residues within a 5 Å proximity from **3b** and **9e**, exhibiting lower fluctuations (∼0.4–1.4 Å) (ESI† Fig. 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 the nsp10–nsp16 complex. To verify the *in* silico results, compounds **3b** and **9e** were tested *in vitro* for inhibition of the 2′-O-methyltransferase activity of the nsp10–nsp16 complex and nsp14 guanine-N7-methyltransferase activity, and in a cellular assay involving M^{pro}.

Biological target validation

Inhibition of the nsp10–nsp16 complex. The nsp10–nsp16 complex is responsible for catalysing the final step of the coronaviral mRNA capping. Nsp16 is a member of the 2′-O-
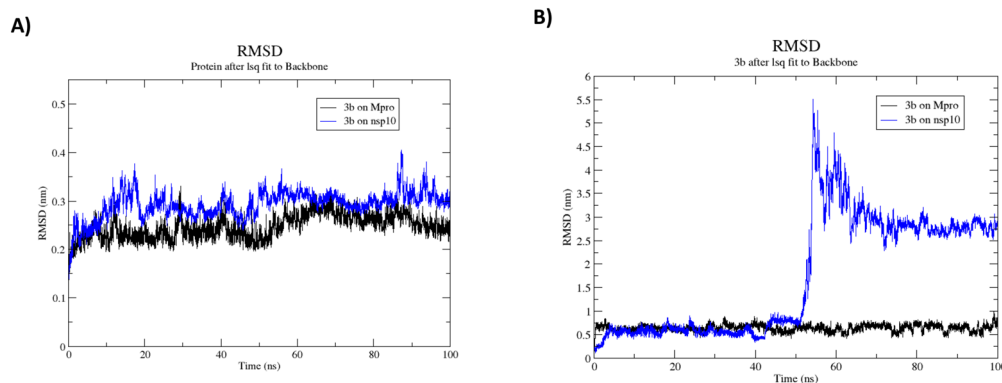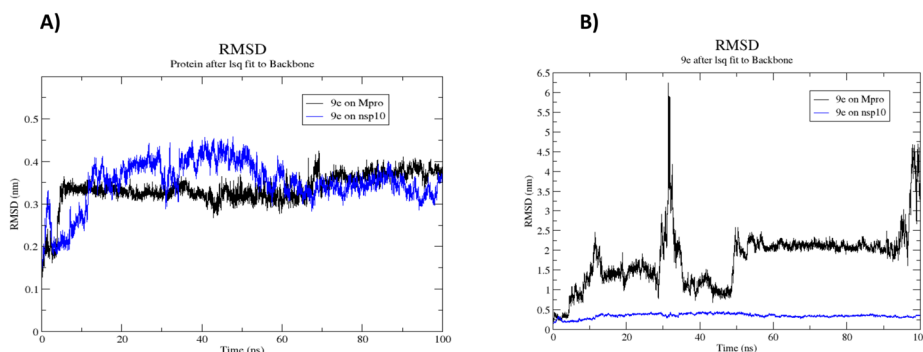

Fig. 4 Structural dynamics of compound **3b** bound to M^{pro} (black curve) and the nsp10–nsp16 complex (blue curve); RMSD of the C$_\alpha$ backbone of the proteins in nm (A) and RMSD of **3b** in nm along the MD trajectory (B).

Fig. 5 Structural dynamics of compound **9e** bound to M^{pro} (black curve) and the nsp10–nsp16 complex (blue curve); RMSD of the C$_\alpha$ backbone of the proteins in nm (A) and RMSD of **9e** in nm along the MD trajectory (B).

Table 2 Effect of **3b** and **9e** on the methyltransferase activities of the nsp10–nsp16 complex and nsp14. *S*-Adenosyl-ʟ-homocysteine (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 Fig. S7.† NI: no inhibition, NA: not applicable.

MTase family, which catalyses the transfer of a methyl group to the RNA substrates from the methyl donor SAM. Nsp16 requires nsp10 for methyltransferase activity and stability. The *in vitro* activity of the SARS-CoV-2 nsp10–nsp16 complex was assessed by monitoring the transfer of ^{3}H-SAM to the biotinylated N7-meGpppACCCCC RNA substrate (ESI†). The subsequently methylated RNA was captured using scintillation proximity assay (SPA) beads followed by quantifying the level of incorporated ^{3}H-methyl by measuring the radioactivity level (counts per minute [CPM]).[41] In dose response experiments (Fig. S7†), commercially available *S*-adenosyl-ʟ-homocysteine (SAH) showed IC$_{50}$ values of 3.9 μM and 0.22 μM for the nsp10–nsp16 complex and nsp14 methyltransferase activities, respectively (ESI† Fig. S7 and Table 2). However, **3b** and **9e** did not show any inhibitory effect on the 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 ESI† Fig. S7).

Inhibition of SARS-Cov-2 main protease (M^{pro}). M^{pro} is a validated target for treating COVID-19 with the recent 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 ESI† 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 (Fig. 6A), while the luciferase signal measured in the presence of the catalytically inactive mutant remained unchanged. Full inhibition of M^{pro} was achieved at 10 μM GC376 and the estimated IC$_{50}$ was in the micromolar range.[37–39] Only compound **3b** showed a specific activity against M^{pro} in the Rev-Nluc assay from 10 μM, as indicated by an increase in the luciferase signal, which was not observed with the catalytically inactive M^{pro} mutant (Fig. 6B). An *in vitro* enzymatic assay on the purified protein confirmed the weak inhibitory activity of compound **3b** (ESI† Fig. S8), while compound **9e** was inactive. This observed activity of **3b** against SARS-CoV-2 M^{pro} was consistent with the computational study.

Conclusions

In conclusion, to explore a new chemical space for the design of inhibitors of SARS-CoV-2, a chemical library of 25 fused-
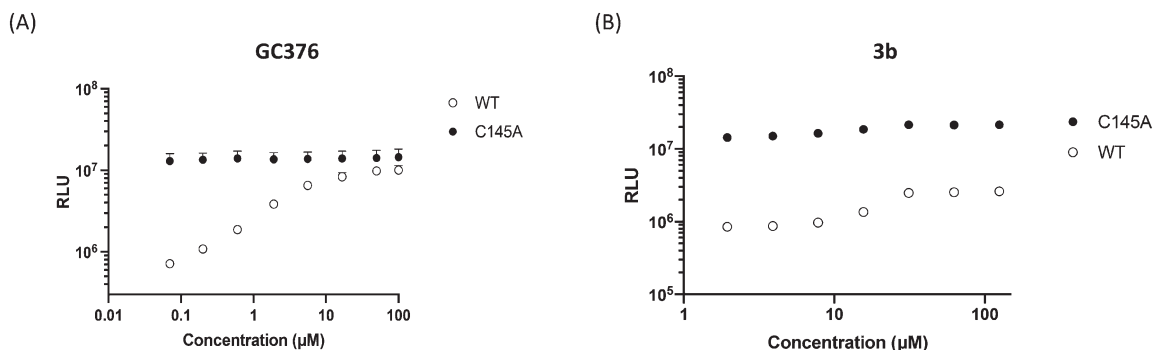
Fig. 6 Cell-based assay luciferase assay to monitor SARS-CoV-2 main protease activity of the wild-type (WT, open symbols) and catalytically inactive C145A mutant (black symbols) of M^{pro}: (A) dose response curve of the positive control GC376. (B) Dose response curve of compound **3b**. The inhibition was assessed using the Rev-Nluc-based assay. The assay was performed as described in the ESI† Methods section. The inhibitor concentrations correspond to 3-fold serial dilutions from 125 to 0.02 μM. The RLUs are shown as the mean ± SD of three independent experiments performed in technical triplicate.

bridged dodecahydro-2a,6-epoxyazepino[3,4,5-c,d]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 antiviral activity against SARS-CoV-2 with an EC$_{50}$ between 3.7 and 1.4 μM with an acceptable cytotoxicity difference. They also showed anti-viral activity against SARS-CoV-2 variant Omicron BA1, which consolidates 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 the 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, an *in vitro* inhibition assay of the 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-c,d]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 allow further modulation of the substituents. 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

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

ANR-18-CE18-0026, ANR-18-CE18-0028 and ANR-10-LABX-62-IBEID).

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. Feng, *N. Engl. J. Med.*, 2020, **382**, 1199.

4 H. Mitsuya and N. Kokudo, *Global Health 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, pp. 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, J. Am. Med. Assoc.*, 2020, **324**, 1041.

9 (*a*) C. D. Spinner, R. L. Gottlieb, G. J. Criner, J. R. A. López, A. M. Cattelan, A. S. 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. Gaggar, D. M. Brainard, M. J. McPhail, S. Bhagani, M. Y. Ahn, A. J. Sanyal, G. Huhn and F. M. Marty, *JAMA, J. Am. Med. Assoc.*, 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 Pharm. Sin. B*, 2021, **11**, 237.

10 S. P. Kaur and V. Gupta, *Virus Res.*, 2020, **288**, 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. Clevel 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. Lécine, *J. 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, *J. Cheminf.*, 2019, **11**, 55.

29 C. Chang, K. Michalska, R. Jedrzejczak, N. Maltseva, M. Endres, A. Godzik, Y. Kim and A. Joachimiak, *Crystal structure of RNA binding domain of nucleocapsid phosphoprotein from SARS coronavirus 2*, to be published.

30 M. Rosas-Lemus, G. Minasov, L. Shuvalova, N. L. Inniss, O. Kiryukhina, J. Brunzelle and K. J. F. Satchell, *Sci. Signaling*, 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, *J. Cheminf.*, 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*, 2021, DOI: **10.5281/zenodo.5053220**.

35 J. C. Ferreira, S. Fadl, A. J. Villanueva and W. M. Rabeh, *Front. Chem.*, 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 Res.*, 2022, **201**, 105272.

37 H. M. Froggatt, B. E. Heaton and N. S. Heaton, *J. Virol.*, 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. Dufloo, 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, *EMBO J.*, 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.