# The structure-based design of SARS-CoV-2 nsp14 methyltransferase ligands yields nanomolar inhibitors

**Tomáš Otava[a], Michal Šála[a], Fengling Li[b], Jindřich Fanfrlík[a], Kanchan Devkota[b], Paknoosh Pakarian[b], Pavel Hobza[a], Masoud Vedadi\*[b,c], Evzen Boura\*[a], Radim Nencka\*[a]**

[a]Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, v.v.i., Gilead Sciences Research Centre at the IOCB Prague, Flemingovo nam. 2., 166 10 Prague 6, Czech Republic.
[b]Structural Genomics Consortium, University of Toronto, Toronto, ON M5G 1L7, Canada
[c]Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON M5G 1L7, Canada.
[†]Electronic Supplementary Information (ESI) available. See DOI: 10.1039/x0xx00000x

\*To whom correspondence should be addressed:

Radim Nencka; Tel.: +420-220-183-265; E-mail: radim.nencka@uochb.cas.cz

Evzen Boura; Tel.: +420-220-183-465; E-mail: evzen.boura@uochb.cas.cz

Masoud Vedadi; Tel.: 416-432-1980; E-mail: m.vedadi@utoronto.ca

**We have focused on the structure-based design of the inhibitors of one of the two SARS-CoV-2 methyltransferases (MTases), nsp14. We have synthesized and tested the designed compounds in vitro and shown that these derivatives exert unprecedented inhibitory activity against this crucial viral enzyme.**

COVID-19 and its causative agent, the SARS-CoV-2 virus, dominated the public debate last year.[1] The virus first emerged in Wuhan, China, and quickly spread around the world, causing a world-wide pandemic.[2] The vaccines have been developed at an amazing pace and vaccination has already started in several countries.[3] However, only one small-molecule drug, remdesivir,[4,5] a broad-spectrum RNA-dependent RNA polymerase inhibitor,[6-8] is currently fully approved for treatment of COVID-19, but small-molecule drugs are urgently needed for people who, for various reasons, cannot be vaccinated or become sick before or despite vaccination. SARS-CoV-2 and other coronaviruses have an unusually large genome (~30,000 bp) for single-stranded positive-sense RNA viruses.[2,9] They encode four structural, 16 non-structural (nsp1–16) and several accessory proteins.[10,11] Among these, enzymes are the best potential drug targets. Coronaviruses encode two proteases, RNA-dependent RNA polymerase (RdRp), helicase, endonuclease and two MTases.[12,13] Each of these is a valid drug target as inhibition of any of them would be lethal for the virus. Several inhibitors of SARS-CoV-2 RdRp and the main protease that have drug-like properties have recently been described.[14-16] Here, we have focused on the nsp14 protein with two enzymatic activities: the N-terminal domain has 3′-to-5′ exoribonuclease (ExoN) activity, while the C-terminal domain has N7-MTase activity.[10,17,18] The ExoN activity is critical for the maintenance of the large RNA genome,[17] whereas the MTase activity is important for the stability of viral RNAs. Nsp14 uses *S*-adenosylmethionine (SAM) as a methyl donor to methylate the N7 position of 5' guanine, one of the steps to build the cap that is essential for mRNA stability and translation in human cells. We have used a structure-guided approach to design molecules with the potential to inhibit the nsp14 MTase activity. Several of the compounds exhibited single-digit nanomolar activities in biochemical assays.

Most of SARS-CoV-2 enzymes have been quickly structurally characterized, but nsp14 still resists the crystallographic analysis. Therefore, we have constructed a model of SARS-CoV-2 in complex with SAM based on its high homology (94.9% sequence identity) with the crystallized SARS-CoV nsp14.[19] We have based the design of our inhibitors on the molecular structure of SAH, a by-product of the methylation reaction catalysed by this protein. We have focused in particular on the interaction of such derivatives with lipophilic residues and residues potentially forming non-covalent interactions with π-systems within a lateral cavity revealed by manual inspection of the homology model. The cavity is defined by several hydrophobic residues (Val287, Phe367, Val389, Pro429) and Arg289, the last of which nicely closes the cavity by an interaction with the benzene ring of Phe367 (Fig. 1). The size of the cavity in this region suggests that it is able to accommodate a relatively large planar substituent, which can significantly reduce the hydrophilic character of the ligand:nsp14 interaction and lead to a considerable increase in this interaction. We thus decided to use aromatic systems connected by appropriate linkers to position 7 of the 7-deazaadenine derivative of SAH (Fig. 1). We selected methylene, ethynediyl and ethylene as suitable linkers for the connection of aromatic parts to the nucleobase. We hypothesized that selected aromatic systems could significantly interact with both aromatic amino-acid residues and Arg289 through the cation-π interaction, which had been observed for numerous interactions between ligands and proteins. Therefore, in most cases, we chose aromatic systems composed of two aromatic rings that provided large enough π electron clouds. Various aromatic

systems bearing both hydrophilic and hydrophobic substituents were prepared in our study, and the most active derivatives were selected for presentation.
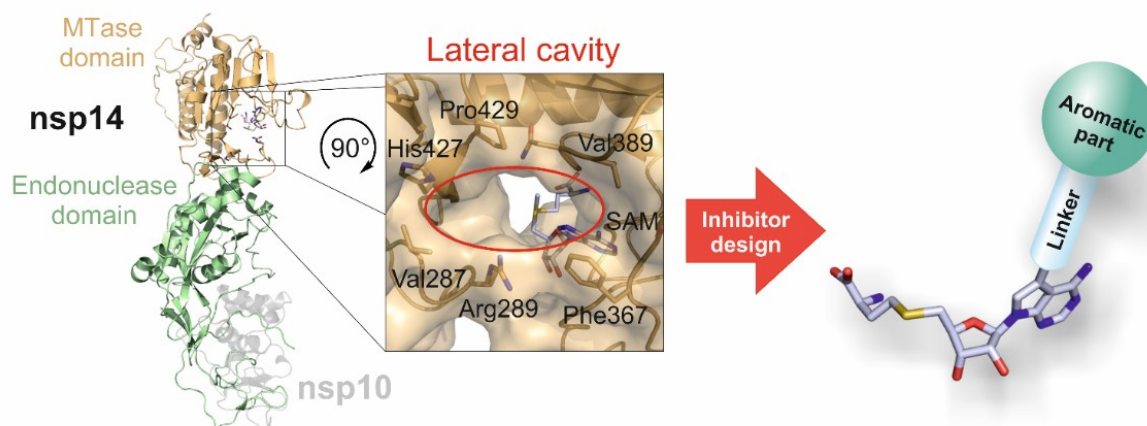


**Fig. 1** The homology model of nsp14 in complex with its activator protein nsp10 shows a well-defined lateral cavity in close proximity to the SAM/SAH binding site of the nsp14 MTase domain. This cavity is adjacent to the N7 position of the purine nucleobase and thus can be targeted by SAH derivatives bearing an aromatic substituent attached to this position by various linkers.

The starting compound for the synthesis of the designed nsp14 inhibitors was persilylated 7-deaza-7-iodoadenosine **1**.[20,21] First, the silyl-protecting group attached at the 5′-position was selectively removed by treatment with trichloroacetic acid affording nucleoside **2** in 72 % yield.[22] Activation of the 5′-hydroxy group of the compound **2** with MsCl in the presence of Et$_3$N allowed to introduce an amino-acid moiety by nucleophilic substitution with freshly prepared protected L-homocysteine salt[23,24] yielding SAH analogue **3** (82 % after two steps). This key intermediate **3** was then subjected to cross-coupling reactions to attach hetaryl and aryl substituents through various linkers. For the preparation of the analogue with a methylene bridge, the Negishi cross-coupling reaction with benzylzinc bromide in the presence of Pd$_2$dba$_3$ and Xantphos was used affording compound **4** in 81 % yield. Derivatives **5**–**10** with an ethynediyl bridge were prepared via Sonogashira cross-coupling reactions with various (het)aryl acetylenes (for the synthesis see ESI†) using a combination of Pd(PPh$_3$)$_4$ and CuI as a catalyst. In the case of compound **8** bearing quinoline substituent, its analogue **11** with ethylene linker was synthesized via reduction of the ethynediyl bridge by catalytic hydrogenation on Pd/C. Final deprotection of 7-substituted SAH analogues (**4**–**11**) under acidic conditions (TFA/water 9:1) afforded the target compounds **12**–**19** in very good to excellent yields (83–97 %, Scheme 1).

A recently developed radioactivity-based nsp14 MTase assay[25] was employed to evaluate the inhibitory effect of compounds **12**–**19**. To our delight, all eight compounds inhibited nsp14 MTase activity with IC$_{50}$ values ranging from 3 to 37 nM (Table 1 and ESI†, Fig. S1a). Compound **16** was the most potent with the IC$_{50}$ value of $3 \pm 0.5$ nM (Fig. 2a, Table 1). To confirm the binding of these potent inhibitors to nsp14 orthogonally, we employed surface plasmon resonance (SPR). Strong binding of all eight compounds to nsp14 was observed (Table 1, ESI†, Fig. S1b); the calculated K$_D$ values from steady-state affinity fitting were in the range of 11–116 nM (Table 1). Binding affinities were in good agreement with the pattern of their inhibitory effect on nsp14 MTase activity (Table 1).
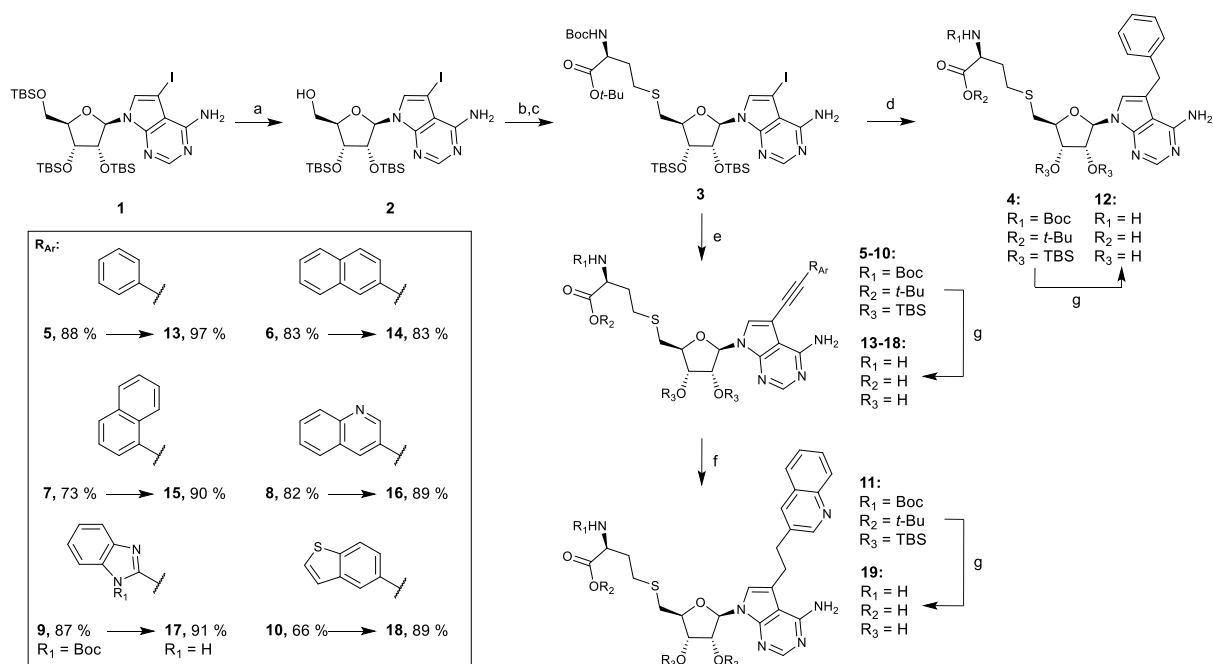
**Scheme 1** Reagents and conditions: (a) CCl$_3$COOH, THF/H$_2$O, 0 °C, 72 %. (b) MsCl, Et$_3$N, CH$_2$Cl$_2$, 0 °C–r.t. (c) *tert*-butyl (*tert*-butoxycarbonyl)-L-homocysteinate, *t*-BuOK, NMP, r.t., 82 % after 2 steps. (d) Benzylzinc bromide, Pd$_2$(dba)$_3$, Xantphos, THF, 60 °C. (e) (Het)aryl acetylene, Pd(PPh$_3$)$_4$, CuI, Et$_3$N, THF, r.t. (f) Prepared from **8**, Pd/C-H$_2$, EtOAc/EtOH, r.t. (g) TFA/H$_2$O (9:1), r.t.
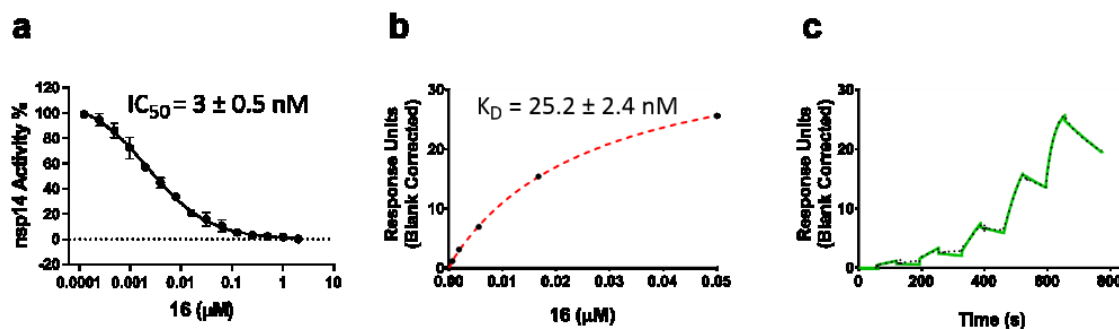


**Fig. 2** The characterization of compound **16**. (a) Concentration-dependent nsp14 MTase inhibition by compound **16** was assessed using the radioactivity-based assay. The binding of compound **16** to nsp14 was tested by SPR. (b) The steady-state response (black circles) with the steady-state 1:1 binding model fitting (red dashed line) and (c) the sensorgram (solid green) with the kinetic fit (black dots) are shown. The experiments were performed in triplicate.

**Table 1** The evaluation of the binding affinity and inhibitory effect of the designed compounds. All values come from the data presented in Fig. 2 (n=3) and, ESI† Fig. S1.

| Compound | $IC_{50}$ (nM) | Hill Slope | $K_D$ (nM) |
|:---:|:---:|:---:|:---:|
| 12 | 37 ± 7 | −0.7 | 31.2 ± 0.9 |
| 13 | 11 ± 1 | −0.8 | 14.1 ± 0.7 |
| 14 | 17 ± 3 | −0.8 | 116 ± 7.7 |
| 15 | 6 ± 1 | −0.8 | 21.6 ± 1.3 |
| 16 | 3 ± 0.5 | −0.9 | 25.2 ± 2.4 |
| 17 | 4 ± 0.5 | −0.9 | 15.1 ± 0.3 |
| 18 | 15 ± 3 | −0.7 | 11.3 ± 0.9 |
| 19 | 12 ± 1 | −0.8 | 23.1 ± 1.0 |

In order to corroborate our observations in enzymatic assays, we used our homology model and performed molecular docking of selected compounds from our study. The homology model was prepared using SWISS-MODEL server[26] based on the crystal structure of nsp14 in complex with nsp10 obtained for SARS-CoV by Ma et al. (PDB ID:5C8T).[19] The SAM-binding site and the position of the ligand were optimized and four water molecules in the immediate vicinity of the amino-acid part of the SAM molecule were modelled in order to stabilize this flexible part of our SAH analogues during the docking experiments. The 3D structures of the ligands were optimized using MOPAC 2016[27] prior to docking. The homology model of the protein was processed using the standard procedure for the GOLD software[28], which was selected as the most suitable software for the docking of SAM/SAH analogues based on our previous experience and initial docking experiments. In all cases, the docking experiments in the GOLD software resulted in the highest scores ranging from 95.9 (for compound **12**) to 110.9 (compound **15**) for the poses that orient the aromatic parts into the lateral cavity defined above (ESI†, Table S1 and Fig. S2). We have decided to present the binding patterns of the compounds that are of interest and may lead to further improvements in these inhibitors (Fig. 3). Interestingly, the compound with the lowest $IC_{50}$ (derivative **16**) gave only an average score (99.0, Fig. 3, A and B) as did the compound **18**, which exhibited the highest affinity in the SPR assay (102.0, Fig. 3, C and D). In contrast, the compound with the third lowest $IC_{50}$ value, derivative **15**, scored the best in GOLD (110.9, Fig. 3, E and F). The 1-naphtyl moiety of this compound seems to occupy very nicely the lateral cavity of nsp14. We have also obtained an interesting docking result for the compound **13**, which exhibits relatively high activity in both $IC_{50}$ and SPR assays. The docking results show that the aromatic core can also be located in the lateral cavity perpendicular to the rotation of this part in other compounds. This can be attributed to the less lipophilic nature of the benzimidazole moiety in comparison with the other substituents used (Fig. 3, G and H). We can only speculate that these discrepancies between the actual activity and the predicted docking scores can be attributed to the small differences between the obtained activities and affinities of the inhibitors, imperfections in the homology model, flexibilities of lateral-cavity residues (e.g. Arg289, Phe367 and Asn388) and/or further water-mediated contacts with the nsp14 protein. We tried to address this problem by molecular modelling at the semiempirical quantum mechanical level of theory, which showed that compound **16** had considerably more favorable binding free energy than SAH, due to the interaction with key Phe367, Asn388 and Pro429 residues in the lateral pocket (see ESI†).
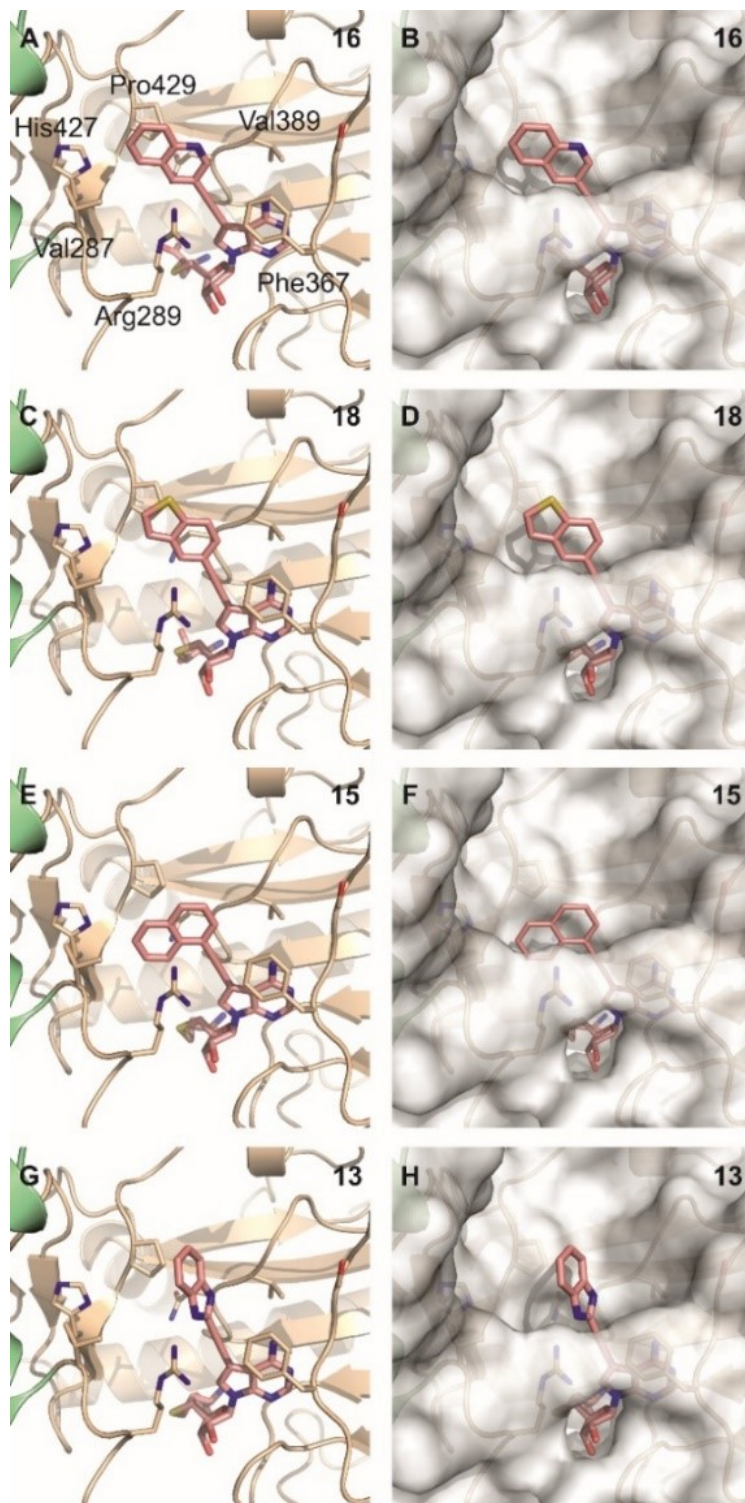
**Fig. 3** The results of the docking studies of compounds **16**, **18**, **15** and **13** using the homology model of SARS-CoV-2 nsp14 MTase.

In conclusion, this study describes the first rationally designed inhibitors of nsp14 MTase, based on modified derivatives of the SAH molecule. In order to target effectively the lateral cavity of the enzyme, which directly communicates with the SAM/SAH-binding site, we used the 7-deazaadenine analogue of SAH. This allowed us to attach aromatic substituents to position 7 using appropriate linkers. These derivatives showed considerable inhibitory activity against nsp14 (the $IC_{50}$ values were between 3 and 37 nM) and strong binding to the nsp14 protein by SPR (with the $K_D$ values being in the range of 11–

31 nM). The most active compound in the study was compound **16** (TO507) with $IC_{50}= 3 \pm 0.5$ nM. Our docking and molecular modelling studies have clearly shown that these compounds can effectively exploit the space within the lateral cavity, which leads to effective binding to the SAM-binding site of the protein. Although the compounds are inherently zwitterionic, and hence their penetration into cells and their activity in cell cultures are rather unlikely, we believe that our study provides a significant springboard for the design of drugs targeting the nsp14 protein.

## Conflicts of interest

A part of the study was also supported by Gilead Sciences Inc.

## Notes and references

1   C. C. Lai, T. P. Shih, W. C. Ko, H. J. Tang and P. R. Hsueh, *Int. J. Antimicrob. Agents*, 2020, **55** (3), 105924.
2   J. F. W. Chan, K. H. Kok, Z. Zhu, H. Chu, K. K. W. To, S. F. Yuan and K. Y. Yuen, *Emerg. Microbes & Infect.*, 2020, **9** (1), 221–236.
3   D. D. Li and Q. H. Li, *Mil. Med. Res.* 2021, **8** (1).
4   A. J. Pruijssers, A. S. George, A. Schafer, S. R. Leist, L. E. Gralinksi, K. H. Dinnon, B. L. Yount, M. L. Agostini, L. J. Stevens, J. D. Chappell, X. T. Lu, T. M. Hughes, K. Gully, D. R. Martinez, A. J. Brown, R. L. Graham, J. K. Perry, V. Du Pont, J. Pitts, B. Ma, D. Babusis, E. Murakami, J. Y. Feng, J. P. Bilello, D. P. Porter, T. Cihlar, R. S. Baric, M. R. Denison and T. P. Sheahan, *Cell Rep.* 2020, **32** (3), 107940.
5   E. de Wit, F. Feldmann, J. Cronin, R. Jordan, A. Okumura, T. Thomas, D. Scott, T. Cihlar and H. Feldmann, *Proc. Natl. Acad. Sci. U.S.A.*, 2020, **117** (12), 6771–6776.
6   M. L. Agostini, E. L. Andres, A. C. Sims, R. L. Graham, T. P. Sheahan, X. T. Lu, E. C. Smith, J. B. Case, J. Y. Feng, R. Jordan, A. S. Ray, T. Cihlar, D. Siegel, R. L. Mackman, M. O. Clarke, R. S. Baric and M. R. Denison, *mBio* 2018, **9** (2).
7   A. J. Brown, J. J. Won, R. L. Graham, K. H. Dinnon, A. C. Sims, J. Y. Feng, T. Cihlar, M. R. Denison, R. S. Baric and T. P. Sheahan, *Antivir. Res.* 2019, **169**, 104541.
8   E. Konkolova, M. Dejmek, H. Hrebabecky, M. Sala, J. Boserle, R. Nencka and E. Boura, *Antivir. Res.* 2020, **182**, 104899.
9   M. Romano, A. Ruggiero, F. Squeglia, G. Maga and R. Berisio, *Cells*, 2020, **9** (5), 1267.
10  E. J. Snijder, E. Decroly and J. Ziebuhr, *Adv. Virus Res.*, 2016, **96**, 59–126.
11  J. Ziebuhr, *Diseases* 2006, **581**, 3–11.

12   Y. U. Wu, Z. Li, Y. S. Zhao, Y. Y. Huang, M. Y. Jiang and H. B. Luo, *Med. Res. Rev.*, 2021, **41**.

13   K. Sarkar, P. C. Sil, S. F. Nabavi, I. Berindan-Neagoe, C. A. Cismaru, S. M. Nabavi and S. Habtemariam, *Mini Rev. Med. Chem.*, 2020, **20** (18), 1900–1907.

14   H. Yousefi, L. Mashouri, S. C. Okpechi, N. Alahari and S. K. Alahari, *Biochem. Pharmacol.* 2021, **183**, 114296.

15   A. Frediansyah, R. Tiwari, K. Sharun, K. Dhama and H. Harapan, *Clin. Epidemiol. Glob. Health*, 2021, **9**, 90–98.

16   B. D. Kevadiya, J. Machhi, J. Herskovitz, M. D. Oleynikov, W. R. Blomberg, N. Bajwa, D. Soni, S. Das, M. Hasan, M. Patel, A. M. Senan, S. Gorantla, J. McMillan, B. Edagwa, R. Eisenberg, C. B. Gurumurthy, S. M. Reid, C. Punyadeera, L. D. Chang and H. E: Gendelman, *J. Neuroimmune Pharmacol.* 2021, **16**, 1–26.

17   E. Minskaia, T. Hertzig, A. E. Gorbalenya, V. Campanacci, C. Cambillau, B. Canard and J. Ziebuhr, *Proc. Natl. Acad. Sci. U.S.A.*, 2006, **103** (13), 5108–5113.

18   N. S. Ogando, J. C. Zevenhoven-Dobbe, Y. van der Meer, P. J. Bredenbeek, C. C. Posthuma and E. J. Snijder, *J. Virol.*, 2020, **94**.

19   Y. Y. Ma, L. J. Wu, N. Shaw, Y. Gao, J. Wang, Y. N. Sun, Z. Y. Lou, L. M. Yan, R. G. Zhang and Z. H. Rao, *Proc. Natl. Acad. Sci. U.S.A.*, 2015, **112** (30), 9436–9441.

20   L. Zhang, Y. Zhang, X. Li and L. Zhang, *Bioorg. Med. Chem.* 2002, **10** (4), 907–912.

21   A. Bourderioux, P. Nauš, P. Perlíková, R. Pohl, I. Pichová, I. Votruba, P. Džubák, P. Konečný, M.Hajdúch, K. M. Stray,T. Wang,A. S. Ray, J. Y. Feng, G. Birkus, T. Cihlar and M. Hocek, *J. Med. Chem.* 2011, **54** (15), 5498–5507.

22   X.-F. Zhu, H. J. Williams and A. Ian Scott, *Synth. Commun.*, 2003, **33** (12), 2011–2016.

23   J. Zhu, X. Hu, E. Dizin and D. Pei, *J. Am. Chem. Soc.*, 2003, **125** (44), 13379–13381.

24   T. Bourdier, C. J. R. Fookes, T. Q. Pham, I. Gre-guric and A. Katsifis, *J. Labelled Compd. Rad.*, 2008, **51** (11), 369–373.

25   K. Devkota, M. Schapira, S. Perveen, A. K. Yazdi, F. Li, I. Chau, P. Ghiabi, T. Hajian, P. Loppnau, A. Bolotokova, K. J. F. Satchell, K. Wang, D. Li, J. Liu, D. Smil, M. Luo, J. Jin, P. V. Fish, P. J. Brown and M. Vedadi, *bioRxiv*, 2021.

26   A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, R. Lepore and T. Schwede, *Nucleic Acids Res.* 2018, **46** (W1), W296–W303.

27   MOPAC2012, James J. P. Stewart, Stewart Computational Chemistry, Version 7.263W web: HTTP://OpenMOPAC.net.

28   G. Jones, P. Willett, R. C. Glen, A. R. Leach and R. Taylor, *J. Mol. Biol.*, 1997, **267** (3), 727–748.

29   Schrodinger, LLC. 2010. The PyMOL Molecular Graphics System, Version 1.8.