Synthesis, biological evaluation and molecular modeling of a novel series of fused 1,2,3-triazoles as potential anti-coronavirus agents

Konstantina Karypidou\textsuperscript{a}, Sergio R. Ribone\textsuperscript{b}, Mario A. Quevedo\textsuperscript{b}, Leentje Persoons\textsuperscript{c}, Christophe Pannecouque\textsuperscript{c}, Christine Helsen\textsuperscript{d}, Frank Claessens\textsuperscript{d} and Wim Dehaen\textsuperscript{a}\ast

\textsuperscript{a} Molecular Design and Synthesis, Department of Chemistry, KU Leuven, Celestijnenlaan 200F, 3001, Leuven, Belgium
\textsuperscript{b} Unidad de Investigación y Desarrollo en Tecnología Farmacéutica (UNITEFA, CONICET), Dpto. Farmacia, Fac. Ciencias Químicas, Universidad Nacional de Córdoba, Córdoba, X5000HUA, Argentina
\textsuperscript{c} Department of Microbiology and Immunology, Laboratory of Virology and Chemotherapy, Rega Institute for Medical Research, KU Leuven, Herestraat 49, B-3000 Leuven, Belgium
\textsuperscript{d} Laboratory of Molecular Endocrinology, Department of Cellular and Molecular Medicine, KU Leuven, Herestraat 49, B-3000 Leuven, Belgium
\textsuperscript{*wim.dehaen@kuleuven.be , konstantina.karypidou@kuleuven.be,}
\textsuperscript{alfredog@fcq.unc.edu.ar, sribone@fcq.unc.edu.ar}
\textsuperscript{christophe.pannecouque@kuleuven.be, leentje.persoons@kuleuven.be}
\textsuperscript{frank.claessens@kuleuven.be, christine.helsen@kuleuven.be}

Abstract

Synthesis and biological evaluation of a novel library of fused 1,2,3-triazole derivatives are described. The in-house developed multicomponent reaction based on commercially available starting materials was applied and broad biological screening against various viruses was performed, showing promising antiviral properties for compounds \textsuperscript{14d, 14n, 14q, 18f} and \textsuperscript{18i} against human coronavirus 229E. Further in silico studies identified the key molecular interactions between those compounds and the 3-chymotrypsin-like protease, which is essential to the intracellular replication of the virus, supporting the hypothesis that the protease is the target molecule of the potential antiviral derivatives.

Keywords: Respiratory Syndrome, Coronavirus, 3CL protease, 1,2,3-triazole, biological evaluation

Coronaviruses are single-stranded RNA viruses associated with mild to severe respiratory symptoms. Human coronaviruses (HCoV) strains HCoV-229E and HCoV-OC43 were first described in the 1960s as causes for respiratory tract infections in humans, including common cold and pneumonia.\textsuperscript{1} In 2002-2003, a new human coronavirus, named SARS-CoV, was identified as the etiological agent for the global outbreak of severe acute respiratory syndrome (SARS), which caused the death of over 800 individuals among 8000 cases worldwide, representing a fatality rate of almost 10%.\textsuperscript{2,3,4} Since then three additional coronaviruses have been recognized. Initially, HCoV-NL63 \textsuperscript{5} and HCoV-HKU1,\textsuperscript{6} were reported causing acute respiratory diseases of lower severity compared to the SARS-CoV and more recently, Middle East respiratory syndrome (MERS-CoV) causing lethal respiratory diseases.\textsuperscript{7,8} To date, there are no approved antiviral drugs or vaccines available for the prevention and/or treatment of SARS-like viruses making the development of effective antiviral agents an imperative need.\textsuperscript{9}

Coronaviruses express two proteases, a papain-like protease (PL\textsubscript{pro}) and a 3-chymotrypsin-like protease (3CL\textsubscript{pro}). The 3CL\textsubscript{pro} enzyme, also referred to as Main protease (M\textsubscript{pro}), is essential to the intracellular viral replication, making it an attractive target for the development of novel inhibitors.\textsuperscript{10} Reports in the literature
classify potential antiviral compounds into two main categories: i) the peptidomimetics (Fig. 1A) and ii) small molecule-based inhibitors (Fig. 1B) presenting both activities in μM and nM range. Despite the satisfying initial results, the majority of those promising compounds did not proceed to clinical studies due to nonideal physicochemical properties.9

References on non-peptidic inhibitors accentuate the presence of the benzotriazole group. The importance of this benzotriazole motif relies on the formation of key interactions with the catalytic dyad, Cys145 and His41, of the 3CLpro active site.11 Considering our interest in the chemistry of 1,2,3-triazole bioactive molecules and their interesting binding mode, we wanted to prepare a novel library of fused 1,2,3-triazoles and subsequently determine their biological activity.

Here, we reported our preliminary results on the development of a novel series of fused 1,2,3-triazole compounds and their potential antiviral activity. For this purpose, we implemented the multicomponent reaction developed within our group, which resulted in a plethora of 1,2,3-triazole derivatives in a single step starting from readily available enolizable carbonyl compounds, primary amines and 4-nitrophenyl azide. In general, the method proceeds via an equilibrium of imine/enamine followed by [3+2] cycloaddition with the azide.12 This leads to a triazoline intermediate which after elimination of 4-nitroaniline results in the final fused 1,2,3-triazole analogues.

**Figure 1.** Representative A) peptidic13 and B) small molecule SARS-CoV 3CLpro inhibitors14-19

The synthesis, presented in Scheme 1, began by following a general method to generate the oxopiperidine carboxylate intermediate which involves Michael addition of aniline onto ethyl acrylate followed
by an intramolecular Dieckmann condensation. Subsequent nucleophilic substitution with benzyl bromide provided the starting material 11 in 65% overall yield. Once 11 was obtained, we proceeded for the fused triazole formation using a collection of primary amines and 4-nitrophenyl azide (PNA, 13).

Scheme 1. Synthetic pathway towards derivatives 14a-r. Reagents and conditions: (a) AcOH, CuCl, 24h, 110 °C, 55%, (b) NaH, Toluene, EtOH, 6h, 100 °C, 90%, (c) BnBr, K₂CO₃, THF, 6h, 70 °C, 65%, (d) Toluene, 18h, 100 °C.

We commenced our investigations with benzylic amines bearing both electron-donating functional groups, (Table 1, 14b-c) and electron-withdrawing groups, (Table 1, 14d-j). Both families were obtained in moderate to good yields.

Table 1. Molecular structures, yield and biological activity against coronavirus (229E) of derivatives 14a-r.
Concentration required to reduce virus-induced cytopathogenicity by 50% further analysis using heterocyclic and aliphatic derivatives 14k-p proved the applicability of these substituents under the high temperature conditions used in this reaction (Table 1). Furthermore, we examined the influence of the introduction of a fluoro group on the benzyl bromide on the reactivity (Table 1, entries 7 and 8), which demonstrated that there is no effect on the outcome of the reaction.

| 14a | 14b | 14c | 14d | 14e | 14f | 14g | 14h | 14i | 14j | 14k | 14l | 14m | 14n | 14o | 14p | 14q | 14r |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ![Chemical Structure](image1.png) | ![Chemical Structure](image2.png) | ![Chemical Structure](image3.png) | ![Chemical Structure](image4.png) | ![Chemical Structure](image5.png) | ![Chemical Structure](image6.png) | ![Chemical Structure](image7.png) | ![Chemical Structure](image8.png) | ![Chemical Structure](image9.png) | ![Chemical Structure](image10.png) | ![Chemical Structure](image11.png) | ![Chemical Structure](image12.png) | ![Chemical Structure](image13.png) | ![Chemical Structure](image14.png) | ![Chemical Structure](image15.png) | ![Chemical Structure](image16.png) | ![Chemical Structure](image17.png) | ![Chemical Structure](image18.png) |
| 80% | 85% | 65% | 80% | 58% | 80% | 67% | 48% | 30% | 73% | 55% | 56% | 85% | 80% | 57% | 51% | 50% | 60% |

*Concentration required to reduce virus-induced cytopathogenicity by 50%*
A second series of fused 1,2,3-triazole analogues was prepared based on the previously described multicomponent method, starting from N-phenyl-4-piperidone material 17 which in turn was synthesized according to Buchwald-Hartwig Pd-catalyzed amination of different aryl bromides, presented in Scheme 2. We initially examined the scope of this reaction with respect to primary amines. As we would expect based on the aforementioned results, we managed to synthesize an adequate number of examples in good to excellent yields (18a-d). Furthermore, the presence of electron withdrawing substituents on the aryl bromides was explored (18e-l). The results are summarized in Table 2.

Scheme 2. Reagents and conditions: (a) Pd(OAc)$_2$, XPhos, NaO-t-Bu, Toluene / t-BuOH (5:1), 18 h, 120 °C, 65%, (b) aq. HCl (5N), 3 h, 100 °C, 45%, (c) Toluene, 18 h, 100 °C.

Table 2. Molecular structures, yield and biological activity against coronavirus (229E) of derivatives 18a-l.
The compounds were evaluated against a broad variety of viruses including HIV-1 (strain IIIb), HIV-2 (strain ROD) in MT-4 cells, herpes simplex virus type 1 (strain KOS), herpes simplex virus type 2 (strain G), herpes simplex virus type 1 TK- (KOS) ACV res, vaccinia virus, adeno virus-2 and coronavirus (229E) in HEL cells and their inhibitory activity was compared to that of reference compounds as zidovudine, brivudine, cidofovir, acyclovir, ganciclovir, zalcitabine, alovudine and Urtica dioica agglutinin (UDA), respectively. All compounds were inactive towards all tested viruses, while 14d (EC50 = 8.95 µM), 14n (EC50 = 9.45 µM), 14q (EC50 = 9.45 µM), 18f (EC50 = 8.90 µM) and 18i (EC50 = 11.95 µM) showed moderate activity against human coronavirus (229E), but all approximately 50 fold lower than the activity observed with UDA (EC50 = 0.2 µM). No alterations of the normal cell morphology in confluent HEL cell

<table>
<thead>
<tr>
<th>Entry</th>
<th>Compound ID</th>
<th>R</th>
<th>R'</th>
<th>Yield</th>
<th>EC50 (µM)⁹</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18a</td>
<td>MeO</td>
<td>H</td>
<td>85%</td>
<td>&gt;100</td>
</tr>
<tr>
<td>2</td>
<td>18b</td>
<td>OMe</td>
<td></td>
<td>74%</td>
<td>&gt;100</td>
</tr>
<tr>
<td>3</td>
<td>18c</td>
<td></td>
<td></td>
<td>90%</td>
<td>&gt;100</td>
</tr>
<tr>
<td>4</td>
<td>18d</td>
<td></td>
<td></td>
<td>48%</td>
<td>&gt;100</td>
</tr>
<tr>
<td>5</td>
<td>18e</td>
<td></td>
<td>-CF₃</td>
<td>90%</td>
<td>&gt;100</td>
</tr>
<tr>
<td>6</td>
<td>18f</td>
<td></td>
<td></td>
<td>85%</td>
<td>8.9</td>
</tr>
<tr>
<td>7</td>
<td>18g</td>
<td></td>
<td></td>
<td>85%</td>
<td>&gt;100</td>
</tr>
<tr>
<td>8</td>
<td>18h</td>
<td></td>
<td></td>
<td>89%</td>
<td>&gt;100</td>
</tr>
<tr>
<td>9</td>
<td>18i</td>
<td></td>
<td></td>
<td>85%</td>
<td>11.95</td>
</tr>
<tr>
<td>10</td>
<td>18j</td>
<td></td>
<td></td>
<td>89%</td>
<td>&gt;100</td>
</tr>
<tr>
<td>11</td>
<td>18k</td>
<td></td>
<td>-F</td>
<td>91%</td>
<td>&gt;100</td>
</tr>
<tr>
<td>12</td>
<td>18l</td>
<td></td>
<td></td>
<td>57%</td>
<td>&gt;100</td>
</tr>
</tbody>
</table>

⁹Concentration required to reduce virus-induced cytopathogenicity by 50%.
cultures was observed at concentrations up to 100 µM (data not shown). The selectivity index (SI) (MCC/EC$_{50}$ ratio) was greater than 8 for all active compounds.

To determine the structure-activity relationship (SAR) for the synthesized library of fused 1,2,3-triazoles, it is of utmost importance to identify the key intermolecular contacts involved in the non-covalent interaction between 3CL$_{pro}$ and the structurally diverse inhibitors. With this aim, several molecular modeling techniques, including molecular docking, molecular dynamics, free energy of binding analyses and intermolecular interaction scanning, were applied in the search of such SAR knowledge. In a first stage, previously reported 3CL$_{pro}$ non-covalent inhibitors were modeled in order to define the key intermolecular interactions required for enzyme inhibition, while in a second stage the in silico analysis was extended to the family of fused 1,2,3-triazoles presented in this report.

The catalytic activity of 3CL$_{pro}$ has been well characterized, with extensive details regarding the structure of this enzyme and the corresponding catalytic site. In this respect, it is known that the active site is located within domains I and II, in which a catalytic dyad consisting of residues Cys145 and His41 is located.$^{11}$ It has also been previously reported that the catalytic site of 3CL$_{pro}$ exhibits a stereoselective recognition of non-covalent inhibitors.$^{18}$ In this context, our molecular modeling protocols initiated with the exploration of the binding mode, intermolecular interaction pattern and stereoselectivity of the non-covalent inhibitor of 3CL$_{pro}$ deposited in the Protein Databank under the code 3V3M. Both enantiomers of the bound ligand were docked within the catalytic site, and as can be seen in Fig S1.a,b, the $R$ enantiomer resulted in an identical interaction pattern to that observed in the experimentally obtained crystal. The ligand binding is stabilized by several hydrophobic and hydrogen bond (HB) interaction, of which those with Glu166, His163 and Gly143 are of particular relevance. Noteworthy, the $S$ enantiomer was also able to establish several hydrophobic interactions, but was not able to establish HB with Glu166 and His163. Taking into account that it has been previously reported that only the $R$ enantiomer is active,$^{18}$ this finding suggests that the establishment of interactions with Glu166 and His163 constitutes a critical feature to inhibit the catalytic activity of the enzyme. To further study the persistence of these HB interactions, the intermolecular complexes were subjected to molecular dynamics (MD) analysis, with Table S1 presenting the persistence value (%) for each HB as calculated from the MD trajectories. As can be seen in Table S1,entry 1, the $R$ enantiomer was able to maintain in relatively high frequencies the HB interactions with Glu166, His163 and Gly143, while the $S$ enantiomer rearranged its binding to contact only Glu166 (60 % persistence). Our findings not only validate the molecular modeling workflow we developed (i.e. the crystallographic binding pose was reproduced), but also strongly suggests that at least two stable electrostatic interactions within the active site are required for effective 3CL$_{pro}$ inhibition.

To further study this hypothesis, a set of 12 previously reported 3CL$_{pro}$ inhibitors containing a fused 1,2,3 triazole ring and exhibiting a wide range of inhibitory activities (i.e. between 51-26000 nM, Tables S3S5), was subjected to our molecular modeling workflow. The lowest energy binding mode obtained for TS8 was in agreement with the reported crystallographic structure (Fig. S3.d and pdb code 4MDS, respectively), further supporting an adequate parametrization and simulation conditions of the molecular modeling protocol. From inspection of the corresponding lowest energy docked poses for the whole set of compounds (Figures S2-S4), we observed that all these inhibitors were indeed able to establish the two HB previously described for 3V3M, i.e. one with the backbone of Glu166 and another one with the side
chain of His163. In these binding modes, the triazole ring is positioned in the proximity of the catalytic dyad. From the MD simulations and the quantification of the persistence of the above mentioned hydrogen bond interactions (Table S1, entries 3-14), we observed that all the compounds maintained the hydrogen bonds required for the inhibition of 3CL\textsuperscript{pro} in high frequencies, and in particular that with Glu166. It is noteworthy that inhibitors that are enantiomerically pure and exhibiting submicromolar activities maintained high frequencies of HB with Glu166 and His163. In particular, TS-1, which is by far the most potent compound within the training set, exhibited also a single cluster of docked poses, suggesting not only an efficient pharmacodynamic interaction with 3CL\textsuperscript{pro}, but also an adequate conformational preorganization that is compliant with the corresponding bioactive conformation.

Compounds 14a-r and 18a-l were subjected to the molecular modeling workflows in order to study whether 3CL\textsuperscript{pro} may represent a plausible molecular target for their observed antiviral activity. Figures S5-S9 shows the lowest energy binding modes to 3CL\textsuperscript{pro} found for compounds 14a-r. When the docked poses corresponding to the bioactive derivatives are observed (i.e. 14d and 14n, Figures S5.d, S8.b, respectively), it can be seen that they establish the two HB interactions with Glu166 and His163, positioning also the triazole ring in the proximities of the catalytic dyad. As it was discussed in previous sections, this interaction pattern is required for the inhibition of 3CL\textsuperscript{pro} catalytic activity, suggesting that this enzyme may be the antiviral target of these fused 1,2,3-triazoles. Further analysis by MD also showed that these interactions are maintained throughout the simulation, further supporting their bioactivities (Table S2, entries 4, 14 and 17). In contrast, the fused triazole derivatives that did not exhibited antiviral activities in infected cells, did not establish these two HB interactions in the lowest energy binding pose, or were not able to form them during the simulation trajectory when subjected to MD assays. Compounds 14c and 14m constitute two exceptions to this behavior, both of them bearing a methoxy substituent on the para position of the phenyl ring substituting the triazole central scaffold. These two compounds were able to establish and maintain the HB interaction pattern required for 3CL\textsuperscript{pro} inhibition, (Table S2, entries 3 and 13) but did not exhibit antiviral activity. This fact may be due to a disfavorable entropic contribution upon binding to the 3CL\textsuperscript{pro} catalytic site, since in order to maintain the interaction with His163, the rotation of the methoxy group is constrained within a dihedral angle of 10 Å. Clearly, this entropic cost is not present for the bioactive analogue bearing a fluorine atom in the para position (14d).

When derivatives 18a-l were analyzed (Figures S10-S12), a similar behavior was observed, with the derivatives exhibiting antiviral activity (i.e. 18f and 18i) being able to establish two stable HB interactions within the catalytic site of 3CL\textsuperscript{pro} (Figures S11.b and S12.a). In particular, derivative 18f is anchored within the catalytic site through a stable interaction with His163, with two additional interactions being established with Glu166 and Thr25. Regarding the interaction of 18i, this compound establishes stable HB contacts with both Glu166 and His163, while further anchoring to residue Gln189. As a result, both compounds positioned the fused triazole scaffold in the vicinity of the catalytic dyad which is consistent with blocking the protease activity of the enzyme. Finally, when the interaction patterns of the inactive compounds within the series 18a-l were analyzed, we found that all of them failed to establish the two required HB interactions within the catalytic site of the enzyme.

In conclusion, we have succeeded in synthesizing a novel library of fused 1,2,3-triazoles using the in-house developed multicomponent reaction. The library was subjected to in vitro analysis using a broad
variety of viruses to determine their antiviral properties and to in silico studies to determine the interactions with 3CL\textsuperscript{pro}. Compounds 14d, 14n, 14q, 18f and 18i showed moderate activity against coronavirus 229E. A molecular modeling work flow was developed based on previously reported 3CL\textsuperscript{pro} non-covalent inhibitors and helped the identification of key molecular interactions. Application of this model to our library, supports that the antiviral activity is mediated through the inhibition of 3CL\textsuperscript{pro}. Additional studies on the structure-activity relationship will enable us to prepare new fused 1,2,3-triazole derivatives with enhanced antiviral properties.

Acknowledgments

This work was supported by Katholieke Universiteit Leuven (KU Leuven), grants C32/15/033 and ISPLA2/15/03. In addition, the authors gratefully acknowledge financial support from the Secretaria de Ciencia y Técnica of the Universidad Nacional de Córdoba (SECYT-UNC), the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), and the Agencia Nacional de Promoción Científica y Técnica (ANPCyT). The authors would also like to thank the GPGPU Computing Group from the Facultad de Matemática, Astronomía y Física (FAMAF), Universidad Nacional de Córdoba, Argentina, for providing access to computing resources. The authors also gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research. Mario A. Quevedo wishes to thank OpenEye Scientific Software and their Free Academic Licensing program for providing him with licenses to use the corresponding software packages.

References


