SARS-CoV-2 main protease design themes based on 6y2f X-ray crystal structure

Peter W. Kenny (email: pwk.pub.2008@gmail.com)

Keywords: SARS-CoV-2, COVID-19, coronavirus, 3CLpro, main protease, Mpro, design theme, inhibitor design, covalent inhibitor, warhead, structure-based design,

Overview of X-ray crystal structure

Some of the design implications of the recently published (PDB:6y2f) X-ray crystal structure of SARS-CoV-2 $3CL^{pro}$ (main protease) in complex with the α-ketoamide inhibitor, **1** (pIC_{50} = 6.2) have been discussed previously. The deposited structure has been further refined at Global Phasing and a number of design themes based on this structure are discussed in the current article. The inhibitors **2** (pIC_{50} = 7.3) and **3** (pIC_{50} = 7.4) have also been reported and each incorporates an aldehyde warhead. Medicinal chemistry approaches against SARS-CoV, MERS-CoV and COVID-19 targets were covered in a recent review and a number of relevant main protease inhibitors were discussed.

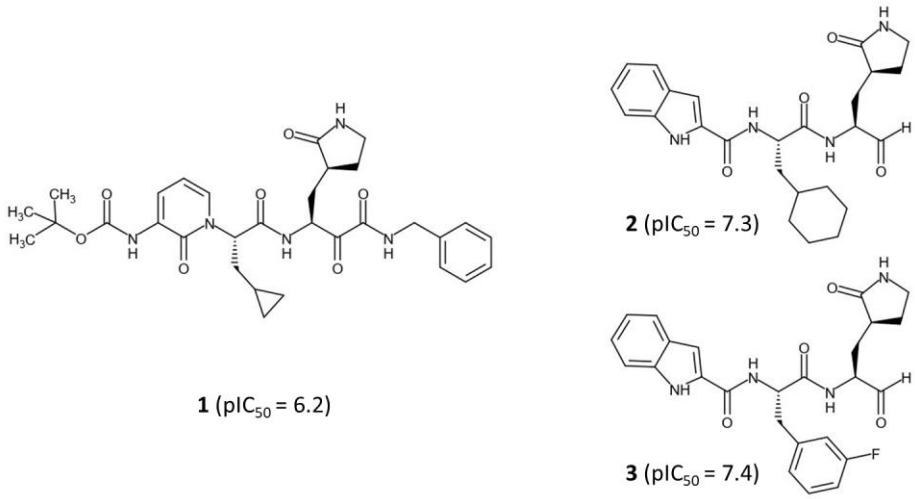


Chart 1. Inhibitors of SARS-CoV-2 main protease; $pIC_{50} = -\log_{10}(IC_{50}/M)$

Drug design is frequently seen by computational scientists and informaticians as an exercise in prediction. However, it may be more appropriate to see drug design as a process of generating information as efficiently as possible (i.e. design of experiments) and it can be helpful to think in terms of design themes when optimizing from chemical starting points. Design themes can be used both to map structure-activity relationships and to provide a framework for prediction-driven design. The design themes presented in this article are intended to illustrate potential directions for design rather than as recommendations that specific compounds be synthesized.

Simple modelling of molecular geometry (e.g. to assess likelihood that a specific molecular interaction forms) can be useful when exploring design themes and the geometric constraints imposed by a covalent bond between protein and ligand increase confidence in predictions. In the early stages of projects, there is not typically enough data for using machine learning approaches for modelling biological activity and, in any case, these approaches do not currently appear well-suited for the prediction of activity differences between structurally-related compounds. Methods for calculating relative free energy of binding fit naturally into a lead

optimization framework although these approaches do not appear to be used routinely for design of reversible, covalent inhibitors of cysteine proteases.

The X-ray crystal structure of the SARS-CoV-2 main protease in complex with **1** is shown in Figure 1. A molecule of dimethylsulfoxide (DMSO) was also observed to bind.

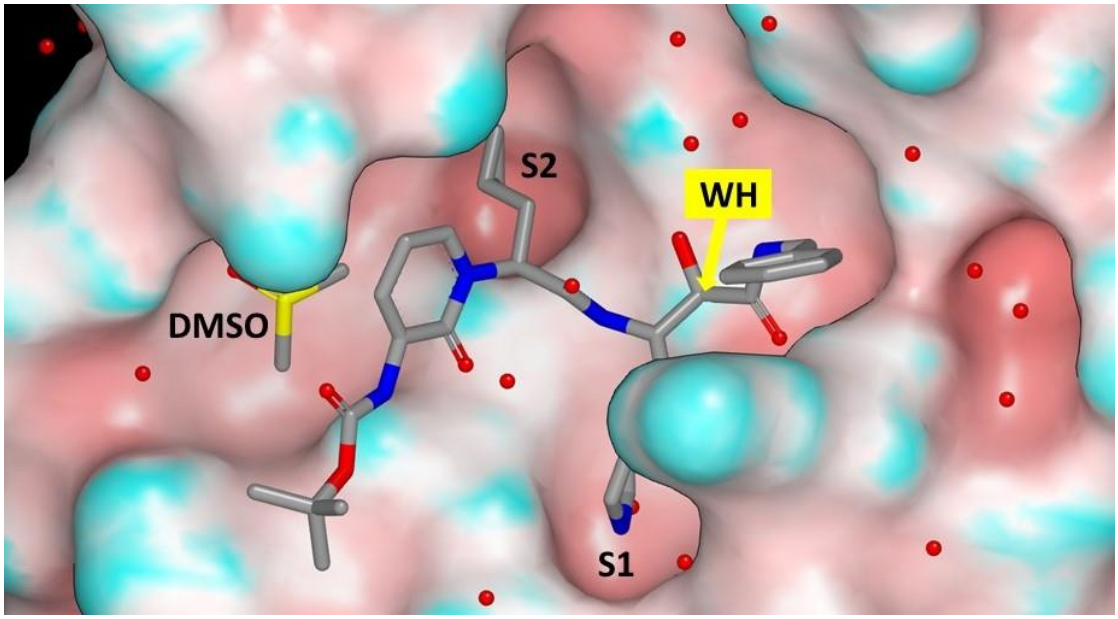


Figure 1. Ligands and water molecules of PDB:6y2f crystal structure (this has been further refined at Global Phasing) shown with molecular surface of protein coloured by curvature (created using VIDA from OpenEye). Labels refer to subsites (S1 and S2), warhead (WH) and dimethylsulfoxide (DMSO) have been labelled.

The warhead

A common tactic in design of cysteine protease inhibitors is to incorporate an electrophilic moiety (warhead) that can form a covalent bond with the catalytic cysteine. Reversibility has important implications for design and the nature of the warhead typically determines whether a covalent inhibitor is reversible or irreversible. This comparative study of warheads for design of cysteine protease inhibitors may be relevant.

The warhead in **1** is an α-ketoamide and the benzyl group on the amide nitrogen would be considered by some to be a potential pharmacokinetic liability. One way forward would be to map the structure-activity relationship (**5**, **6**, **7**) to determine the importance of the benzyl group for activity. Given that **2** and **3** both incorporate the aldehyde warhead, inclusion of **4** in the initial mapping of the structure-activity relationship would also be informative. If the benzyl group proves to be important for activity then aza-substitution (**8**, **9**) may address the potential pharmacokinetic liability. Mapping of the structure-activity relationship could be extended to 1,2-diketones (**10**) and even thiones (**11**) which might be expected to be more electrophilic than the corresponding ketones.

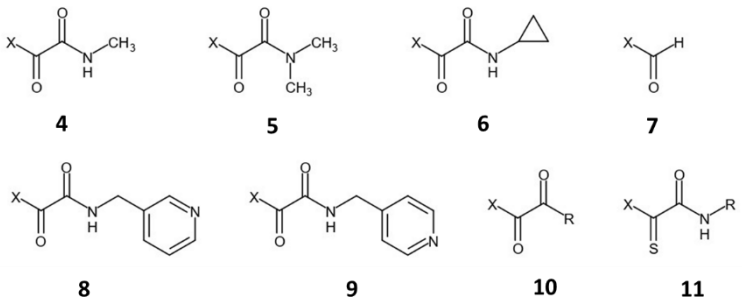


Chart 2. Carbonyl-based warhead design themes

The nitrile warhead that features in the molecular structure of odanacatib has not been included in this design theme. Nitriles are typically less potent than aldehydes as inhibitors of cysteine proteases by about two orders of magnitude and it is important to first assess aldehydes versus 1,2-dicarbonyl compounds before considering nitriles. Azapeptide nitriles are typically more potent cysteine protease inhibitors than the corresponding peptide aldehydes. The potency differences between azapeptide nitriles and the corresponding peptide nitriles may reflect differences in pK_a between the respective covalently-bound adducts. A P1 substituent could be linked to an aza peptide nitrile through the nitrogen that bears the nitrile.

The P1 substituent

A pyrrolidinone P1 substituent (linked at C3) is a structural feature that is shared by the inhibitors **1**, **2** and **3**. While the carbonyl oxygen of the pyrrolidinone accepts a hydrogen bond from the sidechain of H163, the amidic hydrogen of the pyrrolidinone does not function as a hydrogen bond donor, despite being in contact with the molecular surface of the protein. This suggests it may be beneficial to present a tertiary amide carbonyl oxygen to H163 since this will not incur the penalty associated with desolvation of a secondary amide hydrogen bond donor. However, it should be noted that most of the energetic cost of transferring a secondary amide from water to saturated hydrocarbon appears to result from desolvation of the amide carbonyl oxygen rather than the amide NH (see this article).

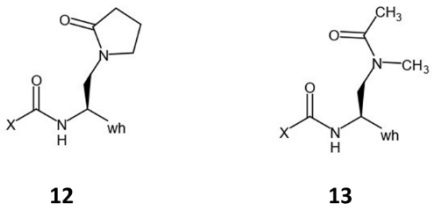


Chart 3. P1 substituent design themes

The P2 substituent

The binding mode of **1** is shown from the perspective of the bound DMSO in Figure 2. The pocket occupied by the DMSO molecule is partly formed from the pyridone ring and its substituent. In the absence of the DMSO, it would be anticipated that transfer of water from bulk phase to the pocket will lead to an increase in free energy. Energetically unfavourable enclosure of water would also be expected to adversely affect the binding affinity of **1**. DMSO is a particularly strong hydrogen bond acceptor (pK_{BHX} = 2.54; see data and article) and the oxygen atom of DMSO accepts hydrogen bonds from both the backbone amide of T190 and sidechain amide of Q192.

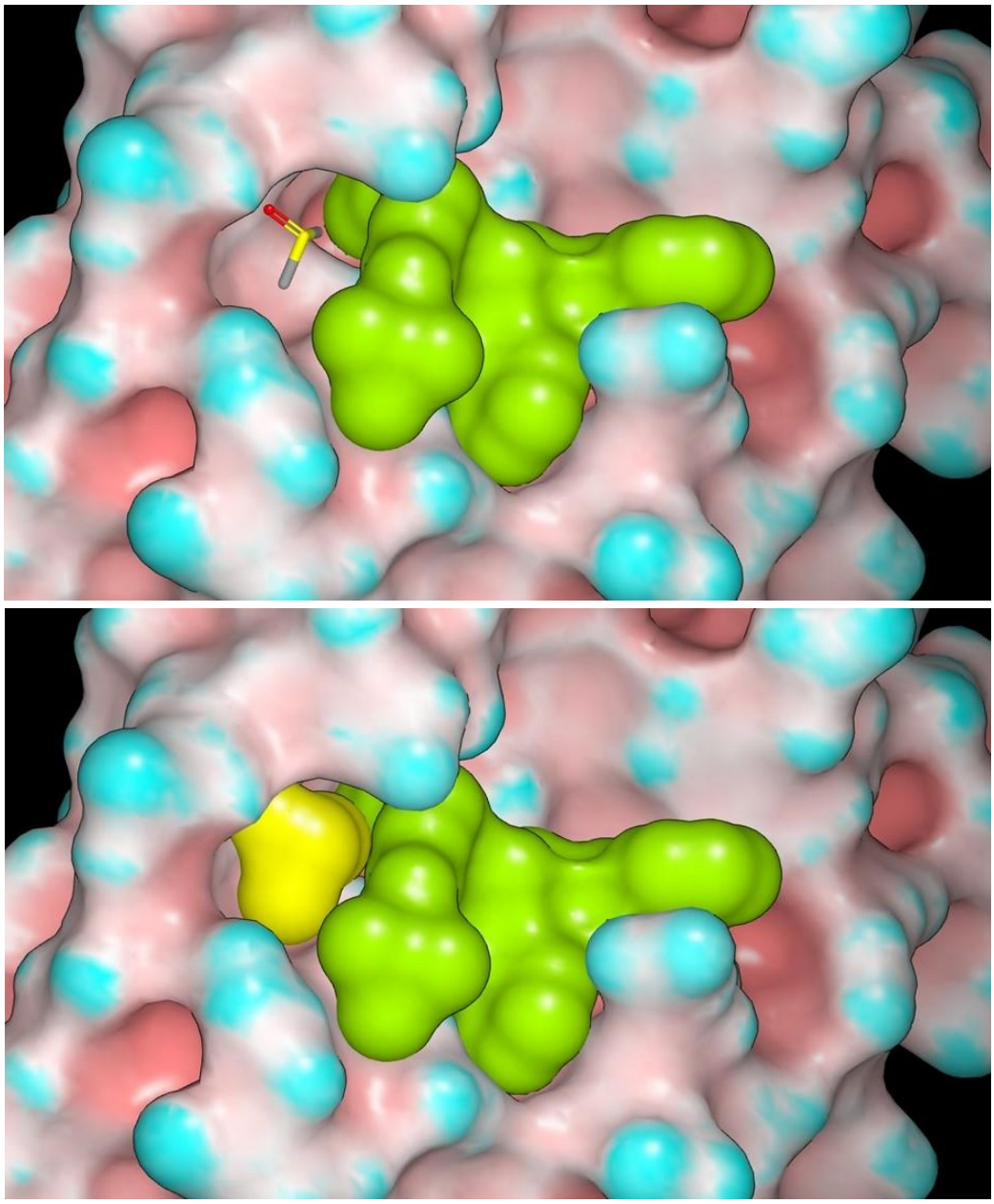
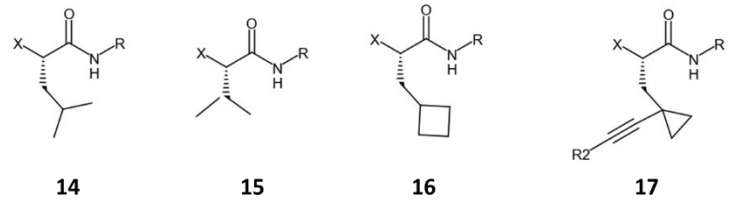


Figure 2. Compound **1** and DMSO in PDB:6y2f crystal structure (this has been further refined at Global Phasing) shown with molecular surface of protein coloured by curvature (created using VIDA from OpenEye).

Design themes are presented in Chart 4 and **14** to **16** map the structure-activity relationship for the P2 substituent. The region occupied by the DMSO molecule might be accessed using an alkyne spacer (**17**) which is relatively undemanding from the steric perspective and is equivalent to a long single bond without torsional bias. One design theme would be to link a substituent capable of mimicking the hydrogen bond acceptor of the DMSO molecule. This article illustrates potential benefits of exploiting hydrogen bond donors within the S2 subsites of cysteine proteases (especially cathepsin S).

As argued in this article, it is desirable to investigate whether the pyridone ring in **1** can be replaced with piperidinone or pyrrolidinone since being able to do so creates design opportunities. For example, the region occupied by the DMSO molecule would be more easily accessed by a substituent linked to piperidinone or pyrrolidinone than by a substituent linked

to pyridone. The P2 substituent could also be linked to piperidinone or pyrrolinone to form a macrocycle.



14 **15** **16** **17**

Chart 4. P2 substituent design themes