Using serpins cysteine protease cross-specificity to possibly trap SARS-CoV-2 Mpro with reactive center loop chimera

Mohamad Aman Jairajpuri and Shoyab Ansari
Department of Biosciences, Jamia Millia Islamia, New Delhi 110025, India
Correspondence: Mohamad Aman Jairajpuri (mjairajpuri@jmi.ac.in)

Human serine protease inhibitors (serpins) are the main inhibitors of serine proteases, but some of them also have the capability to effectively inhibit cysteine proteases. Severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2) main protease (Mpro) is a chymotrypsin-type cysteine protease that is needed to produce functional proteins essential for virus replication and transcription. Serpin traps its target proteases by presenting a reactive center loop (RCL) as protease-specific cleavage site, resulting in protease inactivation. Mpro target sites with its active site serine and other flanking residues can possibly interact with serpins. Alternatively, RCL cleavage site of serpins with known evidence of inhibition of cysteine proteases can be replaced by Mpro target site to make chimeric proteins. Purified chimeric serpin can possibly inhibit Mpro that can be assessed indirectly by observing the decrease in ability of Mpro to cleave its chromogenic substrate. Chimeric serpins with best interaction and active site binding and with ability to form 1:1 serpin–Mpro complex in human plasma can be assessed by using SDS/PAGE and Western blot analysis with serpin antibody. Trapping SARS-CoV-2 Mpro cysteine protease using cross-class serpin cysteine protease inhibition activity is a novel idea with significant therapeutic potential.

Introduction
The outbreak of coronavirus disease (Covid-19) caused by severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2) is now a global pandemic with rapidly expanding mortality rate and with no apparent cure in sight. SARS-CoV-2 main protease (Mpro) is absolutely critical for producing functional proteins involved in transcription and replication of viral coded gene [1]. Mpro is a chymotrypsin-type cysteine protease that cleaves a long peptide coded by viral RNA to give functional protein, and has emerged as one of the main therapeutic target [2]. Serine protease inhibitors (serpins) are the main class of inhibitors of serine proteases in humans that are involved in coagulation, complementation, fibrinolytic, inflammatory and angiogenic pathways [3]. A number of serpins are also potent cross-class inhibitors of cysteine proteases [4]. Serpins follow suicide-substrate inhibition mechanism by presenting its solvent exposed reactive center loop (RCL) with cleavage site (P1–P1’) and forms 1:1 complex with protease leading to its inactivation [5]. Serpin-like antitrypsin has been switched to inhibit cysteine protease (papain) instead of its natural serine protease substrate (elastase) by replacing the RCL residues [6]. Recently solved crystal of Mpro (PDB ID: 6LU7) gives indication of a protease that has an active site cleft and target site that can be trapped through serpins [7]. Frantic efforts are on in several countries to counter the virus by using already available antimalarial, antibacterial and antiviral drugs or by designing vaccine and also using plasma from cured patients [8–10]. Novel approaches will be needed to counter this virus and an ever growing threat of new viruses with extremely high infectivity and virulence.
Mpro structural features and its cysteine protease activity

SARS-CoV-2 viral gene encodes for 29 proteins, its RNA is translated into a long peptide that is cleaved to give functional proteins important for its pathogenicity [1]. Mpro digests 11 conserved sites (Q-S/A/G-GFRK) on the long polypeptide and has emerged as an attractive drug target [2,11,12]. Several synthetic and natural compounds as leads have been predicted to interact with Mpro but functional validation and in vivo testing is still in its infancy, although some repurposing drugs show attractive binding to its active site [13,14]. Mpro from SARS-CoV-2 is a 306 amino acid protein that dimerises into two identical subunits to give a functional protein. The recently solved crystal structure gave a good idea about its active site that forms between two distinct domain termed domain I (1–99 residues) and domain II (100–182 amino acids) [7]. These domains form two \( \beta \) barrel folds that is linked to domain III (199–306) through a short loop (186–190 amino acids), and forms five \( \alpha \) helices (Figure 1A). Recent evidences points that amino acids 183–185 are part of the hinge region that connects mobile and rigid regions with implications in controlling regulatory movements which may affect enzymatic activity [15]. Residues from 190 to 198 have also been shown previously to help in substrate binding with implications in cleavage of C-terminal auto-processing site [16]. Mpro substrate binding site is located in a deep cleft between domain I and domain II, that has at its center H41 and Cys145 that are capable of cleaving multiple sites on the substrate. Domain I and domain II contributes to the substrate binding site (H41, M49, G143, S144 and residues 163–167), but all the three domains are involved in dimerization [17]. Viral genome (\( \sim 30 \) kb) codes for the two overlapping polyprotein that are cleaved by Mpro to give a functional replication product. N-terminal residue (1–7) plays an important role in dimerization and catalytic

Figure 1. Structural representation of serpin and Mpro

(A) The crystal structure of SARS-CoV-2 Mpro (PDB ID: 6LU7). The active site of Mpro is between two distinct domains termed domain I and domain II that forms two \( \beta \)-barrel folds that is linked to domain III that forms five \( \alpha \) helices. Substrate binding site is located in deep cleft between domain I and domain II and is capable of cleaving multiple sites on the target polyprotein encoded by viral RNA. (B) Serpin conformation: native (PDB ID: 3FGQ) serpin structure showing RCL and strand 3A and 5A of \( \beta \)-sheet A, where RCL inserts as strand 4A after protease cleavage inactivating it in the process. Picture was created by using PyMOL molecular graphic system.
Table 1 Representing the serpins with cross-class cysteine protease activity along with their serine protease targets

| S. No. | Serpin with cysteine protease activity | Serine protease target | Cysteine protease target | Serpin RCL (P1–P5′) replaced with Mpro target sites* |
|-------|--------------------------------------|------------------------|--------------------------|-----------------------------------------------------|
| 1.    | Antichymotrypsin                     | Chymase                | Staphopain               | Q-S                                                 |
| 2.    | Antithrombin                         | Factor Xa, thrombin    | Papain                   | Q-S-G-F-R                                           |
| 3.    | Antitrypsin                          | Elastase               | Cathepsin L/K/V          | Q-S-G-F-R-K                                         |
| 4.    | Heparin cofactor II                  | Thrombin               | Cathepsin G              | Q-A                                                 |
| 5.    | Protease Inhibitor 9                 | Granzyme B             | Caspase 1                | Q-A-G-F-R                                           |
| 6.    | Serpin 18                            | ——                     | Fibroinase               | Q-A-G-F-R-K                                         |
| 7.    | Serpin CrmA                          | ——                     | Caspase 1                | Q-G                                                 |
| 8.    | Serpin B1                            | Elastase               | Cathepsin G              | Q-G-G                                               |
| 9.    | SCCA1                                | ——                     | Cathepsin K/L            | Q-G-G-F                                             |
| 10.   | Neuroserpin                          | Tissue plasminogen activator | Control                  | Q-G-G-F-R-K                                         |

*Target sequences that can be incorporated into each serpin RCL.

activity and is auto-cleaved [18]. Mpro role in the maturation of replicase (79 kDa) indicates that inhibiting its activity will directly block viral replication [19,20].

Serpin structural features and its cross-class cysteine protease inhibition activity

Serpins play an important role in a number of physiological processes in humans like coagulation, complementation, fibrinolytic, angiogenesis and tumor metastasis, and are associated with diseases like thrombosis, emphysema, cirrhosis, dementia, epilepsy and angioedema [21]. Serpins usually contains ~400 amino acids and consists of eight to nine helices, three β-sheets and RCL (Figure 1B). Serpin RCL acts as a bait (P1–P1′) that is cleaved on binding to target protease forming a covalent complex, that is transported to the opposite side were RCL inserts in β sheet A as strand 4A inactivating the protease [3,5,22]. Most of the members of this family inhibits serine proteases, although several are known to inhibit cysteine proteinases, while some are non-inhibitory [23,24]. Several serpins shown in Table 1 inhibit both serine and cysteine protease, while some like viral serpin cytokine response modifier protein A (CrmA) and serpin inhibitor 9 can inhibit caspases, others like serpin squamous cell carcinoma antigen (SCCA1) inhibits Cathepsin K and L [23,25–27]. Serpins with cysteine protease activity have serine or cysteine at their target site in the RCL (Figure 2A). The inhibition and recognition mechanism of cysteine proteases is similar to serine proteases although the cleaved forms are more evident as compared with the complex [6]. Antitrypsin RCL replaced by furin target sites switched its inhibition to furin and was not able to inhibit elastase [28]. Similar serpin chimeras with shift in protease specificity has been designed between antitrypsin, antithrombin and antichymotrypsin [29,30]. Antitrypsin chimera replaced with SCCA-1 RCL, converted it into an inhibitor of cysteine protease papain and not its natural serine protease target elastase [6]. In addition to the RCL recognition, serpins also possesses auxiliary residues termed as exosite residues that are essential in protease recognition and inhibition [31]. RCL of serpins with cross-class cysteine protease activity when replaced by Mpro target residues is expected to have exosite specificity to interact and trap Mpro reducing its active presence in producing viral-coded functional proteins.

Mpro direct inhibition of serpin as a basis of increased blood clots in Covid-19 patients

A number of recent reports indicate that Covid-19 patient show significantly increased clotting, with parameters that indicate a procoagulant environment, that directly causes thrombosis [32]. A study of 5700 hospitalized patients in New York at advance stage of disease, showed that 88% of deceased patients did not need ventilators as breathing aid indicating a non-respiratory cause of mortality [33]. Autopsies of 150 patients in a French study showed universal distribution of blood clots [34]. Consequently, with no sign of bleeding, disseminated intravascular coagulation
Figure 2. Sequence alignment of serpin RCL and Mpro target site

(A) Multiple sequence alignment of the RCL regions (yellow) of different serpins. Conserved residue with >70% homology of serpins with cysteine protease inhibition activity is highlighted with red letter, P1 and P1' sites of the RCL region are indicated by green color. The amino acid sequence alignment of the different serpin RCL was performed using the Clustal Omega program. The figure was produced using ESPript3.0. (B) Overlap of the Mpro target site was created manually and labeled using MS powerpoint. Predicted SARS CoV-2 Mpro cleavage site: 11 different sites of the replicase polyprotein cleaved by Mpro and the recognition sequence at most sites was found to be Gln↓(Ser/Ala/Gly) (↓ shows the cleavage site).

(DIC) was also ruled out as the fibrinogen levels were normal [32]. Several reports of Covid-19 patient showing clot formation while drawing of blood and the catheters for kidney dialysis becomes clogged with blood indicated a hypercoagulated phenotype, independent of hypoxia or inflammation based secondary response [35–37]. Indeed a five- to six-times increased d-Dimer levels and a significantly increased coagulation rates as indicated by APTT and PT assays points to a direct role on the coagulation cascade [38,39]. Low antithrombin levels detected consistently (~80%) in Covid-19 patients points to a direct inactivation [37]. One of the Mpro cleavage site is Gln–Ser, and with Ser also at the P1' site of cysteine protease recognizing serpins, it is possible that Mpro may cleave serine protease (Figure 2B). P2 and P3' are shown to be essential for Mpro recognition [11], matching with the predominantly hydrophobic residues at P2 and charged residue at P3' of serpins with cysteine protease inhibition activity (Figure 2A). Several serpins like antithrombin, heparin cofactor II, plasminogen activator inhibitors 1, protein C inhibitor and antiplasmin are part of the anticoagulant control of coagulation proteases. Glycosaminoglycans plays significant role
Serpins cysteine protease inhibition activity can be used to recognize Mpro that is a cysteine protease. However, this indirect inhibition of Mpro can be achieved if we progressively replace the serpin RCL residues with the Mpro protein target site. It is expected that this chimeric serpin will form Mpro–chimeric serpin complex thereby reducing the active concentration of Mpro and the rate of cleavage of SARS-CoV-2 polypeptide. Additionally, it is hypothesized that Mpro may be directly binding and inactivating serpins involved in blood coagulation cascade, which may vary the levels of serpin to induce hypercoagulation, as observed in many Covid-19 patients.

Conclusions and perspectives

Most of the efforts to counter Covid-19 are concentrated on developing DNA- and protein-based vaccines or targeting viral spike protein and protease through natural and synthetic small compounds [7]. However, due to lack of effective therapy or vaccines against existing coronaviruses, the high cost and the enormity of vaccination task, there is a necessity for discovering alternative strategies for countering SARS-CoV-2 in vivo. Covid-19 crisis has changed the perception of research, and more than ever collective efforts are needed to evolve and integrate strategies to counter the virus. Therefore, it has become pertinent to share ideas and concepts that can be tested for therapeutic potential at a wider scale in a short time.
There is persuading evidence that a number of serpins are cross-class inhibitors of cysteine proteases, therefore serpin may have specificity for Mpro that is a cysteine protease. Alternatively, evaluation of Mpro target sites indicates that it may inhibit the serpins involved in coagulation cascade directly, possibly forming the basis of high rate of clotting observed in Covid-19 patients. In the absence of a human analog of Mpro, making a serpin chimera by replacing its RCL residues with Mpro cleavage site residue will significantly increase specificity (Figure 3). Additionally it will remove any cross specificity with serpins natural protease targets in human due to replacement of its recognition site. Furthermore, from the mechanistic standpoint, an effective chimeric serpin by trapping Mpro will counter the disease at a faster rate due to the diffusion rate limit efficiency of serpins. Serpins can be engineered to improve function for treating diseases, especially as low cost alternative to expensive drugs [41]. Although vaccine are first line of defense for disease causing viruses, recombinant protein can still be effective over the counter drug that are needed for symptomatic patients and will also be available and affordable.

Competing Interests
The authors declare that there are no competing interests associated with the manuscript.

Abbreviations
APTT, activated partial thromboplastin time; CoV, coronavirus; Covid-19, coronavirus disease; Mpro, main protease; PT, prothrombin time; RCL, reactive center loop; SARS-CoV-2, severe acute respiratory syndrome-coronavirus 2; serpin, serine protease inhibitor.

References
1 Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J.R. and Hilgenfeld, R. (2003) Coronavirus Main Proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* **300**, 1763–1767
2 Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L. et al. (2020) Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved ketoamide inhibitors. *Science* **24**, 409–412
3 Gettins, P.G., Patston, P.A. and Olson, S.T. (1996) Serpins: Structure, Function and Biology, R.G. Landes Company, New York.
4 Irving, J.A., Pike, R.N., Dai, W., Brömme, D., Margaret Worrall, D., Silverman, G.A. et al. (2002) Evidence that serpin architecture intrinsically supports papain-like cysteine protease inhibition: engineering alpha(1)-antitrypsin to inhibit cathepsin proteases. *Biochemistry* **41**, 4998–5004, https://doi.org/10.1021/bi0159885
5 Gettins, P.G. (2002) Serpin structure, mechanism, and function. *Chem. Rev.* **102**, 4751–4804, https://doi.org/10.1021/cr010170+
6 Scott, B.M. and Sheffield, W.P. (2020) Engineering the serpin α1-antitrypsin: a diversity of goals and techniques. *Protein Sci.* **29**, 856–871, https://doi.org/10.1002/pro.3794
7 Jin, Z., Du, X., Xu, T. et al. (2020) Structure of Mpro from COVID-19 virus and discovery of its inhibitors. *Nature* **582**, 289–293, https://doi.org/10.1038/s41586-020-2223-y
8 Devaux, C.A., Rolain, J.M., Colson, P. and Raoult, D. (2020) New insights on the antiviral effects of chloroquine against coronavirus: what to expect for COVID-19? *Int. J. Antimicrob. Agents* **55**, 10593
9 Amanat, F. and Krammer, F. (2020) SARS-CoV-2 vaccines: status report. *Immunity* **52**, 583–589, https://doi.org/10.1016/j.immuni.2020.03.007
10 Zhang, B., Liu, S., Tan, T., Huang, W., Dong, Y., Chen, L. et al. (2020) Treatment with convalescent plasma for critically ill patients with SARS-CoV-2 infection. *Chest* **20**, 30571–30577, https://doi.org/10.1016/j.chest.2020.03.039
11 Bózóki, M., Mitsuishiška, K., Raczyński, A., Samol, A., Tuszyński, J.A. and Góra, A. (2020) Structural and evolutionary analysis indicate that the SARS-CoV-2 Mpro is a challenging target for small-molecule inhibitor design. *Int. J. Mol. Sci.* **21**, 3099, https://doi.org/10.3390/ijms21093099
12 Macchiagodena, M., Pagliai, M. and Procacci, P. (2020) Identification of potential binders of the main protease 3CLpro of the COVID-19 via structure-based ligand design and molecular modeling. *Chem. Phys. Lett.* **750**, 137489, https://doi.org/10.1016/j.cplett.2020.137489
13 Ngo, S.T., Pham, N.Q.A., Le, L.T., Pham, D.-H. and Vu, V. (2020) Computational determination of potential inhibitors of SARS-CoV-2 main protease. *ChemReV* **21**, 13
14 Sharma, S. and Deep, S. In-silico drug repurposing for targeting SARS-CoV-2. *ChemReV* **8**, 57
15 Ahammad, O.S., Verkhivker, G.M. and Bishop, O.T. (2020) Impact of emerging mutations on the dynamic properties the SARS-CoV-2 main protease: an in silico investigation. *bioRxiv*, https://doi.org/10.1101/2020.05.29.123190
16 Muramatsu, T., Takemoto, C., Kim, Y.-T., Wang, H., Nishii, W., Terada, T. et al. (2016) SARS-CoV 3CL protease cleaves its C-terminal autoprocessing site by novel substrate cooperativity. *Proc. Natl. Acad. Sci. U.S.A.* **133**, 12997–13002
17 Xue, X., Yu, H., Yang, H., Xue, F., Wu, Z., Shen, W. et al. (2008) Structures of two coronavirus main proteases: implications for substrate binding and antiviral drug design. *J. Virol.* **82**, 2515–2527, https://doi.org/10.1128/JVI.01114-07
18 Krichel, B., Falke, S., Hilgenfeld, R., Redecke, L. and Uetrecht, C. (2020) Processing of the SARS-CoV pp1a/ab nsp7–10 region. *Biochem. J.* **477**, 1009–1019, https://doi.org/10.1042/BJ20200029
19 Pillayart, T., Manickam, M., Namasiavayam, V., Hayashi, Y. and Jung, S.H. (2016) An overview of severe acute respiratory syndrome-coronavirus (SARS-CoV) 3CL protease inhibitors: peptidomimetics and small molecule chemotherapy. *J. Med. Chem.* **59**, 6595–6628, https://doi.org/10.1021/acs.jmedchem.5b01461
