Binding mechanism of coronavirus main proteinase with ligands and its implication to drug design against SARS

Kuo-Chen Chou,a,b,* Dong-Qing Wei,b and Wei-Zhu Zhonga,b

a Gordon Life Science Institute, 7088 Arbor Valley, Kalamazoo, MI 49009, USA
b Tianjin Institute of Bioinformatics and Drug Discovery (TIBDD), Tianjin Normal University, Tianjin, China

Abstract

In order to stimulate the development of drugs against severe acute respiratory syndrome (SARS), based on the atomic coordinates of the SARS coronavirus main proteinase determined recently [Science 13 (May) (2003) (online)], studies of docking KZ7088 (a derivative of AG7088) and the AVLQSGFR octapeptide to the enzyme were conducted. It has been observed that both the above compounds interact with the active site of the SARS enzyme through six hydrogen bonds. Also, a clear definition of the binding pocket for KZ7088 has been presented. These findings may provide a solid basis for subsite analysis and mutagenesis relative to rational design of highly selective inhibitors for therapeutic application. Meanwhile, the idea of how to develop inhibitors of the SARS enzyme based on the knowledge of its own peptide substrates (the so-called “distorted key” approach) was also briefly elucidated.

© 2003 Elsevier Inc. All rights reserved.

Keywords: SARS; Coronavirus proteinase; KZ7088; AG7088; Binding pocket; Octapeptide substrate; “Distorted key” mechanism
Materials and methods

The coordinates of SARS-coronavirus main proteinase, or SARS-CoV M\(^{pro}\), recently determined by Anand et al. [1] were used as a targeted basis to conduct docking studies. The proteinase contains 304 amino acids, folded into three domains (Fig. 1): domain 1 (res. 8–99) is coloured magenta, domain 2 (res. 100–183) yellow, and domain 3 (res. 200–300) light blue. The N-terminal (res. 1–7) is coloured green, while the C-terminal (res. 301–304) red. The loop (res. 184–199) linking domain 2 and 3 is coloured purple. The binding cleft is located at the site between domains 1 and 2.

Two ligands were selected for the docking studies. One was KZ7088 (Fig. 2A) and the other the octapeptide AVLQSGFR. As shown in Fig. 2A, KZ7088 is a derivative by modifying AG7088 (Fig. 2B). The latter was developed by Pfizer and is currently in clinical trials for the treatment of rhinovirus, a pathogen that can cause the common cold. Although it is a putative candidate for docking studies related to SARS drug finding, AG7088 has a \(p\)-fluorophenylalanine side chain (\(p\)-fluorobenzyl), which might be too long (or bulky) to fit into the relevant binding pocket [1]. Accordingly, KZ7088 with a modified side chain by removing \(-\text{CH}_2\) (Fig. 2) could well serve as a starting point for modification that will quickly lead to effective drug candidates for the treatment of SARS. The reason to select the octapeptide AVLQSGFR is as follows.

1. The protease-susceptible sites in proteins usually extend to an octapeptide, as generally formulated by \(P_i P_{i-1} P_{i-2} P_{i-3} P_{i-4} P_{i-5} P_{i-6} P_{i-7}\) with the scissile bond located between the subsites \(P_{i-1}\) and \(P_{i-2}\), as generally expressed by \(P_{i-1} \sim P_{i-2}\) [2–4].
2. The SARS coronavirus enzyme and several viral proteinases exhibit Gln \(\sim\) (Ser, Ala, and Gly) specificity [1].
3. According to the “lock-and-key” mechanism in enzymology, the octapeptide cleavable by the SARS proteinase must have a good fit for binding to the active site. However, such a peptide, after a modification of its scissile bond with some simple routine procedure, will completely lose its cleavability but it can still bind to the active site. Actually, the molecule thus modified can be compared to a “distorted key” that can be inserted into a lock but cannot neither open the lock nor be pulled out from it, spontaneously becoming an ideal competitive inhibitor against the SARS proteinase. Therefore, it would provide us with useful insights for drug design by studying the binding mechanism of SARS proteinase with octapeptides that are cleavable by the enzyme itself. For a detailed discussion about the rationale of “distorted key,” see a review by Chou [5] and Fig. 2 therein.

The software package MOE (Chemical Computing Ltd.) was used to conduct the docking of the ligands to the SARS-coronavirus main proteinase. A similar approach was successfully used to investigate the binding interaction of the Cdk5-Nck5a*-ATP [6,7] and that of caspase-9 with Ac-DVAD-fmk [8].

Results and discussion

The overall structure obtained by docking KZ7088 to SARS-CoV M\(^{pro}\) is given in Fig. 3, where the SARS-coronavirus main proteinase is in ribbon drawing, the KZ7088 in ball-and-stick drawing, and the white mesh represents the binding pocket for KZ7088. The constituents of the binding pocket are defined by those residues that have at least one heavy atom (i.e., other than hydrogen) with a distance \(\leq 5\ \text{\AA}\) from a heavy atom of KZ7088 (a similar binding pocket was defined for ATP in the Cdk5–Nck5a*-ATP complex [6] that has later proved quite useful in identifying functional domains and stimulating the relevant truncation experiments [9]). The pocket thus defined consists of 23 residues, as listed in Table 1. It would be intriguing to probe the binding pocket by site-directed mutagenesis of single amino acids as an avenue to find out which of the 23 residues is the most sensitive for the catalytic activity. A close view of the binding interaction is given in Fig. 4. As shown from the figure, KZ7088 is tethered to the enzyme by six hydrogen bonds (blue dotted lines). The
detailed atoms in forming the hydrogen bonds are given in Table 2.

The overall structure obtained by docking the octapeptide AVLQSGFR to SARS-CoV M^pro is given in Fig. 5, where the octapeptide is coloured white (in ball-and-stick drawing). As shown in Fig. 6A, there are six hydrogen bonds (blue dotted lines) formed between the octapeptide and the SARS proteinase. Those residues involved in forming the hydrogen bonds from the enzyme are: Arg-40, His-41, Phe-185, Asp-187, and Gln-189. The detailed atoms in forming the hydrogen bonds

Table 2

<table>
<thead>
<tr>
<th>Atom in SARS proteinase</th>
<th>Atom in KZ7088</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thr-25 N</td>
<td>O-9</td>
</tr>
<tr>
<td>Thr-25 O-γ1</td>
<td>O-20</td>
</tr>
<tr>
<td>Thr-45 O</td>
<td>O-15</td>
</tr>
<tr>
<td>Leu-50 N</td>
<td>O-68</td>
</tr>
<tr>
<td>Tyr-54 O-η</td>
<td>O-67</td>
</tr>
<tr>
<td>Asp-187 O</td>
<td>O-72</td>
</tr>
</tbody>
</table>

detailed atoms in forming the hydrogen bonds are given in Table 2.

The overall structure obtained by docking the octapeptide AVLQSGFR to SARS-CoV M^pro is given in Fig. 5, where the octapeptide is coloured white (in ball-and-stick drawing). As shown in Fig. 6A, there are six hydrogen bonds (blue dotted lines) formed between the octapeptide and the SARS proteinase. Those residues involved in forming the hydrogen bonds from the enzyme are: Arg-40, His-41, Phe-185, Asp-187, and Gln-189. The detailed atoms in forming the hydrogen bonds

Table 1

<table>
<thead>
<tr>
<th>Binding pocket of the SARS proteinase for KZ7088</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cys-22</td>
</tr>
<tr>
<td>Thr-45</td>
</tr>
<tr>
<td>Tyr-54</td>
</tr>
</tbody>
</table>
are given in Table 3, which may provide useful information for in-depth understanding of the mechanism of the nucleophilic attack of the active site on the carbonyl carbon atom of the scissile peptide bond [10]. It is interesting to see that, after a single amino acid mutation for the octapeptide at the position $P_1$, i.e., substitution of Q by E, the relevant hydrogen bonds are reduced from the original six (A) to only three (B).

Conclusions

In addition to the atomic coordinates of the SARS-CoV main proteinase recently provided by Anand et al. [1], the detailed binding interactions of the SARS enzyme with KZ7088 (a derivative of AG7088) and the octapeptide AVLQSGFR obtained through the present docking studies will further provide a solid footing for conducting structure-based drug design against SARS.

References