

Class: SARS-CoV 3C-like protease

Attributes:

Accession #: NP_828863

Synonyms: nsp5, 3Clpro

Molecular weight: 35.8 kDa (1)

Number of amino acids: 601

Structure:

1) Homology:

- has 40 and 44% sequence identity to HCoV 229E and TGEV, respectively (2)
- domains I and II show a 42 to 48% identity with coronavirus group I enzymes (2)
- domain II shows a 36 to 40% identity with coronavirus group I enzymes (2)
- shares a common fold with the human rhinoviral protease SARS 3Clpro has an additional helical domain at the C-term (3)

2) Domain Information:

- has three domains
 - first two domains form a chymotrypsin fold
 - responsible for catalytic reaction
 - the third domain is alpha-helical with unclear biological function
- domains I** (aa 8-99) **and II** (100-183) are six-stranded antiparallel beta barrels and resemble the architecture of chymotrypsin and picornavirus 3C proteinases (2)
- the **ligand binding pocket** is in the cleft between the two beta domains (2)
- a **long loop** (aa 184-199) connects domain III to the C-term domain (domain III, aa 200-300) (2)
- domain III** is a globular cluster of five helices implicated in the proteolytic activity (2)

3) Active Conformation:

- only the dimeric form is active (4)
- the dimeric form has a k_{cat}/K_m of $\sim 1.4 \times 10^3 \text{ mM}^{-1} \text{ min}^{-1}$ (4)
 - tight dimer is predominantly between domain II of molecule A and the N-term residues of molecule B with the two oriented perpendicular to one another (2)
 - dissociation constant of the dimer was estimated to be around 100uM (4)
- the N-term residues are between domains II and III of the parent monomer and domain II of the other monomer (2)
 - mechanism would immediately enable the catalytic site to act on other cleavage sites in the polyprotein (2)

4) Catalytic Site Information

- catalytic dyad** contains residues His41 and Cys144 (3)
 - a buried water molecule is found in the place that would normally be occupied by the third member of the triad (2)
 - this molecule is hydrogen-bonded to His41 N δ 1, Gln163 N ϵ 2, and Asp186 O δ 1 (2)
 - the imidazole of His162 is suitably positioned to interact with the P1 glutamine side chain (2)
 - neutral state of His162 over a broad pH range maintained by two interactions (2)
 - 1) stacking onto the phenyl ring of Phe139 (2)
 - 2) accepting a hydrogen bond from the hydroxyl group of the buried Tyr160 (2)
 - other elements involved in the S1 pocket are the main-chain atoms of Pro51, Met164, Glu165 and His171 (2)
 - hydrophobic S2 subsite is formed by Met164, Pro51, Asp47, His41, Tyr53 (2)
 - an alanine (Ala46) is inserted in the loop between His41 and Pro51 (2)
 - S1' subsite formed by Leu27, His41 and Asp47 (2)

5) Substrate Information

- substrate is cleaved at the Gln-Ser peptide bond (4)
 - three noncanonical cleavage sites with Phe, Val or Met in the P2 position and one noncanonical cleavage site with Asn in the P1' position (4)
 - P1 position has a well conserved Gln residue (4)
 - P2 position of the substrates seems to favor large hydrophobic residues (4)
 - Leu / Ile, Phe, Val and Met are tolerated at P2 position (4)
 - substrates with more beta-sheet-like structure tend to react fast (4)
 - 3Clpro amino acids Arg-40, His-41, Phe-185, Asp-187 and Gln-189 forms hydrogen bonds with AVLQSGFR octapeptide (5)

6) Inhibitor:

- Sabadinine is a natural product isolated originally from the Lily plant Veratrum sabdilla (6)
- Sabadinine docked into the active site of 3CLpro with a docking energy of -11.6 kcal/mol and a clustering of 9 out of 10 conformers (6)
 - distance between the hydroxyl at C(20) of sabadinine and the H bound to N(ϵ) of His44 is 2.93 A (6)

Responsibilities:

Dimerization is necessary for the enzyme to become active;
autoproteolytically cleaves itself from the gene one protein precursor

Processes the gene one protein precursor to yield the mature viral replicase proteins

Collaborators:

3Clpro

SARS-CoV gene one protein precursor

- 1) Sun, H., et al., (2003) Molecular cloning, expression, purification, and mass spectrometric characterization of 3C-like protease of SARS coronavirus, *Protein Expr Purif*, **32(2)**, 302-8.
- 2) Anand, K., et al., (2003) Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs, *Science*, **300(5626)**, 1763-7.
- 3) Yan, L., et al., (2003) Assessment of putative protein targets derived from the SARS genome, *FEBS Lett*, **554(3)**, 257-63.
- 4) Fan, K., et al., (2004) Biosynthesis, Purification, and Substrate Specificity of Severe Acute Respiratory Syndrome Coronavirus 3C-like Proteinase, *J. Biol. Chem.*, **279(3)**, 1637-1642.
- 5) Chou, K., Wei, D., and Zhong, W., (2003) Binding mechanism of coronavirus main proteinase with ligands and its implication to drug design against SARS., *Biochem Biophys Res Commun*, **308(1)**, 148-51.
- 6) Toney, J.H., et al., (2004) Sabadinine: a potential non-peptide anti-severe acute-respiratory-syndrome agent identified using structure-aided design, *J Med Chem*, **47(5)**, 1079-80.