The expanding diversity of serine hydrolases

Istvan Botos[1] and Alexander Wlodawer[2]

Serine hydrolases use a hydroxyl of a serine, assisted by one or more other residues, to cleave peptide bonds. They belong to several different families whose general mechanism is well known. However, the subtle structural differences that have recently been observed across a variety of families shed light on their functional diversity, including variations in mechanism of action, differences in the modes of substrate binding, and substrate-assisted orientation of catalytic residues. Of particular interest are the Rhomboid family serine proteinases that are active within the plasma membrane, for which several new structures have been reported. Because these enzymes are involved in biological and pathological processes, many are becoming important targets of drug design.

Addresses
[1] Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD 20892, USA
[2] Macromolecular Crystallography Laboratory, National Cancer Institute at Frederick, Frederick, MD 21702, USA

Corresponding author: Wlodawer, Alexander (wlodawer@ncifcrf.gov)

Current Opinion in Structural Biology 2007, **17**:683–690

This review comes from a themed issue on
Proteins
Edited by Edward Baker and Guy Dodson

Available online 24th September 2007

0959-440X/$ – see front matter
Published by Elsevier Ltd.

DOI 10.1016/j.sbi.2007.08.003

Introduction

Proteinases (peptidases, proteases) are enzymes that catalyze the hydrolysis of peptide bonds. They are classified into six classes on the basis of their catalytic mechanism: serine, threonine, cysteine, aspartate, glutamate, and metalloproteases. Serine proteinases (SP) cleave the peptide bond by nucleophilic attack of the serine hydroxyl group on the scissile carbonyl bond (some enzymes with nucleophiles such as threonine or cysteine share the fold with serine proteases; they will not be further discussed here). Typical SPs have an active site catalytic triad consisting of the Oγ atom of the serine, the imidazole ring of the histidine, which serves as a general acid/base, and the carboxylate of an aspartic acid, which helps to orient the imidazole ring. In addition, an oxyanion hole provides electrophilic assistance to the nucleophilic attack by the serine Oγ on the carbonyl atom of the scissile bond. The goal of this review is to highlight some key structural features of the serine proteinase structures published in the past two years (Table 1).

Structures of trypsin-like and chymotrypsin-like proteinases

The terms 'trypsin-like proteinases' and 'chymotrypsin-like proteinases' are sometimes used interchangeably to not only identify members of MEROPS [1] (also see http://merops.sanger.ac.uk/) clan PA (mixed catalytic type), family S1 (chymotrypsin), but also less strictly to more distant families in that clan. All these SPs share the same basic fold in their catalytic domains (although they may or may not also have other domains) but vary widely in their specificity and regulation. A number of them will be discussed in this section.

Kallikreins are trypsin-like proteases that cleave kininogen; they are grouped into two major categories, plasma and tissue. Plasma kallikrein is a multidomain, glycosylated protein that is more homologous to factor XI than to tissue kallikrein with a single protease domain. The structural basis for the specificity of plasma kallikrein was revealed by its crystal structure [2]. Human tissue kallikreins (hK) belong to a closely related 15-member family. Its best functionally characterized member, hK1, is involved in blood pressure regulation. Structural studies of hK1 revealed that binding of a peptide substrate induces structural rearrangement of the catalytic Ser that optimally positions it for interactions with the catalytic His [3]. The highly specific hK4, which activates hK3/PSA, is normally expressed in the prostate but is overexpressed in prostate and ovarian cancer [4]. The hK4 structure uncovered a novel metal-binding site formed by a 10-residue loop linked to the N-terminal segment, specific for this enzyme. hK4 can exist in a Zn^{2+}-free high activity form and a Zn^{2+}-bound low activity form [4]. The prostatic fluid contains micromolar concentrations of zinc that can inhibit hK4.

Component B of earthworm fibrinolytic enzyme (EFE-b) is a two-chain glycoprotein that exhibits strong fibrinolytic activity [5]. It has a novel conserved structural motif that is formed by two consecutive Cys residues linked by a *cis* peptide bond and a disulfide bond. This ring may play a similar role to Tyr151 in u-PA, t-PA, or DSPA and may be related to the selectivity for the P2′ of its substrate.

SplC proteinase from *Staphylococcus aureus*, structurally most similar to V8 proteinase, another trypsin-like enzyme [6], is possibly involved in staphylococcal virulence. The protease is latent until cleaved by signal peptidase during secretion. Its structure reveals a V8-like fold, a well-defined S1 pocket, and a possible mechanism of substrate-induced conformational switch [7].

Table 1

Summary of the structures covered in this review

Enzyme	Family	Peptidase type	MEROPS ID	PDB accession code	Reference
Complement factor B	S1A	Chymotrypsin-A	S01.196	2OK5	[13•]
Complement component C2a			S01.194	2ODP,2ODQ	[14]
MASP-2 zymogen R444Q mutant fragment			S01.229	1ZJK	[15]
Human tissue kallikrein 1			S01.160	1SPJ	[3]
Human tissue kallikrein 4			S01.251	2BDG,2BDH,2BDI	[4]
Human plasma kallikrein			S01.212	2ANW,2ANY	[2]
Prophenoloxidase-activating factor I			S01.204	2OLG	[11•]
Prophenoloxidase-activating factor II			S01.960	2B9L	[12]
Earthworm fibrinolytic enzyme component B			S01.243	1YM0	[5]
Factor XI zymogen			S01.213	2F83	[10]
A. actus venom serine proteinase I and II			S01.x	1OP0,1OP2	[8]
A. contortrix contortrix protein C activator			S01.466	2AIP,2AIQ	[9]
SplC proteinase			S01.283	AS9	[7]
Hemoglobin protease	S6	IgA1-specific serine peptidase	S06.003	1WXR	[16••]
NS3 protease	S7	Flavivirin	S07.001	2FOM,2FP7	[17•]
Subtilase cytotoxin protease SubA	S8A	Subtilisin Carlsberg	S08.121	2IY9	[22]
Subtilisin-like protease *Vibrio*			S08.x	1SH7	[24•]
Serratia proteinase K			S08.x	2B6N	[23]
Fibroblast activation protein α	S9A	Dipeptidyl-peptidase IV	S09.007	1Z68	[19]
Prolyl tripeptidyl peptidase			S09.017	2D5L,2DCM	[20]
Prolyl endopeptidase	S9B	Prolyl oligopeptidase	S09.001	1YR2,2BKL	[21]
Archaeoglobus fulgidus LonB	S16	Lon-A	S16.005	1ZOW,1ZOC,1ZOB, 1ZOG,1ZOE	[28•]
Sesbania mosaic virus polyprotein protease	S39A	Sobemovirus peptidase	S39.001	1ZYO	[18]
Membrane protease 1510-N	S49	Signal-peptide peptidase A	S49.005	2DEO	[29]
Birnavirus VP4	S50	Infectious pancreatic necrosis birnavirus Vp4 peptidase	S50.001	2GEF	[30•]
E. coli GlpG	S54	Rhomboid-1	S54.016	2IC8,2NRF,2IRV	[31••,32•,33]
H. influenzae hiGlpG			S54.024	2NR9	[35]

Trypsin-like *Agkistrodon actus* snake venom serine proteinases I and II are identical except for one amino acid. An interesting feature observed in these structures is the conserved *N*-glycosylation site at Asn35. This oligosaccharide, close to the active site, can sterically restrict access of some inhibitors, such as soybean trypsin inhibitor (STI) [8]. *Agkistrodon contortrix* protein C activator activates protein C without relying on thrombomodulin. Structures of the native and inhibited enzyme reveal three carbohydrate moieties positioned around the active site that are involved in substrate recognition [9]. These carbohydrates, together with the positive charge on the interfacial surface, may be crucial for activation—protein C does not bind the macromolecular trypsin inhibitors BPTI and STI.

Factor XI is an essential coagulation protein that circulates as a disulfide-linked homodimer [10] and cleaves factor IX. Each monomer has four homologous apple domains (A1–A4) that mediate essential protein–protein interactions, followed by a chymotrypsin-like serine protease domain (Figure 1). Each apple domain is formed by seven β strands that cradle a single α helix. Thrombin binds the A1 domain through its exosite I, whereas the platelet surface receptor glycoprotein Ib-V-IX (GpIb) binds A3 through its leucine-rich repeats. These two apple domain sites are diametrically positioned in factor XI, separated by the centrally positioned protease domain. Thrombin cleaves after Arg369, and the free N-terminus moves 20 Å and inserts into the activation pocket of factor XIa, generating an oxyanion hole [10].

The easter-type SPs, prophenoloxidase-activating factors (PPAF) I and II are involved in the insect immune response. The structures of the SP domain of PPAF-I [11•] and the whole PPAF-II from *Holotrichia diomphalia* [12] reveal a chymotrypsin-like fold of their proteinase domains (Figure 2). PPAF-I belongs to the catalytic group of clip domain SPs, in contrast to PPAF-II, which belongs to the non-catalytic group (the latter contains Gly353 in place of the catalytic serine). The clip domain of PPAF-II is a protein-interaction module with a novel fold; it is essential to the binding and activation of the 76 kDa phenoloxidase via its central cleft [12]. After cleavage by PPAF-III, the clip domain remains tightly attached to the SP domain, and the enzyme oligomerizes into two hexameric rings that serve as a hub for phenoloxidase binding. In contrast to PPAF-II, the clip domain of PPAF-I is separated from the SP domain during activation.