
Three-Dimensional Structures and Mechanisms of Snake Venom Serine Proteinases, Metalloproteinases, and Phospholipase A₂

11

M. A. Coronado, F. R. de Moraes, A. Ullah, R. Masood, V. S. Santana,
R. Mariutti, H. Brognaro, Dessislava Georgieva, M. T. Murakami,
Christian Betzel, and Raghuvir K. Arni

Contents

Introduction	240
Serine Peptidases	242
Sequence Alignments	242
Overall Structure	243
Active Site	245
Oxyanion Hole and Subsites	245
Mechanism of Catalysis	247
Zymogen Activation	249
Snake Venom Metalloproteinases	249
Domain Organization and Sequence Homology	250
Catalytic Mechanism	253

M.A. Coronado • F.R. de Moraes • A. Ullah • R. Masood • V.S. Santana • R. Mariutti • H. Brognaro
Department of Physics, IBILCE/UNESP, Multi User Center for Biomolecular Innovation,
São Jose do Rio Preto-SP, Brazil
e-mail: monicacoronado@gmail.com; moraesfr@gmail.com; anwar.ms90@yahoo.com;
rehanabiochemist@gmail.com; vds.santana@gmail.com; ricardomariutti@yahoo.com.br;
hbrognao@yahoo.com.br

D. Georgieva • C. Betzel
Institute of Biochemistry and Molecular Biology, Laboratory of Structural Biology of Infection
and Inflammation, University of Hamburg, Hamburg, Germany
e-mail: octopus_dofleini@yahoo.com; christian.betzel@uni-hamburg.de;
betzel.christian@googlemail.com

M.T. Murakami
National Laboratory for Biosciences, National Center for Research in Energy and Materials,
Campinas, Brazil
e-mail: mario.murakami@lnbio.org.br

R.K. Arni (✉)
Department of Physics, Multi User Center for Biomolecular Innovation, Institute of Biosciences,
Humanities and Physical Sciences, IBILCE/UNESP (Universidade Estadual Paulista), São José do
Rio Preto, SP, Brazil
e-mail: arni@sjrp.unesp.br

Disintegrin Domain	256
Phospholipases A ₂	257
Primary Structure	257
Secondary Structure	258
Tertiary Structure	258
The Hydrophobic Channel	260
Catalytic Mechanism	261
Conclusion and Future Directions	261
Cross-References	264
References	264

Abstract

High-resolution crystal structures provide accurate information of the positions of the atoms which can be used to understand substrate specificity, secondary binding sites, and catalytic mechanisms. Detailed structural information and mechanisms of serine proteinases, metalloproteinases, and phospholipases A₂ are presented.

Introduction

Structural information gleaned at the molecular and atomic levels, when correlated with biochemical and biophysical details capable of generating a coherent picture of the salient structural features and interactions that modulate and determine biomolecular recognition, specificity, and hydrolysis, provides us with very powerful tools to decipher, step-by-step, complex biological phenomena, thus permitting a profound understanding of the basic underlying role of molecular architectural diversity in efficiently performing essential, distinct, chemical reactions that are central to life.

The enormous number of protein and DNA sequences currently deposited in the data banks (e.g., <http://www.expasy.org>, <http://www.uniprot.org>) indicate that the one-dimensional representation of protein sequences generally contains a trace of the fingerprint of evolution, and often, only a faint residual of the ancestral protein is retained in the protein linear amino acid sequence. However, upon closer examination, the application of this concept exposes its fundamental limitations, and hence, its utility is strictly limited since (a) proteins that perform the same function or catalyze similar reactions often share only very low-sequence identity, (b) proteins that are about 20 % identical in their primary sequences may still catalyze distinct reactions and modulate different functions, and (c) point mutations in the active sites or cofactor-binding sites often produce proteins that catalyze different reactions or in extreme cases result in enzymatically inactive proteins.

A fundamental conceptual bridge linking the linear protein sequence and its primary biological function is encoded in the three-dimensional fold or, in other words, in the exact positions of the atoms of the protein in three dimensions. Central to this concept is the fact that the three-dimensional (3D) structure of a protein is

more highly conserved during evolution (Bajaj and Blundell 1984; Finkelstein and Ptitsyn 1987) than the linear amino acid sequence of the protein, and consequently, the shape of a protein, the spatial distribution of its atoms, and the surface charge, solvent accessibility, glycosylation, and solvent structure are key factors that determine its biological function. In general, proteins that possess the same three-dimensional structures also perform similar functions and catalyze analogous reactions (Thornton et al. 1991). Basically, only a small subset of amino acids perfectly positioned and conserved in three dimensions, rather than the linear conservation of the primary sequence, forms the core of the protein's biological function (Hasson et al. 1998; Kasuya and Thornton 1999) which is also often dependent on the exact positioning and interaction of the protein with solvent molecules that facilitate charge/proton transfer. Thus, protein structures are more conserved than their sequences (Lesk and Chothia 1980; Chothia and Lesk 1986), and the degree to which an amino acid is evolutionarily conserved in a certain position reflects its structural and functional importance.

Snake venom proteins are present in a wide variety of sizes, amino acid sequences, and consequently three-dimensional structures, which reflects their diverse roles in nearly all envenomation steps. Protein crystallography, nuclear magnetic resonance spectroscopy, and, to a lesser extent, electron microscopy are powerful tools that have been applied very successfully to determine the three-dimensional structures of toxins at the atomic and molecular levels. Structural analysis of snake venom proteins is especially rewarding since they can be obtained with high purity and due to their intrinsic stability, readily form large, well-ordered crystals that are suitable for high-resolution X-ray diffraction analysis.

Snake venom proteins have been reviewed in a number of recent articles (Kang et al. 2011; Kini 2005; Georgieva et al. 2008; Rossetto and Montecucco 2008; Yamazaki and Morita 2007), and here, only the structural features of leading members of select protein families will be presented. Protein crystallography has experienced amazing growth as evidenced by the fact that the protein data bank (www.rcsb.org) contains the atomic coordinates of over 100,000 proteins and excellent recent reviews (Wlodawer et al. 2008, 2013) present the current state of structural biology. Structural data extends our horizon and has provided us with the impetus to decipher complex interactions that define enzyme specificity, mechanism, charge distribution, modes of inhibition, and dynamics and, more importantly, also paves the way for the design of novel inhibitors and drugs with potential clinical and medical applications.

Counterparts of many snake venom proteins are encountered in the mammalian repertoire, and structural comparisons between the snake venom and mammalian proteins indicate the presence of the same basic structural scaffold. The structures of key snake venom proteins and their salient features are compared and highlighted. The catalytic mechanisms of these proteins are described with relation to their structural aspects. Since the structures of a significant number of snake venom proteins are now available, structure-based sequence alignments can be exploited, and this presentation will be limited only to the cases where accurate, high-resolution structural data is available.

Serine Peptidases

Proteinases or peptidases catalyze the hydrolytic cleavage of peptide bonds (Rawlings et al. 2013), and proteinases that contain a highly reactive nucleophilic Ser and are assisted by His and Asp, thus forming the catalytic triad His, Ser, and Asp, are widely referred to as serine peptidases or serine endopeptidases (Hedstrom 2002a, b; Neurath 1984). These enzymes catalyze the cleavage of covalent peptide bonds in proteins and peptides and have been the focus of much research since they participate in a number of diverse, essential biological processes ranging from digestion and blood coagulation to immune response and inflammation (Hedstrom 2002a) by exhibiting high stereospecificity and enantiospecificity. These enzymes, widely encountered in snake venoms (de Oliveira et al. 2013; Serrano 2013; Ullah et al. 2013), probably originated as digestive enzymes and subsequently evolved by gene duplication and sequence modifications to serve specific functions (Birktoft and Blow 1972).

Snake venom serine proteinases (SVSPs) play essential roles in envenomation since they interfere with the functioning, maintenance, and regulation of diverse physiological processes of the prey, specifically the maintenance, regulation, activation, and inhibition of the blood coagulation cascade and the fibrinolytic system, and often additionally trigger blood platelet aggregation (Gempeler et al. 2001; Murakami and Arni 2005; Zhang et al. 1995). Based on their amino acid sequences and functional similarities, SVSPs are classified as belonging to the SA clan and the S1 family (MEROPS classification, <http://merops.sanger.ac.uk>). Despite the fact that they share significant sequence identity (50–70 %), SVSPs display high specificity by binding to and cleaving distinct macromolecular substrates and hence have been classified as activators of the fibrinolytic system, procoagulant, anticoagulant, and platelet-aggregating enzymes.

Sequence Alignments

Structure-based sequence alignment of selected serine peptidases (Fig. 1) serves to highlight the similarities and differences between these enzymes and indicates that the protein chains of the SVSPs contain 245 amino acids and are of approximately the same length. In comparison, the amino acid chain of human α -thrombin which contains insertions in many loop regions is significantly longer. SVSPs differ from the trypsin family of enzymes since they contain a disulfide-linked C-terminal extension (Murakami and Arni 2005; Parry et al. 1998). The position of the three amino acids, His57, Asp102, and Ser195 (sequence numbering based on chymotrypsinogen) (Harley and Shotton 1971), is strictly conserved (marked in red). Only α -thrombin and trypsin bind monovalent and divalent ions (Na^+ and Ca^{2+} ions, respectively, highlighted in light blue). Interestingly, the thrombin exosites I and II, utilized for binding heparin and fibrinogen, are not conserved in SVSPs. Asp189 (marked in dark blue and positioned at the base of the oxyanion hole and which determines specificity) is totally conserved. The glycosylation sites of SVSPs

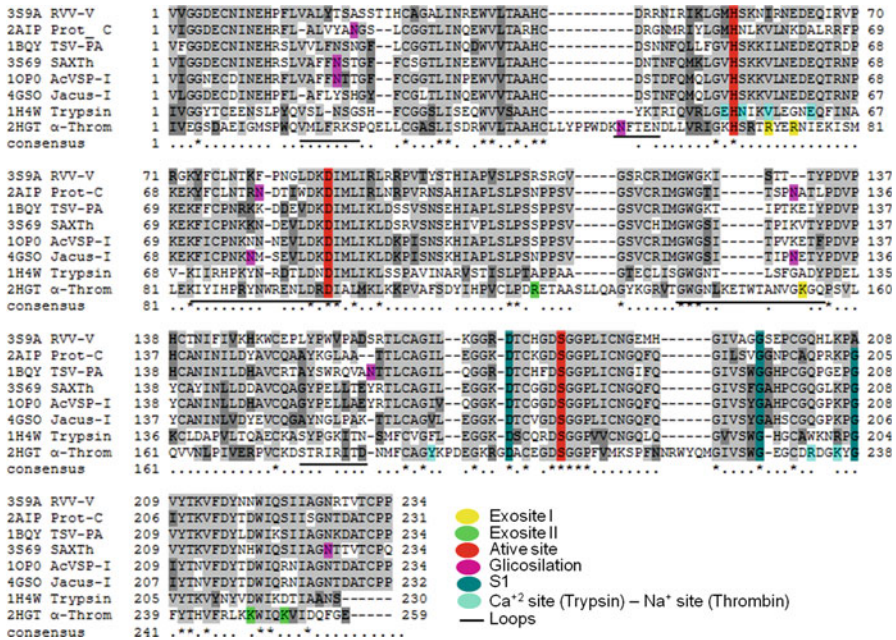


Fig. 1 Structure-based multiple sequence alignment of serine proteinases, whose crystal structures have been determined, performed using the program BoxShade (<http://www.exPASy.org>). RVV-V (*Russel’s viper venom*), Prot_C (*Agkistrodon contortrix contortrix*), TSV-PA (*Trimeresurus stejnegeri*), SAXTh (*Gloydius saxatilis*), AcVSP-I (*Agkistrodon acutus*), Jacus-I (*Bothrops jararacussu*), Trypsin (Human), and α -Thromb (Human). The Protein Data Bank four-letter ids are given on the left

(highlighted in pink) are significantly different. SVSPs are often referred to as thrombin-like enzymes; however, based on the above observations, it can be concluded that SVSPs are structurally more similar to trypsin than to α -thrombin.

Overall Structure

Like their mammalian counterparts, chymotrypsin and trypsin-like serine proteinases, SVSPs consist of about 245 amino acids distributed among the two domains (S and S’); each containing a six-stranded β -barrel and two short α -helices (residues, 165–173 and 235–244, sequence numbering based on chymotrypsinogen) (Harley and Shotton 1971) (Fig. 2a). The N-terminal S domains are stabilized by an intra-chain disulfide bridge (Cys42/Cys58) and two other disulfide bridges Cys22/Cys157 and Cys91/Cys245, the latter are unique to SVSPs (Murakami and Arni 2005; Parry et al. 1998). His57 and Asp102 which form part of the catalytic triad are located in this domain. The C-terminal S’ domain encompasses a six-stranded β -sheet and contains two α -helices, one inserted between strands 8 and 9 and the other located at the C-terminus preceding the extended C-terminal tail; a disulfide bridge interconnects

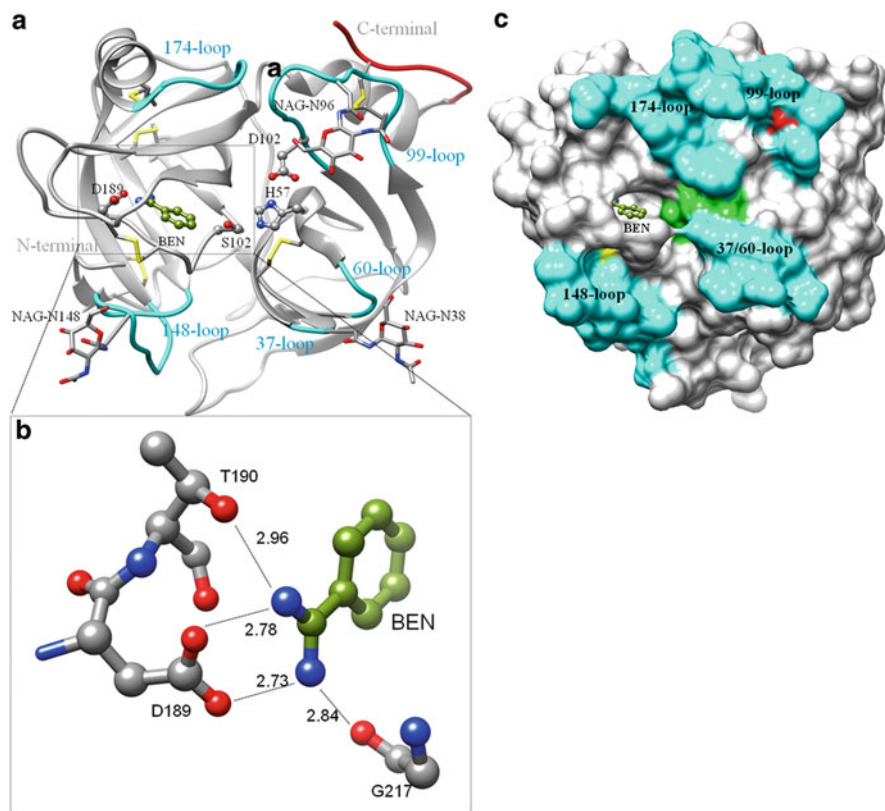


Fig. 2 (a) Ribbon representation of the crystal structure of ACC-C (*Agkistrodon contortrix contortrix*), oriented to indicate the N- and C-terminal lobes and highlighting the loops that are variable in SVSPs (cyan), the amino acids that form the catalytic triad (His57, Asp102, and Ser195) and determine specificity (Asp189) are atom color-coded (white, carbon; red, oxygen; and blue, nitrogen). The benzamidine molecule (BEN) is presented in green. The C-terminal extension that is unique to SVSPs is in red. The positions of the carbohydrate moieties are also indicated. (b) Zoom presenting details of the interaction of benzamidine. Thin lines and numbers indicate hydrogen bonds and hydrogen bond distances are in Å. (c) Space filling representation of ACC-C indicating the positions of the surface loops (cyan), the active site residue (green), benzamidine molecule (BEN), and the N-linked Asn (red). Visualization by UCSF – Chimera (Pettersen et al. 2004)

the tail with the N-terminal subdomain. This subdomain is further stabilized by three disulfide bridges (Cys136/Cys201, Cys168/Cys182, and Cys191/Cys220), and the highly reactive nucleophilic Ser195 is located in this domain. The catalytic triad (His57, Asp102, and Ser195) is located at the junction of both the barrels and is surrounded by the conserved 70-, 148-, and 218-loops and the nonconserved 37-, 60-, 99-, and 174-loops (Fig. 2a and b). As expected from the high sequence identity of SVSPs (about 60 %), superpositioning C α carbon atoms of the structures results in r.m.s. deviations ranging from 0.6 to 0.7 Å, indicating a high degree of structural similarity primarily in the core region. As expected, due to the higher sequence

variation in the surface-loop regions, a greater degree of structural and charge variation is observed in the surface loops (Murakami and Arni 2005).

The principal structural difference between SVSPs and other members of the S1 family is that the SVSPs possess an extended C-terminal tail which is linked by an additional disulfide bridge and is considered important for stability and for allosteric regulation (Murakami and Arni 2005; Parry et al. 1998). Whereas the vitamin-K-dependent mammalian serine proteases require Na^+ ions for optimal catalytic activity and selectivity and for allosteric regulation, this site is not present in SVSPs suggesting a more simple and straightforward mechanism of substrate recognition and binding (Kraut 1977).

SVSPs often contain about 20 % carbohydrates, N-linked to Asn, principally glucosamine, neuraminic acid, and neutral hexoses (Murakami and Arni 2005). In the structures of ACC-C from *Agkistrodon contortrix contortrix* and AaV-SP-I and AaV-SP-II from *Agkistrodon acutus*, these carbohydrate moieties are strategically positioned around the entrance to the active site at the tips of the 37-, 99-, and 148-loops and probably modulate macromolecular selectivity. However, in other SVSPs (e.g., in *T. stejnegeri*-TSV-PA), Asn178 is located on the opposite face and apparently does not participate in recognition, selectivity, or binding of the substrate at the interfacial site.

Active Site

Looking down into the active site cleft (Fig. 2a), the peptide to be cleaved would extend from the north to the south in this cleft (Figs. 2a and c). The catalytic triad is perfectly positioned by the union of the N-terminal lobule containing His57 and Asp102 (Fig. 2a) and the C-terminal lobule containing Ser195. This region is surrounded by the conserved 70-, 148-, and 218-loops as well as the nonconserved 37-, 60-, 99-, and 174-loops (Fig. 2a and c). As in the mammalian counterparts, the catalytic residue, His57, possesses a nonoptimal N δ 1-H tautomeric conformation which is essential for catalysis. The catalytic triad is supported by an extensive hydrogen-bonding network formed between the N δ 1-H of His57 and O δ 1 of Asp102, as well as between the OH of Ser195 and the N ϵ 2-H of His57. The hydrogen bond present between the latter pair is disrupted upon protonation of His57. Recent studies suggest that Ser214, which was once considered essential for catalysis, only plays a secondary role (Epstein and Abeles 1992; McGrath, et al. 1992). Hydrogen bonds formed between O δ 2 of Asp102 and the main-chain NHs of Ala56 and His57 are structurally important to ensure the correct relative orientations of Asp102 and His57.

Oxanion Hole and Subsites

A salient feature observed in the high-resolution structures of trypsin/chymotrypsin-like enzymes is the presence of a narrow channel referred to as an oxanion hole formed by the backbone NHs of Gly193 and Ser195 (Birktoft and Blow 1972;

Birktoft et al. 1976) (Fig. 2b). These atoms contribute to the formation of a positively charged pocket that activates the carbonyl of the susceptible scissile peptide bond, and the residual positive charge additionally stabilizes the negatively charged oxyanion of the tetrahedral reaction intermediate. The oxyanion hole is in close proximity and is structurally linked to the catalytic triad and the Ile16–Asp194 salt bridge via Ser195. The serine proteinases are further characterized by the presence of a number of surface regions or patches referred to as subsites or secondary substrate binding sites that ensure the binding of the substrate and the perfect positioning of the susceptible scissile bond. The primary factors that determine specificity are the interactions at the S1/P1 and S1'/P1' and also the S2/S2' and S3/S3' sites. The specificity of the mammalian chymotrypsin-like serine proteinases is primarily determined by the P1–S1 interaction. The S1 site or oxyanion is a hole located adjacent to the highly reactive Ser195 and is formed by a shallow pocket lined by the residues 189–192, 214–216, and 224–228. Residues 189, 216, and 226 determine specificity. In chymotrypsin, the oxyanion hole contains Ser189, Gly216, and Gly226 which determines its specificity for hydrophobic residues at the S1 subsite. In trypsin-like enzymes, Ser189 is substituted by Asp189, and hence, these enzymes display a marked preference for substrates containing Arg or Lys residues at the S1 subsite.

Due to the presence of Asp189, Thr190, and Gly217, as expected, benzamidine and benzamidine derivatives bind in the specificity pocket of *Agkistrodon contortrix* protein C activator (Murakami and Arni 2005) as illustrated in Fig. 2b.

Based on the above criteria, SVSPs can be classified as trypsin-like enzymes that possess highly conserved S1 subsites. However, since they display high selectivity towards macromolecular substrates, additional structural features need to be taken into consideration. Thus, the observed substrate specificity is not determined entirely by the interactions at the aforementioned subsites but likely involves more distant structural features.

The sequence alignments presented in Fig. 1 indicate that SVSPs are often glycosylated at different amino acid positions (Murakami and Arni 2005; Parry et al. 1998). In Protac (ACC-C), three carbohydrate moieties which are strategically positioned at the tips of the 37-, 99-, and 148-loops that surround the entrance to the active site cleft and extend outward could play a role in restricting access to the active site by macromolecular substrates, and it is speculated that they could play important roles in the modulation and expression of selectivity towards macromolecular substrates (Fig. 2b). Two snake venom serine proteinase isoforms from *Agkistrodon acutus*, AaV-SP-I and AaV-SP-II, also possess an N-linked carbohydrate group (Asn35) that likely interferes with the binding of macromolecular inhibitors and substrates (Zhu et al. 2003). In the case of TSV-PA, the enzyme has a unique glycosylation site at the Asn178 residue located on the opposite face and apparently does not play a role in the binding of macromolecular substrates at the interfacial site (Parry et al. 1998). Another key structural element implicated in the functional differentiation in SVSPs is the electrostatic surface potential, as calculated by PDB2PQR server for charges and radii assignments (Dolinsky et al. 2007) and APBS software for solving the Poisson–Boltzmann equation (Baker et al. 2001).

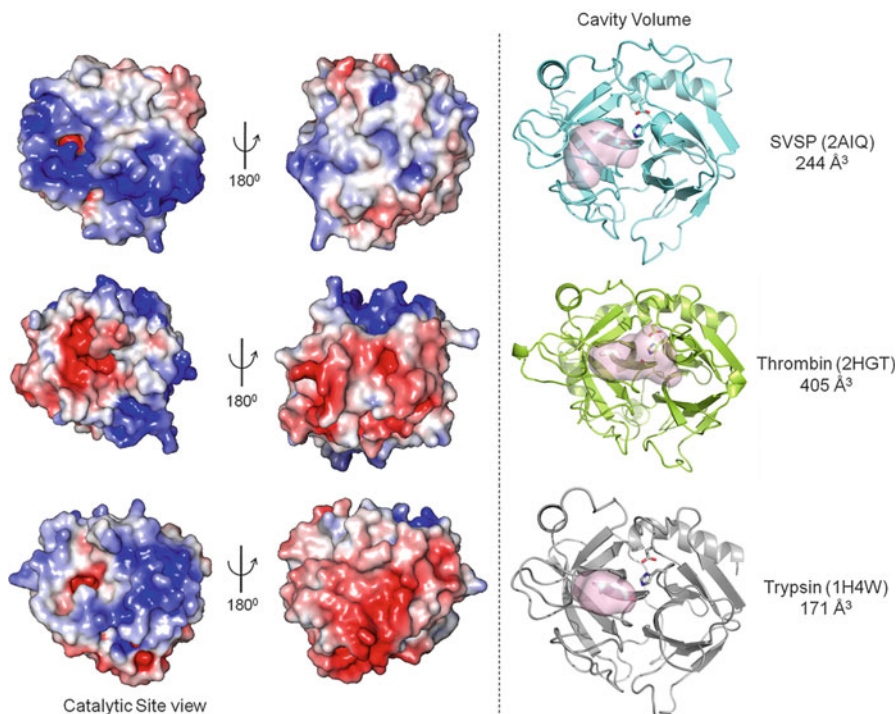


Fig. 3 *Left panel.* Electrostatic surface potential representation when viewed looking into the active site and when rotated 180°. *Right panel,* active site cavity volume of SVSP (PDB ID:2AIQ), human α -thrombin (PDB ID: 2HGT), and trypsin (PDB ID 1H4W). Visualization by The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC

It has been suggested (Murakami and Arni 2005) that the charge around the interfacial surface of Protac mimics the thrombin–thrombomodulin complex presenting high electrostatic affinity for the Asp/Glu propeptide of protein C (Fig. 3). Comparisons of the surface charges and active site cavity volumes of ACC-C (*Agkistrodon contortrix contortrix*), human α -thrombin, and trypsin (Fig. 3) indicate significant differences. The interface that contains the active sites is more negatively charged in thrombin and trypsin than in the SVSP (ACC-C). The active site cavity volume (Fig. 3, right panel) is 405 Å³ in thrombin, 171 Å³ in trypsin, and 244 Å³ in the SVSP (*Agkistrodon contortrix contortrix* protein C activator).

Mechanism of Catalysis

Based on both biochemical and structural data of enzyme substrate and enzyme transition state substrate analogues, the catalytic mechanism of hydrolysis by serine peptidases has been very well characterized (Birktoft et al. 1976), and the steps involved in the acid–base catalytic mechanism can be inferred. In the first step,

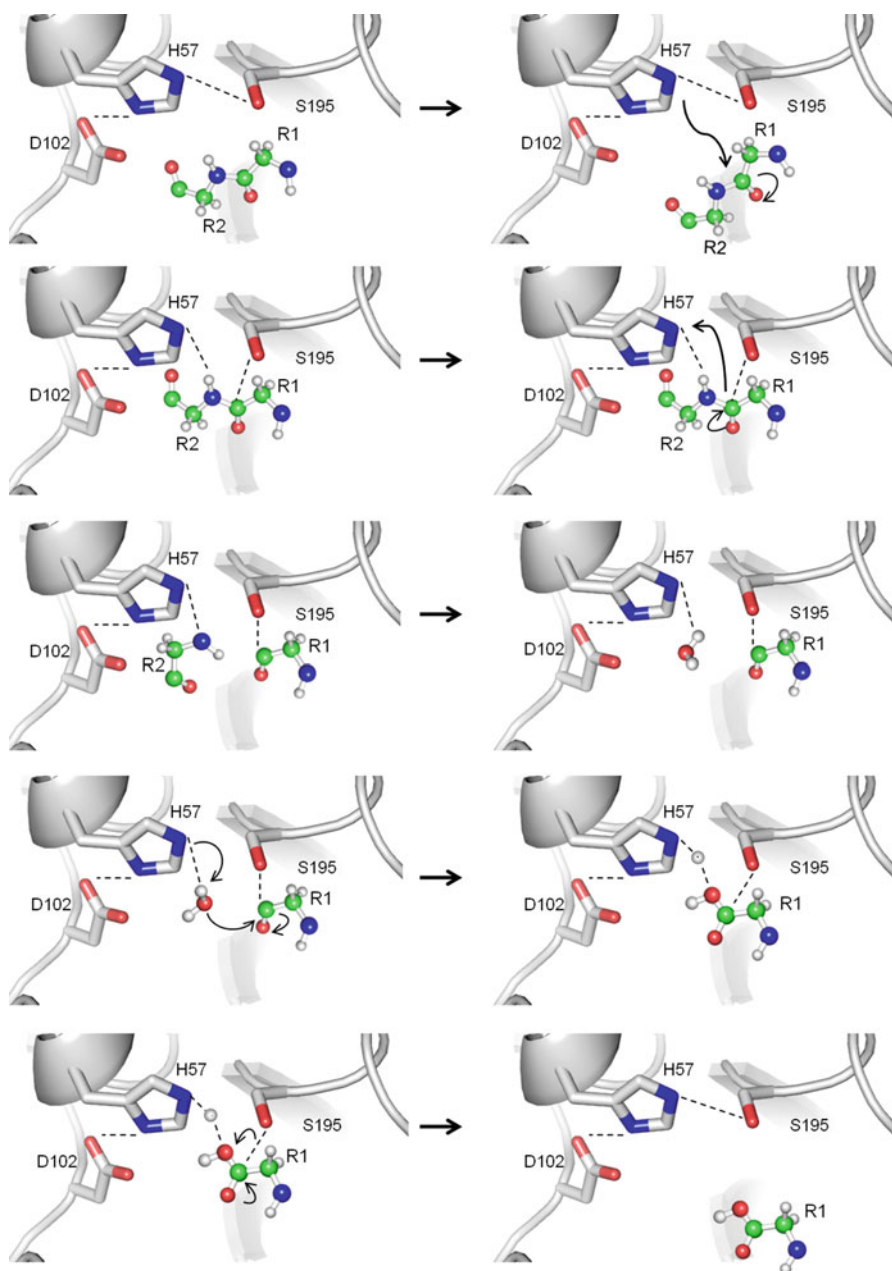


Fig. 4 The steps in the binding and hydrolysis of peptides by serine peptidases (see text for details)

Ser195 initiates the attack on the carboxyl group of the peptide. This reaction is subsequently assisted by His57 which, in turn, plays the role of a general base to form the tetrahedral intermediate. The transition state intermediate is stabilized by interactions formed with the main-chain NHs of the amino acids forming the oxyanion hole. Following the collapse of the tetrahedral intermediate and the expulsion of the leaving group, His57-H⁺ plays the role of a general acid and the acyl-enzyme intermediate is formed. In the second step of the reaction, His57 deprotonates a water molecule which then interacts with the acyl-enzyme complex to yield a second tetrahedral intermediate. The collapse of this second intermediate results in the liberation of the carboxylic acid product.

Zymogen Activation

Mammalian serine proteinases participating in digestion and the blood coagulation cascade are synthesized as inactive zymogens and activation requires the cleavage of the N-terminal peptide and additional cleavages in the regions 142–152, 184–193, and 216–223 (Bode and Huber 1978). This autocatalytic cleavage and subsequent removal of the N-terminal peptide results in the formation of a salt bridge between the newly formed N-terminus and Asp194 and causes dramatic structural changes in both the S1 subsite and the oxyanion hole (Huber and Bode 1978; Bode et al. 1978).

Since neither the activity of SVSP zymogens nor their structures have been determined, only the molecular mechanism involved in the maturation process can be inferred. It is presumed that in the SVSPs, as in the case of trypsin, the S1 subsite and oxyanion hole are only formed upon cleavage and removal of this peptide since the N-terminal portion is conserved in snake and mammalian enzymes. Thus, as in the other serine proteinases, the loss of proteinase activity at high pH probably results from the deprotonation of the N-terminus and the disruption of the salt bridge, shifting the conformational equilibrium to resemble the inactive zymogen-like conformation (Hedstrom et al. 1996).

Snake Venom Metalloproteinases

Based on their primary structures and the configuration of their catalytic sites, zinc proteases are subdivided into the gluzincin, metzincin, inuzincin, carboxypeptidase, and DD carboxypeptidase subgroups (Hooper 1994). The metzincin subgroup is further divided into serralysins, astacins, matrixins, and adamalysins (Stöcker et al. 1995; Bode et al. 1996).

The metzincin superfamily of metalloproteinases which contain a conserved Met residue in a β -turn downstream of the zinc-binding motif includes four protein families, and the snake venom metalloproteinases (SVMs) are leading members of

the repolysin family. These enzymes, probably the most widely distributed venom proteinases, are encountered in both crotalid and viperid venoms and often constitute over 50 % of the total protein in Viperidae venoms (Calvete et al. 2007; Takeda et al. 2012). SVSPs are primarily hemorrhagic but fibrin(ogen)olytic (Retzios and Markland 1988); inhibition of blood platelet aggregation (Kamiguti et al. 1996; Moura da Silva et al. 2008) and other activities have been reported. Snake venom metalloproteinases (SVMPs) are synthesized as inactive precursors in the cytoplasm of secretory cells and are converted into active enzymes by proteolysis of a peptide bond and subsequent liberation of the short, N-terminal propeptide. Central to proteolysis is a metal ion, specifically a Zinc ion, which is coordinated by three His residues and one or two solvent molecules.

Domain Organization and Sequence Homology

Domain Organization

SVMPs range in size from 20 to about 100 kDa, and due to their diversity, a number of classification criteria currently exist. Both from a structural and domain organizational points of view, these enzymes display sequence and domain similarities to the ADAMs (A Disintegrin And Metalloproteinase) family of metalloproteinase proteins and can be grouped into four principal classes (Fig. 5). Apart from the metalloproteinase domain, these enzymes often contain other regulatory domains (Fig. 5). SVMPs belonging to the P-I class are the simplest and only contain a single domain, the zinc-dependent catalytic domain. Members of the P-II class additionally contain a second domain, the small, highly flexible disintegrin domain. The three-domain P-III class contains the aforementioned P-II domains and a cysteine-rich domain. The heterotetrameric P-IV class proteinases are the most complex and consist of the domains of the P-III enzymes linked by a disulfide bridge to a lectin-like domain (Fig. 5). The ADAMs contain additional domains such as the EGF (epidermal growth factor) and transmembrane (TM) domains (Fig. 5).

SVMPs contain a short 18-amino-acid hydrophobic signal peptide. The metalloproteinase or catalytic domains of SVMPs whose structures have been determined contain approximately 215 amino acids (Fig. 6) and indicate a high degree of identity with few, very short insertions. The zinc-binding motif HEBxHxBGBxHD, where B represents a bulky hydrophobic residue and X indicates any residue, is highly conserved (Fig. 6).

The metalloproteinase domain is linked to the disintegrin domain via a 13–15-amino-acid spacer peptide (Fig. 5).

The disintegrin and disintegrin-like domains show a fair degree of variation in the lengths of the polypeptide chain and the loop regions and hence are classified as short (49–51 amino acids), medium (~70 amino acids), and long (~84 amino acids) and contain 4, 6, or 7 disulfide bridges, respectively. The RGD sequence is contained in a short loop that is stabilized by disulfide bridges.

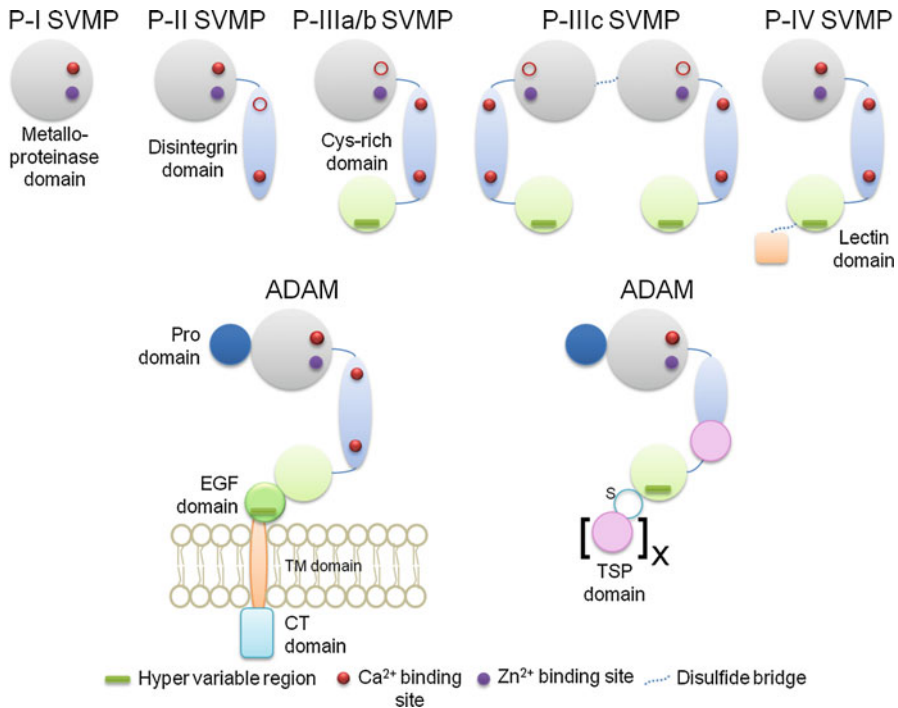


Fig. 5 Domain organization of SVMPs and ADAMs

As the name suggests, the cysteine-rich domain which is about 112 amino acids is rich in cysteines. Since structural information concerning the cysteine-rich and C-type lectin domains is limited or absent, these domains will not be discussed in detail.

Crystal Structure

As expected, due to the high identity, the three-dimensional structures of the metalloproteinase domains are very similar (Fig. 8a). The structure can be described as ellipsoidal, α/β , two-domain structure; the shallow, catalytic cleft is created by the junction of the domains. The major (M, amino acid residues 1–152, sequences refer to *Bothrops asper*, BaP 1) domain consists of four α -helices (A–D) and a five-stranded β -sheet. Strands I, II, III, and V are situated parallel to each other, and strand IV is positioned in an antiparallel orientation relative to the others. The minor, S (residues 153–202) domain is formed by a α -helix and several loops. Three disulfide bridges stabilize the structure, and the bridge Cys 117–Cys 197 links the M and S domains, and Cys 159–Cys 181 and Cys 157–Cys 164 are located within the minor subdomain. Whereas adamalysin II and atrolysin C contain only two disulfide bridges, BaP1 contains a third disulfide bridge

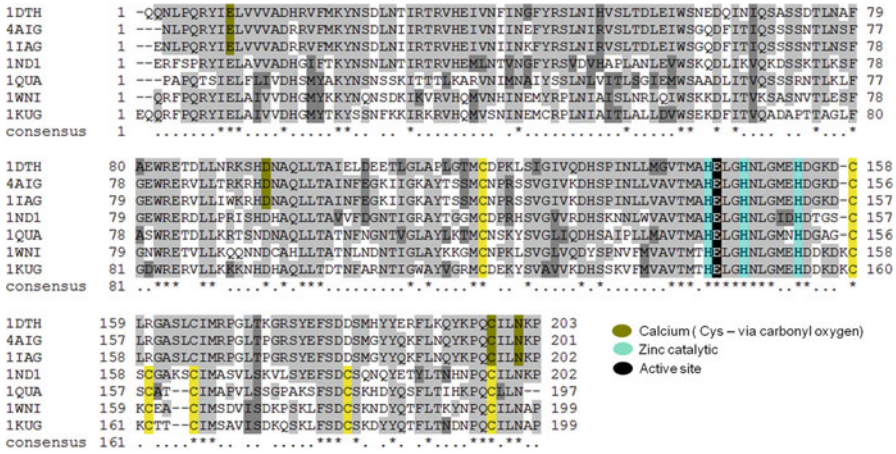


Fig. 6 Multiple sequence alignment of SVMP metalloproteinase domains. PDB id four-character codes are given on the left. Cys residues are in yellow and members of the two-disulfide family are the upper three sequences (1DTH, 4AIG, and 1IAG) and the three-disulfide family are the following four sequences (1ND1, 1QUA, 1WNI, and 1KUG). The active site Glu residue and the Met are boxed in black. The three zinc-binding residues are in blue, whereas the amino acids that bind calcium are in dark yellow. The zinc-binding motif is HEBxHxBGBxHD where B = bulky hydrophobic residue and X indicates any residue

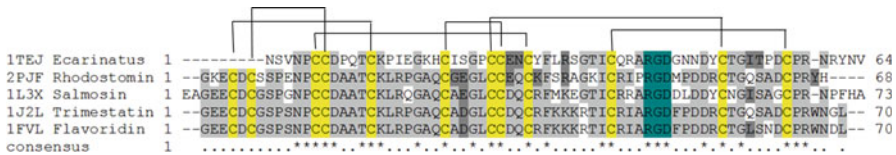


Fig. 7 Multiple sequence alignment of the snake venom disintegrin domains. Cys residues are in yellow and the disulfide bridges are indicated by lines. The RGD sequence is boxed in green. Four-character code on the left indicates PDB id

(Cys 159–Cys 181) (Watanabe et al. 2003). The N- and C-termini are both located on the same side of the molecule and N-terminal amino acid is often a pyroglutamate (Watanabe et al. 2003).

Central to the active site is the well-characterized zinc-binding motif HEXXHXXGXXH (where X is any amino acid) and a methionine-turn (Fig. 4a) located in a shallow depression which also contains the three-catalytic histidines (His142, His146, and His152). This met-turn positions His152 such that it can also coordinate the Zn²⁺ ion and the side chain of Met166 serves as the hydrophobic base for the active site groups. The zinc ion is coordinated tetrahedrally by the Ne2 nitrogen atoms of the three histidines and one water molecule is encountered bound to Glu143 (Fig. 8a, c) in the crystal structure of BaP1 (Watanabe et al. 2003), whereas

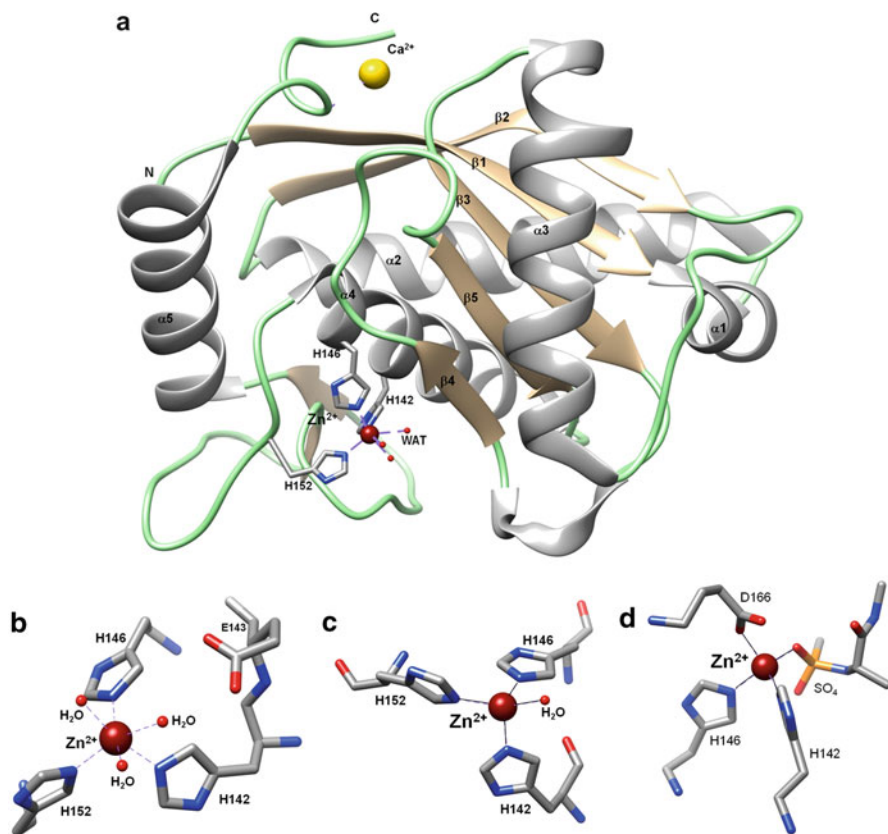


Fig. 8 (a) Ribbon representation of the crystal structure of the metalloproteinase domain from SVMP. (b), (c), and (d) present details of the hydrogen bonds formed with the zinc ion

it is octahedrally coordinated by additional solvent molecules in the crystal structure of BmoMP α -I (Akao et al. 2010) (Fig. 8b). Figure 8c indicates the interactions when BaP1 is complexed with a peptidomimetic inhibitor (Lingott et al. 2009).

Catalytic Mechanism

The catalytic mechanism for bacterial metalloproteinase thermolysin has been proposed, and based on the structural similarities, it can be inferred that (a) the Zn^{2+} ion orchestrates all steps of catalysis from peptide hydrolysis, stabilization of the reaction intermediates to release. In the initial step, Zn^{2+} is tetrahedrally coordinated by the aforementioned His142, His146, and His152 and a solvent water molecule. During catalysis, Zn^{2+} is pentahedrally coordinated with the participation of the substrate carbonyl oxygen atom (Fig. 9).

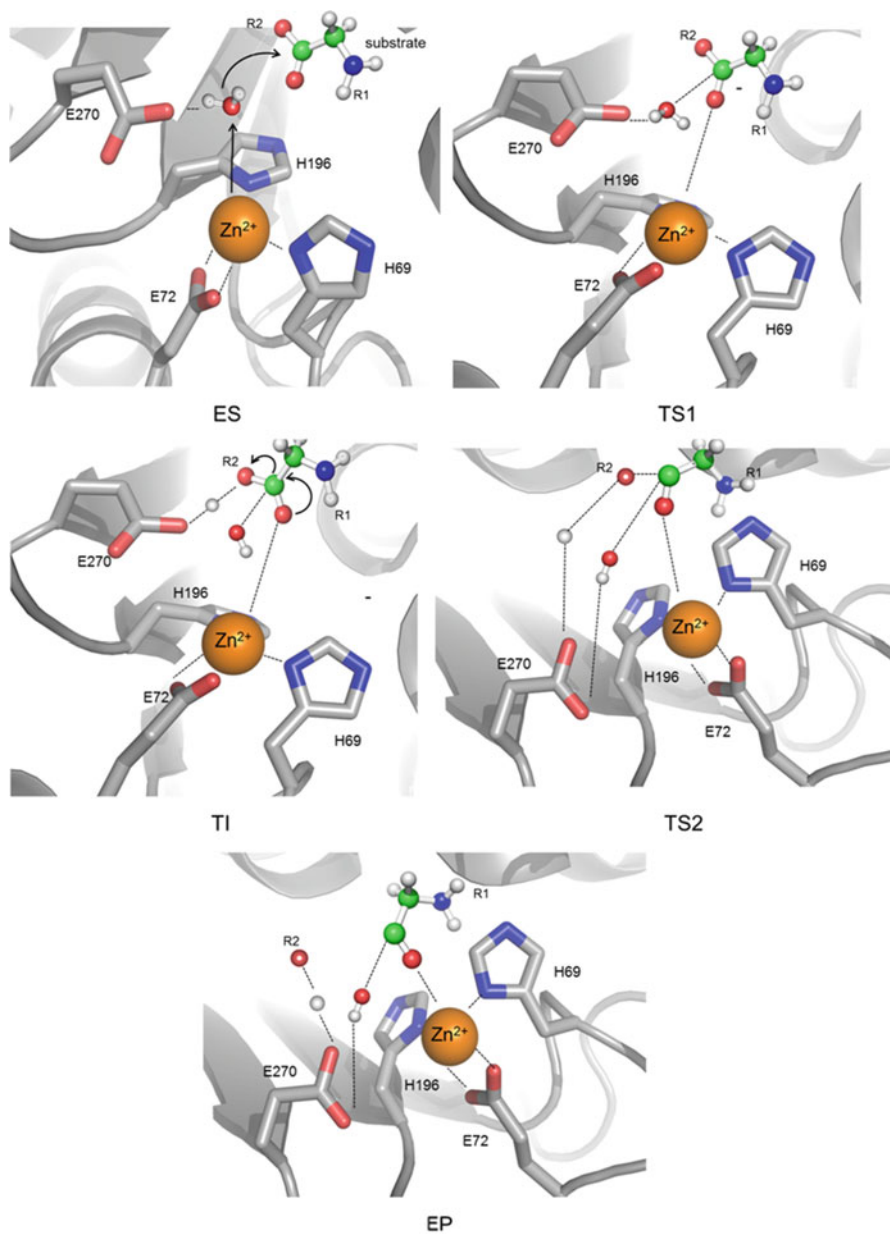


Fig. 9 Steps in the catalysis catalytic mechanism of Zn²⁺ metalloproteinases. ES enzyme + substrate, TS1 first transition state, TI tetrahedral intermediate, TS2 second transition state, and EP enzyme product complex

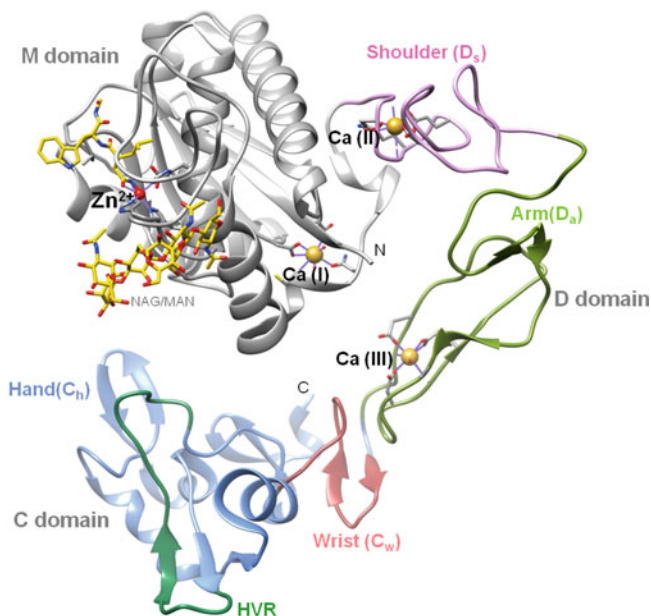


Fig. 10 The domain structure of SVMPs. NAG/NAM = carbohydrate moiety

Glu143 polarizes the catalytic water molecule by abstracting a proton, and this polarized water molecule, in turn, initiates a nucleophilic attack on the carbonyl carbon of the susceptible scissile peptide substrate bond. Glu143 transfers the proton abstracted from a water molecule to the amide leaving group. Site-directed mutagenesis demonstrates that the Glu143Asp substitution negates activity.

The hydrophobic S1' site is conserved in many SVMPs although its depth might vary and in the structure of Bap1, it is lined with Phe178, Val138, Ala141, Tyr176, and Ile165.

The calcium ion located on the surface is implicated in the stabilization of the structure and is coordinated by the carbonyl oxygen atom of Glu9, carboxylate atoms of Asp93, carbonyl oxygen atom of Cys197, and the carboxamide oxygen of Asn200 and a solvent water molecule. It has been suggested (Gomis-Ruth et al. 1994) that this calcium ion likely plays a role in stabilizing the structures of multi-domain SVMPs.

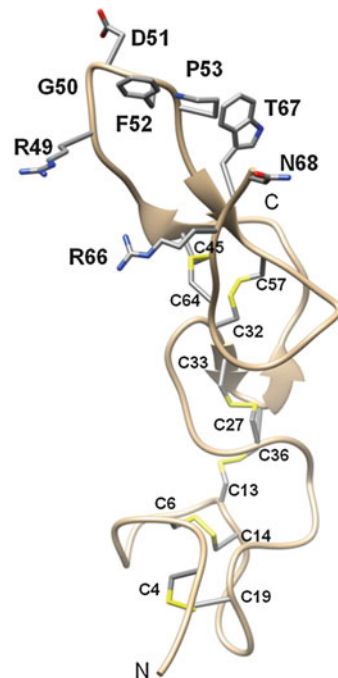
As mentioned earlier, SVMPs often contain additional domains such as the disintegrin and cysteine-rich domain that are linked to the metalloproteinase domain by short peptides (Fig. 10).

Disintegrin Domain

Disintegrins are small, cysteine-rich proteins that vary in mass from 4 to 14 kDa and bind to transmembrane proteins via an RGD motif located at the tip of a protruding loop (Figs. 7 and 11). The RGD-containing disintegrins are potent inhibitors of the platelet fibrinogen receptors, $\alpha_{\text{IIb}}\beta_3$ integrins, whereas the non-RGD disintegrins do not possess this activity. Based on the length of the polypeptide chain and the disulfide bonding pattern, the disintegrins have been classified into the following five groups (Calvete et al. 2003; Calvete 2013): group 1, 41–51 residues and four disulfide bonds; group 2, about 70 amino acids and six disulfide bridges; group 3, about 80 amino acids and seven disulfide bridges; and group 4, contains about 100 amino acids and 8 disulfide bridges and a C-terminal cysteine-rich domain cross-linked by six disulfide bridges (Calvete et al. 2003). The class 4 disintegrins are homodimeric and heterodimeric disintegrins with 67 amino acids and 10 disulfide bridges. The PII and PIII SVMs contain the group 3 and group 4 disintegrins, respectively.

The NMR structures indicate that the loop regions, especially the C-terminus, are very flexible in solution although they contain six disulfide bridges. The disintegrin secondary structure is characterized by β -turns and short antiparallel β -sheets held together by six conserved disulfide bridges. A disulfide bridge determines the orientation of the long, irregular, hairpin loop that contains the RGD sequence at its tip. Due to the presence of a Gly, the side chains of Arg and Asp

Fig. 11 Ribbon representation of the structure of a disintegrin (Trimestatin, *Trimeresurus flavoviridis*, Fujii et al. 2003)



point in different directions. Homodimeric disintegrins are stabilized by four intrachain disulfide bonds and two N-terminal interchain disulfide bridges.

Phospholipases A₂

Phospholipases A₂ (PLA₂; phosphatide sn-2 acylhydrolase, EC 3.1.14) specifically catalyze the hydrolysis of the ester bond at the sn-2 position (sequentially numbered position 2) of glycerophospholipids generating fatty acids and lysophospholipids; their catalytic activity results in the release of arachidonic acid, a precursor of eicosanoids, which is implicated in triggering inflammatory reactions (Kudo et al. 1993). These enzymes are very widely distributed, display enhanced activity towards lipids in lamellar and micellar aggregates both in membranes and at other lipid–water interfaces (Jain et al. 1995; Ramirez and Jain 1991), and have been extremely well studied from the structural point of view, and over a hundred crystal structures of the class I and class II enzymes both with and without substrates, substrate analogues, inhibitors, fatty acid analogues, and other molecules have been determined, and a significant number of excellent, detailed reviews are currently available so this discussion will only present the salient structural features that are necessary to understand the protein fold and the steric and charge characteristics of the catalytic and calcium-binding sites.

Relevant to the study of snake venoms are the small, highly homologous, calcium-dependent secreted (sPLA₂s) PLA₂s with molecular masses ranging from 12 to 14 kDa (119 to 143 amino acids) that are divided into three classes (classes I, II, and III), based on their amino acid sequence and disulfide bonding patterns (Renetseder et al. 1985). The class I enzymes are encountered in *Elapidae* and *Hydrophiidae* snake venoms and mammalian pancreas; the class II PLA₂s are encountered in *Crotalidae* and *Viperidae* venoms and mammalian non-pancreatic tissues. The class III enzymes are almost exclusively found in lizard and bee (*Apis mellifera*) venoms. The interest in these enzymes stems from the fact that apart from their primary catalytic function, snake venom PLA₂s often display additional pharmacological activities such as hemorrhagic, myotoxic, hemolytic, edema-inducing, hypotensive, presynaptic, and postsynaptic neurotoxicity activities (Georgieva et al. 2008). Considering the vast amount of biochemical and structural data available, it is certainly very tempting to attempt to correlate biological activity with structural features, and indeed, a number of attempts have been made to delineate the sequence and structural elements or regions responsible for the specific activities cited earlier (Perbandt et al. 2003; Georgieva et al. 2004; Georgieva et al. 2012).

Primary Structure

The protein sequence databases contain a large number of partial and complete PLA₂ amino acid sequences, so the sequence comparisons in this article will be limited to a few selected examples of class I and II PLA₂s, whose crystal structures

have also been determined. The number scheme adopted is based on the homology numbering scheme with reference to the sequence of the bovine pancreatic PLA₂ (Dufton and Hider 1983; Renetseder et al. 1985).

The two structural criteria used to classify these enzymes as belonging to either class I or II are:

1. The positions of the seven disulfide bridges, class I PLA₂s have a disulfide bridge Cys11–Cys77, whereas class II PLA₂s have a disulfide bridge Cys51–Cys133. The other six disulfide bridges are generally conserved in the two classes.
2. The elapid loop, a two- or three-amino-acid insertion and the five-amino-acid insertion in the pancreatic loop of mammalian pancreatic PLA₂s are features of class I PLA₂s. On the other hand, the C-terminus of class II enzymes has 5–7 additional amino acids and the aforementioned Cys51–Cys133 disulfide bridge links this extension to the main body of the protein.

Whereas the class I and II enzymes share high sequence identity, the class III enzymes show greater sequence diversity, and since the class III enzymes are not encountered in snake venoms, it will not be further discussed except to point out that although the structures are radically different, their catalytic sites are similar.

Secondary Structure

Since a number of manuscripts have presented the results of the sequence alignments of PLA₂s and attempts have been made to correlate sequence information with biological activity, here only the alignment as a means of understanding the structure has been presented. The amino acids His47 and Asp99 that participate in catalysis are highly conserved (Fig. 12). However, Asp49, a key amino acid essential for calcium ion binding, is often replaced by Lys, Ser, or Arg. Thus, these natural mutants are unable to bind calcium and hence are catalytically inactive (Arni and Ward 1996).

Tertiary Structure

The predominant structural feature of the class I/II enzymes is a motif or platform formed by the relative positions of the two long, antiparallel helices (helices 2 and 3, residues 37–54 and 90–109) α -helices (Arni and Ward 1996). Two disulfide bridges ensure the maintenance of the helix axial distance of about 10 Å. An analysis of the hydrophobicity of the amino acids in these two helices does not indicate the existence of a clear amphipathic character. However, in general, the hydrophilic amino acids are exposed to the solvent, whereas the hydrophobic amino

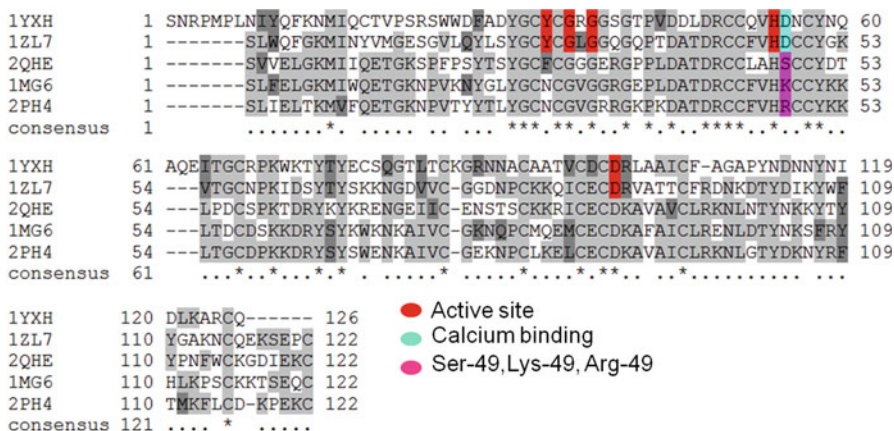


Fig. 12 Structure-based sequence alignment of PLA₂s amino acids forming the active site, calcium-binding site, and the position of Ser/Lys/Arg49 are indicated in red, cyan, and pink, respectively. Note: Although His47 and Asp99 are conserved in the analytically inactive enzymes, they are not marked in red. The four-character PDB ids are given on the left, 1YXH (*Naja naja sagittifera*), 1ZL7 (*Bothrops jararacussu*), 2QHE (*Echis carinatus sochureki*), 1MG6 (*Agkistrodon acutus*), and 2PH4 (*Zhaermia mangshanensis*)

acids are sequestered from the solvent. This rigid platform includes the amino acids forming the catalytic network and calcium ion-binding sites, His48, Asp49, Tyr52, and Asp99. This structural motif is highly conserved both within and between the class I/II PLA₂s. The short antiparallel β-sheet, referred to as the β-wing, is rigidly held in position by a disulfide bridge in class I enzymes. Since this disulfide bridge is absent in the class II enzymes, this region is flexible and can adopt multiple conformations. The extended loop following helix 2 is referred to as the “elapid” or “pancreatic” loop (Fig. 13). This loop is much shorter in the class II enzymes.

Although the class I and II enzymes show some structural differences (Fig. 13a, b), upon superpositioning the atomic coordinates, it is observed that the active sites are analogous (Fig. 14).

Calcium is an essential cofactor (Teshima et al. 1989) and is coordinated by the carboxylate oxygen atoms of Asp49 (Fig. 15a) and three main-chain carbonyl oxygen atoms from the residues forming the calcium-binding loop (region 25–33). Two structurally conserved solvent water molecules complete the coordination sphere of the Ca²⁺ ion by forming a pentagonal bipyramid. As expected, this calcium-binding loop is perfectly positioned in relation to the catalytic site by a disulfide bridge (Cys27–Cys44) that links the second helix to the calcium-binding loop (Fig. 15a).

Calcium binding in this region is only possible when position 49 is occupied by Asp. Substitutions of Asp49 by Glu, Ala, Asn, Gln, Lys, Ser, or Arg which are encountered naturally in snake venoms result in inactive enzymes (Arni and Ward 1996).

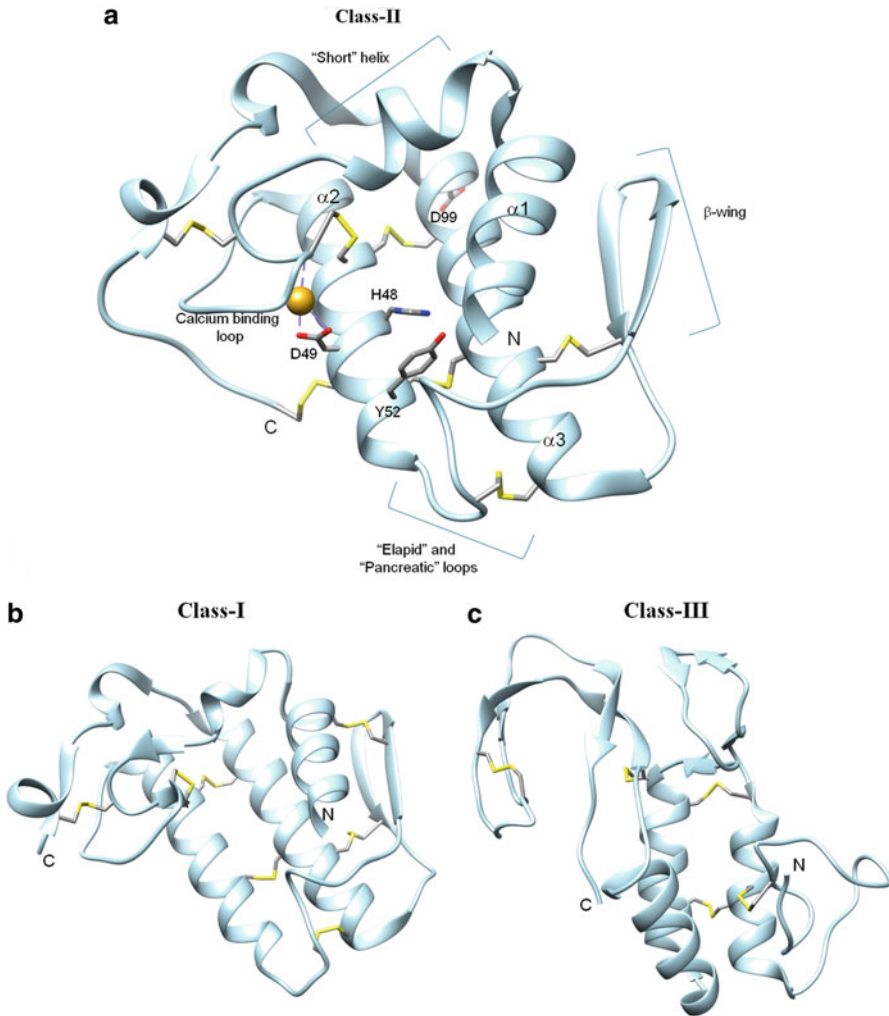


Fig. 13 Ribbon representation of the main structural features of (a) class II, (b) class I, and (c) class III PLA₂s. The positions of the disulfide bridges are indicated. In (a), the main features are labeled and amino acids involved in catalysis and Ca²⁺ binding are included

The Hydrophobic Channel

A hydrophobic channel is necessary to conduct the lipid molecule to the catalytic site. The amino acids in the short N-terminal amphiphilic helix (Helix 1) are highly conserved and stabilized by a disulfide bridge (class I) or hydrogen bonds (class II) form part of the protein core (Arni and Ward 1996). The inner surface of this helix is

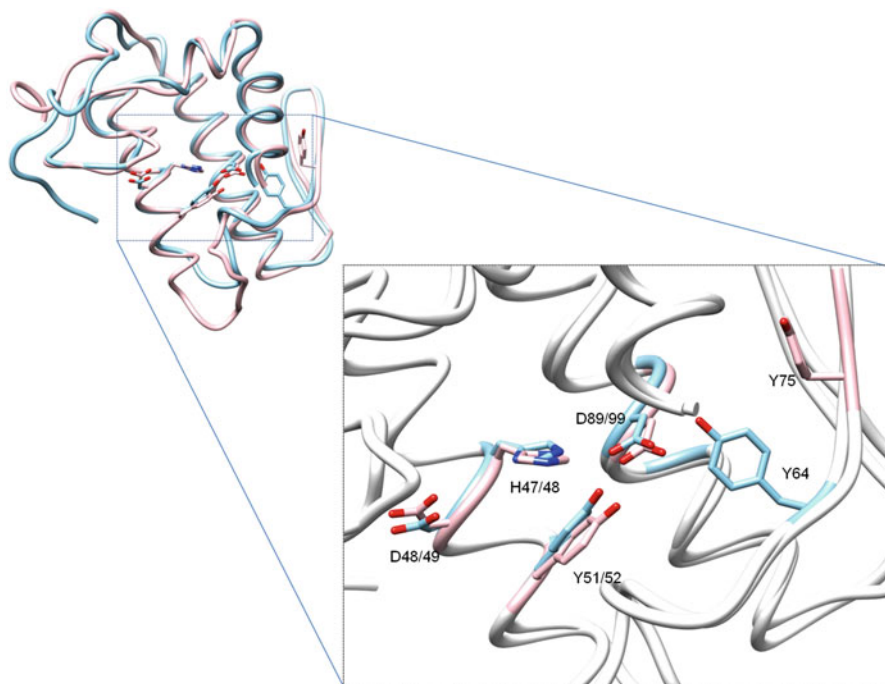


Fig. 14 Superpositioning of the atomic coordinates of class I (pink, PDB id 1YXH) and class II (blue, PDB id 1ZL7)

lined by hydrophobic amino acids that serve to line a sector of the channel and other amino acids from the single helical turn and the β -wing region complete the hydrophobic channel.

Catalytic Mechanism

The catalytic mechanism of the PLA₂s bears a strong resemblance to the mechanism involved in serine proteinases (Fig. 16). In this mechanism, the dyad + ion His48/Asp99/Ca²⁺ is the central component that triggers or initiates catalysis. The Ca²⁺ ion coordinated by a solvent water molecule polarizes the sn-2 carbonyl oxygen; His48 increases the nucleophilicity of the catalytic water via a second bridging water molecule.

Conclusion and Future Directions

Structural information obtained primarily by applying crystallographic techniques to understand the general shape, surface charge and binding pockets have been presented. An attempt has been made to correlate these results with

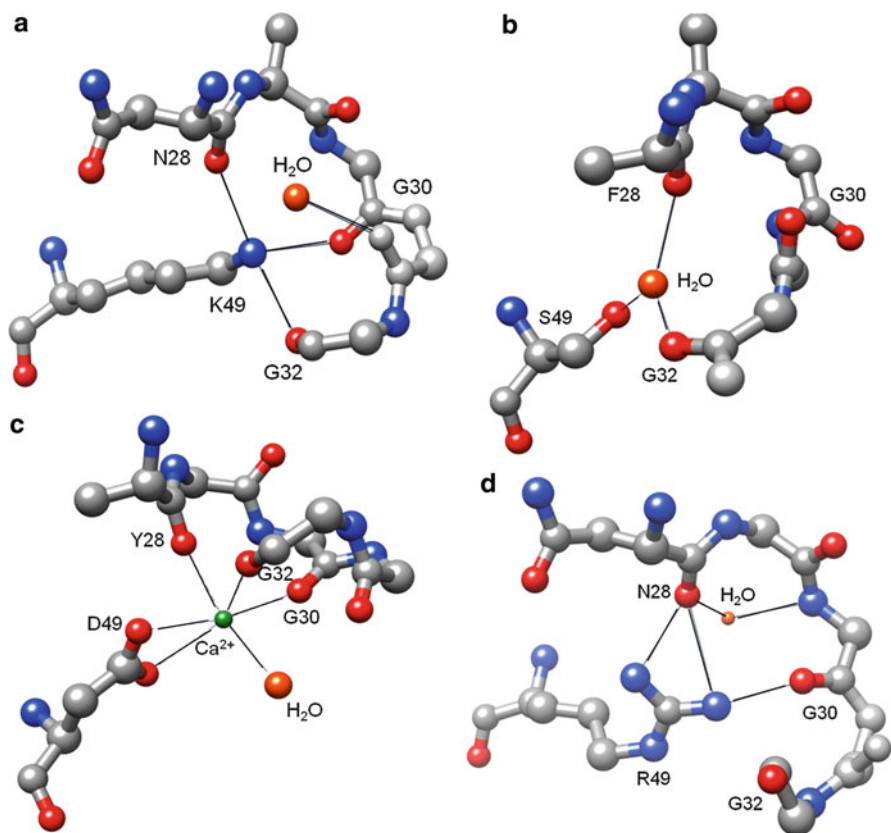
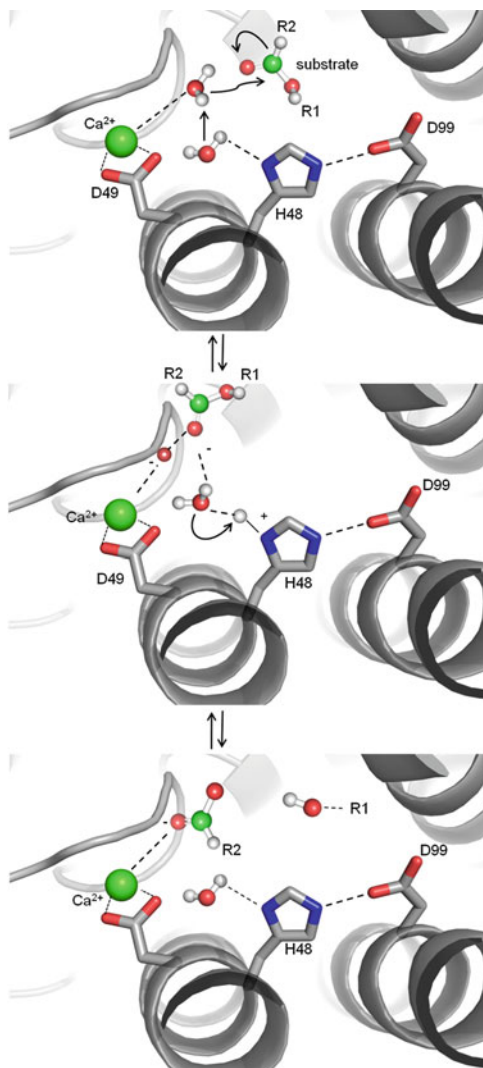


Fig. 15 Conformations of the calcium-binding loops when (c) Asp49 (PDB id 1ZL7) is substituted by (a) Lys (PDB id 1MG6), (b) Ser (PDB id 2QHE), and (d) Arg (PDB id 2PH4)

those obtained by other biophysical and biochemical techniques to understand enzyme specificity and mechanisms. Another powerful, albeit lesser used, technique to study biomolecular structure is nuclear magnetic resonance spectroscopy. The structure of the small (42 amino acids), highly basic (pI 10.3) peptide crotamine (Rádis-Baptista and Kerkis 2011) isolated from the venom of *Crotalus durissus terrificus* has been determined by both in solution (Fadel et al. 2005; Nicastro et al. 2003) and in the crystalline states (Coronado et al. 2012, 2013). Superpositioning the atomic coordinates obtained by these techniques (Fig. 17) indicates that the results obtained are similar and they complement each other since NMR techniques provide us with details of the flexibility of the molecule in solution, whereas crystallographic

Fig. 16 The catalytic mechanism of PLA₂s. See text for details



techniques provide more detailed information regarding the solvent structure and the binding of ions.

The application of novel NMR spectroscopy techniques to study the interactions of macromolecules with peptides, substrates, and other small molecules combined with high-resolution crystallographic techniques should provide us with deeper insights regarding the mode of action of snake venom proteins.

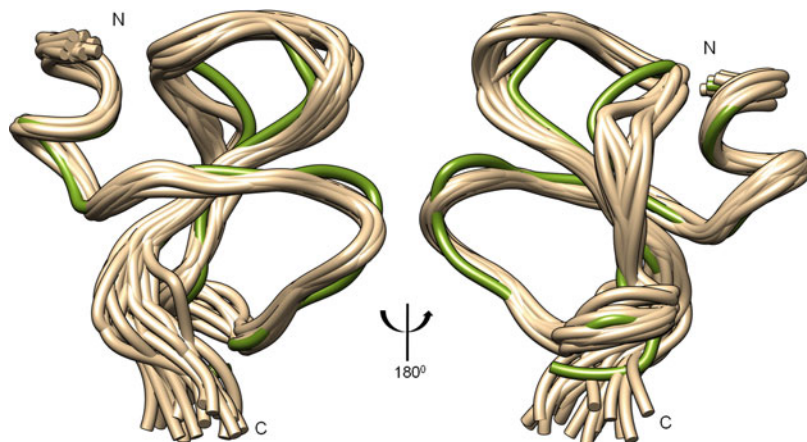


Fig. 17 Results of the superpositioning of the C α atoms of crotoamine (*Crotalus durissus terrificus*) obtained by applying NMR (beige, PDB id 1Z99) and crystallographic techniques (green, PDB id 4GV5)

Cross-References

- ▶ [Biological Activities and Assays of the Snake Venom Metalloproteinases \(SVMPs\)](#)
- ▶ [Structure-Function Relationship in Heterodimeric Neurotoxin PLA2s from Viperidae Snakes Inhabiting Europe, South America, and Asia](#)
- ▶ [Structure-Function Relationship of Modular Domains of P-III Class Snake Venom Metalloproteinases](#)

References

- Akao PK, Tonoli CC, Navarro MS, Cintra AC, Neto JR, Arni RK, Murakami MT. Structural studies of BmooMPalpha-I, a non-hemorrhagic metalloproteinase from Bothrops moojeni venom. *Toxicon*. 2010;55(2–3):361–8.
- Arni RK, Ward RJ. Phospholipase A2 – a structural review. *Toxicon*. 1996;34(8):827–41.
- Bajaj M, Blundell T. Evolution and the tertiary structure of proteins. *Annu Rev Biophys Bioeng*. 1984;13:453–92.
- Baker NA, Sept D, Joseph S, Holst MJ, McCamoon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A*. 2001;28;98(18):10037–41.
- Birktoft JJ, Blow DM. Structure of crystalline α -chymotrypsin. V. The atomic structure of tosyl- α -chymotrypsin at 2 Å resolution. *J Mol Biol*. 1972;68(2):187–240.
- Birktoft JJ, Kraut J, Freer ST. A detailed structural comparison between the charge relay system in chymotrypsinogen and in α -chymotrypsin. *Biochemistry*. 1976;15(20):4481–5.
- Bode W, Huber R. Crystal structure analysis and refinement of two variants of trigonal trypsinogen: trigonal trypsin and PEG (polyethylene glycol) trypsinogen and their comparison with orthorhombic trypsin and trigonal trypsinogen. *FEBS Lett*. 1978;90(2):265–9.

- Bode W, Schwager P, Huber R. The transition of bovine trypsinogen to a trypsin-like state upon strong ligand binding. The refined crystal structures of the bovine trypsinogen-pancreatic trypsin inhibitor complex and of its ternary complex with Ile-Val at 1.9 Å resolution. *J Mol Biol.* 1978;118(1):99–112.
- Bode W, Grams F, Reinemer P, Gomis-Rüth FX, Baumann U, McKay DB, Stöcker W. The metzincin-superfamily of zinc-peptidases. *AdvExp Med Biol.* 1996;389:1–11.
- Calvete JJ. The continuing saga of snake venom disintegrins. *Toxicon.* 2013;62:40–9.
- Calvete JJ, Moreno-Murciano MP, Theakston RD, Kisiel DG, Marcinkiewicz C. Snake venom disintegrins: novel dimeric disintegrins and structural diversification by disulphide bond engineering. *Biochem J.* 2003;372(Pt 3):725–34.
- Calvete JJ, Juárez P, Sanz L. Snake venomomics. Strategy and applications. *J Mass Spectrom.* 2007;42(11):1405–14.
- Chothia C, Lesk AM. The relation between the divergence of sequence and structure in proteins. *EMBO J.* 1986;5(4):823–6.
- Coronado MA, Georgieva D, Buck F, Gabdoulkhakov AH, Ullah A, Spencer PJ, Arni RK, Betzel C. Purification, crystallization and preliminary X-ray diffraction analysis of crotamine, a myotoxic polypeptide from the Brazilian snake *Crotalus durissus terrificus*. *Acta Crystallogr Sect F Struct Biol Cryst Commun.* 2012;68(Pt 9):1052–4.
- Coronado MA, Gabdoulkhakov A, Georgieva D, Sankaran B, Murakami MT, Arni RK, Betzel C. Structure of the polypeptide crotamine from the Brazilian rattlesnake *Crotalus durissus terrificus*. *Acta Crystallogr D Biol Crystallogr.* 2013;69(Pt 10):1958–64.
- Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, Baker NA. PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* 2007;(35):W522–W525.
- Dufton MJ, Hider RC. Classification of phospholipases A2 according to sequence. Evolutionary and pharmacological implications. *Eur J Biochem.* 1983;137(3):545–51.
- Epstein DM, Abeles RH. Role of serine 214 and tyrosine 171, components of the S2 subsite of alpha-lytic protease, in catalysis. *Biochemistry.* 1992;31(45):11216–23.
- Fadel V, Bettendorff P, Hermann T, De Azevedo Jr WF, Oliveira EB, Yamane T, Wüthrich K. Automated NMR structure determination and disulfide bond identification of the myotoxin crotamine from *Crotalus durissus terrificus*. *Toxicon.* 2005;46(7):759–67.
- Fernandes de Oliveira LM, Ullah A, Masood R, Zelanis A, Spencer PJ, Serrano SM, Arni RK. Rapid purification of serine proteinases from *Bothrops alternatus* and *Bothrops moojeni* venoms. *Toxicon.* 2013;76:282–90.
- Finkelstein AV, Ptitsyn OB. Why do globular proteins fit the limited set of folding patterns? *Prog Biophys Mol Biol.* 1987;50:171–90.
- Fujii Y, Okuda D, Fujimoto Z, Horii K, Morita T, Mizuno H. Crystal structure of trimestatin, a disintegrin containing a cell adhesion recognition motif RGD. *J Mol Biol.* 2003;332(5):1115–22.
- Gempeler-Messina PM, Volz K, Bühler B, Müller C. Protein C activators from snake venoms and their diagnostic use. *Haemostasis.* 2001;31(3–6):266–72.
- Georgieva DN, Rypniewski W, Gabdoulkhakov A, Genov N, Betzel C. Asp49 phospholipase A (2)-elaidoylamide complex: a new mode of inhibition. *Biochem Biophys Res Commun.* 2004;319(4):1314–21.
- Georgieva D, Arni RK, Betzel C. Proteome analysis of snake venom toxins: pharmacological insights. *Expert Rev Proteomics.* 2008;5(6):787–97.
- Georgieva D, Coronado M, Oberthür D, Buck F, Duhlov D, Arni RK, Betzel C. Crystal structure of a dimeric Ser49-PLA₂-like myotoxic component of the *Vipera ammodytes meridionalis* venom reveals determinants of myotoxicity and membrane damaging activity. *Mol Biosyst.* 2012;8(5):1405–11.
- Gomis-Ruth FX, Kress LF, Kellermann J, Mayr I, Lee X, Huber R, and Bode, W. Refined 2 Å X-ray crystal structure of the snake venom zinc-endopeptidase adamalysin II. Primary and tertiary structure determination, refinement, molecular structure and comparison with astacin, collagenase and thermolysin. *J Mol Biol.* 1994;(239):513–544.

- Hasson MS, Schlichting I, Moulai J, Taylor K, Barrett W, Kenyon GL, Babbitt PC, Gerlt JA, Petsko GA, Ringe D. Evolution of an enzyme active site: the structure of a new crystal form of muconate lactonizing enzyme compared with mandelate racemase and enolase. *Proc Natl Acad Sci USA*. 1998;95(18):10396–401.
- Hedstrom L. Serine protease mechanism and specificity. *Chem Rev*. 2002a;102(12):4501–24.
- Hedstrom L. An overview of serine proteases. *Curr Protoc Protein Sci*. 2002b; Chapter 21. ps2110s26.
- Hedstrom L, Lin TY, Fast W. Hydrophobic interactions control zymogen activation in the trypsin family of serine proteases. *Biochemistry*. 1996;35(14):4515–23.
- Harley BS, Shotton DM. *The Enzymes*. Academic Press, New York. 1971;1 Ed.
- Hooper NM. Families of zinc metalloproteases. *FEBS Lett*. 1994;354(1):1–6.
- Huber R, Bode W. Structural basis of the activation and action of trypsin. *Acc Chem Res*. 1978;11:114–22.
- Jain MK, Gelb MH, Rogers J, Berg OG. Kinetic basis for interfacial catalysis by phospholipase A2. *Methods Enzymol*. 1995;249:567–614.
- Kamiguti AS, Hay CR, Theakston RD, Zuzel M. Insights into the mechanism of haemorrhage caused by snake venom metalloproteinases. *Toxicon*. 1996;34(6):627–42.
- Kang TS, Georgieva D, Genov N, Murakami MT, Sinha M, Kumar RP, Kaur P, Kumar S, Dey S, Sharma S, Vrieling A, Betzel C, Takeda S, Arni RK, Singh TP, Kini RM. Enzymatic toxins from snake venom: structural characterization and mechanism of catalysis. *FEBS J*. 2011;278(23):4544–76.
- Kasuya A, Thornton JM. Three-dimensional structure analysis of PROSITE patterns. *J Mol Biol*. 1999;286(5):1673–91.
- Kini RM. Serine proteases affecting blood coagulation and fibrinolysis from snake venoms. *Pathophysiol Haemost Thromb*. 2005;34(4–5):200–4.
- Kraut J. Serine proteases: Structure and mechanism of catalysis. *Annu Rev Biochem*. 1977;46:331–58.
- Kudo I, Murakami M, Hara S, Inoue K. Mammalian non-pancreatic phospholipases A2. *Biochim Biophys Acta*. 1993;1170:217–231.
- Lesk AM, Chothia C. How different amino acid sequences determine similar protein structures: the structure and evolutionary dynamics of the globins. *J Mol Biol*. 1980;136(3):225–70.
- Lingott T, Schleberger C, Gutiérrez JM, Merfort I. High-resolution crystal structure of the snake venom metalloproteinase BaP1 complexed with a peptidomimetic: insight into inhibitor binding. *Biochemistry*. 2009;48(26):6166–74.
- McGrath ME, Vásquez JR, Craik CS, Yang AS, Honig B, Fletterick RJ. Perturbing the polar environment of Asp102 in trypsin: consequences of replacing conserved Ser214. *Biochemistry*. 1992;31(12):3059–64.
- Moura-da-Silva AM, Ramos OH, Baldo C, Niland S, Hansen U, Ventura JS, Furlan S, Butera D, Della-Casa MS, Tanjoni I, Clissa PB, Fernandes I, Chudzinski-Tavassi AM, Eble JA. Collagen binding is a key factor for the hemorrhagic activity of snake venom metalloproteinases. *Biochimie*. 2008;90(3):484–92.
- Murakami MT, Arni RK. Thrombomodulin-independent activation of protein C and specificity of hemostatically active snake venom serine proteinases: crystal structures of native and inhibited Agkistrodon contortrix contortrix protein C activator. *J Biol Chem*. 2005;280(47):39309–15.
- Neurath H. Evolution of proteolytic enzymes. *Science*. 1984;224(4647):350–7.
- Nicastro G, Franzoni L, de Chiara C, Mancin AC, Giglio JR, Spisni A. Solution structure of crotonamine, a Na⁺ channel affecting toxin from *Crotalus durissus terrificus* venom. *Eur J Biochem*. 2003;270(9):1969–79.
- Parry MA, Jacob U, Huber R, Wisner A, Bon C, Bode W. The crystal structure of the novel snake venom plasminogen activator TSV-PA: a prototype structure for snake venom serine proteinases. *Structure*. 1998;6(9):1195–206.
- Perbandt M, Tsai IH, Fuchs A, Banumathi S, Rajashankar KR, Georgieva D, Kalkura N, Singh TP, Genov N, Betzel C. Structure of the heterodimeric neurotoxic complex viperotoxin F (RV-4/RV-7) from the venom of *Vipera russelli formosensis* at 1.9 Å resolution. *Acta Crystallogr D Biol Crystallogr*. 2003;59(Pt 10):1679–87.

- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera – a visualization system for the exploratory research and analysis. *J Comput Chem.* 2004;25(13):1605–12.
- Rádis-Baptista G, Kerkis I. Crotamine, a small basic polypeptide myotoxin from rattlesnake venom with cell-penetrating properties. *Curr Pharm Des.* 2011;17(38):4351–61.
- Ramirez F, Jain MK. Phospholipase A2 at the bilayer interface. *Proteins.* 1991;9(4):229–39.
- Rawlings ND, Waller M, Barrett AJ, Bateman A. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 2013.
- Renetseder R, Brunie S, Dijkstra BW, Drenth J, Sigler PB. A comparison of the crystal structures of phospholipase A2 from bovine pancreas and *Crotalus atrox* venom. *J Biol Chem.* 1985;260(21):11627–34.
- Retzios AD, Markland Jr FS. A direct-acting fibrinolytic enzyme from the venom of *Agkistrodon contortrix*: effects on various components of the human blood coagulation and fibrinolysis systems. *Thromb Res.* 1988;52(6):541–52.
- Rossetto O, Montecucco C. Presynaptic neurotoxins with enzymatic activities. *Handb Exp Pharmacol.* 2008;184:129–70.
- Serrano SM. The long road of research on snake venom serine proteinases. *Toxicon.* 2013;62:19–26.
- Stöcker W, Grams F, Baumann U, Reinemer P, Gomis-Rüth FX, McKay DB, Bode W. The metzincins—topological and sequential relations between the astacins, adamalysins, serralsins, and matrixins (collagenases) define a superfamily of zinc-peptidases. *Protein Sci.* 1995;4(5):823–40.
- Takeda S, Takeya H, Iwanaga S. Snake venom metalloproteinases: structure, function and relevance to the mammalian ADAM/ADAMTS family proteins. *Biochim Biophys Acta.* 2012;1824(1):164–76.
- Teshima K, Kitagawa Y, Samejima Y, Kawauchi S, Fujii S, Ikeda K, Hayashi K, Omori-Satoh T. Role of Ca²⁺ in the substrate binding and catalytic functions of snake venom phospholipases A2. *J Biochem.* 1989;106(3):518–27.
- Thornton JM, Flores TP, Jones DT, Swindells MB. Protein structure. Prediction of progress at last. *Nature.* 1991;354:105–6.
- Ullah A, Souza TA, Zanphorlin LM, Mariutti RB, Santana VS, Murakami MT, Arni RK. Crystal structure of Jararacussin-I: the highly negatively charged catalytic interface contributes to macromolecular selectivity in snake venom thrombin-like enzymes. *Protein Sci.* 2013;22(1):128–32.
- Watanabe L, Shannon JD, Valente RH, Rucavado A, Alape-Girón A, Kamiguti AS, Theakston RD, Fox JW, Gutiérrez JM, Arni RK. Amino acid sequence and crystal structure of BaP1, a metalloproteinase from *Bothrops asper* snake venom that exerts multiple tissue-damaging activities. *Protein Sci.* 2003;12(10):2273–81.
- Wlodawer A, Minor W, Dauter Z, Jaskolski M. Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *FEBS J.* 2008;275(1):1–21.
- Wlodawer A, Minor W, Dauter Z, Jaskolski M. Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination. *FEBS J.* 2013;280(22):5705–36.
- Yamazaki Y, Morita T. Snake venom components affecting blood coagulation and the vascular system: structural similarities and marked diversity. *Curr Pharm Des.* 2007;13(28):2872–86.
- Zhang Y, Wisner A, Xiong Y, Bon C. A novel plasminogen activator from snake venom. Purification, characterization, and molecular cloning. *J Biol Chem.* 1995;270(17):10246–55.
- Zhu Z, Gong P, Teng M, Niu L. Purification, N-terminal sequencing, partial characterization, crystallization and preliminary crystallographic analysis of two glycosylated serine proteinases from *Agkistrodon acutus* venom. *Acta Crystallogr D Biol Crystallogr.* 2003;59(Pt 3):547–50.