Pergamon

Bioorganic & Medicinal Chemistry Letters, Vol. 7, No. 24, pp. 3139-3144, 1997

0960-894X/97 $17.00 + 0.00

PII: S0960-894X(97)10168-8

INSIGHTS INTO THE THREE-DIMENSIONAL STRUCTURE OF CROTALASE: IMPLICATIONS FOR BIOLOGICAL ACTIVITY AND SUBSTRATE SPECIFICITY

Irina Massova,[a] Hubert Pirkle,[c] Brian F. P. Edwards,[b] and Shahriar Mobashery[a]*

[a]*Departments of Chemistry, and of* [b]*Biochemistry and Molecular Biology, Wayne State University, Detroit, MI 48202, U.S.A.;* [c]*Department of Pathology, University of California, Irvine, CA 92697, U.S.A.*

Abstract: Crotalase is a serine protease from eastern diamondback rattlesnake (*Crotalus adamanteus*) venom. Crotalase has high amino-acid sequence similarity to three other members of the serine protease family, α-thrombin, β-trypsin and kallikrein A. Their structural information was used to predict the folding of crotalase. The computational structural data were used to explain biochemical properties of this important enzyme.

Crotalase is a serine protease extracted from eastern diamondback rattlesnake (*Crotalus adamanteus*) venom.[1] Snake venoms have found many applications in studies of analgesics, hemostatic agents, and anticoagulants.[2] Crotalase is an endopeptidase with high peptidase specificity for hydrolysis of the -Arg-Gly- peptide bond.[3] It also can hydrolyze basic amino-acid esters, and it shows some activity on non-basic esters as well. Furthermore, *p*-nitrophenyl esters of aromatic and aliphatic amino-acids can serve as substrates for crotalase.[1] In addition, crotalase displays thrombin-like activity. It forms fibrin clots from fibrinogen and at higher concentrations has a fibrinolytic activity.[4] Similarly to some other snake venom enzymes, crotalase would appear to release only fibrinopeptide A. Furthermore, crotalase cleaves the thrombin-sensitive bond in prothrombin.[5] It is intriguing that crotalase displays some features opposite to thrombin. It has no effect on factors V, VIII, and XIII, does not cause platelet aggregation and does not bind to heparin.[4] Crotalase can further be characterized as a trypsin-type serine protease. It is inhibited by known trypsin inhibitors such as chloromethyl ketone of toluenesulfonyl-L-lysine. However, it shows no effect with chloromethyl ketone of toluenesulfonyl-L-phenylalanine.[3] Crotalase also displays kallikrein-like activity. It is inhibited by specific kallikrein inhibitors and displays activity with kallikrein-specific chromogenic substrates.[6] Crotalase catalyzes kinin release from the high-molecular weight kininogen (HMWK) and cleaves the kallikrein-scissile bond in HMWK. However, in contrast to kallikrein, crotalase fails to activate plasminogen.[6] These various properties of crotalase lead us to conclude that in contrast to many enzymes extracted from various animal venoms, crotalase is a multifunctional enzyme.

Crotalase not only shows physiological similarity to other serine proteases, but it displays a relatively high amino-acid homology to other members of the family.[7] As could be expected from the substrate profiles, crotalase shows highest similarity to kallikrein A, β-trypsin, and α-thrombin.[8] These three enzymes have been crystallized, whereas the crotalase structure has not been determined to date.[9] The crotalase three-dimensional structure may shed light on mechanisms of action of the various serine proteases, and would help to find applications in treatment of various diseases, particularly since many examples are known of such use in medicine for other venoms.[2] In this manuscript the computational three-dimensional structure of crotalase is elucidated and the observed properties with substrates and physiological behavior of this enzyme are explained based on the structural information.

As described above, crotalase displays activities similar to α-thrombin, kallikrein A and β-trypsin. Furthermore, these three enzymes were chosen by the COMPOSER program as the ones with the highest

homology to crotalase, among eight serine proteases.[10] Percentages of the sequence identity of crotalase for these three proteins were 42% (kallikrein A), 40% (β-trypsin) and 35% (α-thrombin). These high values for the sequence identity scores indicate evolutionary relationship among them. These three enzymes were used in modeling in order to increase the size of the model structure covered by the Structurally Conserved Regions (SCR)—the areas with the same fold in homologous proteins.

```
16  20        30    -   40        50        60        70      -  80        90
....|.........|.....-....|.........|.........|.........|.......-..|.........|.....
VIGGDECNINEHRFLVALYDYWSQLFLCGGTLINNEWVLTAAHCDRT-HILIYVGVHDRSVQFDKEQRRFPKEKYFFDCSNN

    100       110       120       130       140       150       160       170 ---
....|.........|.........|.........|.........|.........|.........|.........|...---
FTKWDKDIMLIRLNKPVSYSEHIAPLSLPSSP--PIVGSVCRAMGWGQTTSPQETLPDVPHCANINLLDYEVCRTAHPQFR

      180  --      190       200       210       220       230       240
......|....--.....|.........|.........|.........|.........|.-.........|.....------
LPATSRTLCAGVLEGGIDTCNRDSGGPLICNGQ----FQGIVFWGPDPCAQPDKPGLYTKVFDHLDWIQSIIAGEKTVNCP
```

Figure 1. Numbering of the amino-acid sequence of crotalase is according to the "chymotrypsinogen" numbering system,[11] corrected based on results of multiple-sequence alignment of the crotalase sequence with sequences of kallikrein A, β-trypsin and α-thrombin.

To test the predictive power of COMPOSER for the serine proteases, the structure of β-trypsin was predicted based solely on amino-acid sequence of β-trypsin and the three-dimensional structures of α-thrombin and kallikrein A. The predicted and known crystal structures (for β-trypsin) were practically identical, with the rms deviation in positions of the catalytic residues within 0.2 Å (data not shown). The same experiment was repeated for the prediction of the three-dimensional fold for kallikrein A based on structural data available for α-thrombin and β-trypsin. The predicted and X-ray structures of kallikrein A were entirely superimposable, particularly in the active site area, for which the rms deviation was merely 0.14 Å.

Finally, COMPOSER was used to determine the three-dimensional structure of crotalase. Crotalase has 237 amino acids in its sequence (Figure 1). The twelve structurally conserved areas comprise 86% of the length of crotalase. Hence, we expected a high level of confidence in prediction of the structure of crotalase.

Figure 2 shows the backbone of all four proteins superimposed according to their topologically equivalent regions. The regions encompassing the catalytic triads (Ser-195, His-57 and Asp-102) of all four proteins are entirely superimposable.

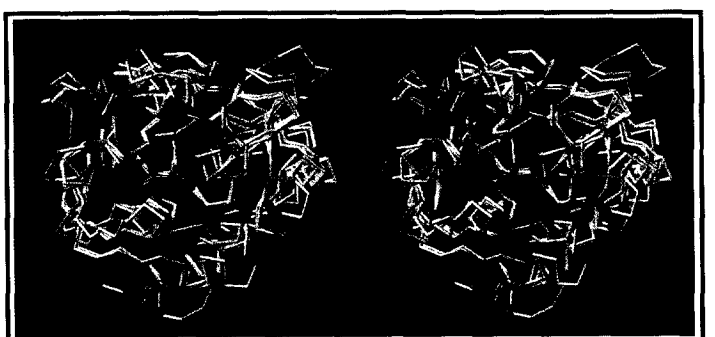

Figure 2. Superimposition of the backbones of the predicted structure of crotalase (in cyan) and the crystal structures (in white) for the kallikrein A, β-trypsin and α-thrombin.

Certain criteria can be used to evaluate the correctness of the folding for the protein of unknown structure. These are factors such as exposure of charged and polar groups to the solvent, spatial location of the catalytic residues, self-consistency of the structure with respect to formation of unconstrained disulfide bridges, and compactness of folding. The predicted structure of

crotalase satisfied all these criteria. The surface area calculated based on molecular weight of crotalase was within 0.1% of the surface area of the predicted structure.[12] The difference may be as much as 20% or more for incorrectly folded proteins. Approximately 10-13% of molecular surface area for a typical globular protein is usually provided by charged groups.[12] Crotalase falls within these limits, having 12% of the surface provided by charged groups. Almost all charged groups are exposed to the solvent. Crotalase has 12 cysteines, all of which exist in the oxidized disulfide form.[1,3] All 12 cysteines in our model were in the range of 4-7 Å from the C_α of another cysteine, a distance range which is favorable for formation of unconstrained disulfide bridges. These pairs are Cys-22 to Cys-157, Cys-42 to Cys-58, Cys-136 to Cys-201, Cys-168 to Cys-182, Cys-191 to Cys-220, and Cys-92 to the cysteine residue near C-terminal end. The disulfide pairs of crotalase are in agreement with those demonstrated in the highly homologous thrombin-like venom enzyme, bilineobin.[13] We used the "chymotrypsinogen" numbering system in this manuscript.[11] This system uses letters to number the amino-acid insertions. For example, α-thrombin has nine-amino-acid insertion after position 60, therefore the following residues are numbered as 60A to 60I. Also, the prime sign is used for the residues of serine protease substrates and inhibitors, to differentiate their residues from those of the enzyme.

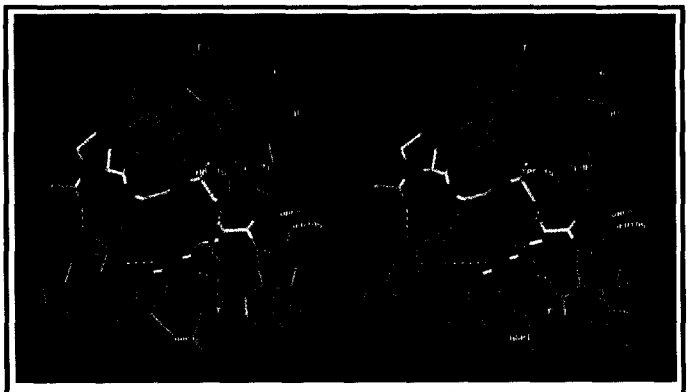

Figure 3. Stereoview of the active site of the energy-minimized structure of the acyl-enzyme intermediate for crotalase. The catalytic triad is shown in green, the remainder of the enzyme is colored in cyan. Substrate atoms are colored according to the atom types (carbons in white, nitrogens in blue, oxygens in red, and hydrogens in cyan). Hydrogen atoms are shown only when they are involved in a few selected hydrogen bonds. Hydrogen bond is depicted as yellow dashed lines.

Water molecules, found at the same positions for the homologous proteins, were added to the folded crotalase structure. Several water molecules were incorporated to cover the active site of the enzyme. A total of 296 water molecules were added to the folded structure of crotalase. Then, the energy of the entire molecule was minimized by the AMBER force field. Energy minimization caused no significant change in folding of crotalase. There is a salt bridge between the N-terminal amino group of Val-16 and the side chain of Asp-194, which precedes the catalytic Ser-195. This salt bridge is common in serine proteases and it is believed to stabilize the structure of the active site. The active site of crotalase has approximately the same dimensions as those for the three homologous proteins used to model the enzyme. However, the presence of Phe-214 near the bulky Trp-215 causes the active site near the catalytic histidine and serine to be slightly narrow. The position 214 is occupied by a smaller serine residue in kallikrein, trypsin and thrombin, compared to the bulkier phenylalanine in crotalase. This change in the size of the residue near the catalytic serine residue can affect substrate preferences and explains increased specificity of crotalase toward Arg-X substrates, where X is preferably Gly.[3] We also modeled a hypothetical acyl-enzyme intermediate for the active site of crotalase modified by

Pro-Gly-Ala-Arg tetrapeptide (Figure 3). The substrate binds in a mode similar to other serine proteases, aligning in an antiparallel manner to the β-sheet of the active site. The arginine residue of the substrate fits to the S_1 pocket, participating in hydrogen bonding to the side chains of Asp-189 and Thr-190. The carbonyl of the acyl-enzyme intermediate is sequestered in the oxyanion hole.

The S_1-binding site. The S_1 pocket of crotalase is the same size as that of kallikrein A. All four enzymes have Asp-189 at the bottom of this pocket. The acidic side chain of aspartate serves as the counter ion in interactions with the basic groups at P_1 of the substrates. Crotalase, β-trypsin, and α-thrombin have Gly-226 located near Asp-189. In contrast, kallikrein has Ser-226 near Asp-189, which makes the S_1 pocket slightly smaller for this enzyme. Another residue found in the S_1 pocket of crotalase is Thr-190, which is also found in kallikrein. The hydroxyl of Thr-190 provides an additional hydrogen bond with the basic P_1 group. Therefore, interaction of the substrates with basic groups at P_1 should be strong. Thrombin has a more hydrophobic alanine residue at this position, and trypsin has a serine. Residue Val-213, located to the side of the S_1 pocket (near its entrance), makes this part of the pocket more hydrophobic. This accounts for enhanced activity with substrates with arginine at P_1 (which has three methylene groups in the side chain) and activity against esters with aromatic and aliphatic groups at that position.[1]

The S_2-binding site. The S_2 binding site is formed by similar residues in the case of crotalase, kallikrein, and trypsin, and to some extent in thrombin. Crotalase, kallikrein, and trypsin lack the nine-residue insertion, a loop 60A-60I, that is found in thrombin. This causes the S_2 site to be more open in these enzymes than in α-thrombin. The S_2 binding site is bordered by the catalytic His-57 and hydrophobic ring systems of Trp-99 and Trp-215 in crotalase. Apart from the catalytic histidine residue, kallikrein, trypsin, and thrombin also have hydrophobic residues such as tyrosine or leucine and tryptophan at positions 99 and 215, respectively. The hydrophobic environment at the S_2 binding site causes preference by all these enzyme for hydrophobic residue at P_2 in substrates.

The "aryl"-binding site. The "aryl"-binding site is formed by residues Leu-99, Ile-174, Trp-215 and Phe-227 in thrombin. It is occupied by the aromatic group of Phe-8' of fibrinogen, during its hydrolysis.[14] This site is also bordered by hydrophobic residues Trp-99, Phe-173B, Leu-174, and Trp-215 in crotalase. These residues could provide favorable hydrophobic interactions with Phe-8' of fibrinogen and may account for the fibrinopeptide A releasing activity observed for this enzyme.

The "fibrinogen-binding" exosite. The "fibrinogen binding" exosite is formed by two loops (Phe-34 to Leu-41 and Lys-70 to Glu-80) in thrombin with affinity for fibrinogen. The counterpart of the Lys-70 to Glu-80 loop in trypsin is referred to as the "calcium loop".[15] This loop and its vicinity are very rich in basic residues in thrombin, but not so in trypsin and kallikrein. These positions are similarly rich in basic residues in crotalase (Arg-73, Lys-78, Arg-81, Arg-82), although to a lesser extent than in α-thrombin. α-Thrombin has seven basic residues (Arg-67, Lys-70, Arg-73, Arg-75, Arg-77A, Arg-78, and Lys-81) in the loop, compared to four in crotalase. Three of the four basic residues of crotalase (Arg-73, Lys-78, and Arg-81) are shared with thrombin, and the side chain of Arg-82 in crotalase is located close to the side chain of Arg-67 of α-thrombin. These residues of crotalase could provide the needed interactions for binding of fibrinopeptide to the enzyme. Kallikrein and β-trypsin both lack basic residues at this loop, and consequently they cannot bind to fibrinopeptide residues 30 to 44 from the αA N-terminus. Kallikrein has only one basic residue in this region, Arg-70. β-Trypsin has also just one basic residue (Arg-66). However, the loop of trypsin and its

vicinity are rich with acidic residues (Glu-70, Asp-71, Glu-77, Glu-80, and residue Asp-153), which favor binding to the calcium ion. Another interesting observation is that the topology of the molecular surface surrounding the active-site cavity of crotalase is unique and different from the other serine proteases. In crotalase two grooves are joined together forming the active-site cleft. One groove is similar to that in the "fibrinogen-binding" exosite of thrombin, and another is unique. This second groove is absent in thrombin and its place is taken by the 60A-60I loop. One of the walls of this second groove is formed by the additional six-amino-acid stretch at the C-terminus of crotalase. Surprisingly, this second groove of crotalase is also lined up with basic residues (Arg-60, Lys-85, Lys-89, Arg-107, and Lys-110). One can see that these two positively charged surfaces "mirror" each other with respect to their position in reference to the active site. Therefore, this second site could be considered as an "alternative fibrinogen recognition exosite", existence of which was proposed previously.[7] Most thrombin-like enzymes have a net negative charge in the area where the regular "fibrinogen-binding" exosite would be located. On the other hand, the basic residues at positions, 85, 89, and 107 are strictly conserved among the seven known thrombin-like enzymes from snake venoms, positions 60 and 110 are basic in six of them, suggesting the existence of such an alternative site.[16]

The "heparin-binding" site. The "heparin-binding" site is unique to thrombin. It is part of the thrombin surface formed by one helix comprised of residues His-230 to Ile-242, and is surrounded by basic residues (Arg-93, Lys-97, Arg-101, Lys-126, Arg-165, Lys-169, Arg-173, Arg-175, Arg-233, Lys-235, Lys-236, Lys-240), which are thought to be involved in heparin binding to thrombin.[15] Crotalase has Ser-93, Thr-97, Lys-101, Ser-126, Tyr-165, Arg-169, Pro-173, Pro-175, Asp-233, Leu-235, Asp-236, Ser-240 at the corresponding positions. These are mostly non-basic residues. Six residues among 13 of the helix (positions 230-242) are the same in thrombin and crotalase. However, all surface residues of this helix are variable in crotalase. They are mostly acidic in nature (Asp-233, Asp-236, and Ser-240). In contrast, α-thrombin has only basic residues—Arg-233, Lys-236, and Lys-240—at these positions, which also contribute to the interaction of α-thrombin to heparin, in contrast to crotalase, which does not bind to heparin, as has been shown experimentally.[4]

The "hirudin-binding" site. Hirudin is a peptide extracted from the leech saliva. It is one of the most potent thrombin inhibitors and the most specific. It has a high content of negatively charged amino-acid residues, and its side chain of Tyr-63' (near the carboxy-terminus) is sulfated. These residues are all necessary for maximum complex formation with thrombin.[15] Among some of the important hirudin-thrombin interactions are ion pairs, Asp-5' to Arg-221A, Glu-17' to Arg-173, Glu-49' to Lys-60F, Asp-55' to Arg-73, Asp-55' to Lys-149E, Glu-58' to Arg-77A, and the carboxy-terminal carboxylate to Lys-36 (all residues are given for human hirudin-thrombin complex; 4htc).[15] Positions 60F and 149E are spatially absent in crotalase. Among seven ion-pairs which are present in the hirudin-thrombin complex, six do not exist if crotalase were to interact with hirudin. The nine-amino-acid-insertion loop (60A-60I) is found only in thrombin, and primarily consists of hydrophobic residues. These residues provide hydrophobic interactions with Leu-13' and Pro-46' of hirudin. This loop is absent in crotalase. Therefore, hirudin does not inhibit crotalase, as has been observed experimentally.[7]

Platelet aggregation. Platelet aggregationis mediated by the thrombin receptor which is activated by cleavageof an arginyl bond. It has been shown that thrombin cleaves the Leu-Asp-Pro-Arg~Ser site of the thrombin receptor on platelets (~ represents the cleavagepoint).[17] The hirudin-like domain of the thrombin

receptor binds to the fibrinogen recognition exosite.[17] However, due to the presence of the bulky tryptophan at position 215, crotalase has specificity toward -Arg-Gly- bonds.[3] Furthermore, as it has been pointed out in a previous section, crotalase may have an alternative fibrinogen recognition exosite which is different from that of thrombin. As a result, crotalase does not bind to hirudin nor, quite possibly, to the hirudin-like domain of thrombin receptor. Therefore, because of all these factors crotalase cannot activate platelet aggregation, as demonstrated experimentally.[4]

Structural analysis reveals that crotalase resembles kallikrein most closely. This is supported by higher crotalase amino-acid sequence identity with kallikrein. The structural data provided herein make available detailed information, which explains many observations on activity of crotalase.[18]

Acknowledgments: IM is the recipient of the Rumble and Heller Fellowships. The research in the Department of Biochemistry and Molecular Biology was supported by the National Institutes of Health grant HL57527 (to BFPE).

References

1. Markland, F. S.; Damus, P. S. *J. Biol. Chem.* **1971**, *21*, 6460.
2. Hutton, R. A.; Warrell, D. A. *Blood Rev.* **1993**, *7*, 176. Marsh, N. A. *Blood Coag. Fibrinol.* **1994**, *5*, 399.
3. Markland, F. S. *Methods in Enzymology.* **1976**, *45*, 223.
4. Damus, P. S.; Markland, F. S.; Davidson, T. M.; Shanley, J. D. *J. Lab. Clin. Med.* **1972**, *79*, 906.
5. Pirkle, H.; Markland, F. S.; Theodor, I. *Thromb. Res.* **1976**, *8*, 619.
6. Markland, F. S.; Kettner, C.; Schiffman, S.; Shaw, E.; Bajwa, S. S.; Reddy, K. N. N.; Kirakossian, H.; Patkos, G. B.; Theodor, I.; Pirkle, H. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 1688.
7. Pirkle, H.; Theodor, I. *Haemostasis* **1996**, *26 (suppl. 3)*, 452.
8. Pirkle, H.; Markland, F. S.; Theodor, I. *Biochem. Biophys. Res. Commun.* **1981**, *99*, 715.
9. Several unsuccessful attempts to crystallize crotalase have been made. Extensive crystallization trials with both commercial and "home-brewed" screening protocols have yielded only highly twinned crystals of crotalase to date (the Edwards laboratory).
10. The COMPOSER software implemented in the TRIPOS product SYBYL, version 6.22, was used for the folding of crotalase. The coordinates for homologous proteins were retrieved from the Brookhaven Protein Database, 1bbr, α-thrombin (Martin, P. D.; Robertson, W.; Turk, D.; Huber, R.; Bode, W.; Edwards, B. F. P. *J. Biol. Chem.* **1992**, *267*, 7911); 2tpc, β-trypsin (Marquart, M.; Walter, J.; Deisenhofer, J.; Bode, W.; Huber, R. *Acta Crystallogr.* sect. B, **1983**, *39*, 480); and 2pka, kallikrein A (Bode, W; Chen, Z.; Bartels, K.; Kutzbach, C.; Schmidt-Kastner, G.; Bartunik, H. *J. Mol. Biol.* **1983**, *164*, 237).
11. Bode, W.; Mayr, I.; Baumann, U.; Huber, R.; Stone, S. R.; Hofsteenge, J. *EMBO J.* **1989**, *8*, 3467.
12. Novotny, J.; Bruccoleri, R.; Karplus, M. *J. Mol. Biol.* **1984**, *177*, 787.
13. Nikai, T.; Ohara, A.; Komori, Y.; Fox, J. W.; Sugihara, H. *Arch. Biochem. Biophys.* **1995**, *318*, 89.
14. Bode, W.; Turk, D.; Stürzebecher, J. *Eur. J. Biochem.* **1990**, *193*, 175.
15. Bode, W.; Huber, R.; Rydel, T. J.; Tulinsky, A. In *Thrombin: Structure and Function;* Berliner, L. J., Ed.; Plenum: NY, 1992; pp 3–61.
16. Pirkle, H. *Thromb. Haemostasis* **1997** (accepted for publication).
17. Coughlin, S. R.; Vu, T.-K. H.; Hung, D. T.; Wheaton, V. I. *J. Clin. Invest.* **1992**, *89*, 351.
18. Full atomic coordinates for the computational structure of crotalase are available at our group Web site (http://www.science.wayne.edu/~somgroup).

(*Received in USA* 1 *October* 1997; *accepted* 5 *November* 1997)