

How to Kill an Enzyme (In More Ways Than One)

The crystal structures of a zymogen and two mutants of the serine-carboxyl protease kumamolisin beautifully describe the mode of inhibition and activation of the proenzyme, while reminding us that our understanding of the enzymatic mechanisms is far from complete.

Proteolytic enzymes are ubiquitous (almost 500 have been identified in the human genome, for example) and very important, since almost all life and death processes depend on their presence and controlled activity. These enzymes have been investigated for so long and in such detail that it might be assumed that not much new information could be gleaned from them. However, as shown by W. Bode and his collaborators in this issue of *Structure* (Comellas-Bigler et al., 2004), unexpected results are still the norm rather than the exception.

For almost 20 years K. Oda and his team have been studying an obscure group of proteases that are active at low pH and often at high temperature (Oda et al., 1987). While they could show that the activity of these enzymes relied on the presence of multiple side chain carboxylates, their exact classification remained uncertain and thus they were named pepstatin-insensitive carboxyl proteases. They were later reclassified as potential serine proteases (Rawlings and Barrett, 1999; Lin et al., 2001); however, a structural fold or the details of the catalytic site were not proposed. A postulate that they might represent an unusual family of serine proteases (Rawlings and Barrett, 1999) was proven by the structure of the first-identified family member, now called sedolisin (Wlodawer et al., 2001), which was soon followed by the structure of kumamolisin (Comellas-Bigler et al., 2002). These structures, which included both apoenzymes and inhibitor complexes, proved that the fold of sedolisins is a superset of the well-studied family of subtilisins (Wlodawer et al., 2003), but although both subtilisins and sedolisins utilize an identical serine residue as the principal nucleophile, other members of the triad are different. A histidine, the second member of the triad in subtilisin, is substituted by a topologically equivalent glutamic acid in sedolisin, while the third residue, although aspartic acid is in both of them, is contributed by topologically different parts of the structure. However, the overall similarity of these two protease families is remarkable, with practically all secondary structure elements found in the smaller subtilisins (~275 residues) also found in sedolisins (~375 residues), although the converse is not true, obviously.

Like many other proteases that have to be expressed as inactive zymogens in order to control their activity, sedolisins are also synthesized as inactive precursors that include a ~200 amino acid long propeptide. This

comparatively large propeptide actually plays a dual role, both assisting in the folding of the enzyme and protecting it from being prematurely activated. Similar, although shorter, propeptides are also found in subtilisins and in many other proteases, such as aspartic proteases that belong to the pepsin family. Although a number of structures of the zymogen forms of proteases have been reported, crystallizing these proteins is not trivial because they are very often unstable under the conditions needed to obtain crystals (i.e., preventing activation during crystallization is not always possible). For instance, a structure of the intact zymogen of subtilisin has yet to be reported, despite many ingenious attempts in that direction. An extensively mutated subtilisin BPN' was crystallized as a complex with a separately expressed 77-residue prosegment, which allowed through model building an analysis of the proenzyme structure at various stages of activation (Gallagher et al., 1995). That work was followed by a study of an autoprocessed Ser221Cys mutant of subtilisin E (Jain et al., 1998). Although mutation of the catalytic residue diminished the activity of the enzyme, it did not abolish it entirely, allowing the initial propeptide cleavage but not further degradation. The complex consisted of two polypeptide chains, one corresponding to the propeptide and the other to the mature enzyme. The C-terminal part of the propeptide was bound to the enzyme in a manner resembling that of a substrate, as judged by comparison with the structures of the inhibitor complexes of subtilisin. However, the distance between the C terminus of the propeptide and the N terminus of the mature enzyme, both well ordered, is ~28 Å, indicating that a large conformational reorganization must have followed the cleavage of the peptide bond.

The new structure of an intact kumamolisin precursor (Comellas-Bigler et al., 2004) shows in atomic-level detail what happens before the autocatalytic cleavage takes place in sedolisins. The enzyme used in this study has been mutated by replacement of the nucleophilic Ser278 by an alanine and is thus completely inactive. The peptide bond between His171 and Phe172 of the propeptide is placed in the active site of the enzyme, with the preceding and following amino acids occupying the substrate binding subsites. This conclusion is strongly supported by a comparison with the structure of sedolisin in which, by serendipity, two inhibitor molecules were found to occupy both the nonprimed and primed sites (Wlodawer et al., 2004). As seen in Figure 1, the main chains and the side chains of the kumamolisin propeptide and the two inhibitors of sedolisin are almost exactly superimposed in the S3 through S3' pockets of the enzyme. In the zymogen of kumamolisin, the propeptide continues uninterrupted until it reaches the well-ordered sequence Gln-Ser*Ala-Ala, where another cut must be made later to expose the N terminus of the mature enzyme. It is quite likely that this second cut is not autocatalytic and involves other proteases, although no unambiguous data are available at this time.

Although the propeptide of kumamolisin is much

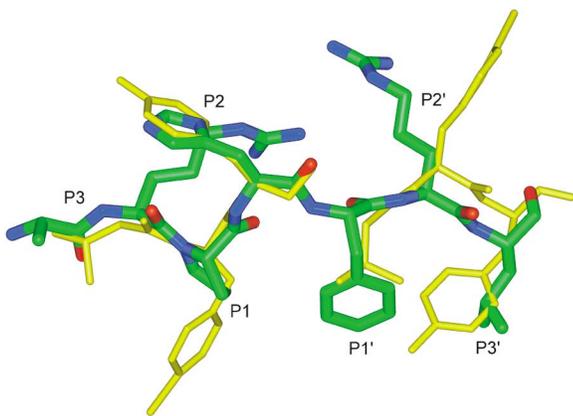


Figure 1. Inhibitors of Sedolisin and the Propeptide of Kumamolisin Overlay of two inhibitors of sedolisin (yellow) and seven residues (Ala168-Leu174, subsites P4-P3') of the propeptide of kumamolisin (carbon atoms, green; oxygens, red; and nitrogens, blue), based on the superposition of the complete molecules of both enzymes. Excellent agreement between the conformations of the main chains and the side chains (with the exception that the P3 side chain of the inhibitor follows the main chain of the propeptide) reinforces the assumption that both types of ligands properly mimic substrate binding.

longer than the corresponding propeptide of subtilisin, the fold of the latter is a subset of the former, again mimicking the situation found in the mature proteins. Thus, it can be concluded that serine-carboxyl proteases are autoinhibited by a peptide that perfectly mimics a substrate and does not lead to substantial modifications of the mature enzyme (with an interesting exception that a positively charged arginine found in the P3 position of the propeptide seems to reorient Asp164, a residue that normally functions as an oxyanion hole during catalysis, perhaps explaining in part the requirement of low pH for the activity of these enzymes). This is in contrast with other known modes of autoinhibition, such as those reported for zymogens of aspartic proteases. These proenzymes are inactivated by insertion of a lysine side chain into the vicinity of the catalytic aspartates (pepsin or phytpepsin) or by large domain movement that completely rearranges the catalytic site in proplasmepsin II (Bernstein and James, 1999).

Two other high-resolution structures of the kumamolisin mutants were presented by Bode and coworkers (Comellas-Bigler et al., 2004) in order to evaluate the differences between the catalytic machinery of mesophilic sedolisin and thermophilic kumamolisin. Hydrogen-bonded interactions in the catalytic triad of the former enzyme involve only Ser, Glu, and Asp, whereas in kumamolisin they are extended through two more residues, Glu32 and Trp129. It was postulated that these additional interactions might facilitate proton delocalization during a nucleophilic attack at high temperature (Comellas-Bigler et al., 2002). Individual mutation of these residues to alanines decreased the catalytic activity of kumamolisin ~ 20 -fold. Removal of the side chain of Glu32 did not lead to significant structural changes and a single water molecule partially assumed the role of the carboxylate moiety. Much larger changes were

found in the structure lacking Trp129, but it appears that the loss of activity is mostly due to the switching of nucleophilic Ser278 between two conformations, one of them inactive. These structures show that the lower activity of the mutants is not a simple result of the replacement of these side chains, but is also a consequence of local structural changes affecting the other catalytic residues. Thus, the differences between the catalytic machinery in sedolisin and kumamolisin still need to be examined further. Actually, no detailed analyses of how a Ser-Glu-Asp triad (with an additional Asp in the oxyanion hole) functions are available in the first place, and it can only be hoped that other experimental approaches and computational investigations will follow structure determination of this family of enzymes.

One other important question that still needs to be answered is the overall significance of this family of proteases. Analyses of the completed genomic sequences of various organisms have shown that sedolisins are present in many, but by no means in all of them (Wlodawer et al., 2003). Either a sedolisin or a kumamolisin is found in many bacteria (with a curious exception of several species of *Burkholderia* that seem to contain both sedolisins and kumamolins), archaea, fungi, and many higher organisms. It can be safely assumed that all readers of this preview have a functioning protease belonging to this family, since it has been shown that mutations leading to the loss of the human enzyme CLN2, a prominent member of the sedolisin family, result in a fatal neurodegenerative disease, classical late-infantile neuronal ceroid lipofuscinosis (Sleat et al., 1997). Although structural studies of sedolisins have not firmly established their biological roles, they have contributed significantly to our understanding of the mechanism of protection and activation of proteases. One may wonder if perhaps the most important function of sedolisins and kumamolins is to yield crystals that diffract to atomic resolution (a number of structures of both proteases have been solved with data extending to 1.0–1.2 Å)! This optimistic assessment is somehow tempered by the failure to obtain diffracting crystals of CLN2 after years of effort by several laboratories. Clearly, the last word has not yet been said about this fascinating family of enzymes.

Alexander Wlodawer

Protein Structure Section
Macromolecular Crystallography Laboratory
National Cancer Institute at Frederick
Frederick, Maryland 21702

Selected Reading

Bernstein, N.K., and James, M.N. (1999). *Curr. Opin. Struct. Biol.* 9, 684–689.

Comellas-Bigler, M., Fuentes-Prior, P., Maskos, K., Huber, R., Oyama, H., Uchida, K., Dunn, B.M., Oda, K., and Bode, W. (2002). *Structure* 10, 865–876.

Comellas-Bigler, M., Maskos, K., Huber, R., Oyama, H., Oda, K., and Bode, W. (2004). *Structure* 12, this issue, 1313–1323.

Gallagher, T., Gilliland, G., Wang, L., and Bryan, P. (1995). *Structure* 3, 907–914.

Jain, S.C., Shinde, U., Li, Y., Inouye, M., and Berman, H.M. (1998). *J. Mol. Biol.* 284, 137–144.

Lin, L., Sohar, I., Lackland, H., and Lobel, P. (2001). *J. Biol. Chem.* 276, 2249–2255.

Oda, K., Sugitani, M., Fukuhara, K., and Murao, S. (1987). *Biochim. Biophys. Acta* 923, 463–469.

Rawlings, N.D., and Barrett, A.J. (1999). *Biochim. Biophys. Acta* 1429, 496–500.

Sleat, D.E., Donnelly, R.J., Lackland, H., Liu, C.G., Sohar, I., Pullarkat, R.K., and Lobel, P. (1997). *Science* 277, 1802–1805.

Wlodawer, A., Li, M., Dauter, Z., Gustchina, A., Uchida, K., Oyama, H., Dunn, B.M., and Oda, K. (2001). *Nat. Struct. Biol.* 8, 442–446.

Wlodawer, A., Li, M., Gustchina, A., Oyama, H., Dunn, B.M., and Oda, K. (2003). *Acta Biochim. Polon.* 50, 81–102.

Wlodawer, A., Li, M., Gustchina, A., Oyama, H., Oda, K., Beyer, B.B., Clemente, J., and Dunn, B.M. (2004). *Biochem. Biophys. Res. Commun.* 314, 638–645.

Structure, Vol. 12, July, 2004, ©2004 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.str.2004.06.009

Visualizing the House from the Brick

In this issue of *Structure*, Dokland et al. (2004) present the crystal structure of West Nile Virus capsid protein and open a new perspective on the possible structure of the flavivirus capsid.

The structures of virus capsid proteins were once fairly predictable. The occurrence of a common fold in the first virus structures solved by X-ray crystallography led to the inference that it might be a consistent feature of virus structures. Membrane viruses, as is so often the case, proved exceptions to this generalization. The prediction that the alphavirus capsid protein would be similar to the nonenveloped ones (Fuller and Argos, 1987) was proven false by the Sindbis capsid protein structure (Choi et al., 1991). Indeed, a serine protease proved a better model for the alphavirus capsid.

Alphavirus and flaviviruses are two icosahedral membrane viruses bearing type two fusion proteins with the same fold (Lescar et al., 2001; Rey et al., 1995). The capsid proteins of the alphaviruses form their icosahedral shells prior to the interaction with the envelope proteins. This interaction and lateral interactions between spikes drives virus budding from the plasma membrane. The nucleocapsid can be isolated from infected cells and from virions. Image reconstructions from cryo-electron micrographs established the organization of the complementary $T = 4$ nucleocapsid and envelope protein layers (Mancini et al., 2000). The capsid protein contains a positively charged tail that interacts with the viral RNA and is toward the center of the nucleocapsid (Choi et al., 1991).

The similarity between the fusion proteins does not extend to the capsid of the flaviviruses. Icosahedral capsids are difficult to isolate from virions or infected cells. Image reconstruction from cryo-electron micrographs of mature or immature viruses show well-ordered envelope proteins (Kuhn et al., 2002; Zhang et al., 2003) but reveal no ordered density for the capsid. Kuhn et al. (2002)

suggested that the flavivirus capsid does not share the organization of the surface proteins. The absence of a structure for the flavivirus capsid makes the interpretation of the capsid protein structure akin to inferring the organization of a house from a brick. Nevertheless, such hypotheses are the only route toward understanding these important structures.

Ma et al. (2004) used an elegant NMR method to determine the structure of a dimer of the Dengue virus capsid protein. They demonstrated that the flavivirus capsid protein is a novel fold. The monomer comprises a core of three helices, $\alpha 1$ – $\alpha 3$, with a fourth, $\alpha 4$, extending from the core. The dimer showed an uneven distribution of charge. They proposed that the dimer would be oriented in the virus with its positively charged regions extending centrally to interact with the RNA and the hydrophobic region interacting with the membrane (Figure 1).

Dokland et al. (2004) determined the crystal structure of the capsid protein of the Kunjin strain of West Nile virus. Their structure confirms the dimer structure presented by Ma et al. (2004) and revealed flexibility in the $\alpha 1$ helix. The structure also revealed that the dimers form tetramers in the crystal. This shields the hydrophobic regions during crystal formation and provides a highly positively charged surface that could interact with the viral genome (Figure 1). They point out that the tetramers form filaments reminiscent of those formed by the HEAT motifs in many nucleic acid interacting proteins. The ribbons formed in the crystal illustrate the types of interactions that could be used to build the more complex network required to construct a capsid. An order-disorder transition of the $\alpha 1$ helix appears to modulate these interactions as is common in virus capsid assembly.

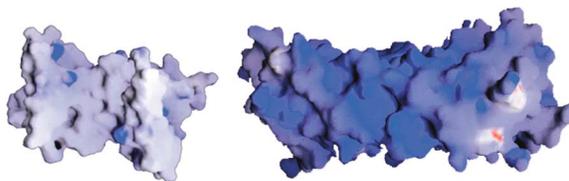


Figure 1. Surface Potential of Flavivirus Capsid Oligomers
The GRASP (Nicholls et al., 1991) representation of the surface potential for the dengue virus capsid protein dimer (left) and the West Nile virus capsid protein tetramer (right). The face proposed to be against the membrane is facing the viewer in both images.