

Mechanism of Action of Serine Proteases

This is the most studied and best known enzymatic mechanism, but new details are still being added, as in the paper by Whiting and Peticolas, *Biochemistry* **33**:552-561 (1994) on the so-called oxyanion hole, and the suggestion that a low-barrier hydrogen bond, which I discussed at the end of the last lecture, plays a role. This description builds on a number of types of evidence, most of which we have covered: the use of active-site-directed inhibitors, the acyl-enzyme kinetic mechanism, X-ray crystallography and related methods for determining structure, comparison of activity on various substrates, and site-directed mutagenesis.

First, some general taxonomy. Proteases belong to one of four broad families: the serine proteases, such as chymotrypsin, in which catalysis involves formation of an acyl-enzyme with the acyl group on the serine; the cysteine proteases, such as papain, with a somewhat similar mechanism involving an acyl-cysteine; the acid proteases, such as pepsin, with an acid pH optimum, and two aspartate residues involved in the mechanism; and the metalloproteases, such as carboxypeptidase, which use a metal ion, usually zinc. There are other hydrolytic enzymes, such as lipases and acetylcholinesterase, with mechanisms closely related to that of the serine proteases.

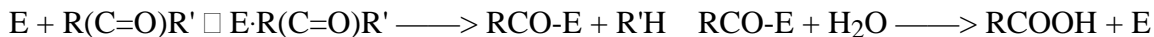
The serine proteases are divided into at least two genetic families, the mammalian serine proteases, such as trypsin, chymotrypsin, elastase, the enzymes of the blood clotting system, and many other proteases with specific roles in control of systems; and the bacterial proteases called subtilisins, which are genetically unrelated to the mammalian enzymes but independently evolved the same mechanism. There are other bacterial serine proteases, such as the α -lytic protease of *Myxobacter* which has been useful for general study because it contains only a single histidine residue; I don't know whether it is more closely related to the mammalian enzymes or to subtilisin. The bacterial enzymes are useful for site-directed mutagenesis studies because they are easily expressed in bacterial cells and because they do not require the activation reactions typical of the mammalian enzymes.

The mammalian enzymes generally belong to two classes defined by their specificity, the chymotryptic enzymes which cleave at the carboxyl side of amino acids with large hydrophobic side chains - phe, tyr, trp, leu, ile - and the tryptic enzymes which cleave at the carboxyl side of basic amino acids, lys and arg. There are also elastases, which attack the structural protein elastin at alanine and valine residues. These enzymes typically are synthesized in inactive forms which require activation by cleavage of a peptide bond near the NH₂-terminus. In chymotrypsin this is between arg¹⁵ and ile¹⁶; the free, protonated amino group of ile¹⁶ is important for the mechanism. Other cleavages remove the dipeptides ser¹⁴arg¹⁵ and thr¹⁴⁷asn¹⁴⁸, yielding the form of chymotrypsin usually worked with, α -chymotrypsin, but these are not important for the mechanism. The amino-terminal peptide with residues 1 through 13 stays attached to the rest of the protein through a disulfide bond. In trypsin the activation cuts off an amino-terminal hexapeptide, which does not remain attached. In most of the blood-clotting proteases such as Factors VII, IX, X, XI and XII, there is a large amino-terminal portion which remains attached by disulfide bonds and is important in keeping the active protease bound on membrane surfaces to act on the next protease, which is similarly bound. The final protease, thrombin, does not have its amino terminal domain attached by a disulfide bond and goes free in the plasma to attack fibrinogen and generate clots.

As already mentioned, these enzymes act by forming and hydrolyzing an ester on a serine residue. This was initially established using the nerve gas diisopropyl fluorophosphate, which inactivates serine proteases as well as acetylcholinesterase. Careful acid or enzymatic hydrolysis of the inactivated enzyme yielded O-phosphoserine, and the serine was identified as residue 195 in the sequence. Chymotrypsin acts on the compound cinnamoylimidazole, yielding a cinnamoylenzyme which hydrolyzes slowly - this was the first active site titrant - and has a spectrum similar but not identical to that of the model compound O-cinnamoylserine, different from that of N-cinnamoylserine or

cinnamoylimidazole; on denaturation of the enzyme in urea the spectrum was identical to that of O-acetylserine.

The serine proteases act on both esters and amides. The kinetic mechanism is that already described as the acyl-enzyme mechanism: the substrate binds noncovalently, the serine displaces the alcohol or amine part of the substrate to form an acyl-enzyme, and water then displaces the serine to yield the acid product and free enzyme.



where $R' =$ either $R'O-$ (ester) or $R''NH-$ (amide). For esters the deacylation step (k_3) is rate-limiting; for amides the acylation step (k_2) is much slower and is rate-limiting. Note that an ester and an amide of the same acid will yield the same acyl-enzyme and have the same rate constant for deacylation; but the acyl-enzyme accumulates only when the substrate is an ester or acyl-imidazole. This was shown originally with *p*-nitrophenyl acetate as substrate of chymotrypsin - the acetyl-enzyme could be isolated at acid pH - and confirmed with more specific substrates such as acetyltryptophan *p*-nitrophenyl ester: one equivalent of *p*-nitrophenol is released at a rate (k_2) much faster than that of the overall catalytic reaction (essentially k_3). When k_3 is small enough so that the overall catalytic reaction is slow even with substrate quantities of enzymes, as with *p*-nitrophenyl guanidinobenzoate and trypsin or *p*-nitrophenyl trimethylacetate and chymotrypsin, you have what I talked about earlier as enzyme titration. The existence of the acyl-enzyme is further confirmed by the formation of amides and methyl esters when the hydrolytic reaction is run in presence of ammonia or methanol (product partitioning). Trypsin and chymotrypsin work well on esters and amides of single amino acids, as long as the α -amino group is blocked, but have a higher k_{cat} for peptides with several residues to the amino side of the residue where cleavage occurs (referred to as the P_2, P_3, P_4 residues, which bind to the corresponding sites S_2, S_3, S_4); in elastase this effect is so strong that the enzyme essentially doesn't work on single amino acid substrates, except for their nitrophenyl esters. I don't know just how this works, but it certainly supports the principle of using binding energy to stabilize the transition state and increase k_{cat} . Binding of the side chain of the P_1 residue - at whose carboxyl cleavage occurs - increases k_{cat} , as shown by increase of the rate of trypsin hydrolysis of acetylglycine ethyl ester by methylguanidine, which binds at the specificity site for the side chain, asp¹⁸⁹.

The V_{max} of the reaction depends on the basic form of a group with $pK_a = 7$ and the acidic form of a group with $pK_a = 9$. The latter is the α -amino group of ile¹⁶, which makes an ion pair with the carboxyl side chain of asp¹⁹⁴ in the active site and causes changes in the position of the peptide chain, which we'll return to. The pH 7 group was guessed to be a histidine, and Elliott Shaw identified this using tosylphenylalanine chloromethylketone as his⁵⁷ in chymotrypsin. This does not act by nucleophilic attack - cinnamoylchymotrypsin does not have the spectrum of cinnamoylimidazole - but by general base catalysis (the reaction is slowed in D_2O). It is generally considered that the histidine pulls the proton off ser¹⁹⁵ as it attacks the substrate carbonyl, facilitating its nucleophilic attack, and then puts it on the leaving oxygen or nitrogen. In the enzyme without substrate bound it is far enough from the serine that only a very weak H bond could be formed, but the serine moves closer to it as substrate binds.

X-ray crystallography of chymotrypsin found an unexpected aspartic acid residue, asp¹⁰², buried in the protein and next to the other side of his⁵⁷ (it was very unexpected, as the initial sequence determination had this as an asparagine, but this was corrected). The group of three residues, asp¹⁰², his⁵⁷ and ser¹⁹⁵, has been referred to as a **catalytic triad** or a charge relay system; it is seen also in subtilisin, and something similar is seen in alcohol dehydrogenase. Considerable argument ensued as to whether the $pK_a = 7$ was that of his⁵⁷, with asp¹⁰² ionized at all reasonable pH (above 2) and H-bonded to a proton on his⁵⁷, or the pK_a was actually that of asp¹⁰², picking up the proton from the inside of his⁵⁷, which would pick up a proton from the medium and remain uncharged. This was settled in the

most direct way, by neutron diffraction studies, like X-ray crystallography but able to see even protons in protein crystals; the proton is on the histidine and the aspartate is ionized at $\text{pH} > 2$. The proton has an unusual chemical shift in nmr, 18 ppm at $\text{pH} 4$ where the histidine is protonated. The shift is 1 ppm less in D_2O at $\text{pD} = 3.5$, as expected for a low barrier H-bond.

Neutral pH , with the histidine unprotonated, is not good for a low-barrier H-bond because neutral imidazole as a proton donor has a pK_a of 13, ionization would form the imidazolide anion, which however is believed to be involved in some enzyme mechanisms. But protonated imidazole has a pK_a of 7 or so, not far from the pK_a of the aspartate.

X-ray crystallography of chymotrypsin with various inhibitors bound has contributed greatly to understanding the mechanism. Initially, as shown in Fig. 12.7 in the hand-out, the hydroxyl of ser^{195} is in an up position, not really H-bonded to his^{57} . When a specific substrate binds, the side chain rotates around the $\text{C}^\alpha\text{-C}^\beta$ bond, the hydroxyl passes by his^{57} and hands off its proton to his^{57} , protonating it and shoving it a bit toward asp^{102} so that a low-barrier H-bond can be formed. This favors formation of ser-O^- , pK_a normally 14. One wonders why his^{57} isn't protonated and H-bonded to asp^{102} in absence of substrate. Probably it is too far away for a low barrier H-bond.

The ser-O^- then attacks the carbonyl carbon, forming a tetrahedral intermediate (Fig. 12.8). This is also favored by the carbonyl oxygen of the substrate H-bonding with the backbone NH of gly^{193} and ser^{195} , a space known as the 'oxyanion hole'; this is improved when the oxygen becomes fully charged and single bonded in the tetrahedral intermediate - the H bonds become shorter and stronger. This H-bonding is shown in Figure 15.9 from Brandén and Tooze. In subtilisin one of these H bonds is to an asparagine amide, rather than to the backbone, and mutation of this residue to something such as glycine greatly slows deacylation of the acyl-enzyme (and presumably also acylation) by not stabilizing the transition state of these reactions as well. The 'oxyanion hole' is not present in the zymogen; it is generated by formation of the ion pair between ile^{16} and asp^{194} , which changes the position of the backbone.

Evidence for the tetrahedral intermediate includes a Hammett ρ constant of +2.1 for the deacylation reaction of substituted benzoylchymotrypsins and the formation of tetrahedral complexes with many inhibitors, such as boronates, sulfonyl fluorides, peptide aldehydes, and peptidyltrifluoromethylketones (structure II of scheme 1). In these last the chemical shift of the imidazole proton is 18.9 ppm, indicating a good low barrier H-bond, and the pK_a of the imidazolium is 12.1, indicating that it is stabilized by 7.3 kcal/mole compared to substrate-free chymotrypsin. The imidazole in effect is a much stronger base, facilitating proton removal from the serine.

But the tetrahedral intermediate does not accumulate in normal catalysis (it can be seen in papain catalysis at very low temperatures); it seems to be a mere blip near the top of the transition state pass. The amine or alcohol receives a proton from his^{57} - perhaps it is shoved close enough to make a temporary low barrier H-bond as a step in proton transfer - and leaves the carbonyl, and the carbonyl oxygen becomes a carbonyl again. The paper by Whiting and Peticolas indicates that in dimethylaminobenzoyl-trypsin and subtilisin, which are stable acyl-enzymes, it remains H-bonded in the oxyanion hole - less so in chymotrypsin. Perhaps whenever a tetrahedral state can be stabilized in other ways, as for instance in phosphate and sulfate esters, addition to boronates and aldehydes, and acyl-enzymes distorted toward tetrahedral geometry, it will yield a stable inhibited complex.) When the carbonyl O shrinks back it interferes with the leaving N or O, which is pushed entirely out of the active site (Fig. 12.9). You now have the acyl-enzyme intermediate.

Deacylation is considered to be essentially the reverse of acylation, and evidence from both is used in constructing the mechanism of both steps. A water molecule would come past his^{57} , have a proton stripped off, and attack the acyl-enzyme carbonyl as something close to OH^- ; the general base catalysis of deacylation is shown by slowing of the reaction in D_2O . This attack again generates a

tetrahedral intermediate - as suggested by the large Hammett ρ constant in deacylation of benzoylated chymotrypsins - which again is stabilized by the oxyanion hole. His⁵⁷ puts the proton back on ser¹⁹⁵, which is thus given enough positive character to leave the tetrahedral intermediate; and when it does so the reaction is formally completed.

Protein inhibitors of serine proteases such as soybean trypsin inhibitor undergo the reaction, including formation of an acyl-enzyme, but have many interactions with the protease, so that the first product does not diffuse away and water has no room to attack the acyl-enzyme. In guanidinobenzoyltrypsin (Mangel et al., *Biochemistry* **29**:8352-7 [1990]) the guanidino group has both charge and H-bond interactions at the bottom of the specificity site, and the benzene ring fills the cleft between specificity and catalytic sites; this holds ser¹⁹⁵ down toward asp¹⁸⁹ to a degree not occurring with natural lysine and arginine side chains, which are longer and more flexible. Also, the distance between the carbonyl oxygen and the oxyanion hole is relatively long (3.3, 3.2 Å to the backbone N) so that it will be harder to form H bonds (although the Raman difference spectra of dimethylaminobenzoyltrypsin indicates that the geometry at the carbonyl C is distorted toward a tetrahedral arrangement and H bonds are formed). Furthermore, in the X-ray structure two water molecules can be seen, relatively fixed in positions which would not allow them to attack the acyl-enzyme carbonyl. The authors suggest that even in solution they are H-bonded in those positions and not able to be deprotonated by his⁵⁷ or attack the carbonyl C.

One question I have and don't know the answer to is suggested by Fig. 15.9: why doesn't the ionized acid product stay in the active site, with the negatively charged carboxylate oxygen in the oxyanion hole, as shown in that figure (apparently that peptide acid, with a proline at the P₂ site to change conformation from that of most peptides, is a good inhibitor)? Of course the product initially is uncharged, since an OH is added to the carbonyl, but ionization of the carboxyl should be fast compared to dissociation of the product from the active site. This inhibitor is actually formed by hydrolysis of the amide in the crystal, and perhaps stays there more in the crystal than in solution; I haven't pursued the reference.