

University of Groningen

## Engineering specificity and activity of thermolysin-like proteases from *Bacillus*

de Kreij, Arno

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2001

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

de Kreij, A. (2001). *Engineering specificity and activity of thermolysin-like proteases from Bacillus*. s.n.

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

## Summary and general discussion

Protein engineering is a multidisciplinary approach to study the properties of proteins, combining computational, physical, biochemical and genetic techniques. The use of protein engineering has dramatically increased the knowledge of protein structure, folding and catalysis and the underlying structure-function relations. In this thesis, protein engineering techniques were used to study the determinants of substrate specificity, structural flexibility needed for catalysis, and activity of thermolysin-like proteases (TLPs) produced by members of the bacterial genus *Bacillus*.

The TLP family includes enzymes from a number of Gram-positive bacteria, including pathogens such as *Legionella*, *Listeria*, *Clostridium*, *Staphylococcus*, *Pseudomonas* and *Vibrio*. The TLPs from these organisms contribute to their pathogenicity. Furthermore, the active site organization of TLPs exhibits similarity to those of a number of eukaryotic metalloproteases, in particular to members of the matrix metalloproteases (MMP's). These latter enzymes have been shown to be involved in a number of important processes in man, including the processing of precursors that play modulatory roles in the formation of tumors.

Several members of the TLP family are applied in industry, e.g. in baking, brewing and leather processing. For example, thermolysin is used for the synthesis of the artificial sweetener aspartame. TLPs are not only interesting from a medical and commercial perspective. The availability of an impressive amount of sequence, structural and kinetic data renders this group of proteases an ideal subject for rational design strategies.

Chapter 1, the introduction, deals with general aspects of enzyme classification,

enzymatic catalysis, substrate specificity and active site electrostatics. The presented literature shows that protein engineering is a valuable method to analyze these aspects.

In Chapter 2 a number of TLPs has, for the first time, been characterized with an identical substrate set and a uniform set of assay conditions. Characterization with peptide substrates, as well as HPLC analysis of  $\beta$ -casein digests, showed that the TLP family is a homogeneous family in terms of catalysis, even though there is a significant degree of amino acid sequence variation. The results of this study showed that differences in substrate specificity within the M4 family do not correlate with overall sequence differences, but depend on a small number of identifiable amino acids. Indeed, molecular modeling, followed by site directed mutagenesis of one of the substrate binding pocket residues of the TLP of *Bacillus stearothermophilus* (TLP-ste) converted the catalytic characteristics of this variant into that of thermolysin.

Chapter 3 shows the importance of conserved glycines in proposed hinge-bending regions by analyzing the effects of Gly $\rightarrow$ Ala mutations on catalytic activity. The apparent importance of conserved glycine residues in proposed hinge-bending regions for TLP activity supports the idea that hinge-bending is an essential part of catalysis. The active site of TLPs is located at the bottom of a cleft between the N-terminal and C-terminal domain. Crystallographic studies have shown that the active site cleft is more closed in ligand-binding TLPs than in ligand-free TLPs. Accordingly, it has been proposed that TLPs undergo a hinge-bending motion during catalysis resulting in "closure" and "opening" of the active site cleft. Two hinge regions have been proposed. One is

## Chapter 6

located around a conserved glycine 78, the second involves residues 135 and 136. The importance of conserved glycine residues in these hinge regions was studied by analyzing the effects of Gly→Ala mutations on catalytic activity. Eight such mutations were made in TLP-ste and their effects on activity towards casein and various peptide substrates were determined. Indeed glycines at positions 78, 135 and 136 turned out to be important for catalytic activity. The apparent importance of conserved glycine residues in proposed hinge-bending regions for TLP activity supports the idea that hinge-bending is an essential part of catalysis.

In Chapter 4 the properties of the major specificity determining hydrophobic  $S_1'$  pocket are discussed. In TLP-ste, the hydrophobic  $S_1'$  subsite is mainly formed by Phe130, Phe133, Val139 and Leu202. The effects on substrate specificity of replacing Leu202 by smaller (Gly, Ala, Val) and larger (Phe, Tyr) hydrophobic residues were studied. The results showed that the wild-type  $S_1'$  pocket is optimal for binding leucine side chains. Reduction of the size of residue 202 resulted in a higher efficiency towards substrates with Phe in the  $P_1'$  position. Rather unexpectedly, the Leu202Phe and Leu202Tyr mutations, which were expected to decrease the size of the  $S_1'$  subsite, resulted in a large increase in activity towards dipeptide substrates with Phe in the  $P_1'$  position. This is probably due to the fact that 202Phe and 202Tyr adopt a second possible rotamer which opens up the subsite compared to Leu202 and which favours interactions with the substrate. To validate these results variants of thermolysin were constructed with changes in the  $S_1'$  subsite. Thermolysin and TLP-ste variants with identical  $S_1'$  subsites were highly similar in terms of their preference for Phe versus Leu in the  $P_1'$  position. The 16-fold increase in activity of the Leu202Tyr mutant towards a  $P_1'$  Phe containing substrate is one of the highest found in the literature for a single mutant.

Chapter 5 describes the possibility of changing the active site electrostatics of TLP-ste by inserting or removing charges on the protein surface by site-directed mutagenesis. Double-mutant cycle analysis was used to determine which of the mutated residues are independent of, and which depend on the amino acids present at the other positions chosen for mutagenesis. The results showed that the effects on the  $k_{cat}/K_m$  of single point mutations were non-additive, even in cases where the point mutations were 10 Å or more from the active site  $Zn^{2+}$  and separated from each other by up to 25 Å. This lack of correlation with electrostatic theory implies that electrostatic networks are probably more complex than previously thought and that other effects, such as active site dynamics, may play an important role in determining the active site electrostatics. Several mutations caused a significant increase in the activity, the most active mutant being almost four times as active as the wild-type. The shape of the pH-activity profile was changed significantly. Remarkably, this was achieved without concomitant large changes of the pH-optimum of the enzyme.

In summary, the work described in this thesis shows that the specificity of thermolysin-like proteases can be altered in a predictable way using protein engineering techniques. The comparison of the various wild-type proteases, in chapter 2, identified the substrate specificity-determining amino acids. The exploitation of this knowledge, in chapter 4, resulted in TLP-ste variants with distinctly altered specificity profiles.

It is widely assumed that enzymes undergo hinge-bending motions during catalysis. However, only a few studies have been described in which hinge-bending was analyzed by rationally designed mutations. The results described in chapter 3 support the idea

that hinge-bending is a critical feature of catalysis in TLPs.

Today, one of the biggest challenges is to rationally influence catalysis. The results in chapter 5 show that the catalysis of TLP-ste can be changed by surface charge mutations. However, it is also clear that the results are difficult to interpret with current electrostatic models of the active site. The surprising observation that many mutations on the surface of an enzyme can influence each other, even though they are far apart, leads to an interesting speculation. The question; why are enzymes so big maybe probably answered as follows: If any residue on the surface influences the electrostatics of the active site, and if all those effects are dependent on each other, then nature optimized more than just the active site during evolution.

In this thesis, the concept of protein engineering, that is the multidisciplinary study of structure-function relationships in proteins, has been successfully applied to elucidate some of the structure-function relationships that determine the activity and specificity of TLPs. However, although the concept of protein engineering has often been quite successful, many challenges remain.

Although the specificity determinants of a number of hydrophobic binding pockets of different enzymes have been elucidated, many attempts at rational design of a novel specificity have met with unexpected problems. One such problem is the existence of previously unknown secondary specificity determining elements. Another problem is that a convenient assay to detect enzymatic activity on an industrially or medically relevant substrate can sometimes not

be developed. Therefore, model substrates are often used during the engineering of a protease. However, these model substrates can have conformations different from those on which the protease acts. As a result, the obtained change in specificity with model substrates could be non-existing on the natural substrate. A partial solution to these problems is to carefully choose the model substrates, such that they resemble the substrate of interest as closely as possible.

Today the electrostatic determinants of both enzymatic activity and pH-optimum are as poorly understood as the thermal stability determinants were fifteen years ago. However, the development of novel random mutagenesis methods and automated high throughput screening should facilitate the elucidation of the electrostatic determinants. The application of these random mutagenesis methods should quickly yield variant enzymes with altered electrostatic profiles. Analyzing these variants by reverse engineering, that is, reverting the mutant back to the wild-type step by step until the novel properties are lost, should yield insight into the electrostatic determinants. In this way, a substantial amount of knowledge could probably be generated in less time than by the conventional protein engineering approach. Subsequent application of protein engineering techniques might eventually lead to sufficient knowledge to rationally modify the activity and pH-optimum of enzymes.

The availability of a large number of primary sequences, various expression and purification systems, a large body of kinetic data, and several tertiary structures solved by X-ray crystallography make TLPs an excellent model system to examine structure-function relationships by protein engineering.