The effect of changing the hydrophobic S₁′ subsite of thermolysin-like proteases on substrate specificity

Arno de Kreij, Bertus van den Burg, Oene R. Veltman, Gert Vriend, Gerard Venema and Vincent G. H. Eijsink

1Department of Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Kerklaan, NN Haren, the Netherlands; 2CMBI, Katholieke Universiteit Nijmegen, Toerooiveld, the Netherlands; 3Agricultural University of Norway, Department of Chemistry and Biotechnology, Ås, Norway

The hydrophobic S₁′ subsite is one of the major determinants of the substrate specificity of thermolysin and related M4 family proteases. In the thermolysin-like protease (TLP) produced by Bacillus stearothermophilus (TLP-ste), the hydrophobic S₁′ subsite is mainly formed by Phe130, Phe133, Val139 and Leu202. In the present study, we have examined the effects of replacing Leu202 by smaller (Gly, Ala, Val) and larger (Phe, Tyr) hydrophobic residues. The mutational effects showed that the wild-type S₁′ 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 P₁′ position. Rather unexpectedly, the Leu202—Phe and Leu202—Tyr mutations, which were expected to decrease the size of the S₁′ subsite, resulted in a large increase in activity towards dipeptide substrates with Phe in the P₁′ position. This is probably due to the fact that 202Phe and 202Tyr adopt a second possible rotamer that opens up the subsite compared to Leu202, and also favours interactions with the substrate. To validate these results, we constructed variants of thermolysin with changes in the S₁′ subsite. Thermolysin and TLP-ste variants with identical S₁′ subsites were highly similar in terms of their preference for Phe vs. Leu in the P₁′ position.

Keywords: metalloendopeptidase; thermolysin; Bacillus stearothermophilus; substrate specificity; hydrophobic binding pocket.

Thermolysin-like proteases (TLPs) are members of the peptidase family M4 [1] of which thermolysin (EC 3.4.24.27) is the prototype. The amino-acid sequences of several TLPS have been determined [1] (also see the Merops database at http://www.merops.co.uk/merops/famcards/m4.htm), and the three-dimensional structures of TLPS isolated from several bacteria have been solved (Bacillus thermoproteolyticus [2], Bacillus cereus [3], Pseudomonas aeruginosa [4] and Staphylococcus aureus [5]). TLPS consist of an α helical C-terminal domain and an N-terminal domain mainly consisting of β strands. The domains are connected by a central α helix. This helix is located at the bottom of the active site cleft and contains several of the catalytically important residues (Fig. 1). Four substrate binding pockets (S₂, S₁′, S₁′′ and S₂′; nomenclature according to Schechter and Berger [6]) have been identified [7]. The S₁′ subsite is a hydrophobic pocket that is considered to be a major determinant of substrate specificity [8,9]. In thermolysin and the TLP produced by B. stearothermophilus, the subjects of this study, the S₁′ subsite is mainly formed by Phe130, Val139, Leu202 and Phe133 (TLP-ste) or Leu133 (TLN).

Crystallographic [2,7,10,11] and modelling studies [7] of thermolysin have indicated that the S₁′ subsite allows efficient binding of a leucine side chain. The notion that the S₁′ subsite in thermolysin is not optimal for binding larger residues, such as phenylalanine [7,11], was experimentally confirmed by Izquierdo and Stein [12]. These authors showed a clear positive correlation between the size of the P₁′ residue and the activity of the enzyme on dipeptide substrates of the 3-(2-furylacryloyl)-L-glycyl-L-X-amide type (FaGXa, where X is a hydrophobic amino acid). Phenylalanine, however, did not conform to this trend as illustrated by the fact that similar kₘ values were obtained for FaGLa and FaGFa. The S₁′ subsite of TLP-ste is similar in structure and character to that of thermolysin, but TLP-ste has a higher preference for substrates with a Phe at P₁′.

In the present study, we have investigated the possibility of modifying the S₁′ subsite in TLPS in order to change the preference of the enzyme for Leu and Phe in the P₁′ position. Our hypothesis is that a limited increase in the size of the S₁′ pocket could result in an enzyme that would retain its catalytic power, while displaying an increased preference for Phe at position P₁′ in the substrate.
As we have previously shown [13] mutating Phe133→Leu results in a decreased specificity for P1′ Phe substrates in TLP-ste and changes its substrate specificity into that of TLN. We therefore chose to mutate residue Leu202 which, together with residue 133, dominates the entrance of the substrate to the S1′ pocket [7,13]. We were particularly interested in TLP-ste as we have previously constructed a highly thermostable variant of this enzyme [14]. To validate our models of TLP-ste and the S1′ pocket of the M4 peptidases in general, we also constructed and characterized a number of mutants of thermolysin.

Our results show that it is indeed possible to increase the preference for a P1′ Phe by mutating residue 202. The results also indicate that the substrate specificity of TLP-ste and TLN will change in a similar manner to mutations at positions 133 and 202. This supports the idea that the specificity of other M4 peptidases can be changed in the same way as TLP-ste and TLN.

**MATERIALS AND METHODS**

**Modelling and mutant design**

A three-dimensional model of TLP-ste was built based on the crystal structure of thermolysin, using the molecular modelling program WHAT-IF [15], as has been previously described [16]. The high sequence identity between thermolysin and TLP-ste (86%) indicates that the TLP-ste model should be sufficiently reliable for prediction and analysis of the effects of most amino-acid substitutions [16,17]. Indeed, the TLP-ste model has been used successfully for the design of various stabilizing mutations [18–21]. Throughout this paper, residues in all TLPs are numbered according to the numbering of the corresponding residues of thermolysin.

**Molecular biological techniques**

The nprT gene encoding TLP-ste [22] was cloned, subcloned, and expressed as previously described [23]. The plasmid pUBTZ2 [24] containing the nprM gene encoding thermolysin [25] was obtained from DSM-HSC (Geleen, the Netherlands). Site-directed mutagenesis was performed either by the PCR-based mega-primer method, essentially as described by Sarkar and Sommer [26], or with the QuikChange™ site-directed mutagenesis kit from Stratagene (La Jolla, CA, USA). The QuikChange™ procedure uses a pair of complementary PCR primers that places the mutation in the middle of the primers. pUC18 containing a subcloned fragment of nprT, or pUBTZ2 containing nprM, was amplified using *Pyrococcus furiosus* DNA polymerase (Pfu Turbo) and these primers for 18 cycles in a DNA thermal cycler. After digestion of the parental DNA with *DpnI*, the amplified DNA incorporated with the nucleotide substitution was transformed into *Escherichia coli* XL1-Blue strain. Mutagenic primers were designed such that mutant clones could be recognized by the
presence or absence of an endonuclease restriction site [23].
The nucleotide sequences coding for the mature part of the proteases were verified by DNA sequence analysis. The mutated fragments of TLP-ste were subsequently cloned into the *Bacillus* expression vector, pGE501 [27], containing the TLP-ste gene with a deletion of the previously subcloned fragment.

### Production and characterization of mutant enzymes

Production and purification of the enzymes were performed as previously described [23] using the *B. subtilis* strain DB117 (ΔaprΔapr) [28]. Before determining the kinetic parameters, protease preparations were desalted using prepacked PD-10 gel filtration columns supplied by Amersham Pharmacia (Uppsala, Sweden). Specific activities of the TLP-ste variants towards casein were determined according to a method adapted from Fujii *et al.* [22]: approximately 0.5 μg of protease was incubated in 1 mL of 50 mM 2-amino-2-(hydroxymethyl)-1,3-propanediol (Tris/HCl), pH 7.5, containing 0.8% (w/v) casein and 5 mM CaCl$_2$, at 37 °C for 1 h. Reactions were quenched by the addition of 1 mL of a solution containing 100 mM trichloroacetic acid (TCA), pH 3.5. One unit of activity was defined as the amount of enzyme required to liberate a quantity of acid-soluble peptide corresponding to an increase in A$_{275}$ of 0.001 s$^{-1}$.

The $k_{cat}/K_m$ and $K_m$ values of the enzymes for furylacryloylated dipeptides and tripeptides were determined at 37 °C in a thermostated PerkinElmer Lambda 11 spectrophotometer. The reaction mixture (1 mL) contained 50 mM Tris, 50 mM Mes, pH 7.0, 5 mM CaCl$_2$, 0.01% Triton X-100, 0.5% 2-propanol, 0.01% Triton X-100 and 100 μM FaGLA as substrate. $K_m$ values were calculated by the method described by Hunter and Downs [30].

### Results

### Mutant design and production of mutant proteins

Position-specific rotamer searches [31] for the residues to be introduced at positions 202 and 133 showed that all new side chains could adopt a favorable rotamer without the introduction of steric overlap in both TLP-ste and thermolysin. Furthermore, the modelling studies indicated that the Leu202$\rightarrow$Gly, Leu202$\rightarrow$Ala and Leu202$\rightarrow$Val mutations would simply increase the size of the pocket by a volume corresponding to approximately four, three and one methyl groups, respectively, whereas Leu202$\rightarrow$Phe and Leu202$\rightarrow$Tyr were expected to lead to some reduction in size.

The various mutant TLPs were constructed, produced, purified and characterized as described in Materials and methods. Production and purification yields were similar to those of the wild-type TLP-ste. All variants had similar thermal stabilities and specific activities towards casein (data not shown).

### Characterization of mutant proteases

To examine the enzymatic properties of the mutant TLP-ste enzymes, kinetic parameters for the reaction with available tripeptide substrates as well as the $K_i$ for phosphoramidon were determined (Table 1). The $k_{cat}/K_m$ values for FaGLA indicate that leucine is the optimal residue at position 202 for substrates with a Leu at the P$_1$ position. The results with FaGFL show that the activity towards substrates with a phenylalanine at the P$_1$ position was increased by replacing Leu202 by a smaller residue. For FaGFL, the Leu202$\rightarrow$Ala mutant was the most active. It is interesting to note that the Leu202$\rightarrow$Phe and Leu202$\rightarrow$Tyr mutations were more deleterious to activity towards FaGLA than to activity towards FaGFL.

#### Table 1. Specificity of *B. stearothermophilus* thermolysin-like protease (TLP-ste) variants for tripeptide substrates.

<table>
<thead>
<tr>
<th>TLP-ste variant</th>
<th>FaGLA $k_{cat}/K_m$ (s$^{-1}$·M$^{-1}$·10$^{-3}$)</th>
<th>FaGFL $k_{cat}/K_m$ (s$^{-1}$·M$^{-1}$·10$^{-3}$)</th>
<th>Phosphoramidon $K_i$ (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L202G</td>
<td>2.6</td>
<td>4.7%</td>
<td>1.3 x 10$^3$</td>
</tr>
<tr>
<td>L202A</td>
<td>9.0</td>
<td>16.1%</td>
<td>2.2 x 10$^4$</td>
</tr>
<tr>
<td>L202V</td>
<td>53.9</td>
<td>96.5%</td>
<td>1.5 x 10$^4$</td>
</tr>
<tr>
<td>TLP-ste</td>
<td>55.8</td>
<td>100.0%</td>
<td>8.3 x 10$^4$</td>
</tr>
<tr>
<td>L202F</td>
<td>9.8</td>
<td>17.6%</td>
<td>4.2 x 10$^3$</td>
</tr>
<tr>
<td>L202Y</td>
<td>18.3</td>
<td>32.7%</td>
<td>1.2 x 10$^3$</td>
</tr>
</tbody>
</table>

*a* Reaction conditions: 50 mM Tris, 50 mM Mes, 5 mM CaCl$_2$, 0.01% Triton X-100, pH 7.0, 37 °C. *b* Standard deviations are less than 15% of the value given.

Standard deviations are less than 10% of the value given.
also occupy the S2 variants. A complication of the use of tripeptides is that they allow determination of the *k*
was more active than the Phe variant for both substrates.

available tripeptide substrates do not have the same P 2
increase in *k*
residue, interpretation of mutational effects in terms of
effects displayed in Table 1 may in part be due to effects on
by the mutation of Leu202 [7], meaning that the mutational
di for TLP-ste variants was conducted using the dipeptide substrates
forward. Therefore, additional characterization of the TLP-
\[ 0 \]
change in the preferences for leucine and phenylalanine at

replacement of Leu202 by Phe or Tyr had only modest

TLP-ste and thermolysin (86%) ensures that the three-dimensional
model of TLP-ste is quite accurate, especially in well-
conserved regions such as the active site (e.g. see [16,17,33]
for a discussion). To verify the presumed similarity of
thermolysin and TLP-ste with respect to the S 1 \[ 0 \]
site, thermolysin was mutated to make its S 1 \[ 0 \]
site identical to that of some of the TLP-ste variants. The subsite in both
enzymes is composed of Phe130, Val139 and Leu202 and

replacements of Leu202 by Phe or Tyr resulted in a

Table 2. Specificity of B. steaerothermophilus thermolysin-like protease (TLP-ste) variants for dipeptide substrates.

<table>
<thead>
<tr>
<th>TLP-ste variant</th>
<th>FaGLa</th>
<th></th>
<th>FaGFa</th>
<th></th>
<th>FaGFa/FaGLa</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>k</em>&lt;sub&gt;cat&lt;/sub&gt;/&lt;K&lt;sub&gt;m&lt;/sub&gt;</td>
<td>Relative</td>
<td><em>k</em>&lt;sub&gt;cat&lt;/sub&gt;/&lt;K&lt;sub&gt;m&lt;/sub&gt;</td>
<td>Relative</td>
<td>Ratio</td>
</tr>
<tr>
<td>L202G</td>
<td>0.2</td>
<td>10%</td>
<td>3.3</td>
<td>97%</td>
<td>14.71</td>
</tr>
<tr>
<td>L202A</td>
<td>0.5</td>
<td>22%</td>
<td>27.8</td>
<td>810%</td>
<td>56.39</td>
</tr>
<tr>
<td>L202V</td>
<td>2.2</td>
<td>97%</td>
<td>12.7</td>
<td>370%</td>
<td>5.88</td>
</tr>
<tr>
<td>TLP-ste</td>
<td>2.2</td>
<td>100%</td>
<td>3.4</td>
<td>100%</td>
<td>1.55</td>
</tr>
<tr>
<td>L202F</td>
<td>0.9</td>
<td>42%</td>
<td>11.8</td>
<td>344%</td>
<td>12.62</td>
</tr>
<tr>
<td>L202Y</td>
<td>3.1</td>
<td>141%</td>
<td>56.6</td>
<td>1648%</td>
<td>18.12</td>
</tr>
</tbody>
</table>

* Reaction conditions: 50 mM Tris, 50 mM Mes, 5 mM CaCl<sub>2</sub>, 0.01% Triton X-100, pH 7.0, 37 °C. b Standard deviations are less than 15% of the value given.

towards FaGFL (with the larger Phe at P 1 \[ 0 \]). The Tyr variant was more active than the Phe variant for both substrates.

The solubility of the tripeptide substrates was sufficient to

leucine and phenylalanine at P 1 \[ 0 \]. This shows the reliability of the TLP-ste model (at least

Table 3. Comparison of the specificity of thermolysin and TLP-ste mutants towards dipeptide substrates. The wild-type enzymes are shown in bold type.

| Position | FaGLa | | FaGFa | | Phe/Leu |
|----------|-------|-----|-------|-----|
| Enzyme   | 133   | 202 | 133   | 202 | ratio |
| TLN      |       |     |       |     |
| Leu      | 12.3  | 3.9 | 0.3   |     |
| Phe      | 10.8  | 15.6| 1.5   |     |
| Leu      | 25.3  | 230.0| 9.1   |     |
| Phe      | 13.0  | 186.4| 14.3  |     |
| TLP-ste  |       |     |       |     |
| Leu      | 17.2  | 6.3 | 0.4   |     |
| Phe      | 2.3   | 3.4 | 1.5   |     |
| Leu      | –     | –   | –     |     |
| Phe      | 3.1   | 56.6| 18.3  |     |

* Reaction conditions: 50 mM Tris, 50 mM Mes, 5 mM CaCl<sub>2</sub>, 0.01% Triton X-100, pH 7.0, 37 °C. b Standard deviations are less than 15% of the value given.
Fig. 2. The P1′ residues preferentially occupy a specific location in the S1′ subsite. A line diagram showing a superposition of several thermolysin structures (PDB accession nos 1TLP, 3TMN, 2TMN, 6TMN, 1TMN and 4TLN). The arrow indicates the position that is always occupied by a ligand-carbon atom (see text for details). The superposition shows the remarkable rigidity of the subsite residues and also shows the variability in the positions occupied by the various ligands.

near the S1′ subsite) and the usefulness of thermolysin crystal structures in discussing the effects of the mutations. The activities of the different enzymes towards each substrate show that thermolysin generally is a more active enzyme than TLP-ste. Although we can only speculate as to the origin of this difference, differences in hinge bending motions [34,35] or in active site electrostatics attributable to the 42 dissimilarities in amino-acid composition may play a role.

Owing to extensive crystallographic studies by Matthews and coworkers [2,10,11], considerable information is available concerning the interaction between thermolysin and a variety of ligands. Superposition of a series of enzyme-ligand complexes shows that the residues making up the S1′ subsite have highly invariant positions (0.15–0.20 Å rmsd), indicating that they hardly adapt to the P1′ residue in the ligand (Fig. 2). Instead, the ligand seems to adapt, thus ensuring that specific ligand–enzyme interactions are preserved, regardless of the type of P1′ residue. Ligands with a leucine chain at P1′ show a prominent preserved interaction involving a C8 atom on the substrate and a Leu133-C6 and Leu202-C8 in the S1′ pocket. In the case of shorter P1′ side chains without C6, the Cα atoms of the ligand shift (by up to 1.2 Å). Consequently, a carbon in the P1′ side chain occupies the position that is occupied by one of the C6 atoms where P1′ is leucine. Therefore, in contrast to what might be expected, larger substrates do not penetrate the pocket more deeply than smaller substrates. In addition to showing that the C8 atoms of Leu202 have important interactions with the P1′ side chain, the crystal structures also suggest that these C6 atoms are a major sterical hindrance for accommodating larger side chains (such as Phe) in the S1′ subsite.

Interestingly, it was indeed possible to increase the activity of TLP-ste towards substrates with Phe at P1′ by reducing the size of residue 202. Of the Val, Ala and Gly mutants, the Ala mutant was most active towards substrates with a Phe at P1′. Modelling studies (not shown) indicated that alanine is the best compromise between creating space in the S1′ pocket (not sufficient in Leu202→Val) and keeping as many contacts as possible with the P1′ side chain (better in Leu202→Ala than in Leu202→Gly).

If Phe and Tyr would adopt the same rotamer as Leu at position 202, then their aromatic rings would partly end up in the S1′ pocket, which would result in a considerable reduction in size and accessibility. This would obviously lead to a drastic reduction in catalytic activity towards all substrates tested. Such reductions were not observed, indicating that Phe and Tyr adopt a second favorable side chain conformation in which the aromatic ring is parallel to the substrate. Molecular modelling indicates that the distance between the aromatic rings of 202Phe/Tyr and the P1′ Phe would be in the range of 4 Å, making steric overlap unlikely. Interestingly, the presence of the aromatic residue at position 202 may permit beneficial aromatic–aromatic interactions [36,37] with substrates that have Phe in the P1′ position. Such interactions may contribute to the activity of the Leu202→Phe and Leu202→Tyr mutants towards the Phe-containing substrates.

The Tyr mutant had clearly anomalous characteristics, such as a highly increased Kcat and higher activity than the Phe mutant for all substrates tested. Most remarkably, the Tyr mutant clearly had the highest activity towards FaGFL as all variants tested in this study. Modelling studies indicate that this anomalous behavior may be caused by a hydrogen bond between the Tyr-OH and the substrate. A detailed explanation of the remarkable effects of Leu202→Tyr awaits crystallographic studies of the mutant.

For substrates with a Phe at P1′, mutational effects were more pronounced for the dipeptides than for the tripeptides. This may be due to the fact that residue 202 also affects the S1′ subsite (illustrated by thermolysin-ligand complexes with PDB accession nos 1TLP, 5TMN and 6TMN [38–40]). It would not be surprising if effects on the S1′ subsite are more notable for FaGFL, as this substrate has a Leu at P1′. As explained above, a Leu at P1′ interacts strongly with S1′ [32] and its interactions with the enzyme are therefore more likely to be affected by mutation of residue 202.

Several mutagenesis studies of proteases have shown that it is possible to manipulate substrate preferences by changing the size and/or character of hydrophobic binding pockets [41–50]. For example, diminishing the space in the S1 subsite of subtilisin YaB [41] and subtilisin E [42] by increasing the size of the subsite residues led to reduced activity towards substrates with large P1 residues while yielding higher activity towards substrates with smaller P1 residues. When it comes to binding of hydrophobic side chains in a subsite, substrate and subsite geometry play a role in addition to subsite size (e.g [43,44]). In the present study we probably see both types of effects. The effects on activity of reducing the size of residue 202 are likely to be caused at least in part by the increase in size in the S1′ subsite. However, as Leu and Phe are not that different in size and shape, it is likely that mutational effects also reflect changes in the quality of the S1′–P1′ interactions. Considering the similarity between Leu and Phe, the changes in...
substrate preferences that were obtained in the present study are remarkably large, especially for the dipeptide substrates. These changes were obtained by considerable increases in activity for substrates with a P_1^I Phe, and not primarily by deterioration of activity towards substrates with a P_1^I Leu (in contrast to [45]). It is interesting to note that the 16-fold increase in activity that the Leu202→Tyr mutant displays towards FaGFa is one of the highest found to date for a single mutant in similar studies.

ACKNOWLEDGEMENT

The authors wish to thank Dr Theo Sonke from DSM-Research for supplying pUBTZ2 and helpful discussions.

REFERENCES


