Structural and functional characterisation of a lytic transglycosylase and two ω-transaminases
Reddem, Eswarreddy

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:
2015

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

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).

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.

Download date: 30-09-2019
Chapter 3

Structure of the cell wall lytic transglycosylase MltF from *Pseudomonas aeruginosa*

Eswar Reddy Reddem, Bauke W. Dijkstra and Andy-Mark W. H. Thunnissen

The native crystal structure reveals the presence of two domains, an N-terminal domain that has a periplasmic solute-binding-protein (SBP) fold, which is closely related to the SBPs of ABC transporters, and a C-terminal domain that has a fold similar to that of the catalytic domains of family 1 lytic transglycosylases. Soaking experiments with amino acids had revealed that enzyme N-terminal domain could bind medium sized amino acids with a lower affinity (nM). The covalent binding of substrate (GlcNAc-MurNAc-L-Ala-D-Glu) and inhibitor (Bulgécin A) to the C-terminal domain indicates that *paMltF* follows the same reaction mechanism as previously described, by using the Glu316 as the catalytic acid/base residues.

Submitted for publication
Chapter 3

ABSTRACT

Bacterial lytic transglycosylases are peptidoglycan-degrading enzymes with roles in bacterial cell wall remodeling and turnover during cell growth and division. Here we report the crystal structure of a truncated but active form of the lytic transglycosylase MltF from *Pseudomonas aeruginosa* (*paMltFΔC*), which lacks the 29 C-terminal amino acid residues. The structure reveals the presence of two domains. The N-terminal domain has a periplasmic solute binding protein (SBP) fold, which is closely similar to that of SBPs of ABC transporters that bind polar and charged amino acids, although the sequence identity is low (15-24%). Structures of *paMltFDC* complexed with leucine, isoleucine, valine or cysteine confirm that the N-terminal domain has the ability to bind amino acids. However, only part of the binding interactions typically observed in amino acid specific SBPs is present in the *paMltFDC*/amino acid structures, and the affinity of *paMltFDC* for these amino acids is very low. The C-terminal domain of *paMltFDC* has a lysozyme-like fold and resembles the catalytic domain found in other lytic transglycosylases, e.g. Slt70, MltC and MltE. A crystal structure of a ternary complex of *paMltFΔC* with a small peptidoglycan fragment and the glycopeptide inhibitor bulgecin identifies the interactions relevant for peptidoglycan binding and cleavage, pointing out Glu316 as the general acid/base in the lytic transglycosylase catalytic mechanism.
Chapter 3  

INTRODUCTION

Lytic transglycosylases (LTs) are N-acetylmuramidases that catalyze the cleavage of the β-1,4-linkage between the N-acetylmuramic acid (MurNAc) and N-acetylg glucosamine (GlcNAc) residues in the bacterial cell wall material peptidoglycan. They act on the same glycosidic bonds as lysozymes do, but instead of hydrolyzing the bond and producing a reducing end MurNAc residue, they catalyze a transglycosylation reaction resulting in the formation of a non-reducing 1,6-anhydromuramoyl product (Höltje et al., 1975; Labischinski et al., 1991; Höltje, 1998; Vollmer et al., 2008). In bacteria, lytic transglycosylases are crucial for cell wall remodeling and recycling during cell growth and division, for the creation of sites for the insertion of flagella, and the creation of pores for secretion systems (Scheurwater et al., 2008).

In *Escherichia coli*, seven LTs have been functionally characterized, one soluble enzyme, Slt70, and six membrane-anchored LTs, MltA, MltB, MltC, MltD, MltE (also called EmtA), and MltF (or YhdD). These enzymes differ in their substrate preferences, in their enzymatic activity, and in their susceptibilities towards certain glycopeptide inhibitors (Romeis et al., 1993; Quintela et al., 1995; Scheurwater et al., 2008; Vollmer et al., 2008; van Heijenoort, 2011).

The first lytic transglycosylase that was successfully crystallized was Slt70 (Rozeboom et al., 1990). Its crystal structure revealed an unexpected super-helical ring of 26 α-helices with a separate domain on top that resembled the structure of goose-type lysozyme, and that contained the active site (Thunnissen et al., 1994; Thunnissen et al., 1995b). Subsequently, crystal structures were elucidated of soluble fragments (without the membrane anchor) of MltB (van Asselt et al., 1999a), MltA (van Straaten et al., 2005), and MltE (Fibransah et al., 2012; Artola-Recolons et al., 2014). In addition, a crystal structure is known of MltC (Artola-Recolons et al., 2011). These structures showed that, except for MltA, all of them contain a lysozyme-like catalytic domain, despite a lack of significant sequence homology. In contrast, MltA has a catalytic domain with a double-psi β-barrel fold, resembling the catalytic domain of endoglu canase V (van Straaten et al., 2005; Powell et al., 2006).
Chapter 3

The only *E. coli* LT that has so far eluded structural characterization is MltF. MltF is an outer-
membrane-attached periplasmic enzyme that contains a typical transglycosylase sequence
motif in its C-terminal domain, in agreement with the C-terminal domain harboring the
enzyme’s peptidoglycan degrading activity (Scheurwater & Clarke, 2008). However, unlike all
other LTs, MltF contains an N-terminal domain that is homologous to the periplasmic solute-
binding protein (SBP) domains of ABC transporters, in particular to those specific for the amino
acids histidine, lysine/arginine/ornithine and glutamine (Scheurwater & Clarke, 2008). The
function of this N-terminal domain of MltF is as yet unknown.

Since our structural studies on *E. coli* MltF were unsuccessful after the initial crystallization of
the protein (Madoori & Thunnissen, 2010; see Chapter 2), we focused our efforts on the MltF
from *Pseudomonas aeruginosa*, which has 33% sequence identity to *E. coli* MltF. As detailed in
Chapter 2, we could crystallize a fully active, recombinant form of the protein, designated as
paMltFΔC, containing residues 27–461 of the gene-derived amino acid sequence with Cys27
replaced by a serine residue, and lacking the C-terminal residues 462-490. Here we describe the
crystal structure of paMltFΔC. In addition, crystal structures were obtained of paMltFΔC with
different amino acids bound (valine, cysteine, leucine and isoleucine) to its N-terminal domain
and of paMltFΔC in a ternary complex with the murodipeptide GlcNAc-MurNAc-L-Ala-D-Glu and
the glycopeptide bulgecin A bound to its C-terminal domain. These structures revealed that
paMltF contains an open, extended peptidoglycan-binding groove, with multiple saccharide-
binding subsites, similar as in the previously determined structure of *E. coli* MltE.
Experimental procedures:-

**Crystal structure elucidation of paMltFΔC.** Details of the cloning, overexpression, purification, buffer optimization and crystallization procedures have been given in Chapter 2. The *paMltFΔC* crystal structure was elucidated by the single-wavelength anomalous dispersion (SAD) method using diffraction data from a Na-iodide-soaked crystal collected in house (see Chapter 2). The final model of *paMltFΔC* was obtained after several cycles of restrained positional and B-factor refinement with the program Refmac5 (Murshudov *et al.*, 2011), alternated by manual model building using Coot (Emsley *et al.*, 2010). Data collection and refinement statistics can be found in Table 1.

**Structure determination of paMltFΔC-ligand complexes**

**Preparation of a ternary paMltFΔC-murodipeptide-bulgecin complex.** Native *paMltFΔC* crystals were briefly soaked for 1 hour at room temperature in mother liquor (0.1 M Tris-buffer, pH 8.5, 0.2 M MgCl₂, 20% PEG 8K) supplemented with 10 mM bulgecin A and 15 mM N-acetylglucosaminyl-N-acetylmuramyl-L-alanyl-D-glutamic acid (Sigma-Aldrich). The soaked crystals were transferred to a cryo-protection solution containing 40% PEG400 in mother liquor, and were directly cryo-cooled in liquid nitrogen. Diffraction data were collected at beam line ID29 of the ESRF (Grenoble, France). Data processing followed a similar procedure as used for *paMltFΔC* (see Chapter 2 and Table 1 for details). Refinement and model building were also done as for *paMltFΔC* (see above and Table 1).

**Binding studies with amino acids.** To analyze the interaction of *paMltFΔC* with amino acids, samples of the protein were subjected to a thermal shift assay analysis (Ericsson *et al.*, 2006) following the same procedures as given in Chapter 2, and using the 20 common amino acids. Crystals were produced by streak seeding (Chapter 2) in mother liquor supplemented with 0.1 M of the amino acid of interest. Diffraction data were collected in-house at 110 K using a Microstar rotating anode X-ray source (Bruker AXS GmbH) in combination with Helios optics (Incoatec GmbH) and a MAR345dtb detector (Marresearch GmbH). The ligand bound structures were built using the native structure as input model and refinement followed the same procedures as for native *paMltFΔC.*
Chapter 3

**Sequence and structure analysis**

Close homologs of *pa*MltF and multiple sequence alignments for sequence conservation analysis were obtained using BLAST (McWilliam *et al.*, 2013), the Consurf server (Celniker *et al.*, 2013), and Weblogo (Crooks *et al.*, 2004).

**Accession numbers**

Atomic coordinates have been deposited with the Protein Data Bank with accession codes 4P11 for the native enzyme and 4P0G for the ternary *pa*MltFΔC/bulgA-GM-Ala-Glu complex. The amino acid bound structures have been deposited with PDB codes 4OXV(valine), 4OYV (leucine), 4OZ9 (isoleucine), and 4OWD (cysteine), respectively.
### Chapter 3

<table>
<thead>
<tr>
<th>Data collection</th>
<th>Apo enzyme</th>
<th>Bulgedin/muro-dipeptide</th>
<th>Valine</th>
<th>Leucine</th>
<th>Isoleucine</th>
<th>Cysteine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beamline</td>
<td>1023-2</td>
<td>1014-1</td>
<td>In house</td>
<td>1n house</td>
<td>In house</td>
<td>In house</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>4.08(87)</td>
<td>4.08(89)</td>
<td>3.9(66)</td>
<td>3.9(66)</td>
<td>2.2(68)</td>
<td>4.3(81)</td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁,2,2,1</td>
<td>P2₁,2,2,1</td>
<td>P2₁,2,2</td>
<td>P2₁,2,2</td>
<td>P2₁,2,2</td>
<td>P2₁,2,2</td>
</tr>
<tr>
<td>Unit cell (a, b, c, Å)</td>
<td>98.85(24)</td>
<td>98.85(24)</td>
<td>98.85(24)</td>
<td>97.029(51)</td>
<td>77.6(17)</td>
<td>97.029(51)</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>12.9 (2.1)</td>
<td>17.3 (2.3)</td>
<td>6.9 (2.4)</td>
<td>7.0 (2.4)</td>
<td>12.9 (2.3)</td>
<td>22.7 (2.2)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>72651 (6641)</td>
<td>52609 (5279)</td>
<td>12588 (1265)</td>
<td>17631 (1679)</td>
<td>19521 (1973)</td>
<td>Z23 (1279)</td>
</tr>
<tr>
<td>Rwork</td>
<td>0.09</td>
<td>0.10</td>
<td>0.09</td>
<td>0.09</td>
<td>0.05</td>
<td>0.04</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>39.30</td>
<td>41.7</td>
<td>39.30</td>
<td>39.30</td>
<td>39.30</td>
<td>39.30</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>19.22</td>
<td>19.22</td>
<td>29.8312</td>
<td>28.2262</td>
<td>24.6240</td>
<td>20.0 (20)</td>
</tr>
<tr>
<td>Mean (I/σ(I))</td>
<td>0.007, 1.1</td>
<td>0.008, 1.2</td>
<td>0.009, 1.2</td>
<td>0.010, 1.2</td>
<td>0.008, 1.1</td>
<td>0.008, 1.1</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>99.0</td>
</tr>
<tr>
<td>R-work, Rfree</td>
<td>99.0</td>
<td>99.0</td>
<td>98.0</td>
<td>98.0</td>
<td>99.0</td>
<td>99.0</td>
</tr>
</tbody>
</table>

**Table 1:** Data collection and refinement statistics of paΔCMltF
Chapter 3

RESULTS AND DISCUSSION

Structure elucidation. The structure of $paMlt\Delta C$ was initially elucidated by the single wavelength anomalous dispersion (SAD) phasing method at 2.45 Å resolution using a single NaI-soaked crystal, and subsequently refined at 1.85 Å resolution to a final crystallographic R-factor of 15.0 %. In addition, structures were obtained of various ligand-bound complexes at resolutions ranging from 1.65 to 2.40 Å (see Table 1 for details). Although the construct contains 435 amino acid residues (residues 27-461), interpretable electron density was only present for residues 32-459.

As already predicted by previous sequence comparisons (Scheurwater & Clarke, 2008) $paMlt\Delta C$ has two well-defined domains, an N-terminal domain with a solute-binding protein (SBP)-like fold, and a C-terminal catalytic domain, which has a lysozyme-like fold and contain the essential catalytic glutamate residue (Figure 1).

![Figure 1](image)

**Figure 1:** Overall structure of $paMlt\Delta C$. (A) Cartoon representation of $paMlt\Delta C$, with the SBP-like N-terminal domain in green (SBP-subdomain I) and red (SBP-subdomain II), and the C-terminal LT-domain in yellow. The catalytic Glu316 residue is shown in red sticks. (B) Secondary structure topology of $paMlt\Delta C$. Colors are as in (A). The secondary structure was assigned with the program DSSP (Andersen et al., 2002).
The N-terminal SBP domain

The structure of the N-terminal domain (NTD) of paMltFΔC comprises residues 32-273. It is built up of two α+β subdomains that are linked together via two β-strands (β7 and β13, Figure 1). Subdomain I, encompassing residues 32-124 and 231-273, consists of six α-helices (α1-α3, α7-α9) packed against a five-stranded mixed parallel/antiparallel β-sheet (β1, β4-β6, β14). The strand order of this β-sheet is β4-β1-β5-β14-β6, with β14, the first β-strand after the crossover from subdomain II to subdomain I, being the only anti-parallel β-strand in the sheet. Subdomain II, encompassing residues 131-223 has a similar five-stranded mixed parallel/antiparallel β-sheet (β8-β12), but is surrounded by three α-helices (α4-α6). The topology of this β-sheet (β10-β9-β11-β8-β12) is identical to that of subdomain I, with β8 running anti-parallel with respect to the other strands and following the crossover from subdomain I to II. The sheet topology of the two subdomains and the linking of the subdomains via two β-strands (β7, β13) identify paMltFΔC-NTD as a Class II SBP (Fukami-Kobayashi et al., 1999; Berntsson et al., 2010).

Comparison of paMltFΔC-NTD with SBPs

To further characterize the structure of paMltFΔC-NTD we carried out a DALI analysis (Holm & Rosenström, 2010). Table 2 summarizes the five structurally most similar proteins, which are all periplasmic solute-binding proteins (SBPs) specific for amino acids. Figure 2 provides a structure-based amino acid sequence alignment of these proteins and paMltFΔC-NTD, also including the sequence of the N-terminal domain of E. coli MltF (ecMltF-NTD).
Table 2: Periplasmic amino-acid-binding proteins with the highest structural similarity to the N-terminal domain of paMltFΔC according to a DALI analysis (Holm & Rosenström, 2010).

<table>
<thead>
<tr>
<th>Ligand specificity</th>
<th>Organism</th>
<th>Seq.LD</th>
<th>PDB code</th>
<th>Z score</th>
<th>RSMD (Å)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histidine</td>
<td>E. coli</td>
<td>14%</td>
<td>1HSL</td>
<td>22.6</td>
<td>2.4</td>
<td>(Yao et al., 1994)</td>
</tr>
<tr>
<td>Glutamine</td>
<td>E. coli</td>
<td>18%</td>
<td>1WDN</td>
<td>22.6</td>
<td>2.2</td>
<td>(Sun et al., 1998)</td>
</tr>
<tr>
<td>Lysine/arginine/histidine</td>
<td>Geobacillus stearothermophilus</td>
<td>18%</td>
<td>2Q2C</td>
<td>22.1</td>
<td>2.2</td>
<td>(Vahedi-Faridi et al., 2008)</td>
</tr>
<tr>
<td>Cysteine</td>
<td>Thermus thermophilus</td>
<td>24%</td>
<td>3VVS</td>
<td>24.6</td>
<td>2.1</td>
<td>(Kanemaru et al., 2013)</td>
</tr>
<tr>
<td>D-Alanine</td>
<td>Salmonella enterica</td>
<td>14%</td>
<td>4DZ1</td>
<td>22.1</td>
<td>2.4</td>
<td>(Osborne et al., 2012)</td>
</tr>
</tbody>
</table>

The DALI analysis revealed that the paMltFΔC-NTD structure most closely resembles SBP structures that have a closed conformation, with an amino acid ligand bound in their binding pocket. Superimposing paMltFΔC-NTD on these structures shows that paMltFΔC-NTD has a similar pocket. It is located between the two subdomains of the SBP domain, and is more spacious than the binding pocket in the SBPs and appears to be solvent accessible. The pocket is rather hydrophobic, lined by the apolar side chains of Leu90, Leu93, Ala108, Leu110, Pro112, Val185, Val202 and Leu229. In addition, several polar side chains are present (Arg50, Tyr94, Thr111, Ser159, Ser160, His161, and Asp203). To investigate whether the pocket of paMltF-NTD contains any conserved amino acid residues that might function in ligand binding, we compared it to the ligand-binding regions of the SBP proteins listed in Table 2. From our comparison we noted that the SBPs have six regions that contribute to ligand binding (regions S1 – S6 in Figure 2). These regions comprise loops and termini of secondary structure elements that face the inter-subdomain cleft. Residues at regions S3 and S5 bind the a-NH$_3^+$ group of the ligand (e.g. Ser, Asp; Figure 4B), while those at regions S3 and S4 bind the a-COOH moiety (e.g. Arg, Thr; Figure 4B). Finally, residues from regions S1, S2, and S6 interact with the side chain of the ligand. In paMltFΔC-NTD only a few of the S1-S6 residues are conserved (Figure 2A). Most conspicuously, the conserved arginine residue in S3, which binds the a-COOH of the ligand in
the SBPs, is absent in paMltFΔC-NTD. Only the Asp from S5 is retained in paMltFΔC-NTD, and might be available to bind a ligand’s amino group.

To analyze whether sequence conservation within the MltF family might give clues on the function of the N-terminal domain we compared the paMltF-NTD sequence with MltF homologues. As can be deduced from Figure 2B the most conserved residues in the MltF family are Ala, Gly, Leu, Arg in S3, Ser in S4, Asp in S5, and Trp in S6. Since the Arg and Asp are also conserved in the SBPs, this may indicate that the MltF-NTD indeed could have an amino-acid-binding function. However, in paMltFΔC-NTD only the Asp is conserved.

Thus, although the overall structure of paMltFΔC-NTD is closely related to Class II periplasmic amino-acid-binding SBPs, the low residue conservation and the absence of the arginine that interacts with the carboxylate group of the ligand question whether paMltFΔC-NTD has an amino-acid-binding function. However, this may be different for other MltF family members that do have a conserved Arg in the S3 region as well as a conserved Asp in the S5 region. In addition, Arg113 in S3 of paMltFΔC-NTD, which has a different sequence position from the conserved Arg in the other proteins, could perhaps take over the carboxylate-binding role.
Figure 2. (A) Structure-based multiple sequence alignment of the N-terminal domains of *P. aeruginosa* and *E. coli* MltF (paMltFΔC-NTD and ecMltF-NTD) with closely related periplasmic solute-binding proteins. 3VV5, crystal structure of (S)-2-aminoethyl-L-cysteine binding protein TTC0807 from *Thermus thermophilus* (Kanemaru et al., 2013); 4GVO, crystal structure of cysteine-binding protein from *Listeria monocytogenes* (unpublished); 2Q2C, crystal structure of arginine-, lysine-, histidine-binding protein Art from *Geobacillus stearothermophilus* (Vahedi-Faridi et al., 2008); 1HSL, crystal structure of histidine-binding protein Hisf from *E. coli* (Yao et al., 1994); 1WDN, crystal structure of glutamine-binding protein GlnBP from *E. coli* (Sun et al., 1998). Boxes indicate the six substrate-binding regions (S1-S6) in SBPs.
Spheres indicate the amino-acid-binding residues in **paMltFΔC** (red, H-bonds with  
-a-carboxylate group of amino acid ligand; blue, H-bonds with a-amino group of amino acid ligand; black, van der Waals contacts 
with side chain of amino acid ligand; see also Figure 4). The secondary structure is based on the crystal 
structure of **paMltFΔC**. Residues in the SBPs sequences which are structurally equivalent to **paMltFΔC** 
[based on 3D structural alignments using PDBeFold (Krissinel, 2007)] are given in upper case, structurally 
non-equivalent residues in lower case. Residue coloring follows the Clustal X coloring scheme. (B) 
Conservation of residues in the **paMltF-NTD** ligand-binding region in the MltF family. This figure was made 
using Weblogo ([Crooks et al., 2004](https://www.ncbi.nlm.nih.gov/pubmed/14735672)).

**Binding experiments with amino acids**

To reveal whether **paMltFΔC-NTD** binds single amino acids we subjected the protein to 
thermal shift assays ([Ericsson et al., 2006](https://www.ncbi.nlm.nih.gov/pubmed/16862555)) (Figure 3). In the absence of amino acids, the 
observed melting temperature (**Tm**) of **paMltFΔC** was 50 ± 1 °C (Figure 3A). Seven of the 20 
common amino acids tested as potential ligands of **paMltFΔC** caused a small but significant 
increase of the **Tm** (the threshold was set at +2 °C): aspartate, arginine, glutamate, glutamine, glycine, threonine and tyrosine. Four amino acids caused the appearance of a second melting 
transition at a **Tm** ranging between 54 °C and 62 °C, in addition to the 50 °C transition: valine, 
leucine, isoleucine and cysteine (Figure 3B). The nature of the double melting transition caused 
by the second group of amino acids is unknown. Unfortunately, the results of the thermal shift 
assays are inconclusive. The differences in thermal melting behavior of **paMltFΔC** caused by the 
amino acids may indicate binding, but this binding appears to be rather aspecific. Also, from 
these experiments alone it is not possible to distinguish whether the effects are due to binding 
to the N-terminal or the C-terminal domain of **paMltFΔC**. Further experiments to obtain 
quantitative data on the affinity of **paMltFΔC** for amino acids revealed that **paMltFΔC** has only 
very low affinity for amino acids (mM range; data not shown), thus casting further doubt on a 
function of **paMltFΔC-NTD** in amino acid binding.

87
Figure 3: (A) Example of the poMltΔC melting transition, as obtained with a thermal shift assay (Ericsson et al., 2006), with and without amino acid (valine). (B) Overview of melting temperatures of poMltΔC (second transition) with the 20 different amino acids of, Native poMltΔC. See text for explanation.
**Figure 4:** Comparison of ligand binding modes in *pa*MltFΔC-NTD and HisJ. (A) Binding modes of cysteine, valine, leucine and isoleucine in *pa*MltFΔC-NTD. (B) Binding mode of histidine in HisP from *E. coli* (PDB entry 1HSL). (C and D) Cut-away surface representation of both amino acid bound *pa*MltF-NTD and histidine-bound HisP.

**Binding mode of amino acids to *pa*MltFΔC-NTD**

To further establish whether *pa*MltF has an amino acid binding potential, co-crystallization screens were carried out of *pa*MltFΔC in the presence of different amino acids. As a result we elucidated crystal structures of *pa*MltFΔC with bound valine, cysteine, leucine, and isoleucine at 2.2 - 2.4 Å resolution. In all cases, well-defined electron density was visible for a bound amino acid in the pocket at the interface of the two subdomains. Its side chain has extensive van der Waals interactions with the surrounding apolar protein side chains. Its α-NH$_3^+$ group forms an ion pair with the carboxylate group of Asp203 and has further polar interactions with the backbone carbonyl oxygen atom of Gly109 and the side chain hydroxyl group of Thr111. Its α-carboxylate group is hydrogen-bonded to the hydroxyl group of Ser160, and, via a water molecule, to the Tyr94 hydroxyl group (see Figure 4); no positive charge is nearby to stabilize the carboxylate of the bound amino acid. The absence of such a charge interaction may explain the low affinity of *pa*MltFΔC for amino acids. Thus, in view of its low affinity, further research will be needed to assess whether amino acid binding is a physiologically important function for *pa*MltFΔC.

**The C-terminal catalytic domain**

The C-terminal catalytic domain consists of residues 274-459, which are organized into 10 a-helices (α10-α19) and three, short β-strands. The domain can be considered to consist of two subdomains, the N-terminal subdomain (residues 274-349) and the C-terminal subdomain (residues 389-459), which are connected by the long a-helix a14 (residues 350-373). The N-terminal subdomain contains four a-helices (α10-α13), with a 20-residue loop inserted between α12 and α13, which forms an irregular sheet of three short antiparallel β-strands. The C-terminal subdomain is made up of five a-helices (α15-α19) (Figure 1B). A deep groove, which
Chapter 3

harbours several conserved amino acid residues characteristic of family 1 LTs (Artola-Recolons et al., 2011; Fibriansah et al., 2012; Artola-Recolons et al., 2014), runs across the surface separating the two subdomains. Most importantly, Glu316, the putative catalytic acid/base residue of \( \text{paMltF} \Delta \text{C} \), is centrally positioned in this groove at the C-terminus of helix a12.

A DALI search (Holm & Rosenström, 2010) revealed that the structure of the \( \text{paMltF} \Delta \text{C} \) catalytic domain is most similar to the fold of the catalytic domains of MltE (Z-score = 18.7, with 161 of the 190 Ca atoms aligning with an RMSD of 2.0 Å; sequence identity 18%; PDB entries 3T36 and 2Y8P (Artola-Recolons et al., 2011; Fibriansah et al., 2012)), MltC (Z-score = 17.1, with 156 of the 173 Ca atoms aligning with an RMSD of 1.4 Å; sequence identity 19.0%; PDB entry 4CFO (Artola-Recolons et al., 2014)) and Slt70 (Z-score = 16.8, with 154 of 173 Ca atoms aligning with an RMSD of 2.4 Å; sequence identity 23% PDB entry 1QSA (van Asselt et al., 1999b)). These catalytic domains have all structures that are very similar to the structure of goose-type lysozyme, as first revealed for the catalytic domain of Slt70 (Thunnissen et al., 1995b).

Comparing the 3D structures of \( \text{paMltF} \Delta \text{C} \)-CTD and its nearest structural homologue \( \text{ecMltE} \) shows that, despite low sequence identity (18 %), the core secondary structure elements (a11, a12, a14, a15, a16, a19, and the three short b-strands (Figure 1B, Figure 5) superpose very well. The first a-helix a10 comes in from a slightly different direction, and helix a13 corresponds to helix a4 of MltE, but is somewhat shifted because of the amino acid insertions in MltE (Figure 6). Residues 346-348 (a5 in MltE) are in a loop leading to a14. Finally, a17 is shifted along its helix axis compared to a10 in MltE, and a18 is a short a-helix in the insertion of residues 419-435. As expected, all the amino acid insertions and deletions occur at the surface of the molecule, and they are far from the active site near Glu316, at the end of a12 (Figure 1).
Figure 5. Structure-based sequence alignment of the C-terminal domains of *P. aeruginosa* and *E. coli* MltF (paMltF-CTD and ecMltF-CTD) with *E. coli* MltE (ecMltE; (Fibriansah et al., 2012)). Spheres indicate the saccharide-binding residues in paMltFΔC-CTD and ecMltE. Secondary structures are based on the crystal structures of paMltFΔC and ecMltE (PDB entry 4HJV). Structurally equivalent residues in paMltFΔC and ecMltE are given in upper case, while structurally non-equivalent residues are given in lower case.

The superposition further shows that Glu316 of paMltFΔC is equivalent to Glu64 of ecMltE. Since in ecMltE Glu64 acts as the catalytic acid/base (Fibriansah et al., 2012), this suggests that Glu316 is the catalytic acid/base in paMltFΔC. Indeed, as shown below, similar to ecMltE, a murodipeptide substrate analogue binds productively near Glu316, and the environment of Glu316 is similarly hydrophobic as in ecMltE, with contributions from Ile312 (Ile60 in ecMltE), Met334 (Gln82), Tyr386 (Tyr146), and Phe442 (Tyr192). Although not all residues are
Chapter 3

conserved, they all contribute to the hydrophobic environment near Glu316, which ensures that the Glu316 side chain is protonated at the beginning of the lytic transglycosylase reaction (Thunnissen et al., 1995a; Artola-Recolons et al., 2011; Fibriansah et al., 2012).

The $pa$MltFΔC-bulgecin-murodipeptide complex

To probe the interactions of $pa$MltFΔC with peptidoglycan a crystal structure was determined at 1.65 Å resolution of a ternary complex of $pa$MltFΔC with the O-sulfonated glycopeptide inhibitor bulgecin A and the murodipeptide GlcNAc-MurNAc-L-Ala-D-Glu (named $pa$MltFΔC/bulgA-GM-Ala-Glu) (Table 1, Figure 6). The compounds bind at non-overlapping sites in the groove between the N- and C-terminal subdomains of the catalytic domain of $pa$MltFΔC, equivalent to situation in MltE (Fibriansah et al., 2012). Bulgecin A occupies subsites -2 and -1, and the murodipeptide binds in subsites +1 and +2. The electron density for the compounds is well defined, except for the D-Glu residue of the dipeptide moiety, which is disordered (data not shown). The bulgecin molecule is bound in an extended conformation with its GlcNAc and L-proline residues occupying subsites -2 and -1, respectively, while the proline-linked taurine group points away from subsite -1 into the solvent. The GlcNAc-linked 4-O-sulfonate group is located in subsite -3, and has no direct interactions with the protein. Both compounds are stabilized by extensive hydrogen bonding and van der Waals interactions with the protein. A structure-based multiple sequence alignment and superposition of $pa$MltFΔC and ecMltE shows that the interactions with the GlcNAc residues at subsites -2 and +1 and with the L-proline derivative at subsite -1 are identical (Figure 5) (Artola-Recolons et al., 2011; Fibriansah et al., 2012), confirming the close relationship between these enzymes. An intriguing question is whether $pa$MltFΔC also recognizes the MurNAc lactyl group and its cross-linking peptide chain. Unfortunately, in our structures the dipeptide chain is disordered, and thus no positive confirmation can be given at this stage of our research.

Endo-/exolytic behaviour

A mass spectroscopy analysis of the activity of $E. coli$ MltF showed that the enzyme is mainly exolytic with minor endolytic activity (Lee et al., 2013). In view of the sequence
homology between the \textit{E. coli} and \textit{P. aeruginosa} enzymes (33\% sequence identity), we assume that \textit{paMltF} is also mainly an exolytic enzyme. Our crystallographic results have provided evidence for the existence of substrate-binding subsites -2, -1, +1, and +2, which are in a groove running across the surface of the C-terminal domain of MltF. However, this groove is clearly large enough to be able to accommodate additional subsites, and we could easily model a (GlcNAc-MurNAc), glycan occupying subsites -4 to +4 (see Figure 7). There are no obstructions in the groove at either end, which would be in agreement with a mainly endolytic activity of the enzyme, as in \textit{ecMltE} (Fibriansah \textit{et al.}, 2012). Nevertheless, lacking detailed kinetics of degradation of murein glycan chains, we cannot exclude that some subsites (i.e., subsites -2, -1, +1, and +2) may have higher affinity for the substrate than others (i.e., subsites +3, +4), resulting in a predominantly exolytic activity (Lee \textit{et al.}, 2013).
Figure 6: Binding mode of bulgecin A and the GlcNAc-MurNAc-L-Ala-D-Glu murodiipeptide in \textit{poMltFΔC}. (A) Overall structure of \textit{paMltFΔC}-CTD in cartoon representation and rainbow coloring (blue to red from N- to C-terminus) showing the bound bulgecin A (magenta) and murodiipeptide (blue). (B) Similar representation as in (A) of \textit{ecMltE} with bound bulgecin A and the murodiipeptide (PDB entry 4HJV). (C) Stereo diagram showing the binding mode of bulgecin A (BLG) and the murodiipeptide (NMP) to \textit{poMltFΔC}. Hydrogen bonds are shown with dashed lines.

Figure 7. Modeling of glycan strand into the peptidoglycan-binding groove of \textit{poMltFΔC}, where the sugar binding subsites are labeled.
Conclusions

In this chapter, we report six crystal structures of MltF from *Pseudomonas aeruginosa* (paMltFΔC). The structures revealed that *paMltFΔC* is a bimodal protein made up of a N-terminal domain, which has a fold similar to that of periplasmic solute binding proteins, and a catalytic C-terminal domain, which has a lytic transglycosylase fold.

Crystal structures with various amino acids (valine, leucine, isoleucine and cysteine) demonstrate that the N-terminal domain of *paMltFΔC* does have the ability to bind medium sized amino acids, although with low affinity (mM range). The low binding affinity is likely due to the absence of a usually conserved arginine residue that binds the carboxylate group of a bound amino acid.

The C-terminal domain has a lytic transglycosylase fold that is very similar to that of lytic transglycosylase family I enzymes, in particular to that of *ecMltE*. The saccharide-binding interactions at subsites -2, -1 and +1 are highly conserved in these family I members, indicating that *paMltF* follows a similar reaction mechanism as previously described for Slt70 and MltE, using Glu 316 as the catalytic acid/base. The peptidoglycan-binding groove is fully accessible from both sides and lacks a steric obstruction near the +2 subsite, allowing *paMltF* to attack peptidoglycan at internal sites. Thus, *paMltF* appears majorly as an endolytic enzyme.
Chapter 3

References


Chapter 3


