The Molecular and Cellular Characterisation of the First Glycocin: Plantaricin KW30

Judith Stepper

2009
The Molecular and Cellular Characterisation of the First Glycocin: Plantaricin KW30

A thesis presented in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biochemistry at Massey University, Palmerston North, New Zealand

Judith Stepper

2009
“What we know is a drop. What we don’t know is an ocean.”

Isaac Newton (1643 – 1727)
ABSTRACT

Bacteriocins, typically secreted by Gram-positive and -negative bacteria, are ribosomally-synthesised antimicrobial peptides which inhibit the growth of competing bacteria. We have purified a 43 amino acid bacteriocin, plantaricin KW30 (PlnKW30) produced by Lactobacillus plantarum KW30, that has little amino acid sequence similarity to any other characterised bacteriocin.

The gene encoding $plnKW30$ is in a cluster with the genes required for maturation and export of, and immunity to, the bacteriocin. This arrangement of genes is similar to the genomic context of bacteriocin genes in other lactic acid bacteria. The $plnKW30$ gene cluster comprises six genes encoding a glycosyltransferase, a proteolytic ABC-transporter, two putative thioredoxins, a response regulator and PlnKW30 itself.

PlnKW30 was found to possess two unusual post-translational modifications: an $O$-glycosylated serine and an unprecedented $S$-glycosylation of the C-terminal cysteine. The modified serine is located on an eight residue loop that is tethered by a disulfide bridge. Both modifications have been identified as $N$-acetylglucosamines (GlcNAc), making PlnKW30 the first described class IV bacteriocin. A post-translational modification with $S$-linked GlcNAc is unprecedented in bacteriocins as well as in all genera. The antimicrobial activity of PlnKW30 on $L.\ plantarum$ ATCC 8014 was analysed using enzymatic dissection coupled with bioassays. It was found to be concentration dependent and both the N- and C-terminal fragments are necessary for activity. Furthermore, reduction of the disulfide bonds results in abolishment of antimicrobial activity and it appears that deglycosylation of the serine 18 decreases the antimicrobial activity by about two thirds. These results show that all post-translational modifications contribute to the antimicrobial activity of PlnKW30. The addition of $N$-acetylglucosamine to cultures of the indicator strain $L.\ plantarum$ ATCC 8014 protects it from the antimicrobial effect of the added PlnKW30. PlnKW30 probably targets an $N$-acetylglucosamine transporter in the target cell membrane, similar to the mannose phosphotransferase system targeted by lactococcin A.
I would like to thank my supervisors Dr Gill Norris and Dr Mark Patchett for all their encouragement, advice and guidance throughout my PhD. Lots of thanks especially to Gill for her positive thinking and support in times of despair and confusion.

I would like to thank everybody in the x-lab and Centre for Structural Biology, especially Jana Filicheva, Meekyung Ahn, Alice Clark and Trevor Loo, for all their help and support and all the good times.

I would like to thank all staff in IMBS for their support and help during my time here. This thesis was financially supported by the Royal Society of New Zealand Marsden Fund.

Big thanks to all my friends for their support and the fun times we had together, especially to my flatmates Anouck de Bonth and Renae Pratt for putting up with my moods in stressful times. Particular thanks go to Francis Stephens for always supporting and believing in me, and cheering me up when things did not go to plan.

Besonders möchte ich meinen Eltern und Schwestern Wibke und Sabrina für ihre immerwährende Unterstützung und Aufmunterung aus der Ferne danken und die schönen, wenn auch kurzen Zeiten, die wir zusammen in Neuseeland verbracht haben. Vielen Dank!
# Table of Contents

Abstract .................................................................................................................................................. I
Acknowledgements ............................................................................................................................... II
Table of Contents .................................................................................................................................... III
List of Figures ........................................................................................................................................ X
List of Tables ......................................................................................................................................... XIV
Abbreviations ......................................................................................................................................... XVI

1 Chapter 1: Introduction ..................................................................................................................... 1
   1.1 Lactic acid bacteria and bacteriocins .......................................................................................... 1
   1.2 Classification of bacteriocins...................................................................................................... 1
   1.3 Genetics and mode of action of bacteriocins .............................................................................. 4
      1.3.1 Class I, lantibiotics .............................................................................................................. 4
      1.3.2 Class II, non-modified bacteriocins .................................................................................... 7
            (a) Anti-listerial pediocin-like bacteriocins ......................................................................... 7
            (b) Two-peptide bacteriocins ................................................................................................. 8
            (c) Cyclic bacteriocins ........................................................................................................... 8
            (d) Genetic organisation and production of class II bacteriocins ........................................ 9
            (e) Regulation of bacteriocin production ............................................................................. 9
            (f) Processing and secretion of bacteriocins .......................................................................... 10
            (g) Mode of action .................................................................................................................. 11
   1.4 Quorum sensing ........................................................................................................................... 13
      1.4.1 Evolution of two-component signalling systems .............................................................. 14
      1.4.2 Peptide signalling in Gram-positive bacteria ..................................................................... 15
            (a) Protein architecture of two-component systems .......................................................... 16
            (b) Quorum sensing in Lactobacillus plantarum C11........................................................... 18
   1.5 Glycosylation ............................................................................................................................... 20
      1.5.1 Glycosylation in eukaryotes .............................................................................................. 20
      1.5.2 Prokaryotic glycosylation .................................................................................................. 24
      1.5.3 S-glycosylation .................................................................................................................. 26
      1.5.4 Glycosyltransferases in prokaryotes ................................................................................. 27
      1.5.5 Classification of glycosyltransferases .............................................................................. 29
Table of Contents

1.5.6 Structure of glycosyltransferases ........................................................................................................................................ 30
1.5.7 Modular architecture of glycosyltransferases ............................................................................................................ 33

1.6 Lactobacillus plantarum .......................................................................................................................................... 34
1.6.1 Plantaricin KW30 ............................................................................................................................................. 34

2 Chapter 2: Material and methods ................................................................................................................................. 37
2.1 General materials and methods ........................................................................................................................................ 37
  2.1.1 Purified water .................................................................................................................................................... 37
  2.1.2 Filter sterilization equipment ................................................................................................................... 37
  2.1.3 Media ...................................................................................................................................................................... 37
     (a) Luria-Bertani (LB) media .................................................................................................................................. 38
     (b) De Man, Rogosa and Sharpe (MRS) media ................................................................................................. 38
     (c) M17 media with 0.5 % glucose ........................................................................................................................ 38
     (d) Yeast media (YM) ................................................................................................................................................... 38
     (e) Chemically defined medium (CDM) ............................................................................................................... 39
  2.1.4 Antibiotic stock solutions ........................................................................................................................................ 39
  2.1.5 Glycerol stocks ................................................................................................................................................... 40
  2.1.6 Cultivation and harvesting of cells .......................................................................................................... 40
  2.1.7 Measurement of optical density (OD) of cultures ........................................................................... 41
  2.2 Methods for Deoxyribonucleic acid (DNA) work ........................................................................................................ 41
     2.2.1 Isolation of DNA ................................................................................................................................................ 41
     2.2.2 Isolation of genomic DNA from L. plantarum KW30 ..................................................................... 41
     2.2.3 Isolation of plasmid DNA from E. coli .................................................................................................... 42
     2.2.4 Isolation of plasmid DNA from L. plantarum ..................................................................................... 43
     2.2.5 Isolation of DNA-fragments ........................................................................................................................ 44
     2.2.6 Isolation of PCR products ............................................................................................................................ 44
     2.2.7 Purification and concentration of DNA ................................................................................................ 44
        (a) Phenol/Chloroform extraction ........................................................................................................................ 44
        (b) Precipitation of DNA ............................................................................................................................................. 45
  2.2.8 DNA sequencing ................................................................................................................................................ 45
  2.2.9 Agarose gel electrophoresis ......................................................................................................................... 45
  2.2.10 DNA hydrolysis with restriction endonucleases ..................................................................................... 46
  2.2.11 Ligation of DNA fragments .......................................................................................................................... 47
  2.2.12 Determination of DNA concentration .................................................................................................... 47
  2.2.13 Size determination of DNA fragments .................................................................................................. 48
### Table of Contents

2.2.14 Transformation of DNA into cells...

- Preparation of competent E. coli cells...
- Transformation and selection of E. coli cells...
- Preparation of electrocompetent E. coli cells...
- Electroporation of E. coli cells...
- Production of electrocompetent cells of Lactobacillus and Lactococcus species...
- Electroporation of Lactobacillus and Lactococcus species...
- Protoplast formation of L. plantarum KW30...
- Protoplast transformation...

2.2.15 Polymerase chain reaction (PCR)...

- Polymerase chain reaction (PCR)...
- Colony PCR...

2.2.16 Targeted gene disruption...

- Targeted gene disruption via homologous recombination in lactobacilli...
- Targeted gene disruption via cre-lox-based homologous recombination system in lactobacilli...

2.3 Methods for ribonucleic acid (RNA) work...

- Isolation of RNA...
- Agarose/formaldehyde gel electrophoresis of RNA...
- Northern blot...
- Preparation of $^{32}$P-labelled probe...
- Northern blotting and hybridisation...

2.4 Protein biochemical methods...

- Determination of protein concentration...
- Bradford method for determination of protein concentration...
- Concentration determination by UV absorption...
- SDS-polyacrylamide gel electrophoresis...
- Detection of proteins in SDS-polyacrylamide gels...
- Coomassie blue staining...
- Silver staining...
- 2D-electrophoresis...
- In-gel tryptic digest for protein identification by mass spectrometry...
- Western blot...
Table of Contents

(a) Passive blotting....................................................................................................................................................... 68
(b) Electroblotting........................................................................................................................................................ 68
(c) Dot blot....................................................................................................................................................................... 69
(d) Standard immunodetection method............................................................................................................. 70
(e) Rapid immunodetection method (Millipore)........................................................................................................ 71
(f) Chemiluminescent protein detection............................................................................................................ 72

2.4.7 Expression and purification of recombinant proteins .................................................................................... 72
(a) Expression of recombinant proteins in E. coli........................................................................................... 72
(b) Expression of recombinant proteins in lactobacilli................................................................................ 73
(c) Expression of recombinant proteins in Lactococcus............................................................................. 73
(d) Cell free expression of proteins........................................................................................................................ 74
(e) Purification of recombinant proteins........................................................................................................... 74
(f) Isolation of inclusion bodies (IB).................................................................................................................... 75

2.4.8 Purification of native PlnKW30................................................................................................................ 75

2.4.9 Biological assay for bacteriocin activity..................................................................................................... 76

2.4.10 Minimum inhibitory concentration of PlnKW30.................................................................................. 77

2.4.11 Tryptic and chymotryptic digest of PlnKW30.................................................................................... 78

2.4.12 Reduction and alkylation of the disulfide bonds of PlnKW30............................................................... 78

2.4.13 O-deglycosylation of PlnKW30.................................................................................................................. 78

2.4.14 Defarnesylation of PlnKW30.................................................................................................................... 79

2.4.15 Farnesyltransferase (FTase) activity assay........................................................................................ 80
(a) Preparation of crude yeast extract............................................................................................................... 80
(b) FTase activity assay ........................................................................................................................................... 80

2.4.16 Pull-downs using wheat germ agglutinin.............................................................................................. 81

2.4.17 Fluorescence microscopy........................................................................................................................... 82

2.5 Biochemical methods............................................................................................................................................. 83

2.5.1 Reverse Phase-High Pressure Liquid Chromatography (RP-HPLC)......................................................... 83

2.5.2 Mass spectrometry.......................................................................................................................................... 84
(a) ESI-MS (Electrospray ionisation mass spectrometry)............................................................................. 84
(b) MALDI-MS (Matrix-assisted laser desorption/ionisation mass spectrometry)................................. 84
(c) Zip tips................................................................................................................................................................ 85
(d) Macrotraps.......................................................................................................................................................... 85

2.5.3 Circular dichroism............................................................................................................................................. 86
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Phylogenetic classification of <em>L. plantarum</em> KW30</td>
</tr>
<tr>
<td>3.2</td>
<td>Analysis of the <em>plnKW30</em> gene cluster</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Bioinformatic analyses of the proteins of the <em>plnKW30</em> gene cluster</td>
</tr>
<tr>
<td>(a)</td>
<td>Analysis of the putative GTase (<em>GccB</em>)</td>
</tr>
<tr>
<td>(b)</td>
<td>Analysis of the putative ABC-transporter (<em>GccC</em>)</td>
</tr>
<tr>
<td>(c)</td>
<td>Analysis of the putative TRX1 (<em>GccD</em>)</td>
</tr>
<tr>
<td>(d)</td>
<td>Analysis of the putative TRX2 (<em>GccE</em>)</td>
</tr>
<tr>
<td>(e)</td>
<td>Analysis of the putative response regulator (<em>GccF</em>)</td>
</tr>
<tr>
<td>(f)</td>
<td>Analysis of the bacteriocin <em>PlnKW30</em> (<em>GccA</em>)</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Brief summary of bioinformatics analysis</td>
</tr>
<tr>
<td>3.3</td>
<td>Transcriptional analysis</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Transcriptional analysis by RT-PCR</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Northern blot analysis</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Growth-phase-dependent gene expression of <em>plnKW30</em></td>
</tr>
<tr>
<td>3.4</td>
<td>Characterisation of the bacteriocin <em>PlnKW30</em> from <em>L. plantarum</em> KW30</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Influence of growth conditions on <em>PlnKW30</em> production</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Purification of native <em>PlnKW30</em></td>
</tr>
<tr>
<td>3.4.3</td>
<td>Minimum inhibitory concentration of <em>PlnKW30</em></td>
</tr>
<tr>
<td>3.4.4</td>
<td>Antimicrobial activity of <em>PlnKW30</em> against the indicator strain <em>L. plantarum</em> ATCC 8014</td>
</tr>
<tr>
<td>3.4.5</td>
<td>Analysis of antimicrobial activity of <em>PlnKW30</em></td>
</tr>
<tr>
<td>3.5</td>
<td>Analysis of N- and C-terminal fragments of <em>PlnKW30</em></td>
</tr>
<tr>
<td>3.5.1</td>
<td>Identification of the C-terminal glycan modification</td>
</tr>
<tr>
<td>(a)</td>
<td>Mass spectrometry of N- and C-terminal fragments</td>
</tr>
<tr>
<td>(b)</td>
<td>Wheat germ agglutinin pull-down experiments</td>
</tr>
<tr>
<td>(c)</td>
<td>Dot blot using anti-O-GlcNAc antibodies</td>
</tr>
<tr>
<td>3.6</td>
<td>Analysis of the post-translational modifications of <em>PlnKW30</em></td>
</tr>
<tr>
<td>3.6.1</td>
<td>Analysis of the disulfide bridges of <em>PlnKW30</em></td>
</tr>
<tr>
<td>3.6.2</td>
<td>Analysis of the glycan modifications of <em>PlnKW30</em></td>
</tr>
<tr>
<td>3.7</td>
<td>Circular dichroism analysis of <em>PlnKW30</em></td>
</tr>
<tr>
<td>3.7.1</td>
<td>Comparison of native, reduced and deglycosylated <em>PlnKW30</em></td>
</tr>
<tr>
<td>3.7.2</td>
<td>Comparison of native, reduced and deglycosylated N-terminal fragment</td>
</tr>
<tr>
<td>3.7.3</td>
<td>Comparison of <em>PlnKW30</em> with the N-terminal fragment</td>
</tr>
</tbody>
</table>
4 Chapter 4: Discussion ..................................................................................................................151
  4.1 Phylogenetic classification of *L. plantarum* KW30 ................................................................. 151
  4.2 Analysis of the *plnKW30* gene cluster .................................................................................. 152
    4.2.1 Bioinformatic analyses of the proteins of the *plnKW30* gene cluster ............................... 162
    4.2.2 Future work .......................................................................................................................... 164
  4.3 Transcriptional analysis ........................................................................................................... 165
    4.3.1 Comparison of Northern blot and RT-PCR analysis ............................................................. 165
    4.3.2 Growth-phase-dependent gene expression of *plnKW30* .................................................... 166
    4.3.3 Future work .......................................................................................................................... 166
  4.4 Characterisation of PlnKW30 of *L. plantarum* KW30 ............................................................. 167
    4.4.1 Antimicrobial activity of PlnKW30 against the indicator strain *L. plantarum* ATCC 8014 ...................................................... 170
    4.4.2 Future work .......................................................................................................................... 175
  4.5 Analyses of N- and C-terminal fragments of PlnKW30 ............................................................. 175
    4.5.1 Discovery of C-terminal glycan modification ....................................................................... 177
  4.6 Analysis of the role of post translational modifications in PlnKW30 activity ............................ 178
    4.6.1 Analysis of the disulfide bridges of PlnKW30 .................................................................... 178
    4.6.2 Analysis of the glycan modifications of PlnKW30 ............................................................... 179
    4.6.3 Future work .......................................................................................................................... 180
  4.7 Circular Dichroism .................................................................................................................... 181
    4.7.1 Future work .......................................................................................................................... 183
  4.8 Conclusions .................................................................................................................................. 184

5 Appendix ........................................................................................................................................ 185
  Appendix 1: The Genetic Code ........................................................................................................ 185
  Appendix 2: Structures and symbols of the standard amino acids ................................................ 186
  Appendix 3: Lists of strains and vectors used in this thesis .............................................................. 187
  Appendix 4: Primer sequences ........................................................................................................ 190
  Appendix 5: Results of early experiments ....................................................................................... 193
  5.1 Introduction into farnesylation ................................................................................................. 193
  5.2 Experimental results ................................................................................................................ 194
    5.2.1 Western blotting using anti-farnesyl antibodies .................................................................. 194
    5.2.2 Two-dimensional gel electrophoresis .................................................................................. 196
    5.2.3 Farnesyltransferase activity assay ...................................................................................... 197
    5.2.4 Removal of farnesyl group by methyliodide treatment ...................................................... 199
List of Figures

Figure 1.1: Proposed classification scheme for bacteriocins................................................................. 3
Figure 1.2: Structure of the lantibiotic nisin.......................................................................................... 5
Figure 1.3: Structure of lipid II................................................................................................................ 5
Figure 1.4: Model for the formation of nisin-lipid II pores in lipid bilayers........................................ 6
Figure 1.5: N-terminal amino acid sequence motif of pediocin-like bacteriocins............................... 7
Figure 1.6: Quorum sensing signalling molecules of prokaryotes......................................................... 14
Figure 1.7: Cartoon representation of the structure of the AgrA-DNA complex............................... 18
Figure 1.8: Genetic map of \textit{pln} locus of different \textit{L. plantarum} strains........................................ 19
Figure 1.9: Glycan linkages to peptide chains.......................................................................................... 21
Figure 1.10: Pentasaccharide inner-core of \textit{N}-linked glycans............................................................. 21
Figure 1.11: Model for the biosynthesis of \textit{N}- and \textit{O}-linked glycoproteins in eukaryotes................ 23
Figure 1.12: Model for the biosynthesis of \textit{N}- and \textit{O}-linked glycoproteins in bacteria..................... 26
Figure 1.13: Mechanism of inverting and retaining glycosyltransferases............................................ 28
Figure 1.14: The hierarchical classification of glycosyltransferases (GT) from folds to clans and families system proposed by Coutinho \textit{et al.} (2003)................................................................. 30
Figure 1.15: Crystal structures of GT-A and GT-B glycosyltransferases............................................. 32
Figure 1.16: The modularity of glycosyltransferases............................................................................. 33
Figure 1.17: Schematic overview of the nested disulfide bonds of the mature PlnKW30.................... 36
Figure 2.1: Diagram of a target plate for MALDI TOF......................................................................... 85
Figure 3.1: Phylogenetic tree of \textit{Lactobacillus} species.......................................................................... 90
Figure 3.2: Schematic bacteriocin gene cluster with putative functions of ORFs identified \textit{in silico}. 91
Figure 3.3: Possible regulatory sites of \textit{gccC} (putativeABC-transporter)............................................ 93
Figure 3.4: Nucleotide sequence and deduced protein of \textit{gccA} (\textit{plnKW30})........................................ 94
Figure 3.5: Schematic view of GccB with glycosyltransferase domain (from cds BLAST)................. 96
Figure 3.6: Schematic view of GccC with individual domains and special motifs (from cds BLAST). 98
Figure 3.7: Schematic view of GccD with TRX and TRX-superfamily domains (from cds BLAST).... 100
List of Figures

Figure 3.8: Schematic view of GccE with TRX and TRX-superfamily domains (from cds BLAST)... 101
Figure 3.9: Schematic view of GccF with LytTR domain (from cds BLAST)................................. 102
Figure 3.10: Results of the RT-PCR of genes and intergenic regions of the plnKW30 gene cluster.. 106
Figure 3.11: Northern blot analysis using a plnKW30 specific probe. ........................................ 108
Figure 3.12: Growth-dependent gene expression of PlnKW30...................................................... 110
Figure 3.13: Effect of growth temperature, aeration and medium on PlnKW30 production.......... 111
Figure 3.14: RP-HPLC chromatogram of PlnKW30 .................................................................. 112
Figure 3.15: Monoisotopic mass spectrometry result of the native PlnKW30........................... 113
Figure 3.16: Minimum inhibitory concentration assay for PlnKW30........................................ 114
Figure 3.17: Growth curve of L. plantarum KW30 (producer strain) and L. plantarum ATCC 8014 (indicator strain). .......................................................................................................................... 115
Figure 3.18: Growth of L. plantarum ATCC 8014 after addition of different concentrations of PlnKW30. ...................................................................................................................................................... 116
Figure 3.19: Antimicrobial effect of PlnKW30 on L. plantarum ATCC 8014 cells....................... 118
Figure 3.20: Concentration-dependent antimicrobial activity of PlnKW30 against L. plantarum ATCC 8014................................................................................................................................. 119
Figure 3.21: Effect of different sugars on antimicrobial activity of PlnKW30............................... 120
Figure 3.22: Growth of L. plantarum ATCC 8014 after addition of 1mg/mL N-acetylglucosamine and PlnKW30. .......................................................................................................................... 121
Figure 3.23: Effect of different concentrations of GlcNAc and glucosamine on antimicrobial activity of PlnKW30................................................................. 122
Figure 3.24: Effect of Man or GalNAc on antimicrobial activity of PlnKW30................................ 123
Figure 3.25: Mature PlnKW30 sequence with trypsin cleavage sites and resulting peptides........ 124
Figure 3.26: RP-HPLC chromatogram of the N- and C-terminal fragments of PlnKW30............. 125
Figure 3.27: Tricine SDS-PAGE of PlnKW30 and its N- and C-terminal fragments overlaid with L. plantarum ATCC 8014.................................................................................................................. 126
Figure 3.28: Antimicrobial effect of N- & C-terminal fragments of PlnKW30.............................. 128
Figure 3.29: Growth of L. plantarum ATCC 8014 after addition of the N-terminal fragment of PlnKW30. ................................................................................................................................. 129
Figure 3.30: Tricine SDS-PAGE of the N-terminal fragment overlaid with *L. plantarum* ATCC 8014 at OD600 of 0.3

Figure 3.31: Monoisotopic mass spectrum of the N-terminal fragment of PlnKW30

Figure 3.32: Monoisotopic mass spectrum of the C-terminal fragment of PlnKW30

Figure 3.33: Monoisotopic mass spectrometry result for the C-terminal HC-X fragment of PlnKW30

Figure 3.34: Cartoon representation of the NMR structure of PlnKW30

Figure 3.35: Tricine SDS-PAGE of lectin pull-downs of PlnKW30

Figure 3.36: Dot blot using anti-\(\beta\)-GlcNAc antibodies to detect PlnKW30 and its N- & C-terminal tryptic fragments

Figure 3.37: Visualisation of N- and C-terminal fragments using ninhydrin

Figure 3.38: RP-HPLC chromatogram of native and reduced/alkylated PlnKW30

Figure 3.39: Monoisotopic mass spectrometry result of the reduced and alkylated PlnKW30

Figure 3.40: Activity of native and reduced PlnKW30 in a bioassay overlay

Figure 3.41: Antimicrobial activity of reduced PlnKW30

Figure 3.42: RP-HPLC chromatogram of native and deglycosylated PlnKW30

Figure 3.43: Monoisotopic mass spectrometry result of the deglycosylated PlnKW30

Figure 3.44: Activity of native and deglycosylated PlnKW30 in a bioassay overlay

Figure 3.45: Antimicrobial effect of deglycosylated PlnKW30

Figure 3.46: Smoothed, base-line corrected spectra of native PlnKW30 (A), deglycosylated PlnKW30 (B) and their corresponding residuals

Figure 3.47: Smoothed, base-line corrected spectra of reduced and alkylated PlnKW30 (A), reduced PlnKW30 (B) and their corresponding residuals

Figure 3.48: Smoothed, base-line corrected spectra of the N-terminal fragment (A), reduced N-terminal fragment (B) and their corresponding residuals

Figure 3.49: Circular dichroism spectra of native, reduced and deglycosylated PlnKW30

Figure 3.50: Circular dichroism spectra of native, reduced and deglycosylated N-terminal fragment

Figure 3.51: Circular dichroism spectra of native PlnKW30 and native N-terminal fragment

Figure 4.1: Comparison of the *plnKW30* gene cluster with similar gene clusters
List of Figures

Figure 4.2: Schematic representation of the mature PlnKW30................................................................. 155
Figure 4.3: Comparison of the bacteriocin amino acid sequences......................................................... 157
Figure 4.4: Comparison of the N-terminal amino acid sequences of bacteriocin GTases.................. 159
Figure 4.5: Comparison of the N-terminal amino acid sequences of the ABC-transporters........... 160
Figure 4.6: Comparison of the C-terminal sequence of the ABC-transporters................................. 161
Figure 4.7: Schematic model of the putative regulation and maturation of PlnKW30......................... 163
Figure 4.8: Schematic overview of the hypothetical interaction between PlnKW30 and a GlcNAc transporter............................................................................................................................... 174
Figure 4.9: Schematic overview of N-terminal and C-terminal tryptic fragments of PlnKW30....... 176
Figure 5.1: SDS-PAGE (A) and Western blot (B) using anti-farnesyl antibodies................................. 195
Figure 5.2: 2D-electrophoresis gel of whole cell extract of *L. plantarum* KW30................................. 196
Figure 5.3: Autoradiography films of FTase activity assays................................................................. 198
Figure 5.4: FTase activity assay separated by RP-HPLC................................................................. 199
Figure 5.5: RP-HPLC chromatogram of native PlnKW30 and CH₃I-treated PlnKW30...................... 200
Figure 5.6: Western blot using anti-GFP antibodies.............................................................................. 204
Figure 5.7: Growth curves of *L. plantarum* KW30 at 30 °C and 43 °C .............................................. 212
Figure 5.8: Amino acid alignment of the PlnKW30 GTase and a putative GTase from *Bacteroides fragilis* (pdb 3bcv) using FUGUE (Shi et al. 2001) (a: alpha helix; b: beta strand; 3:3₁₀ helix; lower case: solvent accessible; UPPER CASE: solvent inaccessible)........................................... 225
Figure 5.9: Amino acid alignment of the plnKW30 GTase and SpsA from *Bacillus subtilis* (pdb 1qg8) using FUGUE (Shi et al. 2001) (a: alpha helix; b: beta strand; 3:3₁₀ helix; lower case: solvent accessible; UPPER CASE: solvent inaccessible)........................................................................................................ 226
List of Tables

Table 1.1: Open reading frames and functions of the nisin gene cluster ............................................................ 4
Table 1.2: Mode of action of class II bacteriocins (adapted from Hechard & Sahl 2002) .........................12
Table 2.1: Chemicals and concentrations for chemically defined medium (Saguir & de Nadra 2007) ..................................................................................................................................................................................................................39
Table 2.2: Antibiotic stock solutions and final concentrations ...........................................................................40
Table 2.3: Coomassie blue staining and destaining solutions ................................................................................61
Table 2.4: Dilutions of PlnKW30 used determine its MIC ................................................................................77
Table 2.5: Experimental conditions for CD scan .............................................................................................................86
Table 2.6: Sample specific parameters for CD scans of native, reduced and deglycosylated PlnKW30 and its N-terminal fragment ......................................................................................................................................................87
Table 3.1: Results from nucleotide BLAST (NCBI) of 16S – 23S rDNA sequence of L. plantarum KW30 .......................................................................................................................................................................89
Table 3.2: Characteristics of the genes and gene products of the plnKW30 cluster of L. plantarum KW30 ...........................................................................................................................................................................92
Table 3.3: Predicted positions of the transmembrane helices of the ABC-transporter ........................................97
Table 3.4: Predicted conserved domains of GccC ..........................................................................................................98
Table 3.5: Predicted positions of the transmembrane helices of GccD .............................................................99
Table 3.6: Results of RT-PCR of genes and intergenic regions of the plnKW30 gene cluster ...............107
Table 5.1: List of strains used in this thesis ................................................................................................................... 187
Table 5.2: Vectors and plasmids used in this thesis .................................................................................................. 188
Table 5.3: Primers used in sequencing and PCR .......................................................................................................... 190
Table 5.4: List of E. coli expression strains and their descriptions ................................................................... 203
Table 5.5: Expression constructs and strains used for expression trials of plnKW30, RR and GT in this thesis ......................................................................................................................................................................................... 205
Table 5.6: Results of disruption attempts with L. plantarum KW30 (pTRK669, pORI28::FTaseKO) at 40, 41 and 43 °C ........................................................................................................................................................................... 213
Table 5.7: MS/MS fragmentation of PlnKW30 .............................................................................................................. 227
Table 5.8: List of polymerases and restriction enzymes ......................................................................................... 228
List of Tables

Table 5.9: List of commercial kits ................................................................. 229
Table 5.10: List of consumables ................................................................. 230
Table 5.11: List of equipment used ............................................................. 231
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Å</td>
<td>Ångström ($10^{-10}$m)</td>
</tr>
<tr>
<td>aa</td>
<td>Amino acid</td>
</tr>
<tr>
<td>ABC-transporter</td>
<td>ATP-binding cassette transporter</td>
</tr>
<tr>
<td>ACN</td>
<td>Acetonitrile</td>
</tr>
<tr>
<td>AHL</td>
<td>Acyl-homoserine lactone</td>
</tr>
<tr>
<td>Amp</td>
<td>Ampicillin</td>
</tr>
<tr>
<td>APS</td>
<td>Ammonium persulfate</td>
</tr>
<tr>
<td>ATCC</td>
<td>American Type Culture Collection</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine-5′-triphosphate</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Search Tool</td>
</tr>
<tr>
<td>bp</td>
<td>Base pairs</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>°C</td>
<td>Degrees Celsius</td>
</tr>
<tr>
<td>C</td>
<td>Carbon</td>
</tr>
<tr>
<td>ca.</td>
<td>Circa</td>
</tr>
<tr>
<td>CAZy</td>
<td>Carbohydrate-Active enZymes</td>
</tr>
<tr>
<td>CBM</td>
<td>Carbohydrate-binding module</td>
</tr>
<tr>
<td>Cds</td>
<td>Conserved domains in sequences</td>
</tr>
<tr>
<td>Cm</td>
<td>Chloramphenicol</td>
</tr>
<tr>
<td>CTP</td>
<td>Cytidine-5′-triphosphate</td>
</tr>
<tr>
<td>ddH$_2$O</td>
<td>Double distilled water</td>
</tr>
<tr>
<td>Dha</td>
<td>Dehydroalanine</td>
</tr>
<tr>
<td>Dhb</td>
<td>Dehydrobutyryline</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulphoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DNase</td>
<td>Deoxyribonuclease</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxyribose nucleotide triphosphate</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylene diamine tetraacetic acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Em</td>
<td>Erythromycin</td>
</tr>
<tr>
<td>et al.</td>
<td>et alteri (and others)</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>EtBr</td>
<td>Ethidium bromide</td>
</tr>
<tr>
<td>ETD</td>
<td>Electron Transfer Dissociation</td>
</tr>
<tr>
<td>EtOH</td>
<td>Ethanol</td>
</tr>
<tr>
<td>ExPASy</td>
<td>Expert Protein Analysis System</td>
</tr>
<tr>
<td>FLP</td>
<td>FNR-like regulatory proteins</td>
</tr>
<tr>
<td>FNR</td>
<td>Fumarate-nitrate reduction</td>
</tr>
<tr>
<td>FPLC</td>
<td>Fast Protein Liquid Chromatography</td>
</tr>
<tr>
<td>FTase</td>
<td>Farnesyltransferase</td>
</tr>
<tr>
<td>g</td>
<td>Gramm; standard gravity (9.81 m/s²)</td>
</tr>
<tr>
<td>GalNAcβ1-Ser/Thr</td>
<td>N-acetyl-β-galactosamine and L-serine or L-threonine</td>
</tr>
<tr>
<td>gDNA</td>
<td>Genomic DNA</td>
</tr>
<tr>
<td>GDP</td>
<td>Guanidine-5’-diphosphate</td>
</tr>
<tr>
<td>G+C %</td>
<td>Percentage of guanine and cytosine</td>
</tr>
<tr>
<td>Gal</td>
<td>Galactose</td>
</tr>
<tr>
<td>GlcNAc</td>
<td>N-acetylglactosamine</td>
</tr>
<tr>
<td>GlcNac</td>
<td>N-acetylglucosamine</td>
</tr>
<tr>
<td>GlcNAcβ1-Asn</td>
<td>N-acetylgulosaminyl-asparagine</td>
</tr>
<tr>
<td>G protein</td>
<td>GTP binding protein</td>
</tr>
<tr>
<td>GT</td>
<td>Glycosyltransferase</td>
</tr>
<tr>
<td>GTP</td>
<td>Guanosine-5’-triphosphate</td>
</tr>
<tr>
<td>hrs</td>
<td>Hours</td>
</tr>
<tr>
<td>HEPES</td>
<td>n-(2-hydroxyethyl)piperizine-n-(2-ethanesulfonic acid)</td>
</tr>
<tr>
<td>HPK</td>
<td>Histidine Protein Kinase</td>
</tr>
<tr>
<td>HPLC</td>
<td>High Pressure Liquid Chromatography</td>
</tr>
<tr>
<td>HTH</td>
<td>Helix-turn-helix</td>
</tr>
<tr>
<td>IEX</td>
<td>Ion Exchange Chromatography</td>
</tr>
<tr>
<td>IM</td>
<td>Inner membrane</td>
</tr>
<tr>
<td>IMAC</td>
<td>Immobilised Metal Affinity Chromatography</td>
</tr>
</tbody>
</table>
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPTG</td>
<td>Isopropyl-β-D-Thiogalactopyranoside</td>
</tr>
<tr>
<td>IR</td>
<td>Infrared radiation</td>
</tr>
<tr>
<td>IUBMB</td>
<td>International Union of Biochemistry and Molecular Biology</td>
</tr>
<tr>
<td>Kan</td>
<td>Kanamycin</td>
</tr>
<tr>
<td>Kbp</td>
<td>Kilo basepairs</td>
</tr>
<tr>
<td>kDa</td>
<td>Kilo Daltons</td>
</tr>
<tr>
<td>L</td>
<td>Liter</td>
</tr>
<tr>
<td>λ</td>
<td>Wavelength</td>
</tr>
<tr>
<td>LAB</td>
<td>Lactic Acid Bacteria</td>
</tr>
<tr>
<td>LB</td>
<td>Luria Bertani broth</td>
</tr>
<tr>
<td>M</td>
<td>Molar (mol/L)</td>
</tr>
<tr>
<td>m</td>
<td>Meter</td>
</tr>
<tr>
<td>Man</td>
<td>Mannose</td>
</tr>
<tr>
<td>man-PTS</td>
<td>Mannose phosphotransferase system</td>
</tr>
<tr>
<td>MCS</td>
<td>Multiple Cloning Site</td>
</tr>
<tr>
<td>min</td>
<td>Minute</td>
</tr>
<tr>
<td>mol</td>
<td>6.023 x 10(^{23}) molecules</td>
</tr>
<tr>
<td>MRS</td>
<td>De Man, Rogosa and Sharpe broth</td>
</tr>
<tr>
<td>MW</td>
<td>Molecular mass</td>
</tr>
<tr>
<td>m/z</td>
<td>Mass-to-charge ratio</td>
</tr>
<tr>
<td>n/a</td>
<td>Not applicable</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium chloride</td>
</tr>
<tr>
<td>NBD</td>
<td>Nucleotide binding domain</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Centre for Biotechnology Information</td>
</tr>
<tr>
<td>ND</td>
<td>Not determined</td>
</tr>
<tr>
<td>NDP</td>
<td>Nucleotide diphosphate</td>
</tr>
<tr>
<td>NEB</td>
<td>New England Biolabs</td>
</tr>
<tr>
<td>NH(_2)</td>
<td>Amine</td>
</tr>
<tr>
<td>(NH(_4))(_2)SO(_4)</td>
<td>Ammonium sulphate</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>No.</td>
<td>Number</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>O</td>
<td>Oxygen</td>
</tr>
<tr>
<td>OD</td>
<td>Optical Density</td>
</tr>
<tr>
<td>OM</td>
<td>Outer membrane</td>
</tr>
<tr>
<td>o/n</td>
<td>overnight</td>
</tr>
<tr>
<td>ORF</td>
<td>Open Reading Frame</td>
</tr>
<tr>
<td>p.a.</td>
<td>pro analysis</td>
</tr>
<tr>
<td>Pa</td>
<td>Pascal (= 10⁻⁵ bar = 145.04 × 10⁻⁶ psi)</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>pH</td>
<td>Negative decadal logarithm of the proton concentration</td>
</tr>
<tr>
<td>PnKw30</td>
<td>plantaricin Kw30</td>
</tr>
<tr>
<td>ppm</td>
<td>Parts per million</td>
</tr>
<tr>
<td>PTase</td>
<td>Prenyltransferase</td>
</tr>
<tr>
<td>PTM</td>
<td>Post-translational modification</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene difluorine</td>
</tr>
<tr>
<td>QS</td>
<td>Quorum sensing</td>
</tr>
<tr>
<td>RBS</td>
<td>Ribosome Binding Sequence</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribose nucleic acid</td>
</tr>
<tr>
<td>RNase</td>
<td>Ribonuclease</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>RT</td>
<td>Room temperature</td>
</tr>
<tr>
<td>RP-HPLC</td>
<td>Reverse-Phase High Pressure Liquid Chromatography</td>
</tr>
<tr>
<td>RR</td>
<td>Response regulator (originally thought to be FTase)</td>
</tr>
<tr>
<td>SAM</td>
<td>S-adenosyl-methionine</td>
</tr>
<tr>
<td>s/sec</td>
<td>Second</td>
</tr>
<tr>
<td>S/D</td>
<td>Shine-Delgarno-Sequence</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>sec pathway</td>
<td>Translocase general secretion pathway</td>
</tr>
<tr>
<td>S-layer</td>
<td>surface-layer</td>
</tr>
<tr>
<td>Sm</td>
<td>Streptomycin</td>
</tr>
<tr>
<td>SOB</td>
<td>Super optimal broth</td>
</tr>
</tbody>
</table>
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOC</td>
<td>SOB with catabolite repression, indicative of the presence of glucose</td>
</tr>
<tr>
<td>Sp.</td>
<td>Species</td>
</tr>
<tr>
<td>SSC</td>
<td>Saline sodium citrate buffer</td>
</tr>
<tr>
<td>Subsp.</td>
<td>Subspecies</td>
</tr>
<tr>
<td>T</td>
<td>Temperature</td>
</tr>
<tr>
<td>t</td>
<td>time</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris-Acetate-EDTA</td>
</tr>
<tr>
<td>TCEP</td>
<td>tris(2-carboxyethyl)phosphine</td>
</tr>
<tr>
<td>TCS</td>
<td>Two-Component System</td>
</tr>
<tr>
<td>Tc</td>
<td>Tetracycline</td>
</tr>
<tr>
<td>TE</td>
<td>Tris-EDTA</td>
</tr>
<tr>
<td>TLC</td>
<td>Thin layer chromatography</td>
</tr>
<tr>
<td>TMD</td>
<td>Transmembrane domain</td>
</tr>
<tr>
<td>TMH</td>
<td>Transmembrane helix</td>
</tr>
<tr>
<td>TRIS</td>
<td>Tris (hydroxymethyl) aminomethane</td>
</tr>
<tr>
<td>TTP</td>
<td>Thymidine-5′-triphosphate</td>
</tr>
<tr>
<td>TRX</td>
<td>Thioredoxin</td>
</tr>
<tr>
<td>UDP</td>
<td>Uridine diphosphate</td>
</tr>
<tr>
<td>UDP-Glc</td>
<td>Uridine diphosphate-α-D-glucose</td>
</tr>
<tr>
<td>USA</td>
<td>United States of America</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet light</td>
</tr>
<tr>
<td>V</td>
<td>Volts</td>
</tr>
<tr>
<td>v/v</td>
<td>Volume per volume</td>
</tr>
<tr>
<td>W</td>
<td>Watt</td>
</tr>
<tr>
<td>w/v</td>
<td>Weight per volume</td>
</tr>
<tr>
<td>w/w</td>
<td>Weight per weight</td>
</tr>
</tbody>
</table>
# Abbreviations of Nucleic Acids

<table>
<thead>
<tr>
<th>One letter code</th>
<th>Base(s) represented</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Adenine</td>
</tr>
<tr>
<td>C</td>
<td>Cytosine</td>
</tr>
<tr>
<td>G</td>
<td>Guanine</td>
</tr>
<tr>
<td>T</td>
<td>Thymine</td>
</tr>
<tr>
<td>U</td>
<td>Uracil</td>
</tr>
<tr>
<td>R</td>
<td>GA</td>
</tr>
<tr>
<td>Y</td>
<td>TC</td>
</tr>
<tr>
<td>K</td>
<td>GT</td>
</tr>
<tr>
<td>M</td>
<td>AC</td>
</tr>
<tr>
<td>S</td>
<td>GC</td>
</tr>
<tr>
<td>W</td>
<td>AT</td>
</tr>
<tr>
<td>B</td>
<td>GTC</td>
</tr>
<tr>
<td>D</td>
<td>GAT</td>
</tr>
<tr>
<td>H</td>
<td>ACT</td>
</tr>
<tr>
<td>V</td>
<td>G or C or A</td>
</tr>
<tr>
<td>N</td>
<td>ANY</td>
</tr>
</tbody>
</table>
## Abbreviations of Amino Acids

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>3-letter code</th>
<th>1-letter code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>Ala</td>
<td>A</td>
</tr>
<tr>
<td>Arginine</td>
<td>Arg</td>
<td>R</td>
</tr>
<tr>
<td>Asparagine</td>
<td>Asn</td>
<td>N</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>Asp</td>
<td>D</td>
</tr>
<tr>
<td>Cysteine</td>
<td>Cys</td>
<td>C</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>Glu</td>
<td>E</td>
</tr>
<tr>
<td>Glutamine</td>
<td>Gln</td>
<td>Q</td>
</tr>
<tr>
<td>Glycine</td>
<td>Gly</td>
<td>G</td>
</tr>
<tr>
<td>Histidine</td>
<td>His</td>
<td>H</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>Ile</td>
<td>I</td>
</tr>
<tr>
<td>Leucine</td>
<td>Leu</td>
<td>L</td>
</tr>
<tr>
<td>Lysine</td>
<td>Lys</td>
<td>K</td>
</tr>
<tr>
<td>Methionine</td>
<td>Met</td>
<td>M</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>Phe</td>
<td>F</td>
</tr>
<tr>
<td>Proline</td>
<td>Pro</td>
<td>P</td>
</tr>
<tr>
<td>Serine</td>
<td>Ser</td>
<td>S</td>
</tr>
<tr>
<td>Threonine</td>
<td>Thr</td>
<td>T</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>Trp</td>
<td>W</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>Tyr</td>
<td>Y</td>
</tr>
<tr>
<td>Valine</td>
<td>Val</td>
<td>V</td>
</tr>
</tbody>
</table>
1 Chapter 1: Introduction

1.1 Lactic Acid Bacteria and Bacteriocins

Lactic acid bacteria (LAB) are a phylogenetically diverse group of bacteria related by their ability to produce lactic acid during homo- or heterofermentative metabolism. They are Gram-positive, nonsporulating, coccus- or rod-shaped and have a low percentage of G and C bases in their DNA. LAB inhabit a wide range of ecological niches including certain foods, the mouth, the gastrointestinal tract and the urogenital tract of both humans and other animals. The growth of LAB is accompanied not only by acidification and enzymatic processes that give flavour and texture to a variety of fermented food, but also by the production of antimicrobial peptides and proteins called bacteriocins. The term ‘bacteriocins’ was coined more than 50 years ago by Jacob et al. (1953) as a general descriptor for colicin-like antimicrobials produced by bacteria. Bacteriocins, especially nisin, have been the subject of intensive research since 1925 when they were found to inhibit the growth of food spoilage bacteria and human pathogens such as Staphylococcus aureus, Listeria monocytogenes and Enterococcus species. Despite extensive use of bacteriocins as food additives since the 1950s, no increase in resistance to them has been detected in the bacteria against which they are active. This has resulted in a renewed interest in their antimicrobial mechanisms and in particular the molecular biology of bacteriocin production (Bauer & Dicks 2005; Breukink & de Kruijff 2006).

1.2 Classification of Bacteriocins

As bacteriocins are very varied in both their biochemical and physical properties several attempts have been made to categorise them according to shared properties. Klaenhammer (1993) defined four major classes of LAB bacteriocins (Figure 1.1A): Class I, lantibiotics; class II, small, heat-stable non-lantibiotics; class III, large heat-labile proteins and class IV, complex bacteriocins. The division of class I and class II bacteriocins has been retained in
current classifications (Nes et al. 2007). Class I are the lantibiotics, which are small (<5 kDa), heat stable peptides that contain serine and threonine residues that have been post-translationally modified to form lanthionine and β-methyl lanthionine residues. The formation of covalent bridges between these residues and cysteine residues is enzymatically catalysed and results in internal cyclic thioethers, which give the bacteriocins their characteristic structural features (for details see section 1.3.1). Class I bacteriocins are subdivided into type A, elongated molecules with a flexible structure in solution (e.g. nisin; figure 1.2) and type B which have a more rigid, globular structure.

Class II bacteriocins are small (< 10 kDa), heat-stable, unmodified peptides, typically with a diglycine processing site in the bacteriocin prepeptide. Class II can be subdivided into three groups. Class IIa are the *Listeria*-active bacteriocins (e.g. pediocin PA-1/AcH), containing an N-terminal sequence motif with at least one disulfide bridge. Class IIb are made up of two polypeptide chains that require the presence of both peptides for optimal activity. Some two component bacteriocins display no, or only some individual activity, but the greatest activity is always displayed by the dipeptide working synergistically.

Nes et al. (2007) divided class II bacteriocins into four subgroups: IIa, pediocin-like, anti-listeria bacteriocins, IIb, two peptide bacteriocins, IIc, peptide bacteriocins which lack a signal sequence and IId, cyclic bacteriocins. Cyclic bacteriocins had previously been suggested by several research groups as deserving of a separate class (Kemperman et al. 2003; Kawai et al. 2004; Maqueda et al. 2004; Heng & Tagg 2006; Heng et al. 2007).

The suggested third class of heat-labile bacteriocins was renamed bacteriolysins by Cotter et al. (2005) (Figure 1.1B), who reasoned that these bacteriocins were really enzymes. Heng et al. (2007) proposed a class III for large bacteriocins and subdivided it further into IIIa, lytic-bacteriocins, and IIIb, non-lytic bacteriocins (Figure 1.1).

The fourth major class proposed by Klaenhammer et al. (1993), the complex bacteriocins, were supposed to be modified by lipid and/or carbohydrate moieties. This was subsequently disregarded as no solid evidence had been found to substantiate the class. One example is staphylococcin 1580 from *Staphylococcus epidermidis* (Jetten et al. 1972; Jetten & Vogels 1972a, 1972b) that was reported to consist of subunits with a molecular
weight of ~20 kDa that formed multimeric structures of 300–400 kDa in conjunction with carbohydrate and lipid. However, a recent investigation revealed the antimicrobial activity to be due to epidermin, a lantibiotic bacteriocin (2164.6 Da), which co-purified with the higher molecular mass species (Sahl 1994). Several other bacteriocins were thought to be modified with carbohydrate moieties (Upreti & Hinsdill 1973, 1975; Lewus et al. 1992; Jimenezdiaz et al. 1993; Schved et al. 1993; Gilbreth & Somkuti 2005), mainly because of their sensitivity to amylase. However, these findings were not further investigated nor substantiated.

No doubt there will be further discussions about the best classification of LAB bacteriocins as new information is obtained about these antimicrobial peptides.
1.3  GENETICS AND MODE OF ACTION OF BACTERIOCINS

1.3.1  CLASS I, LANTIBIOTICS

The best studied lantibiotic is nisin, which is produced by *Lactococcus lactis* subsp. *lactis* (Figure 1.2). The term lantibiotic arose because these small antibacterial peptides were found to contain the modified amino acids lanthionine or β-methyl-lanthionine. Nisin is an autoregulated bacteriocin coded for in a gene cluster, which is made up of a number of open reading frames (ORFs) responsible for the maturation of the bacteriocin (Table 1.1).

**Table 1.1: Open reading frames and functions of the nisin gene cluster.**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>nisA</td>
<td>prepronisin</td>
<td>(Koponen <em>et al.</em> 2002; Kuipers <em>et al.</em> 2006; Li <em>et al.</em> 2006)</td>
</tr>
<tr>
<td>nisB</td>
<td>dehydration of pronisin</td>
<td>(Koponen <em>et al.</em> 2002; Kuipers <em>et al.</em> 2006; Li <em>et al.</em> 2006)</td>
</tr>
<tr>
<td>nisC</td>
<td>cyclisation of pronisin</td>
<td>(Koponen <em>et al.</em> 2002; Kuipers <em>et al.</em> 2006; Li <em>et al.</em> 2006)</td>
</tr>
<tr>
<td>nisT</td>
<td>secretion of the modified peptide</td>
<td>(Kuipers <em>et al.</em> 2004)</td>
</tr>
<tr>
<td>nisP</td>
<td>cleavage of leader sequence from modified peptide</td>
<td>(Kuipers <em>et al.</em> 2004)</td>
</tr>
<tr>
<td>nisR &amp; nisK</td>
<td>regulation of nisin production</td>
<td>(Kuipers <em>et al.</em> 1995; Ruyter <em>et al.</em> 1996)</td>
</tr>
<tr>
<td>nisI, nisF, nisE &amp; nisG</td>
<td>immunity</td>
<td>(Siegers &amp; Entian 1995)</td>
</tr>
</tbody>
</table>

The post-translational modifications (PTM) of nisin include the dehydration of serine and threonine residues to dehydroalanine (Dha) and dehydrobutyrine (Dhb) (Ingram 1969; Bauer & Dicks 2005). Subsequently, the thiols of appropriately located cysteine residues react with the C=C double bond of Dha or Dhb to form cyclic thioethers. Lanthionine originates from the reaction between Dha and cysteine and β-methyl-lanthionine originates from the reaction between Dhb and cysteine. These PTMs result in the formation of five ring structures of varying size within the nisin molecule (Figure 1.2) (Gross & Morell 1971; Hechard & Sahl 2002).
The primary mode of action of nisin is the formation of large non-specific pores in the cytoplasmic membrane of sensitive cells (Hsu et al. 2004), which results in the release of monovalent cations and adenosine-5’-triphosphate (ATP) (Ruhr & Sahl 1985). Nisin specifically uses lipid II (Figure 1.3) (Breukink et al. 1999) in the target cell membrane as a docking molecule for pore formation. Thus, cell wall biosynthesis is inhibited by blocking the incorporation of lipid II into the membrane structure (Brotz et al. 1998). In addition, nisin can form non-targeted pores and, particularly in staphylococci, activate cell wall hydrolysing enzymes (Bierbaum & Sahl 1985). Nisin has also been shown to inhibit the outgrowth of spores (Liu & Hansen 1993; Nissen et al. 2001).

The proposed mechanism for the nisin-lipid II pore formation is shown in figure 1.4. Initially, the N-terminal part binds to the carbohydrate moiety of lipid II in the target membrane. Then, the C-terminal part of nisin inserts into the lipid phase of the membrane, assuming a transmembrane orientation. Loss of flexibility of the hinge region between the rings C and D (Figure 1.2) impairs the insertion of nisin into the membrane and therefore...
its activity (Demel et al. 1996; Breukink et al. 1999; Wiedemann et al. 2001; Hasper et al. 2004; Hsu et al. 2004). In the transmembrane complex of nisin and lipid II, the pyrophosphate moiety of lipid II serves as anchoring point for the N-terminal amide groups of nisin (Wiedemann et al. 2001; Hsu et al. 2004). At first, the binding complex consists of one nisin molecule and one lipid II molecule (Figure 1.4A). Then, a prepore complex is formed, in which multiple nisin and lipid II molecules are assembled at the interface of the membrane (Figure 1.4B). The final pore complex formed by nisin and lipid II has the stoichiometry of eight nisins and four lipid IIs (Figure 1.4C) (Hasper et al. 2004).

Breukink and co-workers found evidence for an alternative mode of action of lantibiotics using lipid II as target (Hasper et al. 2006). Several lantibiotics, including nisin, have an N-terminal pyrophosphate cage, a baseball glove-like structure formed by the lanthionine rings A and B (Figure 1.2), binding the pyrophosphate of lipid II (Hsu et al. 2004). Some lantibiotics targeting lipid II are too short to span the lipid bilayer in order to form pores. These lantibiotics display a bactericidal activity, by removing lipid II from its functional location in the membrane as a cell wall precursor and thereby blocking cell wall synthesis (Hasper et al. 2006).
1.3 Genetics and mode of action of bacteriocins

1.3.2 **CLASS II, NON-MODIFIED BACTERIOCINS**

(a) **Anti-listerial pediocin-like bacteriocins**

Pediocin PA-1/AcH is produced by *Pediococcus acidilactici* and consists of 44 amino acids, four of which are cysteines (Henderson *et al.* 1992; Marugg *et al.* 1992; Nieto Lozano *et al.* 1992) (pediocin PA-1 is identical to pediocin AcH; Motlagh *et al.* 1992). The cysteines are linked by two disulfide bridges in the mature bacteriocin, the formation of which appears to give a broader spectrum of activity compared to those bacteriocins with just one disulfide bridge (Guyonnet *et al.* 2000; Richard *et al.* 2006). Intact disulfides are essential for activity as reduction with DTT (dithiothreitol) markedly decreases activity (Fimland *et al.* 2000). Pediocin PA-1/AcH production genes are located within an operon on a single plasmid (Miller *et al.* 2005).

Mesentericin Y105 produced by *Leuconostoc mesenteroides Y105* is a 37 amino acid bacteriocin, characterised by one disulfide bridge (Fremaux *et al.* 1995).

Pediocin PA-1/AcH and Mesentericin Y105 both exhibit anti-*Listeria* activity and contain the N-terminal YGNGVXC amino acid motif (Figure 1.5) which appears to be important for their antibacterial activity, since modifications and/or deletions in this motif result in a decrease of the anti-listerial activity (Fleury *et al.* 1996; Miller *et al.* 1998a; Miller *et al.* 1998b; Ennahar *et al.* 2000). The group of bacteriocins that contains this N-terminal motif are all active against *Listeria*, and are referred to as “pediocin-like” bacteriocins.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pediocin PA-1/AcH</td>
<td>KYYGNVVTCKHCSVDWNGKATTCCIINNGAMAWATGHQGNIKC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mesentericin Y105</td>
<td>KYYGNVHCSTKSQGCSVNNGEASAEGHIHLANGNGFW</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carnobacteriocin B2</td>
<td>VNYGNGVSCSKTCSVNWQAPAFQERTAGINSFVSFGVAGGSGRPP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carnobacteriocin BM1</td>
<td>ISYGNVYCNKECSVNVKAENQAITGIVGGWASLAMGHP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Consensus: ..YGNGVXC...

**Figure 1.5: N-terminal amino acid sequence motif of pediocin-like bacteriocins.**

Sequences of pediocin-like bacteriocins are aligned by the consensus motif (adapted from Eijsink *et al.* 2002).
(b) Two-peptide bacteriocins

Plantaricin EF and JK, lactococcin G, and lactacin F are two-peptide bacteriocins that exhibit a relatively narrow antibacterial spectrum limited to closely related species. Plantaricin EF and JK are produced by Lactobacillus plantarum C11 (Diep et al. 1996; Anderssen et al. 1998). Lactococcus lactis produces lactococcin G, which consists of an α and β subunit with 39 and 35 amino acids, respectively (Nissen-Meyer et al. 1992). Lactobacillus johnsonii produces two peptides lactacin A (57 amino acids) and LafX (48 amino acids) which form the two-peptide bacteriocin lactacin F (Abee et al. 1994).

Two-peptide bacteriocins consist of two separate peptides, which can be type E (enhancing) where one peptide enhances the antibacterial activity of the other (e.g. lactacin F) or type S (synergistic) where the individual peptides have little or no activity (e.g. lactococcin G, plantaricin EF and JK). Two-component bacteriocins are encoded adjacent to each other on the same operon and there is only one immunity gene present for each two-peptide bacteriocin, which supports the finding that the peptides work together as one unit. Two-peptide bacteriocins are typically cationic and contain hydrophobic and/or amphiphilic regions, which when exposed to a membrane interact with each other in a structured manner (Oppegard et al. 2007). The two peptides of lactococcin G interact in the membrane and form a helix-helix structure that involves GxxxG-motifs in both peptides (Rogne et al. 2008; Nissen-Meyer et al. 2009).

(c) Cyclic bacteriocins

Cyclic bacteriocins differ from linear bacteriocins only in that they are post-translationally modified through a peptide linkage between their N- and C-termini to form a cyclic peptide. The circularisation stabilises the peptide conformation and increases resistance to proteolysis (Maqueda et al. 2008). Lactobacillus gasseri LA39 is the producer strain of the cyclic gassericin A (Kawai et al. 1998). It is synthesised as a 91 amino acid prepeptide that after cleavage forms a 58 amino acid mature peptide, exhibiting a relatively broad spectrum of activity.
(d) Genetic organisation and production of class II bacteriocins

Gene clusters associated with class II bacteriocins contain a minimum four genes, which include the genes encoding the bacteriocin prepeptide, the cognate immunity protein, the dedicated cell-membrane associated ATP-binding cassette (ABC)-transporter and a membrane-bound accessory protein. It is also common to have genes encoding a histidine protein kinase and a response regulator (two-component signalling system (TCS)), and other genes responsible for the maturation of the bacteriocin, clustered in one operon. Most commonly the N-terminal extension of class II bacteriocins is a diglycine leader peptide that is cleaved by a cysteine peptidase forming part of the N-terminal domain of the ABC-transporter. Some bacteriocins, such as enterocin P (Herranz & Driessen 2005), are secreted by the translocase general secretion (Sec) pathway, rather than ABC-transporters. They are also produced as prepeptides, but their N-terminal signal peptide is generally longer than that of bacteriocins with diglycine leader peptides.

Class II bacteriocins are ribosomally synthesised as prebacteriocins and the N-terminal leader peptide is subsequently cleaved to form the active peptide (Riley & Wertz 2002a, 2002b; Papagianni 2003). The leader sequence has two functions; for protection of the host from its own bacteriocin before export and as a signal sequence to direct the prebacteriocin to the correct ABC-transporter protein (Drider et al. 2006).

(e) Regulation of bacteriocin production

Production of class II bacteriocins is often regulated in a cell-density dependent manner, a process that is also referred to as quorum sensing (see section 1.4 for more details). The gene that codes for the induction peptide or pheromone is normally co-transcribed with genes encoding both a histidine protein kinase and a response regulator, which form a two-component signalling system (Diep et al. 1996; West & Stock 2001). Induction peptides are produced as precursors with diglycine-type leader peptides at the N-terminus, which are processed by the same dedicated ABC-transporter as the bacteriocins. Pheromones are generally shorter than bacteriocins, with polypeptide chain lengths varying between 19 and 26 amino acids. In general, these induction peptides do not have antimicrobial activity, although there are some exceptions. For example, nisin (Kuipers et al. 1995), plantaricin A.
(Diep et al. 1995) and carnobacteriocin B2 (Worobo et al. 1994; Franz et al. 2000b) act as both pheromone and bacteriocin. The binding of the pheromone to the sensor domain of the histidine protein kinase results in autophosphorylation of a conserved histidine residue in the kinase domain. The response regulator then catalyzes the transfer of the phosphoryl group from the histidine (His) to a conserved aspartate (Asp) in its own receiver domain, bringing about a conformational change that is thought to modify the activity of the associated output domain, resulting in a specific response. By binding to specific promoter regions of target genes, the activated response regulator induces the transcription of the bacteriocin/pheromone gene cluster.

(f) Processing and secretion of bacteriocins

The activation and secretion of class II bacteriocins is carried out by dedicated ABC-transporters (Havarstein et al. 1995), which are integral membrane proteins. They are produced as a single polypeptide chain, which has a conserved C-terminal nucleotide-binding domain (NBD) (Fath & Kolter 1993). The NBD binds and hydrolyses ATP. The ABC-transporter has most commonly a bacteriocin-processing peptidase domain at the N-terminus that is probably cytoplasmic (Franke et al. 1999). This peptidase domain belongs to the cysteine peptidase family C39 and contains two conserved sequence motifs: a cysteine motif (QX₄D/ECX₂AX₃MX₄Y/FGX₄I/L) and a histidine motif (HY/FY/VVX₁₀I/LXDP). Accessory proteins are commonly found associated with ABC-transporters and have an average size of 470 amino acids. Their hydrophobic N-terminal domains usually span the membrane whereas the large C-terminal cytoplasmic domain is mainly hydrophilic. The exact function(s) of these domains is (are) unknown, but it is suspected that they facilitate translocation across the membrane and/or help in the processing of the leader peptide (Franke et al. 1996; Nes et al. 1996; van Belkum et al. 1997).
(g) Mode of action

Class II bacteriocins mainly induce the formation of pores in the cell membrane, which results in loss of nutrients and metabolites from the cell (Table 1.2). In contrast to leakage from pores formed by lantibiotics, there is no evidence of ATP leakage. Most of the mechanisms of target cell recognition have not been fully characterised. The sensitivity of a target cell may depend on the composition and/or type of phospholipids found in the cytoplasmic membrane or the presence of specific receptors in that membrane. For pediocin-like bacteriocins, the initial electrostatic interaction is mediated by the cationic N-terminal domain (Fimland et al. 2005). Subsequent pore formation takes place regardless of the state of energisation of the cell. The C-terminal domain of pediocin-like bacteriocins penetrates the membrane of target cells and in this way confers specificity to the target cell (Fimland et al. 2005).

There is some evidence that a number of the class II bacteriocins, including pediocin-like bacteriocins, target transmembrane components of the mannose phosphotransferase system (man-PTS) of target cells as receptors (Dalet et al. 2001; Hechard et al. 2001; Ramnath et al. 2004; Diep et al. 2007). Lactococcin A (LcnA) is a class II non-modified bacteriocin from *Lactococcus lactis* that forms a strong complex with EIIC and EIID components of the man-PTS, resulting in permeabilisation of the membrane, leakage of cellular components followed by cell death (Diep et al. 2007). The immunity protein (LciA) confers immunity to the producer strain by binding to the same membrane components of the man-PTS and the bacteriocin to form a strong complex, thereby blocking the action of the bacteriocin. LciA only binds strongly to the man-PTS components when LcnA is present (Diep et al. 2007). Recently, it has been shown that a phylogenetically defined subgroup of the man-PTSs is targeted (Kjos et al. 2009).
Table 1.2: Mode of action of class II bacteriocins (adapted from Hechard & Sahl 2002)

<table>
<thead>
<tr>
<th>Bacteriocin</th>
<th>Mode of action</th>
<th>ΔpH</th>
<th>PMF</th>
<th>ATP</th>
<th>Efflux</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mesentericin Y105</td>
<td>Pore formation</td>
<td>ND</td>
<td>dissipation</td>
<td>ND</td>
<td>aa</td>
<td>(Hechard et al. 1992; Maftah et al. 1993; Fleury et al. 1996; Robichon et al. 1997; Dalet et al. 2000; Dalet et al. 2001)</td>
</tr>
<tr>
<td>Pediocin PA-1/AcH</td>
<td>Pore formation</td>
<td>dissipation</td>
<td>dissipation</td>
<td>Depletion of intracellular ATP</td>
<td>aa, K⁺</td>
<td>(Blunia et al. 1991; Christensen &amp; Hutkins 1992; Chikindas et al. 1993; Chen et al. 1997a; Chen et al. 1997b; Chen et al. 1998; Finland et al. 1998; Waite et al. 1998)</td>
</tr>
<tr>
<td>Lactacin F</td>
<td>Pore formation</td>
<td>ND</td>
<td>dissipation</td>
<td>Depletion of intracellular ATP</td>
<td>K⁺ and P₁</td>
<td>(Abee et al. 1994)</td>
</tr>
<tr>
<td>Lactococcin G</td>
<td>Cation pores</td>
<td>No effect</td>
<td>dissipation</td>
<td>Depletion of intracellular ATP</td>
<td>aa, Na⁺, K⁺, Li⁺, Cs⁺, Rb⁺</td>
<td>(Nissen-Meyer et al. 1992; Moll et al. 1996; Moll et al. 1998)</td>
</tr>
<tr>
<td>Plantaricin EF</td>
<td>Cation pores</td>
<td>dissipation</td>
<td>dissipation</td>
<td>ND</td>
<td>Cations</td>
<td>(Moll et al. 1999)</td>
</tr>
<tr>
<td>Plantaricin JK</td>
<td>Anion pores</td>
<td>dissipation</td>
<td>dissipation</td>
<td>ND</td>
<td>Anions</td>
<td>(Moll et al. 1999)</td>
</tr>
</tbody>
</table>

aa, amino acids; ND, not determined; P₁, phosphate; PMF, proton motive force

(h) Immunity of producer organism

LAB producing bacteriocins are immune to the antimicrobial activity of their own bacteriocins because they express cognate immunity genes (Diep et al. 1995, 1996; Eijsink et al. 1996). The gene encoding the immunity protein is often located directly downstream of the bacteriocin gene and both genes are frequently co-transcribed (Nes et al. 1996). Immunity proteins vary in length between 51 and 154 amino acids and often contain hydrophobic regions. At least 20 putative immunity proteins have been identified and can be divided into three subgroups according to their sequence similarities (Drider et al.
2006). Generally, immunity proteins provide resistance against just one bacteriocin; a specificity that has been shown to be determined by the C-terminal parts of the proteins (Johnsen et al. 2004; Johnsen et al. 2005). It has been reported that some immunity proteins provide immunity to several bacteriocins (Eijsink et al. 1998; Franz et al. 2000b).

There are several proposed modes of action for immunity proteins: they can interact with the bacteriocin-induced pore and physically block it. Alternatively, they can interact with the bacteriocin receptor, thereby altering the receptor conformation so it can no longer bind bacteriocins and/or masking the bacteriocin-binding site (Drider et al. 2006; Martin-Visscher et al. 2008). Another mechanism conferring immunity is where the immunity protein directly interacts with the bacteriocin and the receptor (Diep et al. 2007).

1.4 QUORUM SENSING

Quorum sensing (QS) is a cell-cell signalling process by which bacteria communicate with one another, allowing them to synchronize their behavioural responses, using a chemical language coded by molecules called autoinducers or pheromones. Prokaryotic quorum sensing signals have a variety of structures; generally Gram-negative bacteria use lipid steroids, such as N-3-oxooctanoyl-L-homoserine lactone (AHL; figure 1.6A), whereas Gram-positive bacteria predominantly use peptides as signalling molecules (Figure 1.6B).

Accumulation of an autoinducer over a threshold concentration results in its detection by the bacteria which respond by modifying the expression of specific genes. Quorum sensing was first described for Vibrio fischeri and Vibrio harveyi (Nealson et al. 1970; Nealson & Hastings 1979) where it was found to regulate bioluminescence. It has since been shown that the use of quorum sensing to regulate gene expression is widespread among bacteria (Reading & Sperandio 2006). It is possible for one organism to have multiple quorum sensing systems, which can be organised in different ways. V. harveyi uses a parallel arrangement where all signals are channelled into a shared regulatory pathway. Alternatively, quorum sensing regulatory systems can operate in series to control a wide
range of target genes. An example of this system is present in *Pseudomonas aeruginosa* (Gambello & Iglewski 1991; Ochsner et al. 1994). The quorum sensing system of *Bacillus subtilis* responds to two antagonistic autoinducers that allow the bacterium to commit to one of the mutually exclusive states, either competence or sporulation (Solomon et al. 1995).

![Figure 1.6: Quorum sensing signalling molecules of prokaryotes.](image)

**A:** Gram-negative lipid steroid signal N-3-oxooctanoyl-L-homoserine lactone (AHL; adapted from Hughes & Sperandio, 2008) **B:** Gram-positive peptide signals encoded on the *L. plantarum* WCFS1 genome. Triangles indicate the cleavage site between the diglycine leader peptide and the (predicted) mature peptide in PlnA and PltA (shown in bold). The underlined bold residues in the LamD precursor peptide are processed to the mature thiolactone peptide shown on the right. The mature peptide sequence and structure of the putative autoinducing peptide encoded by lp_3089 are unknown (adapted from Sturme et al. 2007).

### 1.4.1 Evolution of Two-Component Signalling Systems

There is evidence that two-component signalling systems (TCS), consisting of a histidine protein kinase (HPK) and a response regulator protein (RR), evolved from simpler one-component systems (Ulrich et al. 2005) that consisted of an input domain and an output domain combined in one protein. They included some of the same input and output domains typically found in TCSs, and therefore may have detected similar stimuli, resulting in similar responses in the cell.

Some examples of one-component systems have been examined experimentally. The fumarate-nitrate reduction (FNR) system in *E. coli* controls, in combination with the two-component signalling system ArcBA (Iuchi & Lin 1988; Iuchi et al. 1990), gene expression
under anaerobic conditions (Spiro & Guest 1990; Unden et al. 1994; Guest et al. 1995). *Lactobacillus casei* and *Lactococcus lactis* each contain FNR-like regulatory proteins (FLP). They possess FNR-like DNA-recognition motifs but retain only two conserved cysteine residues (one N-terminal and one central) (Guest et al. 1995; Gostick et al. 1998; Gostick et al. 1999).

Compared to TCSs the simpler one-component systems are more common in prokaryotes and have a greater diversity in their domain structure (Ulrich et al. 2005). With few exceptions TCSs are absent in archaea, and it is thought that the few that do exist were probably acquired by horizontal gene transfer from bacteria (Koretke et al. 2000). An advantage TCSs have, is their ability to detect extracellular signals via HPKs with transmembrane domains. In contrast, most one-component regulators are predicted to be cytosolic and can therefore only sense intracellular signals. Because the majority of signalling pathways involve DNA sequence recognition, it is advantageous to divide the tasks of sensing extracellular signals at the cell membrane and the interaction with targets in the genome between two separate proteins.

Phosphorelays are more complex signalling systems, involving multiple phosphorylation steps of successive histidine and aspartic acid residues in more than one protein by more than two kinases. Hybrid kinases are often part of phosphorelays and consist of both His- and Asp-containing domains combined in one protein. Phosphorelays provide a greater number of possible regulation sites and allow a greater versatility of signalling strategies. These phosphorelay systems, together with their associated hybrid kinases are primarily found in eukaryotes, while the simpler two-component signalling systems are predominately found in prokaryotes.

### 1.4.2 Peptide Signalling in Gram-positive Bacteria

Gram-positive bacteria use peptide pheromone signalling pathways to regulate various cellular activities. These include competence development (*com*) in *Bacillus subtilis* and streptococci (Havarstein & Morrison 1999; Lazazzera et al. 1999), secretion of staphylococcal toxins and proteases during stationary growth (*arg*) (Novick 2003), and the

Class I and II bacteriocins are regulated by peptide pheromone regulatory systems. The lantibiotic nisin also functions as peptide pheromone, autoregulating its own biosynthesis (Kuipers et al. 1995). In contrast, the synthesis of class II bacteriocins is regulated by separate pheromones interacting with the TCSs that regulate transcription of the relevant bacteriocin gene cluster (Diep et al. 2000; Ennahar et al. 2000; Franz et al. 2000a; reviewed in Quadri 2002). Many of these pheromones are cleaved from larger precursor peptides with bacteriocin-like diglycine-type leader peptides (Nes et al. 1996). Peptide pheromones cannot diffuse across membranes, so that processing and modification, as well as export, are mediated by dedicated ABC-transporters. Pheromones are recognized specifically by the histidine protein kinases of TCSs, which are autophosphorylated when the pheromone concentration reaches a threshold, resulting ultimately in expression of the bacteriocin gene cluster.

(a) Protein architecture of two-component systems

In general, two-component signalling systems have a modular architecture consisting of a histidine protein kinase and a response regulator. Only some TCSs contain only two proteins; many contain auxiliary proteins in addition to the HPK and RR proteins. Histidine protein kinases are usually made up of two functionally and structurally distinct segments: an N-terminal sensor domain and a conserved C-terminal kinase domain. The C-terminal ATP-binding domain contains a conserved motif (H box) that includes the phosphoryl-accepting histidine residue as well as several highly conserved clusters of residues called homology boxes N, D, F and G (Parkinson & Kofoid 1992; Grebe & Stock 1999; Wolanin et al. 2002) that play crucial roles in substrate binding, catalysis and/or structure. Grebe and co-workers developed a classification of HPKs based on the presence and structure of these homology boxes (Grebe & Stock 1999) that showed that most histidine protein kinases in peptide-based TCSs belong to the HPK10 subfamily. This subfamily does not contain a D box, which is normally part of the nucleotide-binding domain, but it usually possesses five to seven N-terminal transmembrane segments (Sturme et al. 2007).
Generally, response regulators (RR) contain two functional domains: an N-terminal receiver domain and a C-terminal output domain. The receiver domain includes the phosphoryl-accepting aspartate residue that interacts with the kinase domain of its cognate HPK. The most common output domains are transcription factors that bind DNA, but a small number of output domains show enzymatic activity and about 25% of RRs appear to have no output domain. Similar to the histidine protein kinases, response regulators have been classified based on their receiver and DNA-binding domains (Grebe & Stock 1999). RRs related to histidine protein kinases of the HPK\textsubscript{10} subfamily belong to the R\textsubscript{0}/ComE subfamily (Sturme \textit{et al.} 2007). The majority of response regulators of the R\textsubscript{0}/ComE subfamily were classified as LytTR family of response regulators (PF04397) (Nikolskaya & Galperin 2002). LytTR-containing proteins can have a number of domain architectures, but the majority of them are RRs in a quorum sensing two-component signalling system (Gao \textit{et al.} 2007). A recent study showed that about 2.7% of all prokaryotic RRs are LytTR-containing proteins (Galperin 2006).

The first structure of a LytTR domain was recently solved by Stock and co-workers and revealed a novel structure that is not helix-turn-helix (HTH) (Sidote \textit{et al.} 2008). The C-terminal DNA-binding domain of \textit{Staphylococcus aureus} AgrA was co-crystallized with a 15 base pair (bp) DNA duplex containing a 9 bp consensus binding sequence. The LytTR domain has 10 β-strands organized in three antiparallel β-sheets and a two-turn α-helix as shown in figure 1.7 (Sidote \textit{et al.} 2008).

Base-specific contacts are made between residues located in the loops between the β-sheets and two adjacent major grooves and the intervening minor groove of the DNA (Figure 1.7). Only two amino acids have direct base-specific interactions with the DNA, although there are another 10 non-specific contacts. Even though the LytTR domain DNA binding sequence is highly conserved, the residues forming the base-specific interactions are poorly conserved in the LytTR family, which probably explains the diversity of their target DNA sequences. The binding of the LytTR domain results in significant bending of the target DNA, which promotes the successful binding of the RNA polymerase (Galperin 2008).
The aim of all regulatory strategies is to vary the level of phosphorylation of the response regulator, which determines the output response. Typically, RRs have autophosphatase activity limiting the lifetime of the phosphorylated state; the half-lives of response regulators can range from seconds to hours. Some auxiliary proteins enhance autophosphatase activity of the response regulators or exhibit phosphatase activity themselves, resulting in accelerated dephosphorylation of the RR. The phosphatase activity can also be provided by the histidine protein kinase itself. In the case of transmembrane HPKs, regulation is promoted directly by stimuli (such as pheromones) or, for cytoplasmic HPKs, indirectly by interaction with auxiliary proteins.

(b) Quorum sensing in Lactobacillus plantarum C11

Bacteriocin production in *L. plantarum* C11 involves five operons and is regulated by a pheromone induced two-component signalling system (Diep *et al.* 1994; Diep *et al.* 1995, 1996; Anderssen *et al.* 1998). The operons *plnEFI* and *plnJKLR* code for the bacteriocins and immunity proteins and the operon *plnGHSTUV* encodes for an ABC-transporter,
processing and exporting peptides with diglycine leader sequences (Havarstein et al. 1995). The operon plnABCD codes for the two-component signalling system and the genes in the last operon plnMNOP have unknown functions in the bacteriocin production (Diep et al. 1996; Risoen et al. 1998; Risoen et al. 2000; Diep et al. 2001; Risoen et al. 2001; Diep et al. 2003) (Figure 1.8).

The antimicrobial peptide pheromone PlnA (Diep et al. 1995, 1996) induces bacteriocin production by binding to a histidine protein kinase (PlnB), which activates two antagonizing response regulators PlnC and PlnD, both containing C-terminal LytTR binding domains (Risoen et al. 1998; Risoen et al. 2000; Diep et al. 2001; Risoen et al. 2001; Diep et al. 2003). PlnC acts as an activating regulator, whereas PlnD acts as repressor. All promoters of the pln-operon contain a pair of conserved direct repeats that serve as binding sites for the two response regulators. The five operons induced by PlnA are differentially expressed in terms of timing and strength. Early on during induction, the pheromone PlnA causes the strong autoactivation of the plnABCD operon (Diep et al. 2003), which gradually leads to enhanced activation of the other operons.

The pln locus has also been described in several other L. plantarum strains such as WCFS1 (Kleerebezem et al. 2003), NC8 (Maldonado et al. 2004), J23 (Rojo-Bezares et al. 2008)

---

Figure 1.8: Genetic map of pln locus of different L. plantarum strains.
A: L. plantarum C11 (GenBank accession number X94434); B: L. plantarum WCFS1 (GenBank accession no. AL939253); C: L. plantarum NC8 (GenBank accession no. AF522077); D: L. plantarum J23 (GenBank accession no. DQ323671). The pln genes are represented by arrows with different colours corresponding to each operon. The promoter sequences are indicated by small black arrows (from Rojo-Bezares et al. 2008).
(Figure 1.8) and J51 (Navarro et al. 2008). The \textit{pln} locus in these \textit{L. plantarum} strains appears to have a modular structure and has undergone various reorganisations (recently reviewed by Diep et al. 2009). Some regions are highly conserved e.g. the regions related to transport (\textit{plnGH}), bacteriocin maturation (\textit{plnXY}) and variable regions related to bacteriocin regulation, production and immunity (\textit{plnABCD; plnIEF; plnJKL}).

1.5 GLYCOSYLATION

1.5.1 \textit{GLYCOXYLATION IN EUKARYOTES}

Glycoproteins with an impressive variety of carbohydrate linkages have been found in essentially all organisms: eukaryotes, archaea and bacteria. Often, multiple glycans and linkages are found in the same protein, depending on the amino acid sequence and conformation as well as available glycosyltransferases. \textit{N}-glycosyl and \textit{O}-glycosyl linkages occur in all three kingdoms, whereas \textit{C}-mannosylation and phosphoglycosylation are only found in eukaryotes (reviewed in Spiro 2002).

In eukaryotes the majority of secreted and membrane proteins are glycosylated, resulting in glycosylation being the most abundant type of post-translational modification. The glycan moiety carries out several important physicochemical and biological roles. These include the protection of the protein against proteolytic attack, extremes of temperature and increased solubility. In the latter role glycans have been shown to direct protein folding, to act as receptors, as determinants of immuno-recognition and in intercellular recognition and adhesion.

Eukaryotic glycoproteins can be divided into three classes: \textit{N}-glycosylproteins, \textit{O}-glycosylproteins and \textit{N,O}-glycosylproteins. In \textit{N}-glycosylated proteins the sugar chains are attached to the protein via the side chain amide nitrogen of an asparagine residue, while in \textit{O}-glycosylated proteins they are linked via the hydroxyl groups of mainly serine or threonine, and occasionally to tyrosine, hydroxy proline, hydroxy lysine or xylose residues (Figure 1.9).
Chapter 1: Introduction

1.5 Glycosylation

The typical sequence motif (sequon) for N-glycosylated proteins is Asn-X-Ser/Thr, where X can be any amino acid, except Pro, although not every sequence motif will be glycosylated. In eukaryotic proteins 70-90% of the Asn-X-Ser/Thr sites are N-glycosylated. O-glycosidic linkages seem to require no specific amino acid sequence motif, although they are often found in sequences with a higher proportion of proline residues than normal.

*N*-linked glycans have the common pentasaccharide inner-core of Man$_3$GlcNAc$_2$ (Figure 1.10), which is followed by three different types: the high-mannose-type (or oligomannosidic-type), the complex-type, and the hybrid-type.

![Figure 1.9: Glycan linkages to peptide chains.](image)

A: N-glycosyl linkage via the side chain amide nitrogen of asparagine; B: O-glycosyl linkage via the hydroxyl group of serine or threonine (adapted from Niederhoffer 2007).

The outer parts of the glycans are called “antenna” because of their flexibility and their roles in cell-cell, cell-protein and protein-protein recognition. In addition to mutations in their polypeptide chains, almost all glycoproteins are differentiated by number, location and sequence of the bound glycans, to produce a number of different glycoforms of the protein. Such diversity is called “microheterogeneity”. O-Glycosidically linked oligosaccharides vary in size from a single galactose residue in collagen to chains of up to 1000 disaccharide units in proteoglycans.

![Figure 1.10: Pentasaccharide inner-core of N-linked glycans.](image)

Asn, asparagines; GlcNAc, N-acetylglucosamine; Man, mannose.
In eukaryotes, the biosynthesis of \( N \)-glycans occurs via sequential addition of oligosaccharides to the lipid moiety, dolichol monophosphate (a long-chain polyisoprenol; figure 1.11a). The addition of each monosaccharide is carried out by a unique glycosyltransferase. The complete sugar chain is then transferred ‘en bloc’ to the nascent polypeptide chain by a complex of proteins collectively known as oligosaccharyltransferase (OTase) and is only observed in the presence of the sequon for \( N \)-linked glycans. The newly synthesised glycoproteins undergo several cycles of glycosylation and trimming in the endoplasmic reticulum (ER) by enzymes, which play key roles in controlling folding and ensuring that only correctly folded glycoproteins can exit the ER to the Golgi apparatus. Glycoproteins with \( N \)-linked glycans are extensively modified and elaborated in the Golgi apparatus in a process involving further trimming and sequential elongation. Finally, the glycoproteins are sorted and transferred to their destinations via specific membranous vesicles.

\( O \)-linked glycans in eukaryotes are generally synthesised in the Golgi apparatus (Figure 1.11b), although synthesis can also occur in the ER and cytoplasm. The synthesis of \( O \)-linked oligosaccharides is completely sequential, where the product of one glycosyltransferase is used as the acceptor substrate for the next glycosyltransferase. After completion, the \( O \)-glycosylated proteins are sorted and transported in vesicles to their destinations. The position of \( O \)-glycosylation sites is believed to be determined by the secondary or tertiary structure of the protein (Young et al. 1979; Gooley et al. 1991; Muller et al. 1997; Yoshida et al. 1997)
Modification of peptides and proteins with a single O-linked β-N-acetylglucosamine (O-GlcNAc) has emerged as important regulatory mechanism in eukaryotes (Zachara & Hart 2002). It is thought to regulate the function and/or activity of proteins in cells and to have an analogous role to protein phosphorylation. The dynamic and fast modification with O-GlcNAc is carried out by two enzymes; an O-GlcNAc transferase (OGT; Haltiwanger et al. 1990) and a β-N-acetylglucosaminidase (OGA; Dong & Hart 1994). Evidence suggests a complex interaction between O-GlcNAc modifications and phosphorylation at serine and threonine residues. These interactions can be reciprocal at the same site, which occurs in the transcription factor c-Myc (Cheng & Hart 2001) or at adjacent sites, as observed in the tumour suppressor p53 (Yang et al. 2006). Recently, the structure of a bacterial OGT orthologue of Xanthomonas campestris was solved in complex with UDP (Martinez-Fleites et al. 2008) and with UDP-GlcNAc phosphonate analogue (Stimmel et al. 1990).
1.5.2 Prokaryotic Glycosylation

Until the mid-1970s prokaryotes were considered to be unable to glycosylate proteins, which was due to the fact that in the most studied prokaryotes, e.g. *E. coli*, *Salmonella* sp. and *Bacillus subtilis*, no glycosylated proteins had been identified. The first prokaryotic glycoprotein, the surface-layer (S-layer) glycoprotein of the archaeon *Halobacterium salinarium*, was identified and thoroughly described by Strominger and co-workers in 1976 (Mescher & Strominger 1976). The S-layer is part of the cell envelope of bacteria and archaea, where it offers protection and resistance against e.g. bacteriophage and lytic enzymes. The interest in prokaryotic glycosylated proteins was limited, because the then known S-layer glycoproteins originated from non-pathogenic organisms with no medical significance. That changed when protein glycosylation in bacteria was recognised as being associated with virulence factors of medically significant pathogens (Schmidt et al. 2003). The S-layer glycoproteins are the best studied examples of prokaryotic glycoproteins, but glycosylation of enzymes (especially polysaccharide-degrading enzymes), antigens, membrane-associated (e.g. outer membrane) proteins, surface-associated (e.g. flagellins, pilins) and other cell-envelope components (Moens & Vanderleyden 1997; Schaffer et al. 2001) also occurs. It was shown that the pilin proteins of the pathogens *Neisseria meningitidis*, *Neisseria gonorrhoeae* and *Pseudomonas aeruginosa* are all O-glycosylated (Stimson et al. 1995; Castric et al. 2001; Hegge et al. 2004).

Prokaryotic glycoproteins can be located at the S-layer, at the cell-surface, membrane-associated, secreted and intracellular. The functional role of prokaryotic glycosylation needs further investigation, but is believed to be similar to eukaryotic glycoproteins. Glycosylation is an expensive investment; therefore these modifications are likely to be essential. Proposed functions include enzymatic activity, pathogenicity, cell-surface properties, protection against proteases and physicochemical properties (Upreti et al. 2003).

Only N- and O-glycosylation have so far been found in prokaryotes, but the available data shows that the structures of prokaryotic glycoproteins are far more diverse than the structures found in eukaryotes, although they share some characteristics (Moens &
Vanderleyden 1997). The first N-linked glycans were described by Young and co-workers (2002) in Campylobacter jejuni (Figure 1.12). They are attached via the same eukaryotic Asn-X-Ser/Thr consensus. As in eukaryotes, there is no specific sequence motif known for O-glycosylation in prokaryotes. Structural variations of bacterial glycoproteins are found in unusual saccharide constituents and linkages. For example, an exotic sugar constituent 2, 4-diacetamido-2, 4, 6-trideoxyhexose is found on Neisseria meningitidis pilin (Stimson et al. 1995) and a linkage between an asparagine side chain and a rhamnose monosaccharide is seen in the S-layer protein of Bacillus stearothermophilus (Messner & Sleytr 1988).

In a scenario reminiscent of eukaryal protein glycosylation, prokaryal glycosyltransferases catalyse the transfer of a sugar residue from an activated donor substrate to an acceptor substrate (Sumper 1987). However, because of their different cell structure, prokaryotes have to use different mechanisms to glycosylate their proteins. Nevertheless, some similarities to the eukaryotic system exist.

In archaea and bacteria, nucleotide- and lipid-linked oligosaccharide precursors have been found (Doherty et al. 1982; Hartmann & Konig 1989; Lechner & Wieland 1989; Hartmann et al. 1993; Zhu et al. 1995) and the lipid seems to be a dolichol phosphate as in eukaryotes. Their precursors include, however, nucleotide-activated oligosaccharides (Hartmann & Konig 1989; Hartmann et al. 1993) that do not occur in eukaryotic glycosylation processes. The processes of N- and O-linked glycosylation in prokaryotes are shown in figure 1.12 for Campylobacter jejuni (Szymanski & Wren 2005). In C. jejuni the N-linked glycans are built up through the sequential addition of nucleotide-activated sugars to a lipid-linked precursor (Figure 1.12c). The entire glycan is then flipped across the inner membrane into the periplasm by a putative ABC-transporter, where it remains unmodified. In C. jejuni, PglB is the only enzyme necessary for the transfer of the oligosaccharide on to the asparagine residue.

The proposed mechanism of O-glycosylation in C. jejuni is shown in figure 1.12d. As the bipolar flagella of C. jejuni span both the inner and outer membranes, O-linked glycosylation of flagellin monomers is proposed to occur in the cytoplasm/inner membrane where nucleotide-activated sugars are individually added to serine or threonine residues that are surface exposed.
1.5 Glycosylation

An S-glycosidic linkage to a peptide was identified in 1971 by Lote and Weiss (Weiss et al. 1971; Lote & Weiss 1971a, 1971b). They isolated an octapeptide from normal concentrated human urine, which contained two galactose molecules S-glycosidically linked to an N-terminal cysteine. However, this research has never been corroborated using modern mass spectrometry and there have been no further reports of it in the literature.

The glycosylation of a cysteine residue in a polypeptide chain as we found in PlnKW30 (see section 3.5.1), is totally new in biology and no natural examples of protein or peptide S-glycosylation exist (Taylor 1998; Thibodeaux et al. 2007).
1.5.4 Glycosyltransferases in prokaryotes

Carbohydrate-active enzymes are widespread within most organisms, with glycosyltransferases and glycoside hydrolases encoded by 1-3% of ORFs in sequenced genomes (Davies et al. 2005). The vast majority of glycosyltransferases are uncharacterised ORFs and they have been shown to be extremely hard to characterise biochemically because of their intracellular or membrane associated localization (Henrissat et al. 2008). Glycosyltransferases are a ubiquitous group of enzymes that catalyse the transfer of mono-, di- or oligosaccharides from an activated donor onto specific acceptors, forming glycosidic bonds. By using a large number of nucleotide-sugar donors and a vast variety of acceptors they can produce an almost infinite number of products. The possible acceptors include almost any class of molecule, e.g. proteins, sugars, lipids, steroids, nucleic acids and antibiotics. However, most glycosyltransferases are involved in transferring one sugar on to another sugar.

The reactions catalyzed by glycosyltransferases can result in a glycosyl bond with the same stereochemistry to that of the sugar donor or they can instead make glycosidic bonds in which the stereochemistry has been inverted compared to that of the sugar donor. For inverting glycosyltransferases an in-line SN2 attack is supported by structural (Charnock & Davies 1999a; Tarbouriech et al. 2001; Pedersen et al. 2002) and mechanistic data (Murray et al. 1996; Qiao et al. 1996) (Figure 1.13a).

The mechanism for retaining glycosyltransferases is less certain. Originally it was thought that a double displacement reaction involving the formation of a covalently bound glycosyl-enzyme intermediate is utilized analogous to that seen for the retaining glycosidases (Zechel & Withers 2000; Lairson et al. 2004; Lairson et al. 2008) (Figure 1.13b).
Several studies (Boix et al. 2001; Persson et al. 2001; Boix et al. 2002; Ly et al. 2002; Pedersen et al. 2003) of retaining glycosyltransferases could not identify the presence of an active-site nucleophile involved in the formation of the enzyme-sugar adduct. However, a mutant α 1,4-galactosyltransferase C from *Neisseria meningitidis* was identified by Lairson et al. (2004) with a covalently attached sugar to the enzyme. Although some retaining glycosyltransferases may use this double displacement reaction it seems likely that the majority use a mechanism involving the formation of a short-lived ion pair intermediate (Lairson et al. 2008).
1.5.5 **Classification of Glycosyltransferases**

Glycosyltransferases have been conventionally classified on the basis of their donor, acceptor and product specificity, as recommended by the International Union of Biochemistry and Molecular Biology (IUBMB). This classification system does not indicate the intrinsic structural features of the enzymes, nor does it sufficiently accommodate enzymes which act on several distinct substrates. Campbell and co-workers (1997) proposed a classification of glycosyltransferases into families on the basis of amino acid sequence similarities. This classification system currently includes 91 distinct sequence-derived families. The grouping of enzymes with different donor, acceptor and product specificities into polyspecific families offers great insight into the divergent evolution from an ancestral form of glycosyltransferases. Information about the glycosyltransferase families is available from the Carbohydrate-Active enZymes (CAZy) database (http://www.cazy.org/fam/acc_GT.html).

So far, ten x-ray crystal structures of glycosyltransferases have been solved which show that there are only two different folds: GT-A and GT-B (see section 1.5.6). Coutinho and co-workers propose a further division of glycosyltransferase families into clans according to their folds, catalytic apparatus and molecular mechanism. For the inverting enzymes clans I and II were suggested: clan I, with the GT-A fold and clan II, with the GT-B fold (Figure 1.14). Similarly, two clans have been identified for the retaining glycosyltransferases: clan III, with the GT-A fold and clan IV, with the GT-B fold (Figure 1.14).

The GT-A family includes most of the eukaryotic glycosyltransferases required for glycan synthesis in the ER and Golgi apparatus, but also numerous prokaryotic glycosyltransferases. The GT-B family is extremely wide spread and includes most prokaryotic glycosyltransferases that glycosylate secondary metabolites.
Figure 1.14: The hierarchical classification of glycosyltransferases (GT) from folds to clans and families system proposed by Coutinho et al. (2003).
Families are classified into clans on the basis of their fold and activity. GT family numbers belonging to each clan are indicated on the far right. Bona fide families having members with solved 3-D structures are indicated in red. The remaining families are those predicted to adopt either the GT-A or the GT-B fold. Families identified in black with an asterisk are those with structures predicted to adopt either the GT-A or the GT-B fold solely by Liu & Mushegian (2003), and those in black without an asterisk have GT-A or GT-B structures as predicted by both Liu & Mushegian and the CAZYWeb site. This classification system does not include 39 of the 90 glycosyltransferases. Members from 12 (GT22, GT39, GT48, GT50, GT53, GT57, GT58, GT59, GT66, GT83, GT85, GT86) of those families not included were predicted to adopt a proposed GT-C fold. On the basis of a determined 3-D structure, family GT36 was reclassified among the glycosidases as family GH94. Structural characterization of the remaining 26 orphan families (GT11, GT18, GT29, GT37, GT38, GT46, GT51, GT52, GT54, GT61, GT65, GT67, GT68, GT69, GT70, GT71, GT73, GT74, GT75, GT76, GT77, GT79, GT87, GT88, GT89, GT90) will provide insights into the strengths and limitations of predictive bioinformatics tools (from Lairson et al. 2008).

1.5.6 STRUCTURE OF GLYCOSYLTRANSFERASES

Until recently, x-ray crystal structures of glycosyltransferases of different families showed the occurrence of only two different folds (Coutinho et al. 2003). The existence of just two folds could be due to the structural constraints of a nucleotide-binding motif and the potential evolutionary origin from a small number of precursor sequences. Archaea possess only two glycosyltransferase families (GT2 and GT4) from which the others may have evolved.

Liu & Mushegian (2003) predicted a third glycosyltransferase fold named GT-C on the basis of iterative sequence searches using programs such as BLAST (NCBI). The first three dimensional (3D) structure of an enzyme adopting the GT-C fold was determined for the
C-terminal domain of the oligosaccharyltransferase STT3 of *Pyrococcus furiosius* (Igura et al. 2007).

However, Lovering and co-workers (2007) described the structure of a new family of glycosyltransferases that adopts a different fold from both the GT-A and GT-B families, but was not a predicted GT-C member. This bifunctional glycosyltransferase/transpeptidase enzyme is involved in peptidoglycan biosynthesis in *Staphylococcus aureus* and its glycosyltransferase domain (GT51 family) shows a primarily α-helical structure, consisting of two domains: a globular “head” region and a smaller “jaw” region closer to the membrane (Lovering et al. 2007).

The fold of the GT-A family of glycosyltransferases was first described for *Bacillus subtilis* SpsA (Charnock & Davies 1999a) from family GT2 (Figure 1.15a). The glycosyltransferase SpsA is one of the inverting enzymes and is involved in spore coat formation. The GT-B fold was first observed in the bacteriophage T4 β-glucosyltransferase (BGT; Vrielink et al. 1994) (Figure 1.15b). The BGT catalyses the transfer of glucose from uridine diphosphate-α-D-glucose (UDP-Glc) to the 5-hydroxymethylcytosine residues of its own DNA. This protects the bacteriophage DNA from host nucleases.

The GT-A fold contains two Rossmann-like β/α/β domains, consisting of a nucleotide diphosphate (NDP)-sugar binding domain and an acceptor-binding domain. These are tightly associated and tend to form a central sheet of at least eight β-strands. A common feature found in GT-A enzymes is a conserved Asp-X-Asp (DXD) motif or a related one (i.e. NDD etc.; Breton & Imberty 1999) binding a divalent metal ion, which can be either Mg$^{2+}$ or Mn$^{2+}$, although some GT-A enzymes do not need this metal ion for activity (Chiu et al. 2004; Pak et al. 2006). When a metal ion is present, the diphosphate moiety of the nucleotide diphosphate (NDP)-sugar is coordinated by the metal ion and the developing negative charge on the NDP-leaving group is electrostatically stabilised (Murray et al. 1996).

The GT-B fold shows two Rossmann-like β/α/β domains, which are less tightly associated and split by a deep cleft in which the active-site is located (Breton et al. 2001; Hu & Walker 2002). Most characterised members of the GT-B family do not require a metal ion for NDP-sugar binding and catalysis (Bourne & Henrissat 2001; Hu & Walker 2002), but some have
been reported to be metal-dependent (Morera et al. 2001). A highly conserved α/β/α motif is involved in binding the glycosyl donors, suggesting that the donor sugars are held in the same way regardless of the nature of the acceptor (Hu & Walker 2002).

![Figure 1.15: Crystal structures of GT-A and GT-B glycosyltransferases.](image)

The GT-B fold has also been found in the catalytic domain of N-acetylglucosamine (O-GlcNAc) transferase of family GT41 from Xanthomonas campestris. This enzyme catalyses the intracellular modification of serine and threonine residues with O-GlcNAc, which interferes with protein phosphorylation, and thus affects numerous cellular processes (Martinez-Fleites et al. 2008). These O-GlcNAc transferases typically contain a long N-terminal extension made of multiple tetra-tricopeptide (TPR) repeats that mediate the recognition of a broad range of target proteins and is suspected to direct substrates to the active site of the enzyme (Henrissat et al. 2008). TPR domains are not well conserved, but typically consist of 3 – 16 tandem repeats of 34 amino acid residues, although single TPR motifs can be dispersed in the protein sequence. The first solved TPR structure was the three-TPR domain of the protein phosphatase 5 (Das et al. 1998) that showed that TPRs form helix–turn–helix arrangements, where adjacent TPR motifs pack in a parallel fashion, which results in a spiral of repeating anti-parallel α-helices.

In the GT-A enzymes, nucleotide binding has been observed on the N-terminal domain, whereas in GT-B enzymes it was found on the C-terminal domain. Binding of the acceptors occurs in the other domain, accordingly. The stereochemical outcome of the reaction is not
determined by these two folds, as both retaining and inverting glycosyltransferases can have related folds.

1.5.7 MODULAR ARCHITECTURE OF GLYCOSYLTRANSFERASES

Modular architecture is a common feature in many carbohydrate-active enzymes including glycosyltransferases. Numerous glycosyltransferases possess an appended non-catalytic carbohydrate-binding module (CBM), such as CBM-13 or galectin-like (Figure 1.16A). The occurrence of tandem glycosyltransferase modules within one single polypeptide is especially widespread in glycosyltransferases involved in the synthesis of alternating polysaccharides, such as heparin (Figure 1.16B). In a less common category of modular glycosyltransferases the glycosyltransferase module carries an appended glycoside hydrolase/transglycosidase domain (Figure 1.16C).

![Figure 1.16: The modularity of glycosyltransferases.](from Coutinho et al. 2003)
1.6 **LACTOBACILLUS PLANTARUM**

The genus *Lactobacillus* contains over 80 recognised species and the complete genome sequences of 15 different strains are available in public databases (http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi) including that of *Lactobacillus plantarum* WCFS1 (Kleerebezem *et al.* 2003). Lactobacilli are rod-shaped, Gram-positive bacteria that colonise a wide range of habitats including humans, animals, plants, as well as fermenting or spoiling foods and feeds. Non-pathogenic *Lactobacillus* species are used industrially in the production of food products such as cheese, yoghurt, cereal, meat and vegetables. They are also used as probiotics to support the health of humans and animals (e.g. *Lactobacillus acidophilus* or *Lactobacillus rhamnosus* may improve the natural immune response).

The G + C content of lactobacilli DNA varies between 33 - 55 % (Bernardeau *et al.* 2008). *Lactobacillus* species produce more than 20 different bacteriocins which are generally active against other lactobacilli living in the same ecological niche. Bacteriocins produced by *Lactobacillus* species are found in each of the major bacteriocin classes, many of which have been extensively characterised. Numerous bacteriocins have been structurally characterised using nuclear magnetic resonance (NMR) spectroscopy or x-ray crystallography. For example, NMR structures are publicly available for carnobacteriocin B2 (pdb 1cw5; Wang *et al.* 1999), plantaricin A (pdb 1ytr; Kristiansen *et al.* 2005) and sakacin P (pdb 1og7; Uteng *et al.* 2003).

### 1.6.1 Plantaricin KW30

Plantaricin KW30 (PlnKW30) is produced by *Lactobacillus plantarum* KW30, a bacterium isolated from fermented corn (Asmundson *et al.* 1994; Kelly *et al.* 1994; Kelly *et al.* 1996). The activity spectrum of PlnKW30 is limited largely to strains of *L. plantarum* and a few other *Lactobacillus* species (*L. brevis* and *L. delbrueckii* subsp. *lactis*) with no inhibition of other organisms, such as *Listeria* or *Micrococcus* (Kelly *et al.* 1996). The antibacterial activity of the bacteriocin was destroyed by protease treatment but was unaffected by α-amylase, lipase A or lysozyme (Kelly *et al.* 1996). PlnKW30 was shown to be stable over a
pH range from 2 - 10 at 28 °C for 16 hours and at pH 3.5 at 100 °C for 60 minutes, however, all activity was lost after autoclaving. Bacteriocin production was shown to begin and rapidly increase when cell cultures reached stationary phase. Kelly et al. (1996) suggested that the mode of action of PlnKW30 is bactericidal because cells of the indicator strain *L. plantarum* ATCC 8014 in log phase showed a higher sensitivity than those in stationary phase (Kelly et al. 1996). Work by Kelly and co-workers (1996) lead to the assumption that PlnKW30 was not encoded on one of the five plasmids (3.5, 9.5, 10.5, 38, 60 kb) harboured by *L. plantarum* KW30 but on the chromosome, as bacteriocin-negative derivatives were found to still contain all plasmids.

Apart from a diglycine motif, PlnKW30 shows no apparent amino acid similarity to other characterised bacteriocins. Furthermore, the bacteriocin appears to have two unusual post-translational modifications at an internal serine and the C-terminal cysteine (Dr. G.E. Norris & Dr. M.L. Patchett – unpublished results).

The modification of the cysteine at the C-terminus was initially determined by mass spectrometry which showed that the C-terminal peptide HX had a mass of 462 which exceeded the mass of histidine plus any other amino acid. Originally, tandem mass spectrometry experiments showed that the peptide was most likely HC-farnesyl (Dr. G.E. Norris - unpublished data). Subsequently, this finding was supported using Western blotting with anti-farnesyl antibodies (Dr. G.E. Norris & Dr. M.L. Patchett – unpublished results & this thesis). In the course of this work, mass spectrometry data was obtained from an ESI micrOTOF Q, which identified the modification of the C-terminal cysteine as *N*-acetylhexosamine (Section 3.5.1).

PlnKW30 also appears to be glycosylated at an internal serine, serine 18, by mass comparison of the internal peptide AMCGAGYDSGTCDY which had a mass of 203 greater than that of the calculated mass of the amino acid sequence (Moore, C. A.; Mudford, J. & Norris, G. E. – personal communication). Edman sequencing of the peptide failed to show a signal for serine 18, indicating that it is this residue, which is most likely modified. A mass difference of 203 corresponds to an *N*-acytylated hexose.
The disulfide binding pattern of PlnKW30 was also determined by Edman sequencing and shown to consist of two nested disulfide bonds (Figure 1.17). The Edman determined sequence was confirmed by DNA sequencing (Dr. G.E. Norris & Dr. M.L. Patchett – unpublished results).

Figure 1.17: Schematic overview of the nested disulfide bonds of the mature PlnKW30.
2.1 GENERAL MATERIALS AND METHODS

2.1.1 PURIFIED WATER

Purified water was obtained from a Sybron/Barnstead NANOpure II filtration system (Maryland, USA), containing two ion exchange and two organic filter cartridges. This filtered water is referred to as pure H₂O throughout this thesis.

2.1.2 FILTER STERILIZATION EQUIPMENT

Sterile syringes were supplied by Terumo Corporation (Tokyo, Japan) and sterile 0.22 µm syringe filters and filter membranes were obtained from Millipore (MA, USA).

2.1.3 MEDIA

Media were prepared as listed below and autoclaved at 121 °C and 2 x 10⁵ Pa for 20 minutes. Solid media were prepared by addition of 1.5 % (w/v) agar to the culture media. *E. coli* strains were grown in Luria Bertani (LB) media, *Lactobacillus* strains in De Man, Rogosa and Sharpe (MRS) broth or chemically defined media, *S. pyogenes* and *Lactococcus*
strains in M17 media with the addition of 0.5% sterile glucose, *S. cerevisiae* and fungi strains in YM media.

(a) *Luria-Bertani (LB) media*

LB media 25.0 g
Pure H$_2$O made up to 1 L

(b) *De Man, Rogosa and Sharpe (MRS) media*

MRS media 52.2 g
Pure H$_2$O made up to 1 L

(c) *M17 media with 0.5% glucose*

M17 media 42.5 g
10% glucose 50.0 mL
Pure H$_2$O made up to 1 L

(d) *Yeast media (YM)*

Yeast extract 3.0 g
Malt extract 3.0 g
Bacto peptone 5.0 g
Glucose 10.0 g
Pure H$_2$O made up to 1 L
### (e) Chemically defined medium (CDM)

Table 2.1: Chemicals and concentrations for chemically defined medium (Saguir & de Nadra 2007)

<table>
<thead>
<tr>
<th>Chemical</th>
<th>g / L</th>
<th>Chemical</th>
<th>g / L</th>
</tr>
</thead>
<tbody>
<tr>
<td>glucose</td>
<td>10.0</td>
<td>DL-Alanine</td>
<td>0.20</td>
</tr>
<tr>
<td>potassium acetate</td>
<td>10.0</td>
<td>L-Arginine</td>
<td>0.30</td>
</tr>
<tr>
<td>potassium dihydrogen orthophosphate</td>
<td>2.0</td>
<td>L-Asparagine</td>
<td>0.20</td>
</tr>
<tr>
<td>sodium thioglycollate 7H2O</td>
<td>0.5</td>
<td>L-Aspartic acid</td>
<td>0.20</td>
</tr>
<tr>
<td>magnesium sulphate</td>
<td>0.15</td>
<td>L-Cysteine-HCl</td>
<td>0.20</td>
</tr>
<tr>
<td>manganese sulfate 4H2O</td>
<td>0.02</td>
<td>L-Glutamic acid</td>
<td>0.15</td>
</tr>
<tr>
<td>ferrous sulphate 7H2O</td>
<td>0.01</td>
<td>L-Glycine</td>
<td>0.30</td>
</tr>
<tr>
<td>Tween 80</td>
<td>0.001</td>
<td>L-Histidine-Hal</td>
<td>0.20</td>
</tr>
<tr>
<td>adenine</td>
<td>0.05</td>
<td>L-Isoleucine</td>
<td>0.20</td>
</tr>
<tr>
<td>cytidylic acid</td>
<td>0.05</td>
<td>L-Leucine</td>
<td>0.30</td>
</tr>
<tr>
<td>deoxyguanosine</td>
<td>0.05</td>
<td>L-Lysine-HCl</td>
<td>0.30</td>
</tr>
<tr>
<td>guanine HCl</td>
<td>0.05</td>
<td>L-Methionine</td>
<td>0.20</td>
</tr>
<tr>
<td>p-amineobenzoic acid</td>
<td>0.01</td>
<td>L-Phenylalanine</td>
<td>0.20</td>
</tr>
<tr>
<td>Vitamins B12</td>
<td>0.001</td>
<td>L-Proline</td>
<td>0.30</td>
</tr>
<tr>
<td>Calcium pantothenate</td>
<td>0.001</td>
<td>L-Serine</td>
<td>0.30</td>
</tr>
<tr>
<td>D-biotin</td>
<td>0.01</td>
<td>L-Threonine</td>
<td>0.20</td>
</tr>
<tr>
<td>Folic acid</td>
<td>0.001</td>
<td>L-Tryptophan</td>
<td>0.20</td>
</tr>
<tr>
<td>Niacin</td>
<td>0.001</td>
<td>L-Tyrosine</td>
<td>0.30</td>
</tr>
<tr>
<td>Piridoxal ethyl acetate HCl</td>
<td>0.001</td>
<td>L-Valine</td>
<td>0.30</td>
</tr>
<tr>
<td>Riboflavin</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thiamine HCl</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 2.1.4 Antibiotic stock solutions

The antibiotic stock solutions were prepared in pure H2O or ethanol (Table 2.2), filter sterilized after preparation, and then frozen in aliquots at -20 °C. Antibiotics were added to autoclaved and cooled (~50 °C) media to final concentrations.
### Table 2.2: Antibiotic stock solutions and final concentrations

<table>
<thead>
<tr>
<th>Antibiotic</th>
<th>Stock solution (mg/mL)</th>
<th>E. coli</th>
<th>Lb. plantarum</th>
<th>L. lactis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ampicillin (Na-salt)</td>
<td>100 in pure H₂O</td>
<td>100</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tetracycline (-Hydrochloride)</td>
<td>12.5 in ethanol (70 %)</td>
<td>12.5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kanamycin (-Sulfate)</td>
<td>40 in pure H₂O</td>
<td>15</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Chloramphenicol</td>
<td>34 in pure H₂O</td>
<td>30</td>
<td>5 or 10</td>
<td>5 or 10</td>
</tr>
<tr>
<td>Erythromycin</td>
<td>50 in ethanol (absolute)</td>
<td>200</td>
<td>5 or 10</td>
<td>5 or 10</td>
</tr>
<tr>
<td>Streptomycin</td>
<td>500 in pure H₂O</td>
<td>-</td>
<td>500</td>
<td>-</td>
</tr>
</tbody>
</table>

#### 2.1.5 Glycerol stocks

Cultures were stored as 20 % glycerol stocks at -80 °C. To prepare these stocks, cells were streaked onto agar plates (1.5 % agar) made up with the appropriate media and containing the necessary antibiotics. After growth overnight, a single colony was picked with a sterile toothpick and used to inoculate a 2 mL liquid culture, then incubated with shaking at 37 °C overnight. 1 mL aliquots were stored as a 20 % glycerol solution in sterile screw cap tubes (NUNC™ CyroTubes) at -80 °C. The conserved strains could be reactivated by scraping the stock and streaking on agar plates, or by inoculating liquid media, all of which contained the appropriate antibiotics.

#### 2.1.6 Cultivation and Harvesting of Cells

The cultivation of cells in liquid medium was carried out in flasks, with a ratio of volume of flask to liquid of 5:1. The main cultures were inoculated with 0.5 - 2 % (v/v) of an overnight culture and incubated at 37 °C on a rotation shaker Model G25 (New Brunswick Scientific, New Jersey, USA).

Harvesting of cells with a culture volume up to 2 mL was carried out by centrifugation in 1.5 mL tubes in a MiniSpin plus centrifuge (Eppendorf, Hamburg, GER) at 14,000 x g and 4 °C for 5 minutes. Cell cultures with a volume between 10 - 300 mL were harvested by centrifugation in a Sorvall RT7 (Kendro Laboratory Products GmbH, GER) for 15 minutes at
2,700 x g and 4 °C. Cultures with volumes above 300 mL were harvested in the Sorvall Evolution RC (Kendro Laboratory Products GmbH, GER) at 2,800 x g and 4 °C.

Cells from culture media components were washed in PBS to remove any culture media clinging to their surface. This step can be repeated multiple times if necessary.

**2.1.7 Measurement of Optical Density (OD) of Cultures**

The cell density of a culture was determined by measuring the optical density (OD) at a wavelength of 600 nm with a Smart Spec™ Plus Spectrophotometer (BioRad, Milan, I). Samples with an OD$_{600}$ of more than 0.8 were diluted with sterile culture media accordingly. Sterile culture media was used as reference.

**2.2 Methods for Deoxyribonucleic Acid (DNA) Work**

**2.2.1 Isolation of DNA**

All heat-stable solutions and equipment were autoclaved for 20 minutes at 121 °C and 2 x 10$^5$ Pa to deactivate any DNases present. Non autoclavable equipment was treated with 70 % (v/v) ethanol and then rinsed with sterile pure H$_2$O.

**2.2.2 Isolation of Genomic DNA from *L. plantarum KW30***

Genomic DNA from *L. plantarum* KW30 was isolated by a modified method using the Wizard Genomic DNA Purification Kit (Promega, Wisconsin, USA). *L. plantarum* KW30 was grown in 150 mL MRS media with 1.2 % glycine at 30 °C in static culture to an OD$_{600}$ of 1.0. The cells were harvested by centrifugation at 2,700 x g at 4 °C for 10 minutes. Cell pellets were frozen at -80 °C allowing for future use.
The cell pellet from 35 mL of culture was processed in the following way. The cells were thawed on ice and resuspended in 2 mL of a cold solution of 50 mM Na$_3$EDTA (~pH 7.8)/10 mM Tris/HCl (pH 8). The cell suspension was split into two equal aliquots in 1.5 mL tubes and 400 µL of 50 mg/mL lysozyme was added to each sample, mixed by inversion every 10 minutes and incubated for 50 minutes at 37 °C. Cells were separated from the liquid by centrifugation for 60 seconds at 14,000 x g, and the pellets drained well before being rapidly resuspended in 800 µL of room temperature ‘nuclei lysis solution’ (Wizard Genomic DNA purification kit, Promega). Complete lysis of the cells was ensured through three freeze-thaw-cycles; liquid nitrogen for 1 minute followed by immersion in an 80 °C water bath for 2 minutes. After cooling of the tubes to ~35 °C, 5 µL of 4 mg/mL RNase solution was added to each tube, mixed by inversion every 15 minutes and incubated at 37 °C for 45 minutes. Tubes were cooled to room temperature and 270 µL of room temperature protein precipitation solution (Wizard Genomic DNA Purification Kit, Promega) was added, mixed by inversion then incubated on ice for 5 minutes. Samples were centrifuged for 5 minutes at 14,000 x g and the supernatants transferred to fresh tubes. This centrifugation step was repeated until ~800 µL of supernatant free of insoluble material was collected. Isopropanol (620 µL) was layered on top of each sample and mixed by inversion until DNA threads began to form. The samples were given a sharp physical shock by dropping the tubes onto the bench to precipitate more DNA, and the DNA collected by centrifugation for 5 minutes at 14,000 x g. The DNA pellet was washed twice with 800 µL 70 % ethanol, then air dried and resuspended in 50 µL of TE-buffer.

**TE-buffer:**

- 10 mM Tris/HCl
- 1 mM EDTA
- pH 8.5

### 2.2.3 Isolation of Plasmid DNA from *E. coli*

Plasmid-harbouring *E. coli* strains were cultivated overnight at 37 °C in 5 mL LB medium with the appropriate antibiotics. The cells from 1.5 - 4 mL of this overnight culture were harvested in 1.5 mL tubes (5 minutes at 14,000 x g). Plasmids were isolated using the “High Pure Plasmid Isolation Kit” (Roche, Mannheim, GER) according to the manufacturer’s
instructions. The method is based on the adsorption of the plasmid DNA to silica particles then resuspension in an elution solution (e.g. TE-buffer) after several wash steps.

### 2.2.4 Isolation of plasmid DNA from L. plantarum

The method described by O'Sullivan & Kleanhammer (1993) was used to isolate plasmid DNA from *L. plantarum* KW30. An overnight culture of *L. plantarum* KW30, started with 1% inoculum, was harvested by centrifugation for 15 minutes at 2,700 x g and 4 °C. The pellet was resuspended in 200 µL 25% sucrose containing 40 mg/mL lysozyme then transferred to a 1.5 mL Eppendorf tube. After incubation at 37 °C for 30 minutes, 400 µL of alkaline SDS solution was added and mixed by inversion. This mixture was then incubated at room temperature for 7 minutes, and 300 µL of ice-cold 3 M potassium acetate (pH 4.7) was added, mixed by inversion and centrifuged for 15 minutes at 17,000 x g and 4 °C. The supernatant was transferred into a new 1.5 mL Eppendorf tube and 650 µL isopropanol (room temperature) was added, mixed by inversion before being centrifuged for 15 minutes at 17,000 x g and 4 °C. After all liquid was removed, the pellet was resuspended in 320 µL of pure H₂O and 200 µL of 7.5 M ammonium acetate containing 0.5 mg/mL ethidium bromide. DNA was isolated by the addition of 350 µL phenol/chloroform. This was mixed well and centrifuged for 5 minutes at 17,000 x g and room temperature. The upper phase was carefully transferred to a new 1.5 mL Eppendorf tube and 1 mL ethanol (-20 °C) was added, mixed by inversion and centrifuged at 17,000 x g and 4 °C to pellet the DNA. After washing in 70% ethanol at room temperature the pellet was isolated by centrifugation as above, air dried and resuspended in TE-RNase.

<table>
<thead>
<tr>
<th>Alkaline SDS solution:</th>
<th>3.0% SDS</th>
<th>0.2 N NaOH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TE-RNase:</th>
<th>10.0 mM Tris/HCl</th>
<th>1.0 mM EDTA</th>
<th>0.1 mg/mL RNase</th>
<th>pH 8.5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.2.5 **ISOLATION OF DNA-FRAGMENTS**

For purification, preparative isolation and concentration of DNA fragments (Vogelstein & Gillespie 1979) the fragment mixture was separated by electrophoresis in a 0.8 - 2.0 % (w/v) agarose gel in TAE buffer (Section 2.2.9). The gel was stained with ethidium bromide and the DNA fragments to be isolated were cut out of the gel under UV-light. The agarose strips were placed in 1.5 mL tubes and the DNA extracted using the “Perfectprep Gel Cleanup Kit” (Eppendorf, Hamburg, GER), according to the manufacturer's instructions. The method is based on the solubilisation of agarose at 50 °C in a high salt solution, and DNA adsorption to silica particles. The binding is specific for nucleic acids and the DNA remains bound to the silica while impurities are removed by several wash steps. The DNA is then eluted by a solution with low salt content (TE-buffer).

2.2.6 **ISOLATION OF PCR PRODUCTS**

To directly purify and isolate polymerase chain reaction (PCR) products (Vogelstein & Gillespie 1979) up to a size of 5 kbp from a PCR reaction mix the “High Pure PCR Product Purification Kit” (Roche, Mannheim, GER) was used according to the manufacturer's instructions. The method is based on the adsorption of the PCR-DNA fragments to silica particles, followed by several wash steps and then resuspending them in a solution of salt (TE-buffer).

2.2.7 **PURIFICATION AND CONCENTRATION OF DNA**

In order to inactivate any DNases, 1.5 mL Eppendorf tubes or plastic centrifuge tubes were used and all solutions used were autoclaved.

*(a) Phenol/Chloroform extraction*

In order to remove any contaminating protein from isolated DNA a phenol-chloroform extraction was carried out. An equal volume of phenol-chloroform-isoamylalcohol
(25:24:1 v/v/v) was added to the sample and mixed well by inversion. The separation of the phases was assisted by centrifugation 14,000 x g for 5 minutes. The upper aqueous phase containing the DNA was taken off without disturbing the protein layer between the phases. The aqueous layer was transferred into a new tube and the DNA was concentrated by isopropanol precipitation (Section 2.2.7(b)), dried and dissolved in TE-buffer.

(b) Precipitation of DNA

DNA precipitation with isopropanol was used to concentrate the DNA and to remove salts, small oligonucleotides and other contaminants from the DNA. To precipitate the DNA, 0.7 - 1.0 volume (v/v) of isopropanol was added, mixed by inversion and incubated for 15 minutes on ice. DNA was pelleted by centrifugation at 17,000 x g (Biofuge fresco, Heraeus, Thermo Scientific, USA) for 20 minutes. The pellet was washed twice with 70 % (v/v) ethanol, followed by centrifugation, then air dried at room temperature, resuspended in TE-buffer and stored at -20 °C.

2.2.8 DNA SEQUENCING

DNA sequencing was carried out on either an ABI Prism 377-64 sequencer or an ABI Prism 3730 capillary sequencer, using BIGDYE dye labelled dideoxy chain termination chemistry (Applied Biosystems). DNA sequencing was provided by the Massey University Allan Wilson Centre for Molecular Evolution and Ecology Genome Service.

2.2.9 AGAROSE GEL ELECTROPHORESIS

Agarose gel electrophoresis was carried out to separate DNA on the basis of fragment size. Typically, 1 % agarose mini-gels were used with Tris-Acetate-EDTA (TAE) buffer in horizontal electrophoresis chambers. The agarose percentage was varied from 0.8 - 2 % according to the purpose of the gel.
The agarose was dissolved in TAE buffer by heating in a microwave oven until it formed a homogeneous solution. After cooling down to 60 °C the agarose solution was poured in the electrophoresis sledge and a comb was put in to form the wells. The set gel was transferred with the sledge into the electrophoresis chamber, covered with 1 x TAE running buffer and the comb carefully removed. The samples of interest were mixed with 0.2 parts per volume of stop-mix, and then loaded into the wells of the agarose gel. Electrophoresis was carried out at 80 V for 40 minutes using a BioRad power pack (Power Pac 300, BioRad Laboratories, Milan, I). DNA was visualized by staining the agarose gels for 20 minutes in an ethidium bromide solution (0.5 µg/mL), followed by destaining for 5 - 10 minutes in pure H₂O. DNA was illuminated by exposure to UV-light (λ = 254 nm) and recorded using an UV-Trans-Illuminator (BioRad Gel Doc, BioRad Laboratories, Milan, I).

<table>
<thead>
<tr>
<th>TAE buffer (50x):</th>
<th>2.0  M</th>
<th>Tris/HCl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5.0   mM</td>
<td>EDTA</td>
</tr>
<tr>
<td></td>
<td>57.1 mL</td>
<td>Glacial acetic acid</td>
</tr>
<tr>
<td></td>
<td></td>
<td>pH 8.0</td>
</tr>
<tr>
<td></td>
<td>Pure H₂O</td>
<td>up to 1 L</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stop-mix (6x):</th>
<th>60.0 mM</th>
<th>Tris-HCl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>60.0 mM</td>
<td>EDTA</td>
</tr>
<tr>
<td></td>
<td>0.2 % (w/v)</td>
<td>Orange G</td>
</tr>
<tr>
<td></td>
<td>0.05 % (w/v)</td>
<td>Xylene Cyanol FF</td>
</tr>
<tr>
<td></td>
<td>60.0 % (v/v)</td>
<td>Glycerol</td>
</tr>
</tbody>
</table>

**2.2.10 DNA HYDROLYSIS WITH RESTRICTION ENDONUCLEASES**

The sequence-specific hydrolysis of nucleic acids was carried out using restriction endonucleases type II (Invitrogen, CA, USA; Roche, GER), whose recognition sites are palindromes.

For the reaction mix, 0.1 volumes of the appropriate 10 x concentrated buffer, which was supplied by the manufacturer, were added to the DNA (resuspended in TE-buffer). Between 2 and 10 U of restriction enzyme was added to 1 µg of DNA, followed by
incubation for 1 - 4 hours at the appropriate temperature for the particular enzyme. The reaction was stopped by isopropanol precipitation (Section 2.2.7(b)) or addition of 0.2 volumes of stop-mix.

2.2.11 LIGATION OF DNA FRAGMENTS

The in vitro recombination of vector DNA and insert DNA was carried out by ligation. The ATP-dependent T4-DNA Ligase (Invitrogen, USA) was used for the covalent combination of the 5'-phosphate- and 3'-hydroxyl-DNA fragments of compatible DNA-fragment-ends. The vector to insert DNA ratio ranged from 1:3 to 1:10. The ligation buffer was provided by the producer of the T4-DNA Ligase.

**Ligation conditions:**

- Vector DNA: 1.0 µg
- Insert DNA: 5.0 µg
- T4-DNA Ligase: 3.0 U
- T4-DNA Ligase buffer: 1.0 x
- Total volume: 10.0-20.0 µL

2.2.12 DETERMINATION OF DNA CONCENTRATION

The concentration and purity of DNA samples were determined using the Nano drop ND-1000 Spectrophotometer (Thermo Scientific). The optical density of 1 µL of DNA sample was measured at 260 and 280 nm. The software then calculated the DNA concentration in ng/µL based on the correlation of OD$_{260}$ = 1 corresponds to 50 µg/mL dsDNA (Davis et al. 1980; Sambrook et al. 1989). The OD$_{260}$/OD$_{280}$ ratio was indicative of the purity of the DNA and should be 1.8 - 2.0. A value less than 1.8 indicates protein contamination.
2.2.13 SIZE DETERMINATION OF DNA FRAGMENTS

Agarose gel electrophoresis was used to determine the size of DNA fragments. The DNA fragments were separated along with a standard DNA ladder (1 kbp+ DNA Ladder, Invitrogen, appendix 13).

2.2.14 TRANSFORMATION OF DNA INTO CELLS

The transfer of DNA into E. coli cells was carried out either by heat shock or electroporation. Electroporation was the standard method used to transfer DNA into Lactobacillus and Lactococcus species.

(a) Preparation of competent E. coli cells

The method described by Hanahan (1983) was used to prepare competent E. coli cells. The appropriate E. coli strain was grown in LB medium, containing the necessary antibiotics, at 37 °C up to an OD600 of 0.3. The cell cultures were then incubated on ice for 10-15 minutes after which the cells were harvested by centrifugation for 15 minutes at 2,700 x g and 4 °C. The sedimented cells were resuspended in 18 mL of RF1 solution, incubated on ice for 30 minutes, and then pelleted by centrifugation for 15 minutes at 2,700 x g and 4 °C. Subsequently, the sedimented cells were resuspended in 4 mL RF2 solution and the cell suspension distributed in 100 µL aliquots in Eppendorf tubes and quickly frozen in liquid nitrogen at -70 °C for long-term storage.

RF1 solution:  
100.0 mM RbCl  
50.0 mM MnCl₂  
30.0 mM Potassium acetate  
10.0 mM CaCl₂  

pH 5.8 (conc. Acetic Acid)
RF2 solution:
10.0 mM RbCl
10.0 mM MOPS
75.0 mM CaCl$_2$ x 6 H$_2$O
15.0 % (v/v) Glycerin
pH 5.8 (NaOH)

(b) Transformation and selection of E. coli cells

Plasmids were introduced into E. coli cells using a heat shock protocol. A thawed aliquot of 100 µL of competent E. coli cells was mixed well with 50 - 250 ng of DNA. This mixture was incubated on ice for 30 minutes in order to let the DNA adsorb to the cell surface. The transformation mix was then heated for 1 minute at 42 °C to allow the uptake of DNA by the cells. After incubation on ice for 5 minutes, 600 µL of LB medium was added to the transformation mix and the cells incubated at 37 °C for 1 hour to allow the development of the plasmid-borne antibiotic resistances. Transformants were then plated on selective media and incubated overnight at 37 °C to isolate recombinant clones. A single transformant colony was picked for further work, such as growth of a liquid culture for glycerol stocks (Section 2.1.5) or plasmid preparation (Section 2.2.3).

(c) Preparation of electrocompetent E. coli cells

The appropriate E. coli strain was grown overnight at 37 °C in 50 mL LB medium containing the necessary antibiotics. This 50 mL overnight culture was used to inoculate 1 L of LB containing the necessary antibiotics medium in a 5 L flask and grown with shaking at 37 °C until the OD$_{600}$ was between 0.5 - 0.6. This culture was then transferred into two chilled, sterile 500 mL centrifuge bottles and incubated on ice for 30 minutes. The cultures were centrifuged at 2,000 x g for 15 minutes at 4 °C, the supernatant discarded and the bottles placed back on ice. Each cell pellet was then resuspended in 500 mL of cold sterile water and the cells centrifuged again at 2,000 x g for 15 minutes and 4 °C, after which the water was decanted and the bottles once again placed on ice. This process was repeated using 250 mL of cold sterile water. Each cell pellet was resuspended in 20 mL
cold sterile 10 % glycerol using a pre-chilled, sterile 25 mL pipette and transferred to two chilled, sterile 50 mL centrifuge tubes. After centrifugation at 4,000 x g for 15 minutes at 4 °C, the 10 % glycerol was decanted and the bottles placed back on ice. Each cell pellet was now resuspended in 1 mL cold sterile 10 % glycerol and pooled in one of the 50 mL centrifuge tubes. 55 µL aliquots of the cell suspension were placed in 1.5 mL Eppendorf tubes, then frozen in liquid nitrogen and stored at -80 °C.

(d) Electroporation of E. coli cells

The appropriate number of tubes of electrocompetent cells were removed from the -80 °C freezer and thawed on ice. Electroporation cuvettes were also chilled on ice. 1 µL of ligation reaction was added to a tube containing 55 µL competent cells. For the control reactions 1 µL of plasmid DNA was added to a separate tube of 55 µL competent cells. All tubes were incubated for 1 - 2 minutes on ice. The electroporation device (Micro Pulser, BioRad, Milan, I) was set up for electroporation of bacteria. One sample at a time was transferred to an electroporation cuvette and the cuvette placed in the chamber and an electric pulse was discharged. The cuvette was removed and immediately 800 µL room temperature SOC medium was added and transferred to a sterile 1.5 mL tube. The tube was placed on ice and the procedure repeated for all samples. The tubes were incubated with shaking at 37 °C for 60 minutes. Then, 25 µL or 100 µL of the transformation mix were plated on LB plates containing the appropriate antibiotics and incubated at 37 °C overnight.

SOB (per litre):

| 2.0 %  | tryptone            |
| 0.5 %  | yeast extract       |
| 0.05 % | NaCl                |
| 2.5 mM | KCl                 |
| 10.0 mM| MgCl₂               |

1. Dissolve tryptone, yeast extract and NaCl in 950 mL H₂O.
2. Make 250 mM KCl solution by dissolving 1.86 g of KCl in 100 mL H₂O. Add 10 mL of this stock KCl solution to solution in Step 1.
3. Adjust pH to 7.0 with 5 M NaOH, bring volume to 1 L with H₂O.
4. Autoclave, cool to ~55 °C and add 10 mL of sterile 1 M MgCl₂.

SOC (per litre):

SOB
20.0 mM glucose
1. After making SOB medium, add 7.2 mL of 50 % glucose.

(e) Production of electrocompetent cells of Lactobacillus and Lactococcus species

The method described by Kaneko (2000) was used to prepare competent Lactobacillus and Lactococcus cells. An overnight culture of L. plantarum KW30 was used to inoculate fresh 50 mL MRS media containing 1 % glycine to obtain an OD₆₀₀ of about 0.05. The cells were grown at 30 °C for 4 hours until an OD₆₀₀ of 0.3 - 0.4 was reached, then harvested by centrifugation for 15 minutes at 2,700 x g, 4 °C. After washing the cells three times in sterile H₂O, they were resuspended in 0.5 mL 40 % PEG 1000. The cell suspension was divided into 40 µL aliquots and stored at -80 °C for several months.

(f) Electroporation of Lactobacillus and Lactococcus species

A thawed 40 µL aliquot of cells was mixed with 1 - 2 µL (less than 0.1 µg) DNA and transferred to a cold electroporation cuvette (2 mm inter-electrode distance). The following electroporation conditions (Kaneko et al. 2000) were used: electric field strength, 10 kV/cm; capacitance, 25 µF; resistance, 400 Ω. After the pulse, the cuvette was immediately placed on ice, 360 µL MRS media was added to the cuvette, which was incubated at 30 °C for 2 hours. 25, 50 and 100 µL of the cell suspension were plated on MRS agar containing the appropriate antibiotics and incubated at 30 °C for up to 3 days.
(g) Protoplast formation of *L. plantarum* KW30

Protoplast formation was carried out using the method described by Morelli *et al.* (1987). A 10 % inoculum of a mid-log phase culture of *L. plantarum* KW30 was added to 40 mL of prewarmed MRS broth (Section 2.1.3(b)). The cells were grown for 2 hours at 30 °C and then chilled on ice, harvested by centrifugation at 4,400 × g for 10 minutes and washed with sterile saline. The cell pellet was resuspended in 4 mL of protoplast buffer containing 20 µg/mL mutanolysin. After 40 minutes at 37 °C the protoplasts were recovered by centrifugation at 3,500 × g for 10 minutes, washed 3 x in PB and resuspended in 400 µL of the same buffer. The protoplasted cells were confirmed by visual analysis under a microscope.

Protoplast buffer (PB):

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris/HCl (pH 8)</td>
<td>10 mM</td>
</tr>
<tr>
<td>CaCl₂</td>
<td>20 mM</td>
</tr>
<tr>
<td>Sucrose</td>
<td>0.5 M</td>
</tr>
</tbody>
</table>

(h) Protoplast transformation

In order to transform (Morelli *et al.* 1987) the prepared protoplasts, 1 µL plasmid (38 ng/µL; pNZ5319_plnKO 13) and 0.9 mL of 20 % PEG 6000 were added to 100 µL protoplast suspension and incubated for 6 minutes at RT. Then, the cells were centrifuged and washed with 1 mL protoplast buffer (Section 2.2.14(g)), followed by 2 hours incubation at 30 °C in 1 mL MRS, 0.5 M sucrose, 20 mM MgCl₂. After the incubation, the cells were centrifuged again and resuspended in 50 µL protoplast buffer and then plated on RM agar.

Regeneration medium (RM):

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRS</td>
<td>0.5 M</td>
</tr>
<tr>
<td>Sucrose</td>
<td>20 mM</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>2.5 %</td>
</tr>
<tr>
<td>Gelatine</td>
<td>0.5 %</td>
</tr>
<tr>
<td>Heat-inactivated BSA</td>
<td></td>
</tr>
</tbody>
</table>
2.2.15 Polymerase chain reaction (PCR)

(a) Polymerase chain reaction (PCR)

The basic components of a PCR reaction (Innis et al. 1990) are: template DNA, specific primers (oligonucleotides), thermostable DNA polymerase and dNTPs (deoxynucleotide triphosphates). These components were mixed in a PCR tube and placed in a PCR cycler T gradient (Biometra), where it underwent repeated cycles of denaturation (at 94 °C), annealing (at 45-65 °C) and elongation (at 68 °C or 72 °C). This led to the amplification of the desired DNA fragment.

The primers were specifically synthesised for each DNA fragment (Sigma-Aldrich, Australia or Integrated DNA Technologies, USA). Primers are identical or very similar to the desired DNA fragment (ca. 30 bp) and also contain recognition sites for restriction endonucleases at their 5'-end, which could be detected and hydrolysed by the appropriate enzyme. Genomic DNA of L. plantarum KW30 or plasmid DNA was used as template DNA. For the elongation the Taq DNA Polymerase (5 U/µL, Roche, GER) or KOD DNA Polymerase (5 U/µL, Novagen, GER) were used.

(b) Colony PCR

Cell colonies were grown on agar plates containing the appropriate antibiotics. E. coli cells were removed from the plate by touching the surface of each colony using a sterile pipette tip. The pipette tip was then placed into a PCR tube containing prepared PCR mix (dNTPs, buffer, primer and Taq DNA Polymerase made as recommended by the manufacturer), shaken several times to dislodge the bacteria and removed. The PCR reaction (Section 2.2.15(a)) was performed, as usual, with an initial 5 minutes denaturation step to lyse the cells. Products were visualized by agarose gel electrophoresis (Section 2.2.9).
2.2.16 Targeted gene disruption

(a) Targeted gene disruption via homologous recombination in lactobacilli

The site directed mutagenesis method described by Russell & Klaenhammer (2001) is based on a homologous recombination strategy and uses two plasmids. The helper plasmid pTRK669 (ori (pWV01), Cmr, provides repA in trans) is temperature sensitive and becomes unstable above 43 °C. The mutagenesis vector pORI28 (Emr, ori (pWV01)) replicates only with repA in trans, which is provided by the helper plasmid.

The RR (formerly FTase) mutagenesis vector was constructed by cloning an internal fragment of the target gene into the BamHI/EcoRI restriction sites. In the case of the RR a ~400 bp gene fragment was cloned into pORI28. For the cloning, the primers KW30FTaseKOBamF and KW30FTaseKOEcoR (Appendix 4) were used. The correct cloning of the pORI28::RRKO mutagenesis vector was verified by sequencing.

The mutagenesis vector pORI28::RRKO was transformed by electroporation (Section 2.2.14(f)) into L. plantarum KW30 which already contained pTRK669. One transformant, carrying both plasmids, was selected and grown overnight at 37 °C in MRS broth with 5 µg/mL chloramphenicol (Cm) and 10 µg/mL erythromycin (Em). Then, the culture was transferred three times at 43 °C (1% inoculum) in MRS broth plus 10 µg/mL Em. It was necessary to use 10 µg/mL Em, because wild type L. plantarum KW30 is naturally able to grow on 5 µg/mL Em.

As cultures of cells carrying both plasmids were not able to grow at 43 °C or showed only minimal growth it was not possible to continue. However, had the cell cultures grown, they would have been replica plated on MRS plates plus either Em, Cm, or no antibiotic, and incubated for 48 hours at 37 °C. This would have been followed by choosing one Em resistant and Cm sensitive clone for further confirmation by sequencing.
(b) Targeted gene disruption via cre-lox-based homologous recombination system in lactobacilli

The Cre-lox-based system for multiple gene deletions is based on the classic double-crossover strategy (Lambert et al. 2007). Double-crossover gene replacement mutants can be selected based on their antibiotic resistance phenotype. The chloramphenicol resistance cassette in the mutants is flanked by a \textit{lox66-P32-cat-lox71} cassette, which can be removed by the site-specific Cre recombinase. This results in an antibiotic resistance cassette free mutant, where further deletions can be carried out using the same Cre-lox-based system.

The \textit{plnKW30}-specific mutagenesis vector was constructed using a 1 kb-fragment of the upstream sequence and a 1 kb-fragment of the downstream sequence of the \textit{plnKW30} gene. The fragments were amplified by PCR using KOD DNA polymerase (Novagen, GER) and specific primers (cre-lox mutagenesis primer; appendix 4), then cloned into the \textit{Pmel} and \textit{Ecl}136II blunt-end restriction sites of pNZ5319. The primers were designed so that the 5'- and 3'-flanking regions of the \textit{plnKW30} gene encompassed the first and last five codons of this gene. The orientation of the fragments was verified by colony PCR (Section 2.2.15(b)) using a vector-specific primer combined with an insert-specific primer (cre-lox verifying primer; appendix 4).

The \textit{plnKW30}-mutagenesis vector was transferred into \textit{L. plantarum} KW30 by electroporation as described in section 2.2.14(f). Chloramphenicol-resistant (Cm\textsuperscript{r}) transformants were selected and replica plated to check for an erythromycin-sensitive (Em\textsuperscript{s}) phenotype. After a double-crossover the chloramphenicol-resistance gene is located in the genome, whereas the erythromycin resistance gene is lost with the rest of the mutagenesis vector that does not integrate into the genome. This process results in chloramphenicol resistant and erythromycin sensitive clones.

At this stage no further experiments were carried out. However, had more positive results been obtained the next step would have been to analyse candidate double-crossover clones (Cm\textsuperscript{r} Em\textsuperscript{s}) by PCR amplification using chloramphenicol (\textit{cat}) and erythromycin (\textit{ery}) specific gene primers. This would have been followed by confirming the correct integration.
of the \( \text{lox66-P}_{32}\text{-cat-lox71} \) cassette into the genome by PCR analysis using primers annealing to genomic sequences flanking the \( \text{lox66-P}_{32}\text{-cat-lox71} \) cassette.

### 2.3 METHODS FOR RIBONUCLEIC ACID (RNA) WORK

#### 2.3.1 ISOLATION OF RNA

\( L. \text{plantarum KW30} \) was cultivated overnight (about 16 hours) at 30 °C in 1 mL MRS media. This overnight culture was harvested for 5 minutes at 17,000 x g (Biofuge fresco centrifuge). The cell pellet was resuspended in 100 µL TE-buffer containing 40 mg/mL Lysozyme and 100 U/mL Mutanolysin and incubated at 37 °C for 1 hour. The isolation of total RNA was carried out following the instructions of the “Illustra RNAspin mini RNA isolation kit” (GE Healthcare, UK). DNA digestion was performed twice for 30 minutes.

To isolate larger quantities of RNA \( L. \text{plantarum KW30} \) was grown at 30 °C for 6 hours or overnight (about 16 hours) in 10 mL MRS media. The cells were pelleted by centrifugation 15 minutes at 2,700 x g (Sorvall RT centrifuge; Kendro Laboratory Products) and resuspended in 100 µL TE-buffer containing 40 mg/mL Lysozyme and 100 U/mL Mutanolysin and incubated at 37 °C for 30 minutes. Lysis of the cells was achieved using glass beads and a bead mill (Fast prep cell disruptor, Thermo Savant, Qbiogene, USA) for 30 seconds at the highest setting, prior to addition of 1 mL TRlzol\textsuperscript{®} Reagent (Invitrogen, USA). The isolation of total RNA was carried out following the manufacturer's instructions. The removal of contaminating DNA was performed using the TURBO DNA-\textit{free}\textsuperscript{TM} Kit (Ambion, USA) according to the instructions of the manufacturer.

#### 2.3.2 AGAROSE/FORMALDEHYDE GEL ELECTROPHORESIS OF RNA

The RNA formaldehyde gel was prepared following the Current Protocols in Molecular Biology (Brown \textit{et al.} 2004). To prepare a 1.0 % gel, 1.0 g agarose was dissolved in 72 mL water and cooled to 60 °C in a water bath. When the flask had cooled to 60 °C, it was placed
in a fume hood and 10 mL 10x MOPS buffer and 18 mL 12.3 M formaldehyde (37 %) was added. The gel was then poured and allowed to set. After removing the comb, the gel was placed in the gel tank and sufficient 1x MOPS running buffer added to cover the gel.

The volume of each RNA sample was adjusted with H$_2$O to 11 µL, then 5 µL 10x MOPS buffer, 9 µL 12.3 M formaldehyde and 25 µL formamide was added. The samples were mixed by vortexing, centrifuged briefly and incubated for 15 minutes at 55 °C. Then, 10 µL formaldehyde loading buffer was added, mixed and the entire sample loaded onto the gel. The gel was run at 5 V/cm until the bromophenol blue dye had migrated one-half to two-thirds the length of the gel (~3 hours).

Before staining the gel the formaldehyde was removed by soaking the gel in sufficient 0.5 M ammonium acetate to cover it for 20 minutes. The solution was changed and the gel soaked in fresh ammonium acetate for an additional 20 minutes. In order to stain the gel the solution was poured off and replaced with 0.5 µg/mL ethidium bromide in 0.5 M ammonium acetate and allowed to stain overnight. The gel was examined on a UV transilluminator to visualize the RNA and photographed with a ruler laid alongside the gel, so that band positions could later be identified.

**10x MOPS buffer:**

- MOPS 0.4 M, pH 7.0
- Sodium acetate 0.1 M
- EDTA 0.01 M
- Store up to 3 months at 4°C

**Sample loading buffer:**

- bromophenol blue 0.25 % (w/v)
- Glycerol 50 %
- in 1 x MOPS buffer

### 2.3.3 NORTHERN BLOT

Total RNA separated by agarose/formaldehyde gel electrophoresis can be transferred to a membrane and certain mRNAs can be detected using specific, radioactively-labelled DNA probes.
(a) Preparation of $^{32}$P-labelled probe

The preparation of probes was performed following the instructions of the ‘Random primers DNA labelling system’ (Invitrogen, USA) using $\alpha ^{32}$P-deoxycytidine 5'-triphosphate (Perkin Elmer). A PCR product of gccA (PlnKW30) was used as template. The Klenow fragment generates probes from 200 – 1000 nucleotides in length.

The probes were purified from unincorporated labelled nucleotides using ‘ProbeQuant™ G-50 micro columns’ (Pharmacia Biotech, Sweden) following the instructions of the manufacturer.

The activity of the probes was measured using the Liquid scintillation & luminescence counter 1450 Micro Beta, Trilux, Wallac (Perkin Elmer, MA, USA). A 1 µL sample of different dilutions of each probe was counted for 1 minute (cpm/µL).

(b) Northern blotting and hybridisation

The transfer of the RNA from gel to membrane, followed by hybridisation with a radioactively-labelled DNA probe and washing of the membrane were performed following the instructions of ‘Current Protocols in Molecular Biology’ (Unit 4.9: ‘Analysis of RNA by Northern and Slot Blot Hybridisation’, Brown et al. 2004).

2.3.4 Reverse Transcriptase (RT)-PCR

RT-PCR was carried out using the ‘SuperScript™ One-Step RT-PCR with Platinum® Taq kit’ (Invitrogen, USA) or the Tth RT-polymerase (Roche, Mannheim, GER) following the manufacturer’s instructions.
2.4 PROTEIN BIOCHEMICAL METHODS

2.4.1 DETERMINATION OF PROTEIN CONCENTRATION

The concentration of protein samples was determined by Bradford assay (Bradford 1976) or UV absorption (Scopes 1974).

(a) Bradford method for determination of protein concentration

From a BSA stock solution (2 mg/mL BSA) a range of standards were prepared: 1.6, 1.2, 1.0, 0.8, 0.6, 0.4, 0.2, 0.1 mg/µL. 5 µL of each standard and dilutions of the unknown sample were added to separate wells in a 96-well plate. 100 µL of Bradford reagent was added to each well and mixed thoroughly. The absorbance was read at 595 nm and a standard curve was drawn with which the protein concentration of the unknown sample was determined.

Bradford reagent (5x): Coomassie brilliant blue G-250 100 mg
Ethanol (95 %) 50 mL
Phosphoric acid (concentrated) 100 mL
Pure H₂O up to 200 mL

(b) Concentration determination by UV absorption

At 280 nm proteins absorb mainly due to the presence of tyrosine and tryptophan residues. Since the quantity of these two amino acids varies enormously from one protein to another, the absorption at 280 nm can only give a rough indication of the actual protein concentration of the sample. The absorption of the peptide bond is measured at 205 nm and gives an approximate extinction coefficient of 31 for a 1 mg/mL protein solution (error rate is ± 10 %). This value can vary from protein to protein because aromatic and some other residues have some absorption in this range. Secondary structures also have some influence on shape and position of the peptide absorption peak. The correction for the tyrosine and tryptophan content, which are the main contributors to the side-chain absorbance at this wavelength, is made by also measuring the absorbance at 280 nm. The extinction coefficient can be predicted with 2 % error. The following formula is used:
Commonly used buffers contain salts which absorb at 205 nm; the only exceptions are the anions sulphate and perchlorate. A weak phosphate buffer (5 mM, pH 7.0) in the presence of 50 mM sodium sulphate can be used.

### 2.4.2 SDS-POLYACRYLAMIDE GEL ELECTROPHORESIS

Proteins were generally separated on the basis of mass by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). The standard SDS-polyacrylamide gel electrophoresis used the discontinuous buffer method described by Laemmli (1970). The gels were cast and run in a vertical apparatus (Mini-Protean II system, BioRad, Milan, I). Typically, the electrophoresis was performed at room temperature at 200 V for 40 minutes or until the dye front reached the bottom of the gel.

For the separation of smaller peptides the tricine SDS-PAGE method was used, which was developed by Schägger & Jagow (1987). The use of tricine allows a better resolution of small proteins at lower acrylamide concentrations than the glycine SDS-PAGE systems. For these gels, electrophoresis was performed at room temperature. Initially, 30 V was applied across the electrodes, but after 30 minutes the voltage was increased to 100 V and the electrophoresis continued for approximately 1.5 hours or until the dye front reached the bottom of the gel.

### 2.4.3 DETECTION OF PROTEINS IN SDS-POLYACRYLAMIDE GELS

Proteins separated by SDS-PAGE were visualised using different staining methods, e.g. Coomassie blue or silver staining.
(a) Coomassie blue staining

This staining method has a sensitivity range of 300 – 1,000 ng protein per band. The gel is placed in a square petri dish, covered with Coomassie blue staining solution (Table 2.3), and incubated at room temperature with gentle agitation for 30 minutes. Background staining was removed by incubation in destaining solution (Table 2.3) for 1 - 3 hours, depending on the level of background observed.

Table 2.3: Coomassie blue staining and destaining solutions.

<table>
<thead>
<tr>
<th>Amount</th>
<th>Staining solution</th>
<th>Destaining solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1 %</td>
<td>Coomassie blue R-250 (G-250)</td>
<td>-</td>
</tr>
<tr>
<td>40 %</td>
<td>Methanol</td>
<td>Methanol</td>
</tr>
<tr>
<td>10 %</td>
<td>Acetic acid</td>
<td>Acetic acid</td>
</tr>
</tbody>
</table>

(b) Silver staining

Silver staining is a more sensitive staining method than Coomassie Blue staining, with a protein detection limit of 1 – 10 ng per band. Most of the required solutions like the fixing, sensitizing, developing (as 6 % stock solution) and stopping solutions, can be prepared in advance and stored long term. However, the silver nitrate and the 3 % developer (1:1 v/v dilution in H₂O) solutions have to be made up fresh each time.

After separating the proteins by SDS-PAGE (Section 2.4.2) they were fixed in the gel with the fixing solution for 30 minutes. This was followed by 30 minutes incubation in the sensitizing solution, which was washed off thoroughly 3 x 5 minutes with pure H₂O. The silver reaction was carried out for 20 minutes with freshly made up silver nitrate solution, followed by another wash step of 2 x 1 minutes in pure H₂O. The developing takes 2-5 minutes, which depended upon the amount of protein in the gel. At first, the gel was washed with a small amount of developing solution, after discarding that, the rest of the solution is added and the gel incubated with shaking until the desired intensity is reached. The reaction was stopped by incubation for 10 minutes in stopping solution. The gels can be stored in a glycerol/pure H₂O mixture.
2.4 Protein biochemical methods

2.4.4 2D-ELECTROPHORESIS

Two-dimensional (2D) electrophoresis is used to separate proteins in two different dimensions, which results in better separation than in 1D electrophoresis. The first dimension is separation by isoelectric point and the second dimension is separation by the mass of the proteins.

Fixing solution:  
EtOH (40%) 100.0 mL  
Glacial acetic acid (10%) 25.0 mL  
Pure H₂O made up to 200.0 mL

Sensitizing solution:  
EtOH (40%) 75.0 mL  
Sodium thiosulphate 0.5 g  
Sodium acetate 17.0 g  
Pure H₂O made up to 250.0 mL

Silver nitrate solution:  
Silver nitrate 0.4 g  
Formaldehyde (37% w/v) 16.0 µL  
Pure H₂O made up to 40.0 mL

Developing solution (6% stock solution):  
Sodium carbonate x 10 H₂O 6.25 g  
Formaldehyde (37% w/v) 100.0 µL  
Pure H₂O made up to 250.0 mL

Stopping solution:  
EDTA-Na₂ x 2 H₂O 3.65 g  
Pure H₂O made up to 250.0 mL
First dimension: Isoelectric focusing

The cover foil from the Immobiline Dry Strip (GE Healthcare, UK) has to be removed carefully and then the Dry Strip placed into the strip holder with the gel side down. To rehydrate the Immobiline Dry Strips, 125 µL of urea rehydration solution with freshly added DTT and 50 µL of the protein sample (1.4 mg whole cell extract of *L. plantarum* KW30 in sample preparation solution) were distributed evenly into the channel of a ceramic strip holder (GE Healthcare, UK). The solution should be distributed evenly under the strip, without trapping air bubbles. The Dry Strip is then overlaid with Immobiline Dry Strip Cover Fluid and the cover placed on the strip holder. An active rehydration was performed at 30 V for 10 hours. The first dimension, isoelectric focussing, was carried out on an Ettan IPGphor II Isoelectric Focusing Unit (GE Healthcare, UK) using the following program.

1. step & hold (active rehydration) 30 V 10 hours
2. step & hold 300 V 30 minutes
3. gradient 1000 V 30 minutes
4. gradient 5000 V 1:30 hours
5. step & hold 5000 V 30 minutes 13 hours

Reduction and alkylation of focused strips

The strips are placed in individual 15 mL Falcon tubes with the support film towards the wall. 5 mL SDS equilibration buffer solution is needed per tube and strip. First, the strips are equilibrated with gentle shaking for 15 minutes in 5 mL SDS equilibration buffer containing 50 mg DTT. This buffer is poured off, then equilibrated with 5 mL SDS equilibration buffer containing 125 mg Iodoacetamide for 15 minutes with gentle shaking. After pouring off the buffer the strips are washed with SDS running buffer.
Second dimension: SDS-PAGE

The Immobiline Dry Strip is positioned with the gel surface up onto the protruding edge of the longer glass plate of the SDS gel. It is gently pushed down, using a thin plastic ruler, so that the gel of the Dry Strip is in contact with the top of the SDS gel, taking care that all air bubbles are excluded. A molecular weight marker is applied in the left hand corner by mixing 5 µL of marker with 5 µL of hot 1 % agarose solution. The Dry Strip gel is sealed in place with agarose sealing solution and the gel run at 100 V for approximately 85 minutes. When the dye front reaches the bottom of the gel, it is removed from the apparatus; the proteins fixed in the SDS gel by soaking the gel in fixing solution for 30 minutes and then visualized using Colloidal Coomassie Blue G-250 stain or silver stain (Section 2.4.3).

Sample preparation solution:

<table>
<thead>
<tr>
<th>Final concentration</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urea (FW 60.06)</td>
<td>8 M</td>
</tr>
<tr>
<td>CHAPS</td>
<td>4% (w/v)</td>
</tr>
<tr>
<td>IPG Buffer§</td>
<td>2% (v/v)</td>
</tr>
<tr>
<td>DTT (FW 154.2)</td>
<td>40 mM</td>
</tr>
<tr>
<td>Double-distilled water</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Final concentration</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urea (FW 60.06)</td>
<td>8 M</td>
</tr>
<tr>
<td>CHAPS</td>
<td>4% (w/v)</td>
</tr>
<tr>
<td>IPG Buffer (same range as the IPG strip)</td>
<td>2% (v/v)</td>
</tr>
<tr>
<td>1% Bromophenol blue stock solution</td>
<td>0.002%</td>
</tr>
<tr>
<td>Double-distilled water</td>
<td>—</td>
</tr>
</tbody>
</table>

§ Use IPG Buffer in the pH range corresponding to the pH range of the IEF separation to be performed.

Rehydration stock solution:

<table>
<thead>
<tr>
<th>Final concentration</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urea (FW 60.06)</td>
<td>8 M</td>
</tr>
<tr>
<td>CHAPS</td>
<td>2% (w/v)</td>
</tr>
<tr>
<td>IPG Buffer (same range as the IPG strip)</td>
<td>0.5 % (v/v)</td>
</tr>
<tr>
<td>1% Bromophenol blue stock solution</td>
<td>0.002%</td>
</tr>
<tr>
<td>Double-distilled water</td>
<td>—</td>
</tr>
</tbody>
</table>

* DTT is added just prior to use: 7 mg DTT per 2.5-mL aliquot of rehydration stock solution. For rehydration loading, sample is also added to the aliquot of rehydration solution just prior to use.
Chapter 2: Materials and methods

2.4 Protein biochemical methods

**SDS equilibration buffer:**

<table>
<thead>
<tr>
<th>Final concentration</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urea (FW 60.06) 6 M</td>
<td>72.1 g</td>
</tr>
<tr>
<td>Tris-HCl, pH 8.8 75 mM</td>
<td>10.0 mL</td>
</tr>
<tr>
<td>Glycerol (87% w/w) 29.3% (v/v)</td>
<td>69 mL (84.2 g)</td>
</tr>
<tr>
<td>SDS (FW 288.38) 2% (w/v)</td>
<td>4.0 g</td>
</tr>
<tr>
<td>1% Bromophenol blue stock solution</td>
<td>400 µL</td>
</tr>
<tr>
<td>Double-distilled water</td>
<td>to 200 mL</td>
</tr>
</tbody>
</table>

* Just prior to use, DTT or iodoacetamide (for first or second equilibration, respectively) was added.

**1 % bromophenol blue stock solution:**

<table>
<thead>
<tr>
<th>Final concentration</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bromophenol blue 1%</td>
<td>100 mg</td>
</tr>
<tr>
<td>Tris-base 50 mM</td>
<td>60 mg</td>
</tr>
<tr>
<td>Double-distilled water</td>
<td>to 10 mL</td>
</tr>
</tbody>
</table>

**1 % agarose sealing solution** (25 mM Tris base, 192 mM glycine, 0.1 % SDS, 0.5 % agarose, 0.002 % bromophenol blue, 100 mL):

<table>
<thead>
<tr>
<th>Final concentration</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laemmli SDS electrophoresis buffer</td>
<td>100 mL</td>
</tr>
<tr>
<td>Agarose 0.5%</td>
<td>0.5 g</td>
</tr>
<tr>
<td>1% Bromophenol blue stock solution 0.002% (w/v)</td>
<td>200 µL</td>
</tr>
</tbody>
</table>

Ingredients were added to a 500-mL Erlenmeyer flask, and swirled to disperse. The agarose was dissolved by heating in a microwave oven on low and stored as 1.5-mL aliquots at room temperature.

**Fixing solution:**

- 10 % acetic acid
- 40 % ethanol


65
**5 % Coomassie Blue G-250 stock:**

(modified from Neuhoff et al. 1988)

\[
0.5 \text{ g} \quad \text{Coomassie Blue G-250} \\
\text{to } 10.0 \text{ mL} \quad \text{Double-distilled water}
\]

The solution was stirred for a few minutes to disperse the Coomassie Blue G-250 (The dye does not dissolve completely).

**Colloidal Coomassie Blue G-250 dye stock solution:**

(10 % ammonium sulphate, 1 % (w/w) phosphoric acid, 0.1 % Coomassie Blue G-250, 500 mL)

\[
50 \text{ g} \quad \text{Ammonium sulphate (FW 132.1)} \\
6 \text{ mL} \quad \text{Phosphoric acid 85\% (w/w)} \\
10 \text{ mL} \quad 5 \% \text{Coomassie Blue G-250 stock} \\
\text{to } 500 \text{ mL} \quad \text{Double-distilled water}
\]

**Colloidal Coomassie Blue G-250 working solution:**

(8 % ammonium sulphate, 0.8 % phosphoric acid, 0.08 % Coomassie Blue G-250, 20 % methanol, 500 mL)

\[
400 \text{ mL} \quad \text{Colloidal Coomassie Blue G-250 dye stock solution} \\
100 \text{ mL} \quad \text{Methanol}
\]

This solution was always prepared fresh, immediately before staining the gel.

**2.4.5 IN-GEL TRYPTIC DIGEST FOR PROTEIN IDENTIFICATION BY MASS SPECTROMETRY**

SDS-PAGE or 2D gels were run as described in sections 2.4.2 and 2.4.4, respectively. The bands or spots of interest were excised using a clean scalpel, cut into 1 mm pieces and transferred to an Eppendorf tube. They were then incubated in 50 % acetonitrile (ACN) in 200 mM NH\(_4\)HCO\(_3\) for 4 hours to remove the stain. After this the destaining solution was removed using a pipette and the gel pieces dried under vacuum (SpeedVac Concentrator, Savant, USA) for 10 minutes. At this stage they could be stored at -20 °C until required for further processing.
The gel particles were rehydrated and reduced by adding 150 µL 10 mM DTT in 100 mM NH₄HCO₃ after which they were incubated for 1 hour at 56 °C. The tubes were then cooled to RT and the DTT solution replaced with 150 µL 55 mM Iodoacetamide in 100 mM NH₄HCO₃ then incubated for 45 minutes at RT in the dark with occasional vortexing. The iodoacetamide solution was removed by pipette and the gel pieces washed with 3 x 1 mL changes pure H₂O for 30 minutes at RT. After removing the H₂O, the gel pieces were dried under vacuum for 10 minutes (SpeedVac Concentrator, Savant, USA).

Subsequently, the gel pieces were left to rehydrate in 30 µL digestion buffer (0.3 µg final concentration of trypsin) for 30 minutes. Enough 200 mM NH₄HCO₃ was added to cover the gel pieces, which were then incubated o/n at 37 °C, during which time proteolytic cleavage occurred.

After this time, the gel particles were centrifuged for 1 minute at 13,000 x g and the supernatant carefully removed and saved in a separate tube. Peptides were extracted from the gel pieces by incubation in 200 µL of 60 % ACN in 0.1% formic acid for 30 minutes at RT. The supernatant was separated from the gel pieces by centrifugation for 1 minute at 13,000 x g. The supernatant was saved and added to previous supernatant. The gel pieces were extracted with formic acid twice more, using the same method as above, and the extracts added to the previous extractions. The total extracts were then dried under vacuum (SpeedVac Concentrator, Savant, USA).

The extracted peptides were analysed using MALDI TOF (Section 2.5.2) and analysed using Mascot (www.matrixscience.com), Peptident (www.unb.br/cbsp/paginiciais/pepident.htm) or Profound (http://prowl.rockefeller.edu/prowl-cgi/profound.exe).

**Digestion buffer:**

<table>
<thead>
<tr>
<th>20 µg</th>
<th>200 mM</th>
<th>NH₄HCO₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trypsin, sequencing grade (Roche)</td>
<td>(final concentration of trypsin: 10 µg/mL)</td>
<td></td>
</tr>
</tbody>
</table>
2.4.6 Western blot

Proteins separated by SDS-PAGE were transferred to a membrane (either nitrocellulose or polyvinylidene fluoride (PVDF)), and detected using antibodies.

(a) Passive blotting

When electrophoresis was complete, the tricine gel was taken out of the SDS-PAGE chamber and the two glass plates separated. Without dislodging the gel from the plate it was sticking to, surplus buffer was rinsed off with pure H₂O and the stacking gel was removed using a scalpel.

A PVDF membrane (Immobilon-P, 0.45 µm; Millipore) was cut to size and wetted in methanol for 30 sec, then pure water for at least 2 minutes. Excess moisture was removed by gently patting it with lint-free paper, and it was then placed on the gel. Trapped air bubbles were carefully removed and two pieces of dry Whatman 3 mm filter paper were placed on top of the membrane, followed by a glass plate, then a weight on top of the plate.

The protein transfer took 2 - 4 hours at room temperature or 1 - 2 hours at 37 °C. The blot was disassembled and the PVDF membrane washed 3 x for 10 minutes in pure water before air drying it overnight.

(b) Electroblotting

Following electrophoresis, the gel was removed from the gel apparatus and rinsed in transfer buffer to remove any electrophoresis buffer salts and detergents, then soaked in fresh transfer buffer. A nitrocellulose membrane was cut to the dimensions of the gel, labelled on one side with a pencil and one corner of the membrane was cut. It was then slowly immersed in the transfer buffer and left to soak for 15 minutes. Four Whatman 3 mm filter papers were cut to size and, together with two fibre pads, soaked in the transfer buffer.
The buffer tank was placed in the transfer chamber and a stir bar installed at the bottom of the unit. The cooling unit, a frozen insert, was placed in the transfer chamber (Mini Transblot electrophoresis transfer cell, BioRad) shortly before starting the transfer.

To assemble the gel-membrane-sandwich, the black panel of the gel holder cassette was put in a shallow vessel and a pre-soaked fibre pad was placed on it. Two pre-soaked Whatman papers and the equilibrated gel were then placed on the fibre pad, taking care to ensure that no air bubbles were trapped between the paper and the gel. The surface of the gel was wetted with transfer buffer, and the pre-wetted membrane was carefully placed on the gel. Any trapped air bubbles were carefully removed. Then the surface of the membrane was also flooded with transfer buffer. Finally, the last two pre-soaked Whatman papers and the fibre pad were placed on top of the membrane. The gel holder cassette was carefully closed, taking care not to dislodge the sandwich construction, and placed in the buffer tank of the transfer vessel with the black panel facing the black cathode panel of the transfer chamber.

The buffer tank was filled to an appropriate level, the magnetic stirrer turned on and the lid put in place. For mini-gels the running time was 80 minutes at 100 Volts.

Transfer buffer:

- 25 mM Tris
- 192 mM Glycine
- 20% (v/v) Methanol
- pH 8.3 (do not adjust)
- (store at 4 °C)

(c) Dot blot

A dot blot is a simplification of a Western blot, where the samples are applied directly on a membrane as a dot.

An Immobilon P<sup>SQ</sup> PVDF membrane was wetted in methanol for 15 seconds and then carefully placed in pure water to soak for 2 minutes. The transfer stack was assembled by first placing paper towels on the work surface, followed by a dry and a wet sheet of
Whatman 3 mm paper. Lastly, the pre-wetted membrane was placed on top of the wet filter paper and 5 µL of sample were spotted onto the membrane. After the samples are absorbed, the membrane was placed on a clean sheet of Whatman 3 mm paper to dry.

Detection was carried out using the rapid immunodetection method (Section 2.4.6(e)). The primary antibody was diluted 1:5,000 and the secondary antibody 1:20,000 in PBS (phosphate buffered saline) with 1 % BSA (bovine serum albumin).

**Primary antibody:** anti-\(O\)-\(N\)-acetylglucosamine antibody (Sigma-Aldrich, USA)

**Secondary antibody:** anti-mouse antibody conjugated to horse radish peroxidise (Sigma-Aldrich, USA)

**(d) Standard immunodetection method**

The blot was placed in 10 mL of blocking solution to block unspecific sites and incubated for 1 hour at room temperature with agitation. After this time the membrane was washed with three changes of PBST. It was then incubated with the primary antibody solution for 1 hour at room temperature with agitation. The blot was then washed three times for 5 minutes with PBST, before being placed in the secondary antibody solution and incubated for 1 hour with agitation at room temperature. After this time it was once more washed three times for 5 minutes with PBST, before the addition of a chemiluminescent blotting substrate (Roche). Bands were visualised using the Intelligent Dark Box II (Fujifilm, Japan).

**PBST:**

| 10 mM | Sodium phosphate, pH 7.2 |
| 0.9 % (w/v) | NaCl |
| up to 0.1 % | Tween-20 |

**TBST:**

| 10 mM | Tris, pH 7.4 |
| 0.9 % (w/v) | NaCl |
| up to 0.1 % | Tween-20 |
### Blocking solution:
- PBS or TBS (PBST or TBST without Tween-20)
- 5 % (w/v) skim milk or 1 % (w/v) BSA
- 0.05 % Tween-20 (optional)

### Primary antibody solution:
- Blocking solution
- Antibody dilution 1:1,000 (anti-GFP antibody, Sigma)

### Secondary antibody solution:
- Blocking solution
- Antibody dilution 1:2,500 (anti-mouse IgG conjugated to horse radish peroxidase, Sigma)

(e) **Rapid immunodetection method (Millipore)**

The rapid immunodetection method was used for Immobilon-P PVDF transfer membranes (Millipore, USA). The blot had to be completely dry before beginning the rapid immunodetection method.

For the detection of PInKW30 using anti-farnesyl antibodies the primary antibody was diluted 1:1,000 in PBS (phosphate buffered saline) with 1 % BSA (bovine serum albumin). No detergent was added to any of the solutions. The blot was placed in the primary antibody solution and incubated with agitation for 1 hour.

Subsequently, the blot was washed twice with PBS for 5 minutes and then placed in the secondary antibody solution diluted 1:2,500, and incubated with agitation for 30 minutes. After washing the blot in PBS twice for 5 minutes the bands were detected by chemiluminescence (Roche; Intelligent Dark Box II, LAS-1000, Fujifilm, Alphatech).

**PBS (phosphate buffered saline):**
- 10 mM sodium phosphate, pH 7.2
- 0.9 % (w/v) NaCl

**Antibody dilution buffer:**
- PBS
- 1 % (w/v) BSA (bovine serum albumin)
Primary antibody: anti-farnesyl antibody (Sigma-Aldrich, USA)

Secondary antibody: anti-rabbit antibody conjugated to horse radish peroxidase (Sigma-Aldrich, USA)

(f) Chemiluminescent protein detection

The BM Chemiluminescence Blotting Substrate (POD) (Roche, GER) was used for the chemiluminescent detection of the western blots. The detection solution was prepared by mixing 10 mL of solution A (luminescence substrate solution) with 100 µl of solution B (starting solution) (ratio of 1:100) and this was allowed to warm to room temperature (15 - 25 °C). This detection system is based on the oxidative reaction of the horseradish peroxidase (POD or HRP), which is bound to the secondary antibody. In the presence of hydrogen peroxide the peroxidase catalyzes the oxidation of luminol, resulting in an activated intermediate reaction product, which decays to the ground state by emitting light. The light emission is enhanced by 4-iodophenol, which acts as a radical transmitter between the formed oxygen radical and luminol.

The blot was covered completely with substrate solution and incubated for 1 minute, after which excess substrate was drained off and the blot placed on a transparent plastic sheet and covered with a second sheet. Any trapped air bubbles were gently smoothed out and the bands of the protein standard were marked with a phosphorescent marker. The blot was then placed in the dark box (Intelligent Dark Box II, LAS-1000, Fujifilm) and multiple exposures of 10 sec were taken. Images of each interval were recorded for up to 5 minutes and the image (or images) with the best exposure was (were) saved.

2.4.7 Expression and purification of recombinant proteins

(a) Expression of recombinant proteins in E. coli

All genes cloned in expression plasmids used in E. coli were under the control of the T7 promoter and therefore inducible by the addition of Isopropyl-β-D-Thiogalactopyranoside (IPTG). IPTG is a lactose analogue that activates gene expression but cannot be metabolised
by the cell. Cultures were induced at an OD$_{600}$ of $\sim$0.4 by addition of 1 mM IPTG (final concentration). The cultures were then incubated for another 4 hours or overnight.

(b) Expression of recombinant proteins in lactobacilli

To express proteins from *L. plantarum* KW30 in the same species, the pSIP expression system by Sorvig *et al.* (2005) was used. This expression system uses as expression strains *L. plantarum* NC8 and *L. sakei* Lb709. We obtained two expression vectors of which one, pSIP409, could be cloned in *E. coli*, whereas the other one, pSIP412, could only be cloned in *Lactococcus lactis* MG1363. Both plasmids confer erythromycin (Em)-resistance, which was selected for using 200 µg/mL Em for *E. coli* and 10 µg/mL Em for *L. lactis* MG1363.

The full-length GTase, RR and PlnKW30 genes were cloned into the *NcoI* and *HindIII* restriction sites of pSIP409 and transformed into both expression strains using the method by Aukrust & Blom (1992). The correct incorporation of the inserts was verified by colony PCR (Section 2.2.15(b)).

All expression constructs are listed in appendix 6, table 5.5, and expression trials were carried out following the instructions of Sorvig *et al.* (2005). The cultures were grown to an OD$_{600}$ of $\sim$0.3, induced with 50 ng/mL of the sppIP inducing peptide and harvested after approximately 4 hours growth at 30 °C with an OD$_{600}$ of $\sim$2.0. The cell pellet was resuspended in PBS and the cells lysed using glass beads and a bead mill (Fast prep cell disruptor, Thermo Savant, USA). The expression levels were analysed by SDS-PAGE.

(c) Expression of recombinant proteins in Lactococcus

The nisin-controlled expression (NICE) system (Kuipers *et al*. 1995; NIZO Food Research, The Netherlands) uses components of the nisin biosynthesis cluster to control protein expression.

Cloning of the expression plasmids pNZ8148 and pNZ8112 (Table 5.2) was carried out in *L. lactis* MG1363, which was favoured over *E. coli* because of reports of unwanted
recombination events during cloning of these vectors in *E. coli*. The RR and *plnKW30* genes were cloned into the *Ncol* and *HindIII* restriction sites of pNZ8148 and the RR and GTase genes were cloned into the *Nael* and *XbaI* restriction sites of pNZ8112. The expression plasmids (Table 5.5) were transformed into the expression strain *L. lactis* NZ9000 using the method described by Holo & Nes (1989). Expression trials were carried out as described by NIZO (Food Research, The Netherlands).

A 5 mL culture of *L. lactis* NZ9000, harbouring one of the expression plasmids, was grown overnight in M17 broth with 0.5 % glucose and 10 µg/mL chloramphenicol at 30 °C. This culture was diluted 1/25 in 2 x 10 mL fresh medium and grown at 30 °C until the OD$_{600}$ reached approximately 0.4. One of the 10 mL cultures was induced with 1 ng/mL nisin A (supernatant of a full grown culture of NZ9700 containing approximately 10 mg/L nisin A), whereas the second, uninduced 10 mL culture was used as negative control. Both cultures were incubated for 4 hours; the cells were harvested by centrifugation and then resuspended in Tris/HCl buffer. Lysis of the cells was carried out using glass beads and a bead mill (Fast prep cell disruptor, Thermo Savant, USA), then the insoluble fraction was separated from the soluble fraction by centrifugation. Both samples and the uninduced sample were analysed for protein production by SDS-PAGE (Section 2.4.2).

(d) Cell free expression of proteins

The ‘Rapid Translation System (RTS) 100 *E. coli* HY Kit’ (Roche, GER) and the RTS ProteoMaster (Roche, GER) were used for *in vitro* protein synthesis following the instructions of the manufacturer.

(e) Purification of recombinant proteins

Cells were harvested in a GS-3 rotor (Sorvall) at 2,800 x g for 45 minutes at 4 °C (Sorvall Evolution RC). The cells were resuspended in Tris buffer (pH 8.0) and re-pelleted by centrifugation to wash off any residual media. Then, the cells were resuspended in Tris buffer (pH 8.0) containing complete protease inhibitor (Complete, Mini, EDTA-free; Roche)
and reducing agent (10 mM DTT; Sigma) and passed through a French press to lyse the cells (E. coli: 2 passes at 6 kPa; Lactobacillus cells: 30 minutes Lysozyme treatment, followed by 4 passes at 6 kPa). Cellular debris was removed by centrifugation at 17,000 x g for 30 minutes at 4 °C. The cell-free supernatant was decanted and kept on ice until required for further purification steps.

(f) Isolation of inclusion bodies (IB)

The expression of recombinant plnKW30 resulted in the formation of inclusion bodies, which were then isolated. E. coli BL21(DE3)origami (pET32MatBac) was grown in 200 mL LB (Amp 100, Kan 15, Tc 12.5 µg/mL) at 37 °C to an OD₆₀₀ of 0.6 and induced by addition of IPTG to a final concentration of 1.5 mM. After 6 hours growth to an OD₆₀₀ of 1.3, the cells were harvested by centrifugation (10 minutes, 5,700 x g, 4 °C) and resuspended in 10 mL of 40 mM Tris/HCl (pH 7) containing 1 mM EDTA (trisodium salt).

The cell suspension was thoroughly sonicated on ice and then centrifuged for 20 minutes at 2,700 x g and 4 °C. The white layer of the pellet, which lay on top of a firm black layer, was carefully resuspended with the residual liquid and transferred to a new tube. 1.2 mL of 10 mM EDTA (pH 7.8) containing 0.5 % Triton X-100 was added, the tube briefly vortexed and then centrifuged for 3 minutes at 5,700 x g. The supernatant was discarded, the pellet resuspended in 1.2 mL of 10 mM EDTA (pH 7.8) containing 0.5 % Triton X-100 and subjected to centrifugation to repellet the inclusion bodies. This washing step was repeated two more times and was followed by two more washes using the original cell lysis buffer 40 mM Tris/HCl (pH 7) containing 1 mM EDTA (trisodium salt). Finally, the pellet containing the inclusion bodies was resuspended in 200 µL of 40 mM Tris/HCl (pH 7) containing 1 mM EDTA (trisodium salt) and aliquots were frozen at -80 °C.

2.4.8 Purification of native PlnKW30

L. plantarum KW30 was grown in 8 L MRS medium for three days at room temperature (~25 °C) without shaking. The cells were removed by centrifugation and the supernatant
stirred overnight with 1 L phenyl sepharose resin (fast flow, low substitution, GE Healthcare, UK) using an overhead stirrer. The resin was washed with 2 L of 2% NH₄HCO₃, then packed into a 1.5 L glass column and eluted with 2 L of 2% NH₄HCO₃ made in 40% EtOH. 10 mL fractions were collected and tested for activity using a biological assay (Section 2.4.9). The fractions containing active PlnKW30 were pooled into three main fractions and the sample volume was reduced by rotary evaporation. Each pooled sample was analysed for the presence of PlnKW30 using tricine SDS-PAGE (Section 2.4.2). Samples containing the bacteriocin were further purified by RP-HPLC (Section 2.5.1) using a Jupiter 5u C₄ 300 Å column (10 x 250 mm, Phenomenex) using the following solutions to form a linear gradient: A: H₂O, 0.1 % TFA, and B: Acetonitrile, 0.08 % TFA. The gradient started at 20 % B and increased over 25 minutes to 45 % B at a constant flow rate of 5 mL/min. Because of the amount of detergent remaining in the sample, the gradient was increased to 100 % B, and held for 10 minutes before it was returned to the starting conditions, and held for another 10 minutes before starting the next injection. Elution was monitored using a Photodiode Array Detector (Dionex) at 214 nm and 280 nm. Peaks were collected by hand. After testing for the presence of PlnKW30 using the biological assay (Section 2.4.9), active fractions were pooled and concentrated by lyophilisation. The active fractions eluted at 36 – 37 % B.

2.4.9 BIOLOGICAL ASSAY FOR BACTERIOCIN ACTIVITY

To prepare the indicator plates for the biological assay the indicator strain L. plantarum ATCC 8014 was grown overnight in MRS media. Then, 1 mL of this culture was added to ~45 mL of MRS medium with 1 % agar and mixed gently by inversion. This mixture was immediately poured into petri dishes, using 15 mL per plate and the plates were allowed to set at room temperature. The indicator plates could be stored at 4 °C for at least a week.

When required, an indicator plate was warmed to room temperature. Normally, 1 - 2 µL of sample to be tested and purified PlnKW30, as positive control, were spotted onto the top of the indicator plate and placed lid-up in the 30 °C incubator. The indicator plates were incubated overnight at 30 °C or until a clearing was visible for the positive control.
2.4.10 **Minimum inhibitory concentration of PlnKW30**

The minimum inhibitory concentration (MIC) was defined as the minimal amount of bacteriocin that inhibited 50% of the indicator strain growth (50% of the turbidity of the control culture without the bacteriocin). The MIC of PlnKW30 was tested in two ways, using *L. plantarum* ATCC 8014 as indicator strain. First, several dilutions of PlnKW30 in pure water (Table 2.4) were spotted on indicator plates (Section 2.4.9) and grown at 30 °C for 24 hours (Fleury *et al.* 1996). The MIC was determined by visual assessment of the clearing each concentration produced.

A more accurate method used microtitre (96-well) plates where 100 µL of different dilutions of PlnKW30 in pure water (Table 2.4) were added to 100 µL of a 1/100 dilution of an overnight culture of the indicator strain *L. plantarum* ATCC 8014 (modified from Nissen-Meyer *et al.* 1992; modified from Amsterdam 1996). The microtitre plates were incubated at 30 °C and the inhibitory effect of the bacteriocin was detected by measuring the OD595 of the growing cultures. Using this method the MIC was determined in comparison to the growth of the positive control (no addition of PlnKW30).

<table>
<thead>
<tr>
<th>Sample number</th>
<th>PlnKW30 dilutions used for Indicator plate</th>
<th>Sample number</th>
<th>PlnKW30 dilutions used for 96-well plate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1 mg/mL</td>
<td>1</td>
<td>1.0 µg/mL</td>
</tr>
<tr>
<td>2</td>
<td>0.05 mg/mL</td>
<td>2</td>
<td>0.1 µg/mL</td>
</tr>
<tr>
<td>3</td>
<td>0.025 mg/mL</td>
<td>3</td>
<td>50.0 ng/mL</td>
</tr>
<tr>
<td>4</td>
<td>0.0125 mg/mL</td>
<td>4</td>
<td>25.0 ng/mL</td>
</tr>
<tr>
<td>5</td>
<td>6.25 µg/mL</td>
<td>5</td>
<td>12.5 ng/mL</td>
</tr>
<tr>
<td>6</td>
<td>3.1 µg/mL</td>
<td>6</td>
<td>6.0 ng/mL</td>
</tr>
<tr>
<td>7</td>
<td>1.6 µg/mL</td>
<td>7</td>
<td>3.0 ng/mL</td>
</tr>
<tr>
<td>8</td>
<td>0.8 µg/mL</td>
<td>8</td>
<td>1.5 ng/mL</td>
</tr>
<tr>
<td>9</td>
<td>0.4 µg/mL</td>
<td>9</td>
<td>No culture (negative control)</td>
</tr>
<tr>
<td>10</td>
<td>0.2 µg/mL</td>
<td>10</td>
<td>No PlnKW30 (positive control)</td>
</tr>
<tr>
<td>11</td>
<td>0.1 µg/mL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Water only (negative control)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.4.11 Tryptic and chymotryptic digest of PlnKW30

A sample of PlnKW30 was dissolved in pure water and the pH adjusted to 8 with 1% ammonium bicarbonate. 25 µg trypsin in 1 mM HCl was added to the reaction and incubated at room temperature overnight.

The chymotryptic digest was carried out overnight with 50 µL of 2 mg/mL activated chymotrypsin in 1% ammonium bicarbonate at room temperature. Prior to the chymotryptic digest the α-chymotrypsin had to be activated by addition of 20 µL of 25 ng/µL trypsin and incubation for 4.5 hours at room temperature.

2.4.12 Reduction and alkylation of the disulfide bonds of PlnKW30

The two disulfide bonds of PlnKW30 were broken and covalently modified by reduction and alkylation. A sample of PlnKW30 was dissolved in 10 mM DTT (dithiothreitol) in 100 mM NH₄HCO₃ buffer and was incubated for 1 hour at 56 °C. After cooling the sample to room temperature, 55 mM iodacetamide in 100 mM NH₄HCO₃ buffer was added and incubated for 45 minutes at room temperature in the dark with occasional mixing.

The reduced and alkylated PlnKW30 was purified using RP-HPLC with a Jupiter 5µm C₁₈ 300 Å column (250 x 4.6 mm, Phenomenex) and the fractions collected lyophilised.

2.4.13 O-deglycosylation of PlnKW30

First, the deglycosylation of PlnKW30 was carried out by β-elimination under alkaline conditions (Greas et al. 1996), but these harsh conditions resulted not only in deglycosylation of the O-linked GlcNAc but also in degradation of the peptide. Therefore, an enzymatic deglycosylation method was carried out using the N-acetyl-β-D-glucosaminidase (GcnA) from Streptococcus gordonii (Harty et al. 2004). The plasmid containing the GcnA gene was kindly provided by D.B. Langley and D.W.S. Harty. GcnA was expressed in E. coli B314(DE3)(pET28a) in 1 L LB medium containing tetracycline and kanamycin at 37 °C with agitation. After reaching an OD₆₀₀ of about 0.8, the culture was induced with
1 mM IPTG and grown overnight with agitation at 37 °C. The cells were harvested by centrifugation in a GS-3 rotor (Sorvall) at 2,800 x g for 45 minutes at 4 °C (Sorvall Evolution RC), disrupted using a French press and the cell debris pelleted by centrifugation at 17,000 x g for 30 minutes at 4 °C. The supernatant was then subjected to Immobilised Metal Affinity Chromatography (IMAC), because GcnA was fused to a His-tag. IMAC is based on the specific binding of histidine residues to metal ions, such as nickel. The resin (Chelating Sepharose Fast Flow, GE Healthcare) was packed into a column, charged with Nickel(III) chloride hexahydrate dissolved in pure water and washed with 20 mM Tris, 500 mM NaCl (pH 8) buffer prior to sample loading. The supernatant was loaded onto the column via gravity and washed with five column volumes of 20 mM Tris, 500 mM NaCl (pH 8) buffer. GcnA was eluted with 100 mM imidazole in 20 mM Tris, 500 mM NaCl (pH 8), a buffer exchange to 50 mM Tris, 150 mM NaCl (pH 8) was carried out and the enzyme was stored at 4 °C.

The deglycosylation mix contained 50 µL PlnKW30 dissolved in pure water, 5 µL GcnA (145 ng/µL) and 45 µL 50 mM Tris, 150 mM NaCl buffer. The reaction mix was mixed well by pipetting and incubated at 42 °C overnight. Samples were taken over time and analysed using MALDI TOF mass spectrometry to check for the loss of 203 mass units, which indicates the loss of an N-acetylglucosamine. The deglycosylated PlnKW30 was purified from remaining native PlnKW30 by RP-HPLC using a Jupiter C18 column.

2.4.14 Defarnesylation of PlnKW30

The putative farnesyl group of PlnKW30 was chemically removed using methyl iodide (Casey et al. 1989). A sample of PlnKW30 was incubated with CH₃I (16.7 % per volume) for 48 hours in the dark. This chemical reaction was stopped by addition of 18 % NH₄HCO₃ (3 % per volume), shaking for 12 hours in the dark. The sample was then air dried, concentrated using a Speed-vac and finally dissolved in H₂O. Further purification of the CH₃I-treated PlnKW30 sample was achieved by RP-HPLC. The loss of the putative farnesyl group was monitored by MALDI TOF and electrospray mass spectrometry analysis.
2.4.15 *Farnesyltransferase (FTase) Activity Assay*

Crude yeast extract or purified recombinant yeast FTase (Sigma) was used as a source of enzyme and the reactions were carried out using the appropriate controls.

(a) *Preparation of crude yeast extract*

*S. cerevisiae* (yeast hereafter) was grown with shaking in 150 mL of YM broth at 30 °C up to an OD$_{600}$ of 0.25 - 0.75. The cells were harvested by centrifugation for 10 minutes at 9,800 x g and 4 °C and then washed with yeast cell extraction buffer. After re-centrifugation for 10 minutes using the same conditions, the pelleted cells were resuspended in yeast cell extraction buffer, and lysed using glass beads and a bead mill (Fast prep cell disruptor, Thermo Savant, USA) for three times 45 seconds at the highest setting. The lysate was clarified by centrifugation, 13,000 x g for 10 minutes and the protein concentration was determined by Bradford assay (Section 2.4.1(a)).

Yeast cell extraction buffer:  
<table>
<thead>
<tr>
<th>Buffer</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hepes</td>
<td>50.0 mM</td>
</tr>
<tr>
<td>MgCl$_2$</td>
<td>0.1 mM</td>
</tr>
<tr>
<td>EGTA</td>
<td>0.1 mM</td>
</tr>
<tr>
<td>TCEP</td>
<td>5.0 mM</td>
</tr>
<tr>
<td>pH</td>
<td>7.5</td>
</tr>
</tbody>
</table>

(b) *FTase activity assay*

At first, the FTase assay was performed using thin layer chromatography (TLC), but later RP-HPLC was used to separate the samples. Separation by TLC was a common method for FTase assays described in the literature (Vogt *et al.* 1995; Hightower *et al.* 2001). The FTase activity assay had to be discontinuous, because of the unavailability of labelled substrate that could result in a colorimetric product.
The reaction mixture consists of:

3 µL: reaction buffer (5 x; 250 mM Hepes, pH 7.5; 50 mM MgCl₂)
1 µL: TCEP (150 mM)
1 or 3 µL: ^3H-FPP (3.7 µM in 50 % EtOH)
1.5 µL: peptide (GCVLS: specific substrate for yeast FTase, containing the typical CaaX motif; 33.3 µM)

2 µL: purified recombinant yeast FTase (55 µg/mL)
or
6 µL: whole cell extract of yeast
H₂O to a total volume of 15 µL

The composition of the reaction buffer was changed to minimise the amount of ethanol added, although some ethanol derives from the ^3H-FPP which is dissolved in 50 % EtOH and 50 % water. The controls included reaction mixes without peptide, without ^3H-FPP or without yeast extract or purified yeast FTase.

All samples were pre-incubated at 30 °C for 5 minutes before the peptide was added. After 40 minutes at 30 °C the reactions were stopped with an equal volume of isopropanol and spotted onto TLC plates (ALUGRAM® SIL G, Macherey-Nagel). After drying overnight, the TLC plate was run with 5:3:2 (v:v:v, isopropanol/NH₄OH/H₂O) as mobile phase and again dried overnight. The TLC plate was sprayed three times with EN³HANCE® spray (Perkin Elmer), after each spray the plate was left to dry for 15 minutes and then exposed to film. After one week of exposure at -80 °C the radiolabelled components of the reaction were visualised by developing the film.

2.4.16 Pull-downs using wheat germ agglutinin

Wheat germ agglutinin is a plant lectin which is capable of binding to N-acetylglucosaminyl and sialic acid residues. Agarose bound, succinylated wheat germ agglutinin (Vector Laboratories, Inc., USA) is a derivative, which does not bind to sialic acid residues, unlike the native form, but retains its specificity toward N-acetylglucosamine.
Wheat germ agglutinin was used in pull-down experiments with purified PlnKW30. The lectin-resin was washed three times in lectin buffer and then about 100 µL of resin slurry was mixed with 100 µL lectin buffer and incubated with ~20 µg of sample. After 2 hours incubation at room temperature with gentle agitation, the samples were centrifuged to collect the lectin-resin in the bottom of the tube. The supernatant was taken off and the lectin-resin washed three times with lectin buffer. Both the supernatant and lectin-resin were analysed on tricine SDS-PAGE (Section 2.4.2).

Lectin buffer:
- 10.0 mM HEPES
- 1.0 M NaCl
- pH 7.5

2.4.17 Fluorescence Microscopy

1 mL of an overnight culture of *L. plantarum* ATCC 8014 or *L. plantarum* KW30 was used to inoculate 5 mL fresh MRS broth and grown to mid-log phase (OD$_{600}$ ~3.0). From this culture 500 µL aliquots were pipetted into 1.5 mL Eppendorf tubes and the components to be tested were added.

The effect of PlnKW30 on the viability of the cells was visualised using the ‘LIVE/DEAD BacLight Bacterial Viability Kit’ (Invitrogen, USA) according to the manufacturer’s instructions. It is a two-colour fluorescence assay using SYTO9 and propidium iodide stains that both stain nucleic acids. The SYTO9 (green stain) can pass the cell membrane of viable and dead cells, whereas the propidium iodide (red stain) can only penetrate cells with compromised membranes that are dead or dying. In the presence of both stains, the red stain causes a reduction in the SYTO9 fluorescence, resulting in red fluorescence of damaged cells.

Live cell imaging was performed in an imaging chamber (CoverWell, 20-mm diameter, 0.5 mm deep) (Molecular Probes) filled with 800 µL of 2 % agarose in MRS and sealed with a 22 x 22 mm glass cover slip.

Microscopy was carried out at room temperature using an Olympus IX 71 microscope with a x 100 oil immersion lens, NA 1.4. Images were captured using a Hamamatsu ORCA-ER
C4742-80 digital charge-coupled device camera (Hamamatsu Corporation). Counts and measurements were made using METAMORPH software (Molecular Devices Corporation) and downloaded to Microsoft Excel for analysis.

At each time point, four photographs of each sample were taken with filters for red and green emission and phase contrast. The photographs taken with the red and green filters of each time point were overlaid using the METAMORPH software (Molecular Devices Corporation) and a portion of the photo containing between 300 – 600 cells was counted manually. The red cells present within these counted cells, were then counted separately and the percentage of dead cells calculated. Two different sets of photographs were analysed in this way for each time point and sample, and the average and standard deviation of these results was then calculated.

Parallel to the photographs, the OD$_{600}$ of the cell cultures was measured at certain time points and plotted in a graph for comparison with the results from the fluorescence cell assay. Each culture was mixed by pipetting prior to taking a sample.

2.5 **Biochemical methods**

2.5.1 *Reverse Phase-High Pressure Liquid Chromatography (RP-HPLC)*

RP-HPLC is used to separate and purify peptides and proteins over columns in which the matrix is silica beads of a restricted diameter, usually 5 µm, containing pores of a restricted size, in this case 300 Å. Alkyl groups of various carbon chain lengths ($C_2 - C_{18}$; stationary phase) are covalently linked to the silica. The routine purification of PlnKW30 was performed using Jupiter 5µm C$_4$ 300 Å columns (250 x 10 mm, Phenomenex). Other peptide work was done with Jupiter 5µm C$_{18}$ 300 Å columns (250 x 4.6 mm, Phenomenex). RP-HPLC was carried out on Summit (Dionex) or UltiMate 3000 (Dionex) instruments.
2.5.2 **MASS SPECTROMETRY**

To analyse the mass of peptides and proteins two methods are used: Electrospray ionisation mass spectrometry (ESI-MS) (Micromass ZMD) and Matrix-assisted laser desorption/ionisation mass spectrometry (MALDI-MS) (Micromass ZMD). The results of both methods were analysed using the MassLynx software. Samples for mass spectrometry were stored in Maximum™ recovery tubes (Axygen, USA).

**(a) ESI-MS (Electrospray ionisation mass spectrometry)**

Samples for ESI had concentrations of ~10 µmol/L and were purified by RP-HPLC before analysis. Samples were dissolved in 50% acetonitrile:pure water and acidified by addition of 0.1% acetic or formic acid. They were sonicated for 10 - 15 minutes on high (Sonicator, Soniclean Pty. Ltd.) to remove any material adhering to the side of the tubes before being subjected to ESI-MS (Micromass ZMD), using the following settings: source temperature ~95 °C, desolvation temperature ~190 °C, desolvation gas ~200 L/hrs, cone gas ~50 L/hrs. The samples were injected using the bypass of the Waters 2790 HPLC with a flow rate of 0.1 mL/min.

**(b) MALDI-MS (Matrix-assisted laser desorption/ionisation mass spectrometry)**

The matrix solution for MALDI-MS was prepared by dissolving 5 mg of nitrocellulose and 20 mg of α-cyano-4-hydroxy-trans-cinnamic acid (HCCA, Sigma) in 500 µL of acetone with vortexing (1 - 2 minutes) after which 500 µL of isopropanol was added. This matrix solution was prepared fresh for each use. Samples subjected to MALDI-MS had to be purified by RP-HPLC or by the use of C18 Zip tips (Millipore) and acidified by addition of 0.1% acetic or formic acid. The matrix solution (0.6 µL) was spotted on one well of the MALDI target plate (Micromass) and immediately, 0.6 µL of the sample was pipetted onto the same spot. This peptide-matrix mix dried relatively quickly and formed an even crystalline surface. Angiotensin II (MW 1296.686) was used as reference and was applied
to the middle well of a set of five rings (Figure 2.1). The sample(s) of interest was (were) applied to the remaining four wells using the same protocol.

![Figure 2.1: Diagram of a target plate for MALDI TOF.](image)

The sample target plate was then loaded into the instrument (Micromass M@LDI™) and data collected using the following settings: pulse voltage ~2650 Volts, source voltage ~15,000 Volts, reflectron voltage ~2000 Volts. Reflectron mode (path length of 2.3 m) was used for samples up to a mass of 2000 mass units and linear mode (path length of 0.7 m) was used for all samples with a larger mass (up to ~200,000 m/z). The analyser vacuum had to be below 3.90e-6 for operation. A nitrogen UV laser (337 nm) was used and most of the acquisition was carried out with the laser energy at high (20 - 70 %). The data was collected and analysed using MassLynx.

(c) Zip tips

ZipTipC18 pipette tips (Millipore), containing silica beads with a diameter of 15 µm and a pore size of 200 Å with an attached carbon chain of C18, were used to purify and concentrate samples for MALDI-MS. They were used following the manufacturer’s guidelines.

(d) Macrotraps

Desalting of samples was carried out using a Macrotrap, Peptide 6PK (Michrom Bioresources, Inc., USA) and a 250 µL GASTIGHT syringe (Hamilton, USA). The trap was equilibrated with two 5 column volumes (CVs) washes of 0.1 % TFA in pure H2O, before the
sample was loaded by slowly pushing it through the trap. After washing 5 x with 5 CVs of 0.1 % TFA in pure H₂O the sample was eluted with 1 CV 80 % ACN, 0.1 % TFA, this being repeated several times, up to a total of 400 µL. Finally, the sample was dried down using a speed-vac (SpeedVac Concentrator, Savant, USA) and then dissolved in sample buffer.

2.5.3 **CIRCULAR DICHROISM**

Circular Dichroism (CD) spectroscopy was performed on a Chirascan CD spectrometer (Applied Photophysics, UK) according to the instructions of the manufacturer. Samples of PlnKW30 and the N-terminal tryptic fragment were diluted in 2 mM KH₂PO₄, 10 % acetonitrile to a concentration of about 0.2 mg/mL. Approximately 300 µL of each sample was pipetted into a Quartz SUPRASIL® precision cell with a 1 mm path length (Hellma, GER), placed in the cell holder and then into the CD spectrometer.

The experimental conditions used in each scan are listed in table 2.6 and the sample specific parameters are listed in table 2.7. Since all non-reduced samples were dissolved in 2 mM KH₂PO₄, 10 % acetonitrile, the baseline scans were carried out 10 x on 2 mM KH₂PO₄, 10 % acetonitrile without sample added, then averaged and subtracted from the sample scans. The reduced samples were resuspended in 2 mM KH₂PO₄, 10 % acetonitrile containing 10 mM TCEP and measured at 60 °C. 2 mM KH₂PO₄, 10 % acetonitrile with 10 mM TCEP at 60 °C was used for the 10 baseline scans for the reduced samples, which were then averaged and subtracted from the sample scans.

**Table 2.5: Experimental conditions for CD scan**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Path length</td>
<td>1 mm</td>
</tr>
<tr>
<td>Wavelength range</td>
<td>180 nm – 260 nm</td>
</tr>
<tr>
<td>Time per point</td>
<td>0.25 s</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>1 nm</td>
</tr>
<tr>
<td>Step Size</td>
<td>1 nm</td>
</tr>
<tr>
<td>Repeats in set</td>
<td>20</td>
</tr>
</tbody>
</table>
Table 2.6: Sample specific parameters for CD scans of native, reduced and deglycosylated PlnKW30 and its N-terminal fragment.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>PlnKW30</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>N-terminal fragment</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>native</td>
<td>reduced</td>
<td>O-deglyco-</td>
<td>native</td>
<td>reduced</td>
<td>O-deglyco-</td>
<td>native</td>
<td>reduced</td>
<td>O-deglyco-</td>
<td></td>
</tr>
<tr>
<td>Molecular weight [Da]</td>
<td>5199.0561</td>
<td>5203.0561</td>
<td>4996.9827</td>
<td>3827.55043</td>
<td>3831.55043</td>
<td>3624.4726</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of amino acids</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>32</td>
<td>32</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The scan of each sample consisted of 20 repeats, which were averaged using the ProData Viewer (Chirascan, Applied Photophysics, UK) and the baseline was subtracted. Finally, the curves were smoothed with a ‘Window size’ value of 2. The ‘Window size’ value should be as high as possible without distorting the spectrum. The examination of the residual plot shows if any distortion occurred during smoothing. The noise should be randomly distributed around zero.
3  Chapter 3: Experimental results

At the outset of this work, experimental evidence strongly suggested that the C-terminal modification of the bacteriocin PlnKW30 was a farnesyl group (Section 1.6.1). Based on this assumption several experiments were carried out, which are described and discussed in appendix 5. The experiments included:

- Analysis of the cellular content of *L. plantarum* KW30 and purified PlnKW30 using Western blot with anti-farnesyl antibodies (Appendix 5.2.1),
- Two-dimensional gel electrophoresis of the same samples and Western blotting with anti-farnesyl antibodies (Appendix 5.2.2),
- Development of a farnesyltransferase activity assay (Appendix 5.2.3),
- Methyliodide treatment of purified PlnKW30 to remove the putative farnesyl group from the polypeptide chain (Appendix 5.2.4).

The results of these experiments are now largely irrelevant, due to the recent discovery that the C-terminal modification of PlnKW30 is an *S*-linked *N*-acetylglucosamine, a modification not previously reported for a peptide cysteine in any genera (Section 3.5.1).

Attempts were also made to disrupt the *plnKW30* gene and to express soluble recombinant PlnKW30 and the products of other genes of the *plnKW30* gene cluster in heterologous hosts, but neither objective was successful. As there was a substantial investment in time to achieve these objectives, and because of their potential importance to the project, these experiments are described and discussed in Appendices 6 and 7.
3.1 PHYLOGENETIC CLASSIFICATION OF *L. plantarum* KW30

A lactic acid bacterium was isolated and classified by Kelly *et al.* (1994) as *Lactobacillus plantarum* KW30. The genus *Lactobacillus* shows an unusual genetic diversity (Claesson *et al.* 2007), where phylogeny of the species is difficult to associate with their phenotypes. The genes of the small (16S) and large (23S) subunits of the ribosome are often used for phylogenetic classifications. Therefore, a 16S - 23S ribosomal RNA (rRNA) gene fragment with an intergenic spacer region was amplified from *L. plantarum* KW30 genomic DNA using the primers Lacto1 (Chagnaud *et al.* 2001) and Lab16 (Song *et al.* 2000) and sequenced using those primers and the Lacto2 primer (Chagnaud *et al.* 2001) (primers listed in appendix 4). The DNA sequence of this 2,155 bp fragment (Appendix 8) was subsequently used as the query for a nucleotide BLAST (blastn; Altschul *et al.* 1990) search at the NCBI website (http://www.ncbi.nlm.nih.gov/), which showed that the closest related strain was *L. plantarum* WCFS1, followed by *L. brevis* ATCC 367 (Table 3.1).

<table>
<thead>
<tr>
<th>Strain (GI number)</th>
<th>Reference</th>
<th>Query coverage</th>
<th>E-value</th>
<th>Maximum identity</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Lactobacillus plantarum</em> WCFS1 complete genome (30407129)</td>
<td>(Kleerebezem <em>et al.</em> 2003)</td>
<td>100%</td>
<td>0.0</td>
<td>98%</td>
</tr>
<tr>
<td><em>Lactobacillus brevis</em> ATCC 367, complete genome (116098028)</td>
<td>(Makarova <em>et al.</em> 2006)</td>
<td>98%</td>
<td>0.0</td>
<td>93%</td>
</tr>
<tr>
<td><em>Lactobacillus sakei</em> 23K; complete genome (78609255)</td>
<td>(Chaillou <em>et al.</em> 2005)</td>
<td>98%</td>
<td>0.0</td>
<td>92%</td>
</tr>
<tr>
<td><em>Lactobacillus plantarum</em> ZDY36a; 16S ribosomal RNA gene, partial sequence (171850487)</td>
<td></td>
<td>68%</td>
<td>0.0</td>
<td>98%</td>
</tr>
<tr>
<td><em>Lactobacillus pentosus</em> NRIC 1837; gene for 16S rRNA, partial sequence (157907490)</td>
<td></td>
<td>68%</td>
<td>0.0</td>
<td>98%</td>
</tr>
</tbody>
</table>

The sequence was also subjected to a genomic BLAST search including all complete genomes of the genus *Lactobacillus* currently available (as of February 2009). The phylogenetic tree resulting from this genomic BLAST search is shown in figure 3.1 (NCBI; BLAST pair wise alignment using fast minimum evolution (Desper & Gascuel 2004)). In order to root this tree, *Lactococcus lactis* subsp. *lactis* II1403 was included in the genomic BLAST search as an outgroup for the alignment tree.
Chapter 3: Experimental results

3.1 Phylogenetic classification

Figure 3.1: Phylogenetic tree of *Lactobacillus* species. Classification of *L. plantarum* KW30 showing it to be closely related to *L. plantarum* WCFS1.
As expected, *L. plantarum* KW30 proved to be closely related to *L. plantarum* WCFS1, the only completely sequenced *Lactobacillus plantarum* strain (Figure 3.1). The two *L. plantarum* strains are clustered separately, showing their early separation and divergent development from the other *Lactobacillus* strains.

### 3.2 Analysis of the plnKW30 Gene Cluster

The amino acid sequence of mature *plnKW30* was determined by Edman sequencing (Dr. G.E. Norris - unpublished results). The sequencing of the gene cluster started at the *plnKW30* gene (*gccA*), whose DNA sequence was obtained from the reverse-translation of the peptide sequence. This translation was carried out using the preferred codons of *L. plantarum* according to its known codon usage. Further sequencing of the *plnKW30* gene cluster was achieved by primer walking, both up- and downstream from the *plnKW30* gene (*gccA*), using genomic DNA isolated from *L. plantarum* KW30. The primers used for the sequencing of the *plnKW30* gene cluster are listed in appendix 4. The *plnKW30* gene cluster sequence (14,149 bp) can be found in appendix 10.

A schematic representation of the *plnKW30* gene cluster is shown in figure 3.2, where putative open reading frames (ORFs) were identified and putative functions were assigned using the BLASTx (NCBI; BLASTx used with the standard settings) search.

![Figure 3.2: Schematic bacteriocin gene cluster with putative functions of ORFs identified in silico.](image)


From the putative functions of the identified genes (Figure 3.2) and the typical arrangement of individual genes in other bacteriocin gene clusters (Section 1.3.2(d)), the core genes necessary for PInKW30 production were concluded to be, apart from the
**plnKW30** gene itself, the GTase, the ABC-transporter, the two TRXs and the response regulator. Furthermore, a possible immunity gene upstream from the GTase ORF was identified. This ORF contains a CUPIN_2 domain that is not present in other lactobacilli, but has similarity to ORFs in *B. subtilis*. Similarities in the 5' untranslated regions suggest that the CUPIN_2 and *plnKW30* genes may be co-regulated. The presence of a transposase, a helicase and a resolvase probably indicates that this bacteriocin gene cluster might have been acquired by horizontal gene transfer, and therefore these ORFs could be remnants of proteins necessary for this acquisition. Further bioinformatic analyses of each gene and gene product of the *plnKW30* gene cluster was carried out and the results listed in table 3.2.

The name glycocin (gcc) was created for a new class of bacteriocins exemplified by PlnKW30, as it is the first bacteriocin that has been shown to be post-translationally modified with carbohydrate moieties (Section 1.2 and 3.5.1).

**Table 3.2: Characteristics of the genes and gene products of the *plnKW30* cluster of *L. plantarum* KW30**

<table>
<thead>
<tr>
<th>Gene name</th>
<th>ORF (bp)</th>
<th>Protein (aa)</th>
<th>Theoretical molecular weight (Da)(^a)</th>
<th>Theoretical isoelectric point (pI)(^b)</th>
<th>Amino acid sequence identity to the highest scoring orthologue from other species (%)(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>gccB</td>
<td>1269</td>
<td>422</td>
<td>48776.65</td>
<td>7.65</td>
<td>32 (4e-29; Glycosyltransferase, GT2; <em>Enterococcus faecalis</em> V583; GI: 29377896)</td>
</tr>
<tr>
<td>gccC</td>
<td>2121</td>
<td>706</td>
<td>79086.62</td>
<td>9.29</td>
<td>29 (7e-61; ABC-type bacteriocin/ lantibiotic transporter; <em>Lactococcus lactis</em> subsp. cremoris SK11)</td>
</tr>
<tr>
<td>gccD</td>
<td>420</td>
<td>139</td>
<td>15774.37</td>
<td>9.34</td>
<td>30 (7e-10; hypothetical protein; thioredoxin-family; <em>Enterococcus faecalis</em> V583, plasmid pTEF2)</td>
</tr>
<tr>
<td>gccE</td>
<td>534</td>
<td>177</td>
<td>20674.47</td>
<td>9.60</td>
<td>31 (8e-13; putative secreted thioredoxin, TRX-family; <em>Listeria welshimeri</em> serovar 6b strain SLCC5334)</td>
</tr>
<tr>
<td>gccF</td>
<td>777</td>
<td>258</td>
<td>30139.02</td>
<td>9.09</td>
<td>34 (1e-09; putative response regulator of the competence regulon ComE1; <em>Streptococcus bivis</em>)</td>
</tr>
<tr>
<td>gccA</td>
<td>195</td>
<td>64</td>
<td>-</td>
<td>-</td>
<td>Pre-PlnKW30</td>
</tr>
</tbody>
</table>

\(^a\) Theoretical molecular weight and isoelectric point (pI) was predicted using 'Compute pI/MW' (ExPASy).

\(^b\) tBLASTn search translated nucleotide databases using a protein query with standard settings was used to predict identities.
The putative ABC-transporter most likely starts at the alternative start codon shown in figure 3.3, which is supported by protein BLAST searches (NCBI), even though it does not have an optimal ribosome binding site (RBS). The traditional start codon is 12 bp downstream and has a good RBS, but alignments to other ABC-transporter proteins favour the upstream start codon.

```
TCAGGTTATTTTAGAATGTGACATAATAAGTGCCAACTACGTCATCTTTTCCCTTTAATATCAGAAATT
-35       -10    +1(TS)   RBS
CCTAAAAACTTTTC
AGGCGATATAAAAAGTAGTGTCAAATCT
GCAGTGAAACTGTATTCAAA
ATTACAAG
Æ
ABC  traditional start codon
GTGACAGCAAAATGAATAATTTAACAATTGCCAAATGATTGATTGGACCAGCAGCATAGCAACAA
```

Figure 3.3: Possible regulatory sites of gccC (putativeABC-transporter).
Start codons and the transcriptional start (+1; TS) are shown in grey. Promoter -35 and -10 sites, ribosome binding site (RBS) are underlined.

Possible regulatory sites for the plnKW30 gene (gccA) were identified (Diep et al. 1996; Kotelnikova & Gelfand 2002) and are presented in figure 3.4. Potential promoter elements were found upstream from the plnKW30 gene (gccA) and contained plausible -35 and -10 promotor and Shine-Delgarno sequences.

Direct repeats with an 18 nucleotide spacing (AAG-N18-AAG), a motif that is characteristic of the DNA-binding sites for response regulators that regulate bacteriocin production, were also identified. This repeat sequence may be specific for the regulation of gccA expression. Several possible terminator sequences with dyad symmetry were identified, but the terminator indicated in figure 3.4 is relatively strong with a free energy of -16.00 kcal/mol (Dobson et al. 2007; Naterstad et al. 2007). The transcript size of 357 bp from transcription start point (+1; TS) to the end (underlined and in italics in figure 3.4) is consistent with the size of gccA transcripts in Northern blots (Section 3.3.2).
Chapter 3: Experimental results

3.2 Analysis of the plnKW30 gene cluster

Figure 3.4: Nucleotide sequence and deduced protein of gccA (plnKW30).
Start and stop codons and the transcriptional start (+1; TS) are shown in grey. Promoter -35 and -10 sites, ribosome binding site (RBS) and potential direct repeats (L, R) with an 18 nt spacing (AAG-N18-AAG) characteristic of the DNA-binding sites for RRs that regulate bacteriocin production are underlined. Sequences of dyad symmetry with potential to serve as transcription terminators are in italics and underlined.

3.2.1 Bioinformatic analyses of the proteins of the plnKW30 gene cluster

The following programs were used in the bioinformatics analysis of the six core proteins of the plnKW30 gene cluster.

The program 'SignalP 3.0' (Nielsen et al. 1997; Bendtsen et al. 2004) was used to predict the occurrence of a signal peptide and the position of the signal sequence cleavage site in the proteins of the plnKW30 gene cluster.

For the prediction of transmembrane helices (TMHs), and the resulting cellular location of the C- and N-termini, the programs TMHMM 2.0 (CBS; Denmark; ExPASy); TopPred (von Heijne 1992; Deveaud & Schuerer; Institut Pasteur; new implementation of the original toppred program, based on G. von Heijne algorithm; Claros & von Heijne 1994), PredictProtein (Rost et al. 2004) or TOPCONS (Bernsel et al. 2008; Viklund & Elofsson
Chapter 3: Experimental results  3.2 Analysis of the plnKW30 gene cluster

2008; Bernsel et al. 2009), which includes the programs SCAMPI-seq, SCAMPI-msa, PRODIV, PRO and OCTOPUS, were used.

To find conserved domains in the ORFs of the plnKW30 gene cluster the ‘conserved domains in sequences (cds) BLAST’ program (NCBI) and the Simple Modular Architecture Research Tool (SMART; Schultz et al. 1998; Letunic et al. 2006) were used. The Pfam sequence search (Finn et al. 2008) and InterProScan (EMBL-EBI; The InterPro Consortium (Apweiler et al. 2001)) were also used to predict specific domains in protein sequences of the plnKW30 gene cluster.

For structure homology predictions the program GenTHREADER (Jones 1999; McGuffin & Jones 2003) was used, which can be found on the PSIPRED server (Bryson et al. 2005).

(a) Analysis of the putative GTase (GccB)

The SignalP 3.0 analysis of the GTase predicted it to be a non-secreted protein with zero probability of containing a signal peptide.

The TMHMM 2.0 transmembrane prediction program predicted there are no transmembrane helices in the sequence of the putative GTase, with the same result being obtained using TopPred and TOPCONS. The best predicted model of PredictProtein (Rost et al. 1996) detected one transmembrane helix, although the reliability of this model is low. The putative transmembrane helix is between residues 364 and 381, with the N-terminus facing towards the cytosol and the C-terminus outside of the cell membrane. The reliability of the topology prediction is 1, where 0 is low and 9 is high. It is most likely that GccB is a cytosolic protein, as no signal sequence was predicted and only one prediction program found a possible TMH.

The predictions of conserved domains were carried out using different programs, all of which found a glycosyltransferase family 2 (GT2) domain at the N-terminus of the putative GTase (e-value 3e-13). The ‘conserved domains in sequences BLAST’ recognized a 168 amino acid (aa) long GT2 domain (residues 41 - 220), where the residues 116 - 129 are not part of the predicted GT2 domain (Figure 3.5).
Figure 3.5: Schematic view of GccB with glycosyltransferase domain (from cds BLAST).

The ‘SMART’ program predicted the GT2 domain at residues 41 – 222 (e-value 5.40e-15), whereas the ‘InterProScan’ predicted it to be at 41 - 131 aa (e-value 5.5e-18).

GT2 is a diverse family which transfers sugars from a range of donors including UDP-glucose, UDP-N-acetylgalactosamine, GDP-mannose or CDP-abequose to a range of acceptors including cellulose, dolichol phosphate and teichoic acids. Members of the GT2 family normally use nucleotidediphospho-α-D-sugars to generate β-linked products. The best-scoring protein is SpsA from *Bacillus subtilis* (GI: 18158753) and belongs to pfam00535.

SMART also predicted a tetratricopeptide (TPR) repeat between residues 283 and 316. These repeats are normally present in proteins in four or more copies and are a structural motif that mediates protein-protein interactions (D'Andrea & Regan 2003). TPR domains have been shown to be characteristic of O-N-acetylglucosamine transferases (Section 1.5.6; Hurtado-Guerrero et al. 2008; Martinez-Fleites et al. 2008). The amino acid alignments of gccB with the putative glycosyltransferase of *Bacteroides fragilis* and SpsA from *Bacillus subtilis* show the presence of several helices that could form TPR domains. The amino acid alignments can be found in figure 5.8 (Appendix 11).

The structure homology prediction method GenTHREADER predicted GccB to have high similarity (p-value 0.0004) to the structure of the nucleotidediphospho-sugar transferase SpsA from *Bacillus subtilis* (pdb 1qg8; Charnock & Davies 1999b). SpsA is a glycosyltransferase, belonging to GT2, and is implicated in the synthesis of the spore coat of *B. subtilis*.

(b) Analysis of the putative ABC-transporter (GccC)

The analysis of the ABC-transporter using SignalP 3.0 predicted it to be a non-secreted protein with zero probability of containing a signal peptide.
TMHMM and TopPred both predicted five transmembrane helices (TMHs) for GccC, but TOPCONS predicted that the putative ABC-transporter would contain six TMHs, which is the usual number of TMHs found in most ABC-transporters. The specific amino acid sequences of the predicted TMHs are shown in table 3.3.

The cellular location of the protein sequences between the transmembrane helices was also predicted. The amino acid sequence between the helices alternated between being on the inside and on the outside of the membrane, starting with the N-terminus being located in the cytosol and ending with the C-terminus also being inside the cell. It is most likely that GccC contains six transmembrane helices, as most other ABC-transporters have six TMHs (Davidson et al. 2008; Rees et al. 2009) and the ATPase domain has to be located inside the cell (Table 3.4). With five TMHs this would not be possible.

Table 3.3: Predicted positions of the transmembrane helices of the ABC-transporter.

<table>
<thead>
<tr>
<th>Prediction program</th>
<th>TMH1</th>
<th>TMH2</th>
<th>TMH3</th>
<th>TMH4</th>
<th>TMH5</th>
<th>TMH6</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMHMM</td>
<td>171-193</td>
<td>208-230</td>
<td>280-302</td>
<td>306-325</td>
<td>401-423</td>
<td></td>
</tr>
<tr>
<td>TopPred</td>
<td>168-188</td>
<td>202-222</td>
<td>281-301</td>
<td>305-325</td>
<td>397-417</td>
<td></td>
</tr>
<tr>
<td>SCAMPI-seq</td>
<td>165-185</td>
<td>210-230</td>
<td>280-300</td>
<td>306-326</td>
<td>401-421</td>
<td></td>
</tr>
<tr>
<td>SCAMPI-msa</td>
<td>168-188</td>
<td>205-225</td>
<td>280-300</td>
<td>306-326</td>
<td>393-413</td>
<td>415-435</td>
</tr>
<tr>
<td>PRODIV</td>
<td>168-188</td>
<td>205-225</td>
<td>280-300</td>
<td>305-325</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRO</td>
<td>168-188</td>
<td>205-225</td>
<td>280-300</td>
<td>305-325</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OCTOPUS</td>
<td>168-188</td>
<td>204-224</td>
<td>280-300</td>
<td>305-325</td>
<td>397-417</td>
<td>419-439</td>
</tr>
<tr>
<td>TOPCONS</td>
<td>168-188</td>
<td>205-225</td>
<td>280-300</td>
<td>305-325</td>
<td>397-417</td>
<td>419-439</td>
</tr>
</tbody>
</table>

Conserved domains in GccC were predicted using cds BLAST, SMART, InterProScan and Pfam sequence searches and the predicted domains are listed in Table 3.4.
3.2 Analysis of the plnKW30 gene cluster

Table 3.4: Predicted conserved domains of GccC

<table>
<thead>
<tr>
<th>Prediction program</th>
<th>SunT*</th>
<th>Peptidase C39 domain*</th>
<th>ABC-membrane domain*</th>
<th>ABC-transporter-like*</th>
<th>ATPase domain*</th>
</tr>
</thead>
<tbody>
<tr>
<td>cds BLAST</td>
<td>4 - 693 (1e-72)</td>
<td>5 - 136 (cd02418; 2e-14)</td>
<td>169 - 438 (PF00664; 1e-12)</td>
<td>-</td>
<td>480 - 688 (cd03254; 3e-28)</td>
</tr>
<tr>
<td>SMART</td>
<td>-</td>
<td>8 - 141 (PF03412; 7.80e-10)</td>
<td>169 - 438 (PF00664; 7.30e-06)</td>
<td>-</td>
<td>504 - 687 (SM00382; 1.22e-11)</td>
</tr>
<tr>
<td>InterProScan</td>
<td>-</td>
<td>8 - 141 (PF03412; 6.9e-10)</td>
<td>169 - 438 (PF00664; 3.4e-07)</td>
<td>508 - 686 (PF00005; 4e-33)</td>
<td>504 - 687 (SM00382; 5.5e-11)</td>
</tr>
<tr>
<td>Pfam sequence</td>
<td>-</td>
<td>8 - 141 (PF03412; 13.254)</td>
<td>169 - 438 (PF00664; 3.4e-07)</td>
<td>480 - 704 (PS50893; 12.945)</td>
<td>-</td>
</tr>
<tr>
<td>Pfam search</td>
<td>8.7e-10</td>
<td>-</td>
<td>169 - 438 (PF00664; 4.1e-07)</td>
<td>508 - 686 (PF00005; 1.9e-35)</td>
<td>-</td>
</tr>
</tbody>
</table>

*Pfam accession numbers and e-values in parentheses

The putative ABC-transporter is predicted by cds BLAST to belong to the SunT family of multi-domain ABC-type bacteriocin exporters, which contain an N-terminal diglycine peptidase domain (Figure 3.6). This N-terminal peptidase domain belongs to the peptidase superfamily C39, which contains mostly bacteriocin-processing endopeptidases. The cysteine peptidases in this family cleave the diglycine leader peptides from the precursors of various, mainly non-lantibiotic, bacteriocins (Havarstein et al. 1995). The ABC-transporter transmembrane region belongs to the ABC-transporter membrane domain clan and is a unit of six transmembrane helices (Davidson et al. 2008; Rees et al. 2009). The C-terminal sequence of GccC shows homology to an ABC-transporter nucleotide-binding domain (Davidson et al. 2008).

Figure 3.6: Schematic view of GccC with individual domains and special motifs (from cds BLAST).

GenTHREADER predicts GccC has the best similarity (p-value 2e-11) to the multidrug ABC-transporter Sav1866 from *Staphylococcus aureus* (pdb 2hyd; Dawson & Locher 2006), which has the ABC exporter fold containing six TMHs.
(c) Analysis of the putative TRX1 (GccD)

The prediction for GccD made by SignalP 3.0 is that it is a non-secreted protein, even although there is a probability of 0.465 (0.0 is low and 1.0 is high) that it contains a signal peptide. The most likely cleavage site for this signal peptide is between amino acids 21 and 22 (cleavage site probability of 0.432). The SMART prediction includes a SignalP prediction that predicts residues 1 – 21 of GccD comprise a signal peptide.

TMHMM 2.0 and TOPCONS both predict that there is a likelihood of only one TMH in GccD between residues 5 – 22 and 3 – 23, respectively (Table 3.5). The four N-terminal amino acids are predicted to reside inside the cell, and consequently, the C-terminal 116 aa are predicted to be on the outside. TopPred, on the other hand, predicts that GccD contains two possible TMHs, as listed in table 3.5. The first helix (residues 3 - 23) is at a position similar to that predicted by TMHMM 2.0. It is most likely that TRX1 contains only one TMH possibly between residues 3 and 23, as TopPred is the only program that predicts a second TMH, but with much lower probability (Table 3.5).

Table 3.5: Predicted positions of the transmembrane helices of GccD.

<table>
<thead>
<tr>
<th>Prediction program</th>
<th>TMH1</th>
<th>TMH2</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMHMM</td>
<td>5 – 22</td>
<td></td>
</tr>
<tr>
<td>TopPred</td>
<td>3 - 23</td>
<td>93 - 113</td>
</tr>
<tr>
<td>SCAMPI-seq</td>
<td>4 - 24</td>
<td></td>
</tr>
<tr>
<td>SCAMPI-msa</td>
<td>4 - 24</td>
<td></td>
</tr>
<tr>
<td>PRODIV</td>
<td>3 - 23</td>
<td></td>
</tr>
<tr>
<td>PRO</td>
<td>3 - 23</td>
<td></td>
</tr>
<tr>
<td>OCTOPUS</td>
<td>2 - 22</td>
<td></td>
</tr>
<tr>
<td><strong>TOPCONS</strong></td>
<td>3 - 23</td>
<td></td>
</tr>
</tbody>
</table>

Using cds BLAST to predict conserved domains indicates the presence of a thioredoxin family domain (cd02947) between residues 40 and 133 (e-value 2e-09; figure 3.7), which contains a redox active CXXC motif at positions 61 - 64.
Chapter 3: Experimental results  3.2 Analysis of the plnKW30 gene cluster

Figure 3.7: Schematic view of GccD with TRX and TRX-superfamily domains (from cds BLAST).

SMART also predicts a thioredoxin domain (PF00085; e-value 1.10e+00) between residues 34 and 131 and a glutaredoxin domain (PF00462; e-value 1.80e+00) between residues 52 and 116, but the scores for both domains were higher than the required minimal threshold of significance. A BLAST search within SMART predicts a d2trxa domain between residues 37 and 129 (e-value 7.00e-09), which is characteristic of the thioredoxin-like superfamily.

InterProScan predicts it is likely there is a thioredoxin fold between residues 24 and 139 (e-value 1e-07), while the Pfam sequence search found no significant matches. It is most likely that GccD contains a TRX domain between residues 34 and 133, as all other programs predicted it.

GenTHREADER predicts that GccD has some structural homology to a thioredoxin from *E. coli* (pdb 2trx; Katti et al., 1990) and a thioredoxin from *Mycobacterium tuberculosis* (pdb 2i1u; Hall et al. 2006) (p-value 0.004 for both). Both proteins have two possible functions: one as an electron carrier and the second as a protein disulfide oxidoreductase, which is probably located inside the cell to keep intracellular proteins reduced (Arner & Holmgren 2000). GccD might therefore be located on the inside of the cell to keep PlnKW30 reduced, and thus inactive, until it is transported outside.

(d) Analysis of the putative TRX2 (GccE)

The SignalP 3.0 analysis of the second TRX predicted it to be secreted, with the probability of it having a signal peptide being 1.0 with the most likely cleavage site being between amino acids 35 and 36 (p = 0.668). SMART also predicts the presence of a signal peptide between residues 1 and 24 using SignalP.

The TMHMM 2.0 prediction for GccE is very similar to that of GccD, with a single TMH being predicted to be between residues 9 and 31. The N-terminal moiety (residues 1 - 8) is
predicted to be on the inside of the cell membrane and the C-terminal part (residues 32 - 177) on the outside of the cell membrane. Using TopPred, one TMH (residues 9 - 29) is predicted with a score of 2.726 (Cut-off for highly likely transmembrane segments: 1.00; Cut-off for putative transmembrane segments: 0.60). TOPCONS also predicts a TMH at residues 4 – 24 with the N-terminus facing the cytosol and the larger C-terminal part of the protein on the outside of the cell. Considering these predictions, it seems likely that GccE is transported outside the cell, but is anchored in the membrane rather than cleaved.

Cds BLAST predicts GccE contains a TRX domain (cd02947; e-value 5e-07) between residues 67 and 150 with the redox active motif CXXC being positioned between residues 88 and 91 (Figure 3.8).

SMART failed to predict the presence of any domains with confidence, but a transmembrane helix is predicted between residues 9 and 31. When InterProScan was used, a thioredoxin fold was predicted between residues 68 and 163 (e-0.0001; IPR012335). Interestingly, the Pfam sequence search predicted a protein of unknown function (DUF1312) between residues 16 and 105 (e-0.77).

The best structural homologue as determined by GenTHREADER is a bacteriocin transport accessory protein from *Streptococcus pneumoniae* (pdb 1zma) which has a p-value of 0.003. The second best match is the structure of the thioredoxin mutant R82E from *Alicyclobacillus acidocaldarius* (pdb 1nw2; Bartolucci et al. 2003) with a p-value of 0.005. As an accessory protein, GccE would be expected to be located in the membrane as it would be closely associated with the ABC-transporter (Section 1.3.2(f)).
(e) Analysis of the putative response regulator (GccF)

The analysis of GccF using SignalP 3.0 predicted a non-secreted protein, which had zero probability of containing a signal peptide.

TMHMM 2.0, TopPred and TOPCONS predicted there were no transmembrane helices in the sequence of the putative response regulator. PHDhtm (PredictProtein; Rost et al. 1996) on the other hand predicted GccF to contain one TMH, although the reliability level for this prediction was very low. These predictions lead to the conclusion that GccF most likely does not have a TMH and is thus a cytosolic protein.

A LytTR superfamily domain (pfam04397; e-value 1e-10) in the C-terminal part of GccF (Figure 3.9) was predicted using cds BLAST. Domains of this type bind to specific DNA sequences and are found in a variety of bacterial transcriptional regulators (Section 1.4.2(a)). LytTR domains are typically found in proteins regulating the production of important virulence factors, such as fimbriae, extracellular polysaccharides, toxins and many bacteriocins (Galperin 2008). Both response regulators PlnC and PlnD of the bacteriocin producing L. plantarum C11 contain LytTR domains which are involved in the regulation of the bacteriocin production (Section 1.4.2(b); Risoen et al. 1998; Risoen et al. 2000; Section 1.4.2(b); Diep et al. 2001; Risoen et al. 2001; Diep et al. 2003).

![Figure 3.9: Schematic view of GccF with LytTR domain (from cds BLAST).](image)

Similarly, the SMART, InterProScan and Pfam algorithms predicted there to be a LytTR domain between residues 160 and 254 with e-values of 5.90e-08, 2.3e-19 and 4.5e-21, respectively.

The N-terminal part of the putative response regulator has no homology to any known domains. Normally, this part of the protein would contain an input domain that recognises signals from a histidine protein kinase. However, no histidine protein kinase had been found in the vicinity of this ORF.
GenTHREADER predicts the sequence of GccF is likely to have structural homology to the *Staphylococcus aureus* agrA LytTR domain (pdb 3bs1; Sidote *et al.* 2008) (p-value of 5e-05) and to the LytTR DNA-binding domain of a putative methyl-accepting/DNA response regulator from *Bacillus cereus* (pdb 3d6w) (p-value of 0.0003). The cytoplasmic response regulator AgrA activates the expression of secreted virulence factors and down-regulates cell-wall associated proteins (Sidote *et al.* 2008).

---

**(f) Analysis of the bacteriocin PlnKW30 (GccA)**

The prediction made by SignalP 3.0 for the ORF of PlnKW30 is that it is a ‘non-secretory protein’, with zero probability of having a signal peptide or a cleavage site between amino acids 16 and 17. However, PlnKW30 contains a diglycine at positions 20 and 21 which is a typical cleavage site for leader peptides in class II bacteriocins (Section 1.3.2(d)). After cleavage of the putative 21 amino acid leader-sequence, the 43 amino acid mature PlnKW30 remains. SignalP 3.0 is not the optimal program to predict signal sequences in bacteriocins, as it predicts signal sequences for the translocase general secretion (Sec) pathway, which is not generally used for class II bacteriocins (Section 1.3.2(d)).

The transmembrane prediction program TMHMM 2.0 predicted a single 22 amino acid residue transmembrane segment in the GccA ORF; however, TopPred predicts an uncertain transmembrane segment close to the N-terminus of PlnKW30 and TOPCONS does not predict any transmembrane regions at all. In this case this is a false prediction as it is known that this segment of PlnKW30 is a cleaved signal sequence.

A search of the sequence for conserved domains with cds BLAST, InterProScan and Pfam found there was little likelihood of any conserved domains being present in the mature bacteriocin PlnKW30. The SMART domain search found an intrinsically disordered region between residues 32 and 43, which could be a flexible C-terminal region and is supported by the results of the NMR analysis (Figure 3.34).

No definite structural homologues could be found for GccA using GenTHREADER, a structural prediction program. The closest structure (p-value of 0.165) is the carbohydrate-
binding wheat germ agglutinin in complex with N-acetylglucosamine (pdb 2uvo), suggesting the peptide may be involved in binding to N-acetylglucosamine.

3.2.2 BRIEF SUMMARY OF BIOINFORMATICS ANALYSIS

GccB is a non-secreted, cytosolic protein that contains a glycosyltransferase family 2 (GT2) domain at the N-terminus between residues 41 and 220. It has a high similarity to SpsA from B. subtilis which is implicated in the synthesis of the spore coat.

The putative ABC-transporter GccC is not secreted and most likely contains six transmembrane helices, similar to most ABC-transporters. Thus both the N-terminal peptidase C39 domain and C-terminal ATPase domain are likely to be located inside the cell. The transporter appears to be most similar to the multidrug ABC-transporter Sav1866 from Staphylococcus aureus, which also contains six TMHs.

It is most likely that GccD contains only one TMH possibly between residues 3 and 23 and is a non-secreted protein. A thioredoxin domain was predicted to be present between residues 34 - 133 and the GccD is likely to have some structural homology to a thioredoxin from E. coli. It is most likely that GccD is located in the cytosol to keep PlnKW30 reduced, and thus inactive, until it is transported outside the cell.

GccE is probably transported outside the cell without cleavage of the signal sequence, as it is part of the membrane spanning helix. A TRX domain was predicted for GccE between residues 67 and 150 which contains the redox active motif CXXC. The best structural homologue is a bacteriocin transport accessory protein from Streptococcus pneumoniae indicating that GccE might be an accessory protein of unknown function, closely associated to the ABC-transporter GccC.

GccF is not predicted to have a signal sequence or a TMH and is thus most likely to be a cytosolic protein. It contains a LytTR domain that is typically found in proteins regulating the production of many bacteriocins. GccF is likely to have structural homology to the Staphylococcus aureus agrA LytTR domain, which is a cytoplasmic response regulator. No homology to any domains was found for the N-terminal part of GccF.
GccA contains the typical diglycine cleavage site for leader peptides in class II bacteriocins at positions 20 and 21. It does not contain any transmembrane helices or conserved domains, but has some structural similarity to a carbohydrate-binding wheat germ agglutinin in complex with N-acetylglucosamine, suggesting that PlnKW30 may be involved in binding to N-acetylglucosamine.

3.3 TRANSCRIPTIONAL ANALYSIS

The lack of intergenic regions between the ORFs of the plnKW30 gene cluster starting from the GTase through to the response regulator might indicate the presence of a single transcript. However, there is a small 77 bp intergenic region before the plnKW30 ORF, which might point to the possibility of it being transcribed separately from the other ORFs.

3.3.1 TRANSCRIPTIONAL ANALYSIS BY RT-PCR

To determine if the plnKW30 gene cluster is transcribed as one single mRNA or each gene individually, RT-PCR (Section 2.3.4) was performed using gene-specific primers (Appendix 4) (Dobson et al. 2007). The ABC-transporter was analysed with two separate primer sets: one for the N-terminal (ABC1) and the other for the C-terminal (ABC2) part of the protein. To analyse the size of the mRNA transcripts the forward primer of the first gene and the reverse primer of the second gene were used, as done by Dobson et al (2007), for the six core genes of the plnKW30 gene cluster. A product should only occur when both genes are transcribed on the same mRNA. The result of these RT-PCRs is shown in figure 3.10A and table 3.6, from gccB (GTase) through to gccA (plnKW30). The positions of the primers used in this RT-PCR are shown in figure 3.10B.
Figure 3.10: Results of the RT-PCR of genes and intergenic regions of the *plnKW30* gene cluster. 


No PCR product was obtained in the negative control (lane 14), which omits the reverse transcriptase step, indicating that the RNA sample was not contaminated with DNA. The GTase primers were used in the negative control, as they gave reliable results. The RNA was only tested with one primer pair for DNA contamination, as the same RNA sample was used for all PCR reactions in figure 3.10.
Table 3.6: Results of RT-PCR of genes and intergenic regions of the \textit{plnKW30} gene cluster.

<table>
<thead>
<tr>
<th>Lane</th>
<th>Gene(s) analysed</th>
<th>Primers(^a)</th>
<th>Expected size (bp)</th>
<th>Result of RT-PCR (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GTase 1 &amp; 2</td>
<td>GlycoT_KW30_BamHI_F/GlycoT_KW30_HindIII_R</td>
<td>1269</td>
<td>1200</td>
</tr>
<tr>
<td>2</td>
<td>GTase - ABC1 1 &amp; 4</td>
<td>GlycoT_KW30_BamHI_F/ABC1KW30_Not_R</td>
<td>2031</td>
<td>2000</td>
</tr>
<tr>
<td>3</td>
<td>ABC1 3 &amp; 4</td>
<td>ABC1KW30_Bam_F/ABC1KW30_Not_R</td>
<td>762</td>
<td>800</td>
</tr>
<tr>
<td>4</td>
<td>ABC1 - ABC2 3 &amp; 6</td>
<td>ABC1KW30_Bam_F/ABC2KW30_XhoI_R</td>
<td>2106</td>
<td>2100</td>
</tr>
<tr>
<td>5</td>
<td>ABC2 5 &amp; 6</td>
<td>ABC2KW30_NcoI_F/ABC2KW30_XhoI_R</td>
<td>1344</td>
<td>1300</td>
</tr>
<tr>
<td>6</td>
<td>ABC2 - TRX1 5 &amp; 8</td>
<td>ABC2KW30_NcoI_F/Trx1KW30_XhoI_R</td>
<td>1764</td>
<td>1700</td>
</tr>
<tr>
<td>7</td>
<td>TRX1 7 &amp; 8</td>
<td>Trx1KW30_NcoI_F/Trx1KW30_XhoI_R</td>
<td>420</td>
<td>420</td>
</tr>
<tr>
<td>8</td>
<td>TRX1 - TRX2 7 &amp; 10</td>
<td>Trx1KW30_NcoI_F/Trx2KW30_HindIII_R</td>
<td>954</td>
<td>950</td>
</tr>
<tr>
<td>9</td>
<td>TRX2 9 &amp; 10</td>
<td>Trx2KW30_Bam_F/Trx2KW30_HindIII_R</td>
<td>534</td>
<td>530</td>
</tr>
<tr>
<td>10</td>
<td>TRX2 - RR 9 &amp; 12</td>
<td>Trx2KW30_Bam_F/FTase.RRHindIII_R</td>
<td>1311</td>
<td>1300</td>
</tr>
<tr>
<td>11</td>
<td>RR 11 &amp; 12</td>
<td>FTase.RRNcoIF/FTase.RRHindIII_R</td>
<td>777</td>
<td>800</td>
</tr>
<tr>
<td>12</td>
<td>RR plnKW30 11 &amp; 14</td>
<td>FTase.RRNcoIF/Pln_HindIII_R</td>
<td>931</td>
<td>1000</td>
</tr>
<tr>
<td>13</td>
<td>plnKW30 13 &amp; 14</td>
<td>plnMatModNcoI/Pln_HindIII_R</td>
<td>154</td>
<td>200</td>
</tr>
</tbody>
</table>

\(^a\)Numbers refer to figure 3.10B and appendix 4

RT-PCR products were obtained for all intergenic regions, as shown in figure 3.10 and table 3.6, although the band for the RR and PlnKW30 was very faint (arrow in lane 12). The results show that possibly the whole \textit{plnKW30} gene cluster is transcribed on one mRNA, although \textit{gccA} (\textit{plnKW30}) may be transcribed on a separate mRNA.
Several bands of sizes different to those expected were visible in some lanes in figure 3.10, but they were mostly less intense than those of the expected bp size.

### 3.3.2 Northern blot analysis

A formaldehyde RNA gel (Section 2.3.3) was loaded with 40 µg of *L. plantarum* KW30 total RNA, blotted and developed using *plnKW30*-specific $^{32}$P-labelled probes. The labelling of the probes, blotting and hybridisation was carried out as described in section 2.3.3(a).

Figure 3.11 shows a Northern blot using a *plnKW30* specific probe. The *plnKW30* mRNA is visible in lane 1 and 2 (20 and 40 µg total RNA, respectively) with a size of about 330 bp, which was determined by plotting a molecular weight curve using the log molecular weight and the relative mobility of the marker RNA and the positive control. This is similar to the expected size of about 325 bp, which includes the ribosome binding site (RBS), signal sequence and the terminator region. Lane 3 shows the PCR product of the mature PlnKW30, without the diglycine signal sequence, as a positive control with a size of about 154 bp (primers: plnMatModNcoI, Pln_HindIII_R; appendix 4).

![Northern blot analysis using a *plnKW30* specific probe.](image)

**Figure 3.11:** Northern blot analysis using a *plnKW30* specific probe.

1: 20 µg of total RNA of *L. plantarum* KW30; 2: 40 µg of total RNA of *L. plantarum* KW30; 3: *plnKW30* PCR product as positive control.
This result shows that gccA from *L. plantarum* KW30 is transcribed as a monocistronic mRNA, similar to *sunA* encoding sublancin 168 from *Bacillus subtilis* 168 (Serizawa *et al.* 2005).

Efforts to establish the transcript size of the *plnKW30* gene cluster using probes for the response regulator and glycosyltransferase failed due to lack of specificity of the probes.

### 3.3.3 *GROWTH-PHASE-DEPENDENT GENE EXPRESSION OF plnKW30*

A Northern blot analysis (Section 2.3.3) was performed to compare *plnKW30* (gccA) gene expression in *L. plantarum* KW30 grown for 4, 8, 16 and 24 hours. A RNA sample of a *L. plantarum* KW30 mutant without the *plnKW30* cluster (*L. plantarum* KW30Δgcc) was used as negative control.

The lower bands shown in figure 3.12 were obtained using a gccA specific probe (expected size of about 325 bp), whereas the upper bands were obtained using a gap (glyceraldehyde 3-phosphate dehydrogenase; expected size of 1023 bp) specific probe. Northern blotting was carried out as described in Section 2.3.3(b) using *L. plantarum* KW30 RNA normalised to total RNA concentrations. A gradual increase in intensity of the gccA bands is visible, starting with the faintest at 4 hours (lane 1) up to 24 hours (lane 4).

The housekeeping gene gap was used as an internal control for gene expression, as it is supposed to be expressed at the same level throughout the growth of the cells. The upper bands of the gap gene in figure 3.12 should therefore all have similar intensities. The variations in the results of the gap gene could be due to the normalisation of the RNA samples, as the same amount of total RNA was loaded in each lane. It is possible that the relative amount of the gap transcript in the 16 and 24 h samples is much lower due to the large quantities of the *plnKW30* transcript. There was evidence that the *plnKW30* transcript was present in abundance (results not shown).
Figure 3.12: Growth-dependent gene expression of PlnKW30.
Total RNA of *L. plantarum* KW30 was isolated at different time points as indicated and examined with *plnKW30* and *gap* specific probes. 1: 4h; 2: 8h; 3: 10h; 4: 24h; 5: RNA of *L. plantarum* KW30Δgcc.

This would suggest that the intensity of the *gap* transcripts is not a good loading control and that the increase in *plnKW30* transcript is real. Consequently, the production of *plnKW30* mRNA increases with time until it peaks at the end of log phase.

### 3.4 Characterisation of the Bacteriocin PlnKW30 from *L. plantarum* KW30

#### 3.4.1 Influence of growth conditions on PlnKW30 production

To determine the optimal conditions for PlnKW30 production, the effects of growth temperature, aeration and culture medium were examined. *L. plantarum* KW30 was grown for 72 hours at 22, 25, 30 and 37 °C in MRS (Section 2.1.3(b)) or chemically defined medium (CDM; section 2.1.3(e)) in open or closed tubes to restrict oxygen supply. Chemically defined medium was tested, because it might facilitate the purification of PlnKW30 (Section 3.4.2), because of reduced browning during autoclaving (see section 3.4.2). After 72 hours, cells were separated from the culture medium by centrifugation
(14,100 x g) and 2 µL of each supernatant was spotted on an agar plate containing the indicator strain (Section 2.4.9). 2 µL of purified PlnKW30 was spotted on to the same plate to provide a positive control. The plate was then incubated at 25 °C overnight and photographed the next day (Figure 3.13).

![Figure 3.13: Effect of growth temperature, aeration and medium on PlnKW30 production.](image)

L. plantarum KW30 was grown in MRS or chemically defined medium (MM) at 22, 25, 30 and 37 °C in closed (C) or open (O) tubes. 2 µL samples were spotted on indicator plate and incubated overnight at 25 °C. Control: purified PlnKW30.

Figure 3.13 shows that PlnKW30 production was generally better in MRS medium than in the chemically defined medium. The temperature appears to have the greatest influence on the bacteriocin production; there was no PlnKW30 production in the MRS at 37 °C, and only a faint clearing is visible at 30 °C with the best production occurring at 22 °C. The level of aeration does not seem to influence bacteriocin production in MRS media; the clearings being of similar size regardless of the degree of aeration. In contrast to that, the chemically defined medium samples showed a small amount of PlnKW30 production with restricted aeration at 22 °C and 25 °C, but none in open tubes. This might be due to the increased stress of nutrient deprivation and presence of oxygen.

### 3.4.2 Purification of Native PlnKW30

The mature native bacteriocin PlnKW30 was purified from the supernatant of L. plantarum KW30 cultures as described in section 2.4.8. The last step of this purification
is RP-HPLC, in which PlnKW30 is eluted as a sharp peak at 17 minutes, 36 - 37 % B (Figure 3.14).

Concentrated samples were yellow-brown, which did not appear to affect the bacteriocin activity and no impurities were evident from the NMR spectra (H. Venugopal - personal communication). The discolouration appeared to be due to the autoclaved MRS medium as it could be decreased by not autoclaving the MRS medium. Repetition of the RP-HPLC step also resulted in cleaner bacteriocin. The yield varied between 0.25 – 1 mg/L, which is well above the minimal inhibitory concentration (Section 2.4.10). The antimicrobial activity of the bacteriocin PlnKW30 was routinely tested using a biological assay (Section 2.4.9).

The purified PlnKW30 was subjected to mass spectrometry analysis, which showed that its actual molecular mass (m + H) is 5200.06050 (Figure 3.15) compared to the theoretical mass of 4800.36 Da. The mass difference can be explained by the post-translational modifications of PlnKW30, which are described in section 3.6.
3.4 Characterisation of the bacteriocin plnKW30

3.4.3 Minimum inhibitory concentration of PlnKW30

Purified PlnKW30 was used to analyse its antimicrobial activity. The minimum inhibitory concentration (MIC) was defined as the minimal concentration of the bacteriocin that resulted in a reduction of cell growth by 50% (50% of the turbidity of the control culture without the bacteriocin) (Section 2.4.10). The MIC of PlnKW30 was determined using *L. plantarum* ATCC 8014 as the indicator strain in a 96-well microtiter plate assay. An overnight culture of the indicator strain was diluted 1/100 in 100 µL fresh MRS media and different concentrations of PlnKW30 were added. The cultures were grown at 30 °C and the growth monitored over 7 hours, as shown in figure 3.16. The MIC of PlnKW30 was determined to be 12.5 ng/mL (~2.5 nM). A minimum inhibitory concentration in the nanomolar range makes PlnKW30 comparable to the well known potent bacteriocin nisin.
Chapter 3: Experimental results

3.4 Characterisation of the bacteriocin plnKW30

3.4.4 Antimicrobial activity of PlnKW30 against the indicator strain L. plantarum ATCC 8014

The growth of the producer strain L. plantarum KW30 and the indicator strain L. plantarum ATCC 8014 were measured over a period of 24 hours. The cultures were inoculated with 10% of an overnight culture (stationary phase) and grown at 30°C, without agitation. For each strain measurements were taken from three separate cultures at each time point and averaged to generate the growth curves shown in figure 3.17.
Both cultures have similar growth rates, reaching log phase after about 4 hours growth (Figure 3.17). L. plantarum ATCC 8014 reached stationary phase after about 10 hours, whereas the growth of L. plantarum KW30 continues after 10 hours, slowing after 14 hours of growth.

In the following experiments the antimicrobial activity of purified PlnKW30 was examined on mid-log phase cells (OD$_{600}$ of around 3.0) of the indicator strain L. plantarum ATCC 8014, whereas the MIC experiments were performed at OD$_{600}$ of 0.4. Different concentrations of PlnKW30 (final concentrations are shown) were added and the cells incubated at 30 °C. The cell density of all cultures was measured at particular time points and graphed (Figure 3.18).
Figure 3.18: Growth of *L. plantarum* ATCC 8014 after addition of different concentrations of PlnKW30. PlnKW30 was added to mid-log phase *L. plantarum* ATCC 8014 cells. Each data point represents the mean of measurements from at least two separate cultures with standard errors shown.

The cell density of the indicator strain *L. plantarum* ATCC 8014 without addition of PlnKW30 increased up to an OD$_{600}$ of about 14 within 22 hours. In contrast, the addition of 1 µg/mL PlnKW30 to the cell culture resulted in no further OD$_{600}$ increase over a period of 13 hours, after which a slight increase in cell density was noticeable. The slight increase seen in cultures with 250 ng/mL and 1 µg/mL PlnKW30 between 14 and 22 hours could be due to a lower bacteriostatic effect of PlnKW30 after that time, or to a small proportion of indicator cells that have developed resistance. The antimicrobial effect of PlnKW30 is concentration dependent as visible by the reduction of cell growth with increasing bacteriocin concentration.
3.4.5 Analysis of Antimicrobial Activity of PlnKW30

To analyse the antimicrobial activity of PlnKW30 in more detail (e.g. more quantitative results and possible mode of action), a fluorescent Live/Dead bacterial viability stain (Invitrogen; section 2.4.17) was used. This stain shows dead cells in red and live cells in green (Figure 3.19A) and has been previously used to investigate viability and membrane integrity in other lactobacilli (Caldinia et al. 2005; Marco & Kleerebezem 2008). The results of the live/dead cell assays were shown to be consistent with viability as assessed by plate counts. Counting of live and dead cells after exposure to PlnKW30 thus provides a quantitative measure of the antimicrobial activity of PlnKW30. The death of L. plantarum ATCC 8014 cells after addition of PlnKW30 and other additives was analysed.

The addition of 1 µg/mL (~0.2 µM) PlnKW30 to the indicator strain L. plantarum ATCC 8014 (at OD₆₀₀ ~ 3) resulted in an increase of the percentage of dead cells over thirteen hours compared to L. plantarum ATCC 8014 cells with no added PlnKW30 (Figure 3.19B). After 22 hours about 1/3 of L. plantarum ATCC 8014 cells with PlnKW30 addition were dead (Figure 3.19A). In comparison, the L. plantarum ATCC 8014 cells without added bacteriocin showed just a small increase of dead cells overall (Figure 3.19). These results were obtained in MRS medium, but comparable results were obtained in chemically defined medium (results not shown). Both media types were tested to ensure that these results were not due to a component in the complex medium, as the quantities of each ingredient in MRS are not known.
Figure 3.19: Antimicrobial effect of PlnKW30 on *L. plantarum* ATCC 8014 cells.
A: Fluorescence microscopy images of *L. plantarum* ATCC 8014 cells at 0 and 22 hours with and without addition of 1 µg/mL PlnKW30; each data point represents the mean of at least two measurements of separate cultures; B: Percentage of dead cells of *L. plantarum* ATCC 8014 and *L. plantarum* KW30 with and without addition of 1 µg/mL PlnKW30.
The percentage of dead cells of the producer strain *L. plantarum* KW30 stayed relatively low for 22 hours and there was no significant difference between the cultures grown in the presence of PlnKW30 and those grown in its absence (Figure 3.19B), confirming the immunity of the producer strain to its bacteriocin.

The concentration dependence of the antimicrobial activity of PlnKW30 was tested by adding different concentrations to log phase *L. plantarum* ATCC 8014 cells (OD$_{600}$ ~3.0). The results showed that increased PlnKW30 concentration led to an increased percentage of dead *L. plantarum* ATCC 8014 cells (Figure 3.20).

![Figure 3.20: Concentration-dependent antimicrobial activity of PlnKW30 against L. plantarum ATCC 8014.](image)

Percentage of dead cells of *L. plantarum* ATCC 8014 after the addition of different PlnKW30 concentrations.

Lactobacilli populate rich, carbohydrate-containing environments (e.g. milk products, fermenting vegetation), in which sugars are used as carbon sources for growth and for the generation of energy through fermentation (Brock 1999). Carbohydrate metabolism results chiefly in the production of ATP and NADH (Brock 1999). Bacteria have a number of
different carbohydrate transporters, some of which are relatively non-specific, to translocate sugars across the cell membrane (Lorca et al. 2007; Jahreis et al. 2008).

Work by Diep and co-workers (2007) showed that some class II bacteriocins use components of the mannose phosphotransferase system (man-PTS) of susceptible cells as receptors. If this is the case, the addition of N-acetylglucosamine or other carbohydrates to the cell culture medium might have an effect on the antimicrobial activity of PlnKW30, especially as it is modified with GlcNAc (for a detailed description see section 3.5). Hence, the effect of different sugars on the activity of PlnKW30 was tested by adding 1 µg/mL (~0.2 µM) PlnKW30 plus 1 mg/mL (~4.5 mM) N-acetylglucosamine (GlcNAc), glucosamine, mannose (Man) or N-acetylgalactosamine (GalNAc) to L. plantarum ATCC 8014 cells (Figure 3.21).

Figure 3.21: Effect of different sugars on antimicrobial activity of PlnKW30. The percentage of dead cells of L. plantarum ATCC 8014 at time 0, 6 and 22 hours with the addition of 1 µg/mL PlnKW30 and 1 mg/mL of N-acetylglucosamine (GlcNAc), glucosamine, mannose (Man) or N-acetylgalactosamine (GalNAc) is shown.
The addition of either GlcNAc or glucosamine protected the indicator strain from cell death, resulting in reduced levels of dead cells compared to those seen in the absence of PInKW30 (Figure 3.21). In contrast, the addition of 1 mg/mL Man or GalNAc failed to protect the cells from cell death, with the percentage of dead cells being as high as that seen for the control (Figure 3.21). The tests with 1µg/mL PInKW30 and 1mg/mL GlcNAc or Man were also repeated in chemically defined medium and gave similar results as those obtained with MRS medium (data not shown).

The small differences in the zero time samples are probably due to slight differences in the growth of the cell cultures and not to noise from cell counting, as 300 – 600 cells were counted each time.

This protective effect of GlcNAc was also seen in the cell density test, where the addition of 1 mg/mL GlcNAc resulted in almost complete protection of the cells from the effects of PInKW30 (Figure 3.22). The controls in this figure are the same as shown in figure 3.18.

![Figure 3.22: Growth of *L. plantarum* ATCC 8014 after addition of 1mg/mL N-acetylglucosamine and PInKW30. PInKW30 was added to mid-log phase *L. plantarum* ATCC 8014 cells. Each data point represents the average of two measurements of separate cultures.](image-url)
The effect of increasing concentrations of GlcNAc or glucosamine in the presence of 1 µg/mL PlnKW30 is shown in figure 3.23. As a control, the effects of adding the same concentrations GlcNAc or glucosamine to the indicator strain were tested in the absence of PlnKW30.

Figure 3.23: Effect of different concentrations of GlcNAc and glucosamine on antimicrobial activity of PlnKW30.

The percentage of dead cells of L. plantarum ATCC 8014 at times 0, 6 and 22 hours in the presence of 1µg/mL PlnKW30 and different concentrations of N-acetylglucosamine or glucosamine is shown.

The results clearly show that the protective effect of GlcNAc or glucosamine is concentration-dependent; the higher the GlcNAc or glucosamine concentrations, the smaller the percentage of dead cells. The addition of the same concentration of either GlcNAc or glucosamine to a culture of the indicator strain did not have a significant effect.
on the percentage of dead cells in the culture (Figure 3.23), showing that the sugar itself did not adversely affect the culture.

To prove that even a huge excess of either Man or GalNAc did not have a significant protective effect, 10 mg/mL (~45 mM) of either monosaccharide were added to the cell culture in the presence and absence of PlnKW30 (Figure 3.24).

As expected, neither monosaccharide had a significant effect on the percentage of dead cells or on the growth of the indicator strain in the absence of PlnKW30 (Figure 3.24).

![Figure 3.24: Effect of Man or GalNAc on antimicrobial activity of PlnKW30.](image)

The percentage of dead cells of *L. plantarum* ATCC 8014 at time 0, 6 hours and 22 hours in the presence of 1µg/mL PlnKW30 and different concentrations of mannose and N-acetylgalactosamine is shown.

It is clear from these results that the protective effect is related to the amine group on the second carbon of a glucose molecule as mannose, *N*-acetylgalactosamine and glucose have
no effect. While glucose was not specifically tested for, it is present in the culture media at a concentration of 1 g/L. Thus similar tests involving glucose would have been redundant.

### 3.5 Analysis of N- and C-terminal Fragments of PlnKW30

To simplify the analysis of the modifications of serine 18 and cysteine 43, the purified mature PlnKW30 was digested with trypsin (Section 2.4.11). Figure 3.25 shows a schematic representation of the mature PlnKW30 with possible trypsin cleavage sites underlined; post-translationally modified amino acids (Ser 18 and Cys 43) shadowed and the two nested disulfide bonds indicated by brackets.

![Mature PlnKW30 sequence with trypsin cleavage sites and resulting peptides.](image)

The first lysine is not cleaved off by the trypsin probably because it is followed directly by a proline, as seen by the actual mass of the N-terminal peptide (Figure 3.31). The 32 amino acid long N-terminal peptide with intact disulfide bonds (KPAWCWYTLAMCGAGYSCTCDYMYSYHCFGIK) has a theoretical molecular weight of 3626.18 Da and the 11 amino acid long C-terminal peptide (HHSSGSSSYHC) of 1188.2 Da. Both fragments were purified by RP-HPLC, eluting at 9.5 minutes (17 % B) and 22 minutes (44 % B) (Figure 3.26).
The C-terminal peptide is very hydrophilic and elutes earlier than the relatively hydrophobic N-terminal peptide. In comparison, the full length PlnKW30 elutes on a C18 column between both fragments at 40 % B, which is to be expected as it consists of both peptides.

Figure 3.26: RP-HPLC chromatogram of the N- and C-terminal fragments of PlnKW30.
A Jupiter C18 column was used with the mobile phase A: H₂O, 0.1 % TFA and B: Acetonitrile, 0.08 % TFA, with the conditions used as shown in chromatogram. The C-terminal fragment eluted at 9.5 minutes (17 % B) and the N-terminal fragment eluted at 22 min (44 % B).

The RP-HPLC purified N- and C-terminal peptides, along with the synthetic C15 C-terminal peptide (FGIKHSSGSSYHC without GlcNAc; see section 3.5.1) were analysed by tricine SDS-PAGE in reducing and non-reducing sample buffer, then after staining the gels were soaked in water to remove residual SDS. The gels were then overlaid with the indicator strain L. plantarum ATCC 8014 at a cell density of about 3 in MRS (1 % agar) and incubated at 30 °C for 16 hours (Figure 3.27).
Figure 3.27: Tricine SDS-PAGE of PlnKW30 and its N- and C-terminal fragments overlaid with 
*L. plantarum* ATCC 8014.

Samples of PlnKW30, its N- and C-terminal fragments and synthetic C15 C-terminal peptide were prepared in reducing and non-reducing sample buffer and separated on tricine SDS-PAGE (16 %, 6 % bisacrylamide) and then overlaid with the indicator strain *L. plantarum* ATCC 8014 OD600 of 3.

**A:** 1: PlnKW30 in non-reducing buffer; 2: PlnKW30 in reducing buffer;

**B:** 1: N-terminal fragment in non-reducing buffer; 2: N-terminal fragment in reducing buffer; 3: C-terminal fragment in reducing buffer; 4: synthetic C15 C-terminal fragment in reducing buffer; 5: synthetic C15 C-terminal fragment in non-reducing buffer; **M:** protein standard (kDa).

The native PlnKW30 showed strong antimicrobial activity in the non-reducing buffer as shown by the clearing around the PlnKW30 band (Figure 3.27A, lane 1), whereas the N-terminal peptide showed no activity (Figure 3.27B, lane 1). Equal amounts of sample were loaded in reducing and non-reducing buffer, but it seems that reduced samples stain better. Reduction of the disulfide bonds, which tether the two helices together (Section 3.5.1) results in a loss of structure, as indicated by CD (Section 3.7). This loss of structure will enable the dye molecules to access the polypeptide chain more easily, and make side chains available for the binding of dye molecules. The C-terminal peptide could not be visualised on the gel and no clearing was visible (Figure 3.27B, lane 3). This is most likely because of the size of the peptide and the presence of a GlcNAc on the C-terminal end will act as shield, preventing the dye molecules accessing the polypeptide. It is interesting to note the relatively high apparent molecular mass of the synthetic C-terminal peptide (Figure 3.27B, lane 4), which runs just above the three times larger N-terminal fragment.
This peptide is only four amino acids longer, including a lysine, than the C-terminal fragment of PlnKW30 and is not glycosylated, which will make it more hydrophobic than the real C-terminal peptide. The full length PlnKW30 has a molecular mass of about 5.2 kDa (Figure 3.15), but runs at about 7 kDa and the N-terminal fragment with a molecular weight of about 3.8 kDa (Figure 3.31) runs at about 5.5 kDa, which indicates that peptides generally run differently on tricine SDS-PAGE. It is possible that the synthetic peptide dimerises by forming disulfide bonds through the free cysteine at the C-terminal end, which would increase its molecular weight to 3266 Da (2 x 1633 Da), a mass still well short of that indicated by the band. The fact that its retention time in the presence of DTT is higher than in the absence of DTT (Figure 3.27B; lanes 4 and 5) suggests that the free cysteine may be having an effect on the anomalous behaviour of the peptide. However, it may be possible that the peptides are forming strong aggregates. Such a result has been observed before by Chesneau et al (2000) who were investigating the effect of enzyme degradation of amyloid β protein. In this study, a peptide with an expected molecular weight of 2.5 kDa had an observed molecular weight of 7.9 kDa on a tricine SDS-PAGE. This peptide contained two histidines and a tyrosine, residues also contained in the C-terminal peptide of PlnKW30.

Additionally, the N- and C-terminal fragments of PlnKW30 were analysed using the Live/Dead cell assay (Section 2.4.17). Both fragments were added separately (0.2 µM), as well as together to the indicator strain *L. plantarum* ATCC 8014 (Figure 3.28).
3.5 Analysis of N- & C-terminal fragments

Figure 3.28: Antimicrobial effect of N- & C-terminal fragments of PlnKW30.
Percentage of dead cells of *L. plantarum* ATCC 8014 at time 0, 6 and 22 hours with addition of ~0.2 µM PlnKW30, N- or C-terminal fragment separately or together.

Neither the N-terminal fragment, or the C-terminal fragment, or the combination of both fragments displayed any significant bactericidal activity (Figure 3.28).

The effect of the N-terminal fragment on the growth of the indicator strain was also measured by OD (Figure 3.29). The controls in this figure are the same as shown in figure 3.18.
Using this method it was confirmed that the N-terminal fragment of PInKW30 does not have an antimicrobial effect on *L. plantarum* ATCC 8014, as the OD$_{600}$ of this culture reached a similar density to that of a culture to which no PInKW30 had been added (Figure 3.29).

However, it has been observed that the N-terminal fragment shows some antimicrobial activity against the indicator strain at a much lower OD$_{600}$ of about 0.3 (Figure 3.30). This antimicrobial activity is less effective than that of the full length PInKW30, but it supports the observation that younger cells of *L. plantarum* ATCC 8014 at the beginning of log phase are more susceptible than cells in mid-log phase. Another explanation could be that the C-terminal fragment is necessary for bactericidal activity, but not for bacteriostatic activity, as the N-terminal fragment appears to be slightly bacteriostatic.
Figure 3.30: Tricine SDS-PAGE of the N-terminal fragment overlaid with *L. plantarum* ATCC 8014 at OD<sub>600</sub> of 0.3.

Samples of the N-terminal fragment were prepared in reducing and non-reducing sample buffer and separated on tricine SDS-PAGE (16 %, 3 % bisacrylamide) and then overlaid with the indicator strain *L. plantarum* ATCC 8014 OD<sub>600</sub> of 0.3. 1: N-terminal fragment in non-reducing buffer; 2: N-terminal fragment in reducing buffer; M: protein standard (kDa).

In this gel the reduced band did not migrate as far as the non reduced band. This is probably because the sample was not completely reduced. When PlnKW30 was reduced by the addition of increasing concentrations of TCEP, the reduction was monitored by FTMS (Dr. G.E. Norris – personal communication), which showed that one disulfide was relatively easy to reduce, while the second was much harder, requiring higher concentrations of TCEP and heating at 65 C. It is known that the electrophoretic mobility of proteins can vary with the reduction of internal disulfide bonds. This variation appears as a difference in the apparent mass of 2 - 4 kDa (Schägger 2006).

3.5.1 **IDENTIFICATION OF THE C-TERMINAL GLYCAN MODIFICATION**

(a) Mass spectrometry of N- and C-terminal fragments

The purified N-terminal (Figure 3.31) and C-terminal fragments (Figure 3.32) of PlnKW30 were analysed by mass spectrometry. The monoisotopic molecular masses of the N- and C-terminal fragments were measured to be 3827.55043 Da and 1391.52972 Da, respectively.
The purified C-terminal fragment was digested with chymotrypsin (Section 2.4.11), resulting in an HC-X fragment, which was purified by RP-HPLC. Thereafter, it was analysed by MS/MS (ESI positive mode) on a micrTOFQ (Bruker). Figure 3.33 shows the spectra recorded which shows the C-terminal (m + H) fragment had a monoisotopic mass of 462.1626. Collision induced fragmentation (CID) resulted in a major ion with m/z of 259.0849, which is the monoisotopic mass of the dipeptide HC minus an N-acetylglucosamine. The ion with m/z 204.0881 is the monoisotopic mass of the protonated N-acetylglucosamine (m + H). If the modification had been a farnesyl group, the mass difference would have been 203.1794 due to the loss of C_{15}H_{23}. 
As can be seen, the difference between an N-acetylhexosamine and a farnesyl group is very small (about 0.1 mass units) and it is hardly surprising with the other circumstantial evidence that the modification was initially thought to be a farnesyl group. After all, while C-terminal farnesyl groups are common in eukaryotes, S-glycosides have not been previously observed in any genera.

Tandem mass spectrometry (MS/MS) fragmentation using electron transfer dissociation (ETD) was performed on the full length PlnKW30 by Dr. P. Novak (Laboratory of Molecular Structure Characterization, Institute of Microbiology, Academy of Sciences of the Czech Republic). The results are listed in appendix 12 and clearly show the addition of a mass of 203 to both the serine 18 and the cysteine 43.

Recently, the structure of PlnKW30 was solved using $^1$H-NMR spectroscopy (Figure 3.34; H. Venugopal – personal communication).
Figure 3.34: Cartoon representation of the NMR structure of PlnKW30. 
The two helices are linked by disulfide bridges between Cys 12 – Cys 21 and Cys 5 – Cys 28. The two 
N-acetylglucosamines on Ser 18 and Cys 43 are shown in stick representation (PyMol; DeLano 2002).

This structure confirms that both serine 18 and cysteine 43 are modified by β1-linked 
N-acetylglucosamines (Figure 3.34). For serine 18 the link is through the side chain 
hydroxy oxygen; for the cysteine 43 the link is through the side chain thiol sulphur. The 
structure of PlnKW30 is stabilised by two nested disulfide bonds, between residues 5 and 
28 and 12 and 21. These cysteines form the end of the two helices, the first between 
residue 5 and 12, which contains elements of both α- and $3_{10}$-helices. The secondary 
structural characteristics of this helix seem to be shifting dynamically between those of an 
α-helix and a $3_{10}$-helix. A $3_{10}$-helix contains three amino acids within a 360° turn, stabilised 
by hydrogen bonds between residues $i$ and $i + 3$, in contrast to the more usual α-helix 
which has 3.6 residues per turn with hydrogen bonds between residues $i$ and $i + 4$. The 
second helix between residues 21 and 28 has a conformation that is more in line with an 
α-helix. The two helices are connected by a loop, which contains the glycosylated serine. 
This loop appears to be reasonably inflexible, which is unusual, and may reflect its role in 
the activity of PlnKW30.
(b) Wheat germ agglutinin pull-down experiments

Wheat germ agglutinin (Vector Laboratories Inc., USA) is a plant lectin which specifically binds β-N-acetylglucosamine and N-acetylneuraminic acid residues and was used to pull-down purified PlnKW30. The pre-equilibrated lectin-resin was incubated with the purified PlnKW30 for 2 hours at room temperature with gentle agitation (Section 2.4.16). The samples were then centrifuged to collect the lectin-resin, the supernatant removed and the lectin-resin washed three times with lectin buffer (Section 2.4.16). 10 and 20 µL of the supernatant (samples 2 & 3 in figure 3.35) and lectin-resin (samples 4 & 5 in figure 3.35) were then analysed using a tricine SDS-PAGE.

![Figure 3.35: Tricine SDS-PAGE of lectin pull-downs of PlnKW30.](image)

The slight difference in mobility of the PlnKW30 control and the pulled-down PlnKW30 samples can be explained by the higher concentration of the control sample and by incomplete reduction of the PlnKW30 bound to the lectin-resin on the beads. The gel clearly shows that PlnKW30 bound tightly to the resin as it was not present in the supernatant, indicating that PlnKW30 must be modified with N-acetylglucosamine.
(c) Dot blot using anti-\(O\)-GlcNAc antibodies

Dot blots (Section 2.4.6(c)) using monoclonal antibodies specific to \(O\)-linked \(N\)-acetylglucosamine were carried out to confirm the glycosylation linkage present in PlnKW30, where \(\beta\)-\(N\)-acetylglucosamine is attached to the serine residue via an \(O\)-linkage and to the cysteine residue via an \(S\)-linkage. The mature PlnKW30 and its \(N\)- and \(C\)-terminal fragments were the samples used for this analysis (Figure 3.36).

![Figure 3.36: Dot blot using anti-\(O\)-GlcNAc antibodies to detect PlnKW30 and its \(N\)- & \(C\)-terminal tryptic fragments.](image)

1: 4 µg mature PlnKW30; 2: N-terminal fragment; 3: C-terminal fragment.

The blot shows the high specificity of the anti-\(\beta\)-\(O\)-GlcNAc antibodies, which detect the full length PlnKW30 and its \(N\)-terminal fragment containing the \(O\)-GlcNAc-linkage, but not the \(S\)-GlcNAc-linkage of the \(C\)-terminal peptide.

As the \(C\)-terminal fragment was not visible on tricine SDS-PAGE, it was shown to be present, along with the \(N\)-terminal fragment and the synthetic \(C\)-terminal C15 fragment, by using ninhydrin. Ninhydrin reacts with primary and secondary amines, resulting in a purple colour. The samples were spotted on thin layer chromatography (TLC) plates, treated with ninhydrin and baked at 110 °C for about 10 min (Figure 3.37).

![Figure 3.37: Visualisation of \(N\)- and \(C\)-terminal fragments using ninhydrin.](image)

The peptide samples were spotted on TLC plates and sprayed with ninhydrin which reacts with peptides resulting in a purple colour. 1: \(N\)-terminal fragment; 2: \(C\)-terminal fragment; 3: synthetic C15 \(C\)-terminal fragment.
The spot of the synthetic peptide is more diffuse than that of the other two samples. This sample differs from the other two by being diluted in 10 mM HEPES buffer with 10 mM DTT and it lacks carbohydrates, whereas the other samples were diluted in pure H₂O and contain one or two GlcNAcs. The purple colour of the spots indicates the reaction of ninhydrin with the peptides and confirms the presence of protein in all samples.

3.6 Analysis of the post-translational modifications of PlnKW30

3.6.1 Analysis of the disulfide bridges of PlnKW30

The two disulfide bonds of PlnKW30 were broken by reduction and alkylation (Section 2.4.11) and the reduced and alkylated PlnKW30 was subjected to RP-HPLC (Figure 3.38) and mass spectrometry (Figure 3.39).

![Figure 3.38: RP-HPLC chromatogram of native and reduced/alkylated PlnKW30.](image)

A Jupiter C4 column was used with the solutions A: H₂O, 0.1 % TFA and B: Acetonitrile, 0.08 % TFA, conditions used as shown in chromatogram. 1 (red): native PlnKW30 (eluted at 16 min, 36 % B); 2 (black): reduced and alkylated PlnKW30 (eluted at 10 min, 30 % B).
Chapter 3: Experimental results

3.6 Analysis of post-translational modifications

The reduced and alkylated sample of PlnKW30 elutes significantly earlier (10 min; 30 % B) compared to the native PlnKW30 sample (16 min; 36 % B) (Figure 3.38), which is expected due to the addition of seven ethylacetate groups to the protein.

Mass spectrometry results revealed the addition of seven ethylacetate groups (7 x 57 Da). As iodacetamide is not specific for cysteines, and under some conditions acetylates serine residues, it is most likely that in addition to the four cysteines, three of the five serines in the C-terminal tail are acetylated. Whether these additional acetyl groups will affect the antimicrobial activity of the peptide is unknown. Based on comparison with other class II bacteriocins (Section 1.3.2) and the analysis of reduced but not alkylated PlnKW30 (Figure 3.41), it is most likely that it is the disruption of the two disulfide bonds which results in the loss of antimicrobial activity.

![Figure 3.39: Monoisotopic mass spectrometry result of the reduced and alkylated PlnKW30.](image)

The reduced and alkylated bacteriocin PlnKW30 (m + H) has a molecular mass of 5717.29542 Da, which includes the addition of seven ethylacetyl groups (57 mass units each) (ESI, positive mode, Bruker Apex-Q-FTMS (9.4T Dual Source), Laboratory of Molecular Structure Characterization, Academy of Sciences of the Czech Republic).

The reduced and alkylated PlnKW30 was analysed by tricine SDS-PAGE (Section 2.4.2) and overlaid with MRS 1 % agar containing the indicator strain *L. plantarum* ATCC 8014 (Figure 3.40).
3.6 Analysis of post-translational modifications

Figure 3.40: Activity of native and reduced PlnKW30 in a bioassay overlay.
About 1 µg of peptides were loaded. 1: reduced + alkylated; 2: native PlnKW30; both samples were loaded in non-reducing buffer.

Growth of the indicator strain was not inhibited around the reduced and alkylated sample (lane 1), which shows that reduction and alkylation resulted in total loss of antibacterial activity of PlnKW30. In contrast, native PlnKW30 exhibits antimicrobial activity, which is visible by the clearing around the bacteriocin band (lane 2). These results are consistent with those shown in figure 3.27, where PlnKW30 was reduced but not alkylated.

Furthermore, the role of the disulfide bonds of PlnKW30 was analysed using the Live/Dead cell assay (Section 2.4.17). Native PlnKW30 was reduced by incubation with 10 mM TCEP in pure water for 20 minutes at 25 °C and then added to the L. plantarum ATCC 8014 cells (Figure 3.41). As a negative control, 10 mM TCEP only was also added to the cells.

As shown in figure 3.41, no antimicrobial activity of the reduced PlnKW30 was detected after 22 hours. The percentage of dead cells in the no-bacteriocin control was comparable to that found in the culture to which TCEP had been added, as well as the culture to which reduced PlnKW30 had been added (see 22 hours in figure 3.41). The combined evidence therefore strongly suggests that the disulfide bonds in PlnKW30 are required for the antimicrobial activity of PlnKW30.
Figure 3.41: Antimicrobial activity of reduced PlnKW30.
Percentage of dead cells of *L. plantarum* ATCC 8014 at time 0, 6 and 22 hours with addition of 1 µg/mL reduced PlnKW30; each data point represents the mean of at least two measurements of separate cultures.

3.6.2 **ANALYSIS OF THE GLYCAN MODIFICATIONS OF PlnKW30**

Characterisation of both the N- and C-terminal fragments of PlnKW30 is discussed in section 3.5, including the identification of the post-translational modification of both fragments.

To further characterise the antimicrobial activity of PlnKW30, it was deglycosylated using an *N*-acetyl-β-**d**-glucosaminidase (Section 2.4.13) to remove the O-glycosidic bond *N*-acetylhexosamine from the serine. It was not possible to remove the S-linked glycan from the cysteine, as S-glycosidic bonds are much more stable than O-linked (Zhu *et al.* 2004; Liang *et al.* 2009). The treatment resulted in a slightly later elution of the
O-deglycosylated sample from the RP-HPLC column compared to the untreated PnKW30 (Figure 3.42).

Analysis by mass spectrometry confirmed that only one N-acetylglucosamine (GlcNAc) had been removed from PnKW30 by the enzyme-treatment and that the disulfide bonds remained intact (Figure 3.43). The resulting deglycosylated PnKW30 had a mass of 4996.9827 m/z, which is 5200.0605 m/z minus 203.0778 m/z of the GlcNAc.

The deglycosylated PnKW30 was subjected to collision induced fragmentation (CID) and no fragment ions corresponding to serine 18 modified with an O-linked GlcNAc could be detected. In contrast, only ions with a modified C-terminal cysteine residue could be detected (Dr. P. Man – personal communication).
Figure 3.43: Monoisotopic mass spectrometry result of the deglycosylated PlnKW30.
The native PlnKW30 (m + H) has a molecular mass of 5200.0593 and the O-deglycosylated PlnKW30 (m + H) of 4996.9796. Other peaks correspond to sodiated and potassiated molecules. (ESI, positive mode, Bruker Apex FTMS (9.4T Dual Source); Dr. P. Man, Laboratory of Molecular Structure Characterization, Academy of Sciences of the Czech Republic)
The deglycosylated PlnKW30 was subjected to the biological plate assay (Section 2.4.9) as well as to the Live/Dead cell assay (Section 2.4.17). The tricine SDS-PAGE overlaid with the indicator strain (Figure 3.44) consistently showed a decrease in the antimicrobial activity of the deglycosylated PlnKW30 compared to the native bacteriocin.

Figure 3.44: Activity of native and deglycosylated PlnKW30 in a bioassay overlay.
About 1 µg of peptides were loaded in non-reducing buffer on a tricine SDS-PAGE, which was overlaid with the indicator strain. 1: native PlnKW30; 2: deglycosylated PlnKW30.

A decrease in bactericidal activity of the deglycosylated PlnKW30 was confirmed by the Live/Dead cell assay (Figure 3.45). The percentage of dead cells after addition of the native PlnKW30 was at about 23%, whereas the deglycosylated PlnKW30 resulted in only about 7% of dead cells, which is a decrease of about 2/3, and is consistent with what is seen in the biological plate assay (Figure 3.44).
Figure 3.45: Antimicrobial effect of deglycosylated PlnKW30.
Percentage of dead cells of *L. plantarum* ATCC 8014 at time 0, 6 and 22 hours with addition of 1 µg/mL deglycosylated PlnKW30; each data point represents the mean of at least two measurements of separate cultures.
3.7 **CIRCULAR DICHROISM ANALYSIS OF PLNKW30**

Circular dichroism (CD) spectroscopy is a form of light absorption spectroscopy that measures the difference in absorbance of right- and left-circularly polarized light (rather than the commonly used isotropic light) by a substance. It has been shown that CD spectra in the far UV (between 260 and approximately 180 nm) are indicative of the different secondary structural elements present in proteins: α-helices, parallel and antiparallel β-sheets, β-turns and random coils.

The native, reduced and O-deglycosylated PlnKW30, as well as the native, reduced and O-deglycosylated N-terminal fragments, were analysed by CD spectroscopy using a Chirascan CD spectrometer. The experimental conditions used for these scans are described in section 2.5.3. The spectra of the reduced samples were measured using pure H₂O with TCEP as baseline, which was subtracted from the spectra. The residuals for all spectra are shown in figures 3.46, 3.47 and 3.48 and show that the noise is randomly distributed around zero, which indicates that no distortion of the spectra occurred during the smoothing process.

The spectra of the native and O-deglycosylated PlnKW30 samples and their corresponding residuals are shown in figure 3.46.
Chapter 3: Experimental results

3.7 Circular dichroism analysis of PlnKW30

Figure 3.46: Smoothed, base-line corrected spectra of native PlnKW30 (A), deglycosylated PlnKW30 (B) and their corresponding residuals.

The spectra of the reduced and alkylated and reduced (using TCEP) PlnKW30 and their residuals are shown in figure 3.47.
Figure 3.47: Smoothed, base-line corrected spectra of reduced and alkylated PInKW30 (A), reduced PInKW30 (B) and their corresponding residuals.

The spectra of the reduced and alkylated PInKW30 sample and the TCEP-reduced sample are very similar. The only difference is visible between 185 to 190 nm, where the spectra of the reduced PInKW30 sample are uneven, as evident by the residual. This is probably due to the presence of TCEP in the buffer, although the baseline was measured with TCEP.

The spectra of the native, reduced (using TCEP) and O-deglycosylated N-terminal fragment and their residuals are shown in figure 3.48.
Figure 3.48: Smoothed, base-line corrected spectra of the N-terminal fragment (A), reduced N-terminal fragment (B) and their corresponding residuals.
The CD spectra of the native and O-deglycosylated N-terminal fragment both show helical characteristics, whereas the spectrum of the reduced N-terminal fragment shows the total loss of structure. The residual spectra are randomly distributed for all samples, showing that the smoothing process did not modify the spectra. The residual of the reduced N-terminal fragment is slightly uneven between 185 to 190 nm, which is probably due to the TCEP in the buffer, although the baseline also contained TCEP.

3.7.1 **Comparison of native, reduced and deglycosylated PlnKW30**

In order to compare the spectra of the native, reduced and O-deglycosylated PlnKW30, all spectra were plotted on the same graph (Figure 3.49).

The spectra of the native and O-deglycosylated PlnKW30 show the typical spectra of a helical peptide. The spectrum of the reduced PlnKW30 shows a strong decrease in helicity, indicating that both disulfide bonds have been reduced, resulting in loss of secondary structure.
3.7.2 *Comparison of native, reduced and deglycosylated N-terminal fragment*

For comparison of the native, reduced and O-deglycosylated N-terminal fragment all spectra were plotted in the same graph (Figure 3.50).

![Circular dichroism spectra](image)

*Figure 3.50: Circular dichroism spectra of native, reduced and deglycosylated N-terminal fragment.*

The spectra of the native and O-deglycosylated N-terminal fragment show the typical bands of a helical protein. The native N-terminal spectra also shows a slight peak at about +230 nm, which is probably due to the aromatic residues, particular tryptophans which have positive bands in the 220 – 230 nm region (Woody 1994). The reduced N-terminal spectrum shows the total loss of structure, due to the reduction of both disulfide bonds.
3.7.3 *Comparison of PlnKW30 with the N-terminal fragment*

The spectra of the native PlnKW30 and the native N-terminal fragment are compared in figure 3.51.

![Circular dichroism spectra of native PlnKW30 and native N-terminal fragment.](image)

The comparison of the full length native PlnKW30 and its N-terminal fragment shows that both peptides have helical spectra. This was expected, because the only structural features, the two helices, are present in the N-terminal fragment and the small C-terminal fragment is unstructured. The increased peak at about 230 nm in the spectra of the N-terminal fragment is probably due to the large amount of aromatic acids, as already discussed above.
4 CHAPTER 4: DISCUSSION

4.1 PHYLOGENETIC CLASSIFICATION OF L. PLANTARUM KW30

The Lactobacillus plantarum strain designated L. plantarum KW30 was first identified by Kelly and co-workers (1994). In order to phylogenetically identify this strain, a 2,155 bp 16S - 23S gene fragment was amplified and sequenced (Section 3.1) in order to compare it to other sequences, especially other L. plantarum strains, in public databases. The NCBI nucleotide database has 48 entries of 16S – 23S sequences of L. plantarum strains.

The blastn search showed that the L. plantarum KW30 sequence is closely related to the only completely sequenced L. plantarum strain: L. plantarum WCFS1 (Kleerebezem et al. 2003) (Section 3.1), with the next closest homologues being L. brevis ATCC 367 (Makarova et al. 2006), L. sakei 23K (Chaillou et al. 2005), L. plantarum ZDY36a and L. pentosus NRIC 1837. These last two genomes are only partially sequenced (mainly 16S rRNA), in contrast to the L. plantarum WCFS1 genome sequence where the coverage is 100 % with a maximum identity of 98 %. The next closest homologue is L. brevis ATCC 367 which has 98 % query coverage and 93 % maximum identity. None of the other L. plantarum strains in the NCBI nucleotide database showed high identity to L. plantarum KW30. After showing that L. plantarum KW30 is closely related to L. plantarum WCFS1, a genomic blast search was carried out, resulting in the tree shown in figure 3.1. These results confirmed the close relationship between L. plantarum KW30 and L. plantarum WCFS1, followed by a more distant relationship to L. brevis ATCC 367. These three organisms are positioned in a separate branch away from the majority of other Lactobacillus species, indicating an early separation and divergent development from the other species.

Lactococcus lactis subsp. lactis II1403 was used as an outgroup for the tree alignment. It was chosen because it is closely related to lactobacilli, but distant enough to be clearly classified as an outgroup. Any conclusions drawn from this phylogenetic tree should however be treated with caution, because a simple algorithm (BLAST, NCBI; Desper & Gascuel 2004) was used to create it and only a single gene fragment was analysed. It is possible that the analysis of only one gene from a number of different species might not
accurately represent their evolutionary history. For instance, the gene analysed might be one acquired by horizontal gene transfer from a species that is not a close neighbour. However, ribosomal RNA (rRNA) genes are extremely conserved, which is the reason why 16S and 23S rRNA genes are frequently used to identify the taxonomy of organisms (Tannock et al. 1999; Pena et al. 2004; Delfederico et al. 2006).

In this case we did not want to provide a detailed analysis of the genetic evolution of lactobacilli, but rather show that *L. plantarum* KW30 is a *Lactobacillus plantarum* strain. The preliminary analysis of the whole genome sequence data of *L. plantarum* KW30 shows about 89.5% of the *L. plantarum* WCFS1 genome is covered by the Solexa reads (Section 4.22), further confirming that KW30 is closely related to WCFS1.

### 4.2 Analysis of the *plnKW30* gene cluster

Sequencing of the *plnKW30* gene cluster by primer walking was laborious and time consuming, especially because several regions were difficult to sequence and numerous primers at slightly different positions had to be designed. The difficulties encountered were most likely due to secondary structures formed by the DNA, and one region could only be sequenced using a sequencing reaction chemistry specifically designed to minimise template secondary structures. Other reasons could have been low primer affinity and specificity, because of the low G + C content in *L. plantarum*, resulting in large stretches of A and T rich sequences. The final sequence obtained for the *plnKW30* gene cluster and ORFs in its vicinity was validated by sequence obtained for the whole *L. plantarum* KW30 genome using the Solexa Genome Analysis Facility (Allan Wilson Centre Sequencing Service, Palmerston North; unpublished results). Interestingly, the small region that could not be sequenced using conventional sequencing reaction chemistry was also absent from the Solexa data (Dr. M.L. Patchett – personal communication).

Downstream of the *gcc* genes in the *plnKW30* cluster (Figure 3.2) there are open reading frames (ORFs) with homologies to a helicase, resolvase and permease, which are proteins.
involved in separation of DNA strands, mediation of site-specific recombination, and facilitation of diffusion through the membrane, respectively. The putative helicase has a confirmed full length orthologue on pMRC01, a 60 kb \textit{L. lactis} plasmid, which is also adjacent to a bacteriocin locus. While the helicase and resolvase are unlikely to be involved in mediating bacteriocin synthesis and modification directly, they may be important for generating bacteriocin diversity. Furthermore, an ORF was identified upstream from the GTase gene that has sequence homology to a transposase from \textit{Enterococcus hirae}. Transposases are known to mediate the excision and insertion of movable DNA sequences called transposons. In bacteria, transposons usually contain several genes conferring, for example, resistance to antibiotics. The genes for the biosynthesis of the lantibiotic nisin are also located on a 70 kbp conjugative transposon (Dodd \textit{et al.} 1990; Rauch & Devos 1992).

The presence of transposase, helicase and resolvase ORFs flanking the \textit{plnKW30} gene cluster points to the possibility that the functional cluster coalesced as a result of recombination events, and that some or all cluster genes may have been acquired by horizontal gene transfer.

It has been shown that lactic acid bacteria undergo extensive gene loss and gene gain via horizontal gene transfer in order to adapt to their habitats (Makarova \textit{et al.} 2006). Horizontal gene transfer between bacteria usually occurs by natural competence or bacteriophage infection, and although \textit{L. plantarum} has never been reported to be naturally competent, it contains gene encoding for DNA binding and uptake proteins that are similar to those found in \textit{B. subtilis} (Dubnau & Lovett 2002). These findings show that there is a reasonable possibility that the \textit{plnKW30} gene cluster was acquired via horizontal gene transfer, especially when the \textit{plnKW30} gene cluster is compared to clusters from other bacteria that have similar gene arrangements (Figure 4.1). The \textit{plnKW30} gene cluster may have arrived as a functional block of genes or \textit{L. plantarum} KW30 may have acquired one or more genes in separate events and then rearranged them into a functional bacteriocin operon. The lack of a consensus gene order for the \textit{plnKW30}-like clusters in figure 4.1 suggests that these clusters are not evolutionarily related, but have resulted from convergent evolution.
Chapter 4: Discussion 4.2 Analysis of the plnKW30 gene cluster

Figure 4.1: Comparison of the plnKW30 gene cluster with similar gene clusters.

PlnKW30 gene cluster from *L. plantarum* KW30; Gene cluster on pTEF2 plasmid from *Enterococcus faecalis* V583; Sublancin 168 gene cluster from *Bacillus subtilis* 168; Gene cluster on pE33L466 plasmid from *Bacillus cereus* E33L. ABC-transporter: ATP binding cassette transporter; bac: bacteriocin; GTase: glycosyltransferase; ORF: open reading frame; pln: plnKW30; RR: response regulator; SunA: sublancin 168; TRX: thioredoxin.

The similarities between the gene clusters of *Bacillus subtilis* 168 (Kunst et al. 1997), *Bacillus cereus* E33L (Han et al. 2006), *Enterococcus faecalis* V583 (Paulsen et al. 2003) and the plnKW30 gene cluster are obvious (Figure 4.1). All contain ORFs for at least one bacteriocin (four in *B. cereus* E33L), a GTase (except in *B. cereus* E33L), an ABC-transporter and two putative thioredoxins (Figure 4.1). Interestingly, only the plnKW30 cluster contains a response regulator gene, which might indicate that it is not actually a functional gene. The RT-PCR experiments support this, as the response regulator gene was the most difficult transcript to detect (Section 3.3.1). It might be possible that these clusters are not regulated by the typical quorum sensing two-component signalling systems, which is
supported by the fact that no histidine protein kinases have been identified in these clusters. The arrangement of the genes is slightly different in each cluster.

It is interesting to note that the cluster from *E. faecalis* V583 has a set of ORFs with a very similar order to those of the *plnKW30* gene cluster, apart from the insertion of two unidentified open reading frames and a transposase gene between the ABC-transporter and the two TRXs. Sections of these unidentified ORFs, the 3’ end of ORF1 and the 5’ end of ORF2, lie directly adjacent to the transposase gene and have homology to a transposase from the insertion sequence element IS256. It is quite likely that the two ORFs and the transposase gene inserted themselves into the cluster and disrupted its arrangement. The cluster of *B. cereus* E33L lacks a GTase gene, but encodes four putative bacteriocin genes. The plantaricin locus of *L. plantarum* KW30 has also been disrupted by an IS30 transposase between *plnQ* and *plnA* (Dr. M.L. Patchett – personal communication), suggesting that transposase-mediated recombination and disruption of bacteriocin gene clusters is not uncommon.

All the bacteriocins encoded in these four clusters, and putative bacteriocins of *Streptococcus mutans* V1996 and *Streptococcus suis* 89/1591 show a conserved structural motif of nested disulfide bonds, which is illustrated in figure 4.2 using the amino acid sequence of the mature PlnKW30 as an example.

![Figure 4.2: Schematic representation of the mature PlnKW30.](image)

The nested disulfide bonds are indicated by vertical lines between cysteines. The presence of six amino acids between the disulfide bonds, and the loop region, are indicated by braces.

The consensus structural motif contains two nested disulfide bonds, which are divided by a loop region of 7 – 25 amino acids, with the disulfide bond forming cysteines being six amino acids apart from each other. Figure 4.3 shows an alignment of bacteriocins from the
four clusters in figure 4.1, and from *S. mutans* V1996 and *S. suis* 89/1591 that demonstrates the similar positioning of these structural features.

A comparison of the amino acid sequences of the loop regions of PlnKW30, EF_bac, SM_bac and SS_bac in figure 4.3 reveals some weak sequence similarities. This suggests that the most likely position for an *O*-linked glycan is S33 or S34 in the mature EF_bac, T23 or S24 in the mature SM_bac and T25 or S26 in the mature SS_bac. An *O*-linked glycan modification is proposed, because of the presence of glycosyltransferase ORFs in these clusters.

The four bacteriocins encoded by *B. cereus* E33L have almost identical amino acid sequences up to the first cysteine, after which the sequences show some differences, with only L41, K48 and L55, in addition to the four cysteines, invariant. These bacteriocins probably originated from the duplication of one gene; the duplicated genes were then individually modified to produce bacteriocins with different specificities, making it harder for sensitive strains to develop immunity against them.

Sublancin 168 contains two nested disulfide bonds between cysteine 7 and 36, and cysteine 14 and 29 (Paik *et al*. 1998), which is a similar arrangement to the disulfides between cysteine 5 and 28, and cysteine 12 and 21 in PlnKW30. Both bacteriocins contain a loop region, consisting of 8 residues in PlnKW30 and 14 in sublancin 168, and have similar lengths of 43 and 37 amino acids, respectively (Figure 4.3). Dorenbos *et al*. (2002) looked at the potential role of the two thioredoxin genes in the sublancin 168 gene cluster in the formation of the two disulfide bonds of sublancin 168. They found that only TRX2 (BdbB) is essential for sublancin 168 production, but not TRX1 (BdbA). It is interesting to note that TRX2 is encoded at the end of the operon, as it is the essential TRX (Figure 4.1). Amino acid sequence alignments of both TRXs from *L. plantarum* KW30 to each of the TRXs from *B. subtilis* 168 showed that both are more similar to TRX1 of *B. subtilis* 168. This suggests that the two thioredoxins in the PlnKW30 cluster may not be necessary for the disulfide bond formation of the mature PlnKW30; rather they might be implicated in keeping it reduced while in the cytoplasm. The involvement of the KW30 TRX1 and 2 in the formation of the PlnKW30 disulfide bonds could be tested by gene disruption studies similar to those described by (Dorenbos *et al*. 2002) for sublancin 168.
Chapter 4: Discussion

4.2 Analysis of the plnKW30 gene cluster

Figure 4.3: Comparison of the bacteriocin amino acid sequences.
PlnKW30: PlnKW30 of *L. plantarum* KW30; EF_bac: putative bacteriocin of *E. faecalis* V583; SunA: sublancin 168 of *B. subtilis* 168; BC_1-4: putative bacteriocins of *B. cereus* E33L; SM_bac: putative bacteriocin of *Streptococcus mutans* V1996; SS_bac: putative bacteriocin of *Streptococcus suis* 89/1591. The diglycine site is highlighted in green, the disulfide bond forming cysteines in yellow, any additional cysteines in blue and the loop region is underlined.
Although sublancin 168 is annotated as a lantibiotic (Paik et al. 1998), it is possible that the modification may be a sugar residue on either serine 18 and/or cysteine 22 of the loop, similar to the glycosylated serine 18 and cysteine 43 in PlnKW30. Paik et al. identified sublancin 168 as a lantibiotic mainly because of the difficulties that were encountered during N-terminal sequencing using Edman degradation and because they were solved using ethanethiol derivatization. Ethanethiol derivatization prevents the blockage of the Edman degradation reaction caused by the dehydro residues of lantibiotics. However, this reagent can also bring about β-elimination of phosphates and sugars from serines and threonines (Jaffe et al. 1998). Another reason given for classifying sublancin 168 as a lantibiotic was the presence of a diglycine (or GS in SunA) cleavage site between the leader peptide and the mature bacteriocin that occurs in class AII lantibiotics. However, this site is also typical for class II unmodified bacteriocins (Section (d)1.3.2(d)).

Furthermore, Paik et al (1998) did not identify any orthologues of the genes that are deemed to be necessary for lanthionine formation, such as nisB for dehydration and nisC for cyclisation in L. lactis (Section 1.3.1) and these genes are not present in the genome of B. subtilis 168. However, the sunA operon does contain a GTase gene. An alignment of the amino acid sequences of the N-terminal glycosyltransferase domains of the GTase genes of B. subtilis 168, E. faecalis V583 and L. plantarum KW30 is shown in figure 4.4.

Clearly there are strong sequence similarities between the N-terminal domains of the GTases of these three sequences. The GTase of E. faecalis V583 has 32% identity at a protein sequence level to the GTase (GccB) encoded in the plnKW30 gene cluster (Table 3.2), and was identified in public data bank searches as the best sequence match.
Chapter 4: Discussion

4.2 Analysis of the plnKW30 gene cluster

Figure 4.4: Comparison of the N-terminal amino acid sequences of bacteriocin GTases.

BS_GT: sublancin 168 GTase of B. subtilis 168; EF_GT: GTase (EF_B0049) of E. faecalis V583; LP_GT: GTase of L. plantarum KW30. Highlighted in yellow are conserved regions which occur in the N-terminal glycosyltransferase region. Identities are marked *; highly conserved :; conserved .(Clustal W; Larkin et al. 2007).

The comparison of the ABC-transporter amino acid sequence of the SunA cluster (Figure 4.5) to the ABC-transporters of B. cereus E33L, E. faecalis V583 and L. plantarum KW30 shows some homology between the N-terminal proteolytic domains. These include the conserved cysteine and histidine domains that are known to be part of the active site of the C39 peptidase domains characteristic of many bacteriocin ABC-transporters (shadowed in figure 4.5) (Havarstein et al. 1995).
4.2 Analysis of the plnKW30 gene cluster

Paik and co-workers aligned SunT of *B. subtilis* 168 to LcnDR3 (the ABC-transporter that exports the non-lantibiotic lactococcin DR) and PepT (the ABC-transporter that exports the lantibiotic Pep5). The alignment shows conserved regions in the N-terminal sequence similar to those seen in figure 4.5, but the conservation is only between SunT and LcnDR3, which is the ABC-transporter for the non-lantibiotic lactococcin DR (LcnDR). Furthermore, LcnDR contains a diglycine type leader peptide. Paik *et al.* also stated that SunT has strong homology to the C-terminal part of PepT, which is the ABC-transporter for the lantibiotic Pep5, which does not contain a diglycine leader peptide. It is interesting that the C-terminal part of SunT also has significant amino acid sequence similarities to the ABC-transporters of the non-lanthionine bacteriocin producers *B. cereus* E33L, *E. faecalis* V583 and *L. plantarum* KW30, as shown in figure 4.6.

These findings support the idea that sublancin 168 could be glycosylated, but further analyses of sublancin 168 and its gene cluster are necessary to substantiate this hypothesis (work in progress).
Chapter 4: Discussion

4.2 Analysis of the plnKW30 gene cluster

Figure 4.6: Comparison of the C-terminal sequence of the ABC-transporters. BS_SunT: sublancin 168 ABC-transporter of B. subtilis 168; BC_ABC: SunT ABC-transporter of B. cereus E33L; EF_ABC: ABC-transporter (EF_B0050) of E. faecalis V583; LP_ABC: ABC-transporter of L. plantarum KW30. Highlighted in yellow are the conserved regions. Identities are marked *; highly conserved : ; conserved . (Clustal W; Larkin et al., 2007).

The comparison of the plnKW30 gene cluster to those found in E. faecalis V583, B. subtilis 168 and B. cereus E33L show that the plnKW30 cluster could have been obtained via horizontal gene transfer. This could then have been followed by gene rearrangements and modifications, as discussed on page 164. Bacteriocin gene clusters containing GTase and TRX genes are not widespread and despite extensive data mining the only similar clusters that have been found are those in B. subtilis 168 and E. faecalis V583 (Figure 4.1).

Another example for gene rearrangements and modifications in bacteriocin gene clusters is the plantaricin (pln) locus, which is found in most L. plantarum strains; including L. plantarum KW30 (unpublished results). The genes of the pln locus also appear in a variety of different arrangements (Section 1.4.2(b)), but a basic order of genes is maintained.

The development of a separate prediction algorithm that recognises glycosyltransferases common to glycocin clusters would be very useful. This could be integrated into BAGEL (de Jong et al. 2006) to help to identify bacteriocin gene clusters of glycocins.
4.2.1 **Bioinformatic analyses of the proteins of the plnKW30 gene cluster**

Bioinformatic analyses were used to predict the function of the proteins encoded by the *plnKW30* gene cluster. Such analyses are based on the knowledge that proteins with similar amino acid sequences usually carry out similar functions.

The first step was to use the tBLASTn search (NCBI) to identify possible functions of the *gccBCDEFA* gene products (Table 3.2). These results provided initial indications of possible functions and were followed by further bioinformatic investigations. These are described in section 3.2.1 for each of the six core proteins (GTase, ABC-transporter, TRX1, TRX2, RR and PlnKW30) encoded by the *plnKW30* gene cluster. The functions of the two putative TRXs and the putative response regulator are uncertain, as they are based only on bioinformatic analyses. The assumption that GccB is the glycosyltransferase that modifies PlnKW30 (GccA) is reasonable, because of its position in the cluster and the presence of GTase genes in other bacteriocin clusters encoding PlnKW30-like bacteriocins (Section 4.2). The proteolytic ABC-transporter GccC was identified by bioinformatic analysis and this is supported by its similarity to the SunT ABC-transporter of *B. subtilis* 168 (Section 4.2). The proximity of gccC to gccA encoding the bacteriocin PlnKW30, which contains a diglycine leader peptide, also increases the likelihood that it is a bacteriocin-processing ABC-transporter (de Jong *et al.* 2006). The function of GccA (PlnKW30) was confirmed by experimental evidence as described in section 3.4.

The results of the bioinformatic analyses of each protein were used to develop a schematic model of how the proteins might be organised in the cell (Figure 4.7). As it is most likely that these proteins are all necessary for the maturation and export of PlnKW30, it is possible that they form a multi-protein complex *in vivo* or at least work closely together.

Firstly, an extracellular signal, presumably a pheromone, activates the response regulator via an unknown interaction. It is likely that another protein serves as a sensor, such as a histidine protein kinase (Section 1.4), that in turn interacts with the response regulator. A putative histidine protein kinase is not present in the *plnKW30* gene cluster; however the putative response regulator might not be part of a two-component signalling system but act as a cytosolic one-component signalling system (Section 1.4.1). Bioinformatics analyses
did not identify a possible receiver domain for GccF, but recognized a LytTR output domain (Figure 3.9) that is known to regulate gene expression (Section 1.4.2).

Figure 4.7: Schematic model of the putative regulation and maturation of PlnKW30. See text for detailed description.

The bacteriocin PlnKW30 is produced as a prepeptide containing a 21 amino acid diglycine leader peptide, which directs it to the dedicated ABC-transporter (GccC). The C39 peptidase domain of GccC most likely cleaves the signal peptide, after which PlnKW30 is possibly transported through the cell membrane via the ABC-transporter membrane domain (Figure 3.6).

Prior to secretion, PlnKW30 could be glycosylated by the N-terminal glycosyltransferase domain of GccB (Figure 3.5), which belongs to the large GT2 family of inverting glycosyltransferases (Section 1.5.5). It is possible that this GTase glycosylates both serine 18 and cysteine 43, but this is not certain. TRX1 may keep PlnKW30 in a reduced, and therefore inactive, state before it is transported outside the cell, although this is
probably already provided by the reducing conditions inside the cell (Figure 3.7). The role of TRX2 is uncertain, as it is most likely positioned outside the cell, but as it is similar to thioredoxins it is possibly not involved in the disulfide bond formation of PlnKW30. Once PlnKW30 is transported outside of the cell, the two disulfide bonds are formed, resulting in mature and active PlnKW30.

Conformation of the putative molecular functions of the proteins encoded by the plnKW30 gene cluster as determined by bioinformatic analyses was sought using experimental approaches. These included protein expression for x-ray crystallography (Appendix 6) and gene disruptions (Appendix 7), but unfortunately these were unsuccessful.

4.2.2 Future work

The KW30 DNA sequence data obtained by the Solexa Genome Analysis Facility (Allan Wilson Centre Sequencing Service, Palmerston North) could be used for comparison of the pln locus to those of other L. plantarum strains (Section 1.4.2(b)).

Future work for the plnKW30 gene cluster could include the identification of the cellular localisation of each protein via expression of gene fusions with reporter genes, such as β-galactosidase.

Furthermore, using pull-down experiments it might be possible to determine which proteins interact. It is possible that the proteins required for production and maturation of PlnKW30 are located together, most likely associated with the membrane, in a protein maturation complex. Another approach could be the analysis of membrane protein complexes from L. plantarum KW30 membrane fractions using blue native gel electrophoresis.

Of primary interest will be the identification of any signal peptides, e.g. pheromone, and the acceptor protein that transfers the signal to the putative response regulator that may ultimately regulate the expression of the plnKW30 gene cluster.
4.3 TRANSCRIPTIONAL ANALYSIS

4.3.1 COMPARISON OF NORTHERN BLOT AND RT-PCR ANALYSIS

The analysis of the transcriptional pattern of the \textit{plnKW30} gene cluster was carried out using RT-PCR and Northern blotting (Sections 3.3.1 and 3.3.2). The RT-PCR was designed to give information about the length of the messenger RNA (mRNA) of the cluster. Each mRNA transcript was amplified separately with gene specific primers to show that they were transcribed. These amplifications worked well, apart from the response regulator gene (Figure 3.10). It was not always possible to achieve a PCR product for \textit{gccF} and when it could be amplified there was only a very faint band visible on the gel. This could be due to low abundance of this particular mRNA or because of secondary structures formed by the RNA, which might inhibit the reverse transcriptase. The same primers worked well for amplification of the response regulator gene using DNA as template, so the problems are most likely to result from RNA properties.

In order to find out the length of the \textit{gcc} transcripts, gene pairs were amplified using the forward primer of the first gene and the reverse primer of the second gene (Figure 3.10). RT-PCR amplification should only be possible if both genes are located on the same mRNA, therefore if PCR products are visible the two genes are transcribed together. Products are visible for all combined gene regions in figure 3.10, but only a very faint band was obtained for the response regulator-\textit{plnKW30} product. This could again be due to strong secondary structures in the mRNA inhibiting the reverse transcriptase. Another explanation could be that the response regulator transcript is only present in low abundance, because it is actually not a functional gene, or it may not be transcribed. After all the related \textit{E. faecalis} and \textit{B. subtilis} gene clusters do not have a response regulator (Figure 4.1.)

To confirm these findings Northern blot analyses were carried out, using probes specific for the \textit{plnKW30} (Figure 3.11), response regulator and glycosyltransferase genes. The Northern blots carried out using probes for the glycosyltransferase and response regulator genes were difficult to interpret, possibly because the labelled probes lacked specificity. The result of the hybridisation using the \textit{plnKW30} specific probe is very clear: only one band is visible with the size of the pre-\textit{plnKW30} mRNA product (Figure 3.11). This includes the \textit{plnKW30} gene itself (154 bp), the diglycine signal sequence (Section 1.3.2(d)) and the
promoter region, which adds up to about 170 bp. This result shows that \textit{plnKW30} is transcribed on its own and not on a long mRNA transcript together with the rest of the cluster, as predicted by the RT-PCR results. It is possible that the \textit{plnKW30} gene is transcribed in two forms: a short, abundant \textit{gccA} transcript and a longer very low abundance transcript comprising \textit{gccBCDEFA}. Although a longer transcript was not visible on the \textit{plnKW30} blot, even after longer exposure of the blot to film, the existence of such a transcript would be consistent with the RT-PCR product obtained for \textit{gccE-gccF}.

### 4.3.2 Growth-phase-dependent gene expression of \textit{plnKW30}

Bacteriocin production commonly starts when cell densities reach a certain threshold (Section 1.3.2) and Northern blotting was used to determine the start and intensity of \textit{plnKW30} expression. The blot (Figure 3.12) showed that \textit{plnKW30} mRNA was produced from the beginning of log phase (4 hours), and peaked in late log phase (24 hours).

It is important to examine the expression of a housekeeping gene in parallel with the gene of interest to determine if the different intensities are actually due to differences in gene expression. In this case, normalisation was carried out using the concentration of total RNA and not the amount of cells the RNA was isolated from. RNA isolation is easier from log phase \textit{L. plantarum} KW30 cells than from stationary phase cells; hence the amounts of isolated RNA differ and do not represent the actual RNA concentration in the cell. The amount of \textit{plnKW30} transcript increases strongly from 4 to 24 hours and thus probably reduces the relative amount of \textit{gap} transcript in the total RNA sample. This could explain the lower concentrations of \textit{gap} transcript in the 16 and 24 hour samples.

### 4.3.3 Future work

The GTase and RR Northern blots should be carried out using more specific probes. Small gene fragments (about 60 bp) will be used as templates for the preparation of random labelled probes as before. Another option is to use 5'-labelled N- and/or C-terminal specific
oligonucleotide primers or to use end-labelled primers for PCR to produce specifically labelled probes.

At the time of writing this thesis, work was being done to generate mutant *L. plantarum* KW30 strains that lack the *plnKW30* gene cluster. These could serve as negative controls for Northern blots in the future.

The analysis of the growth-phase-dependent gene expression of the genes of the *plnKW30* gene cluster from *L. plantarum* KW30 should be carried out using real-time PCR experiments. The expression levels of *plnKW30*, RR and GTase at different time points during growth of the producer strain should also be analysed using real-time PCR.

To determine the transcription start region of putative transcripts (*gccBCDEF* and/or *gccBCDEFA*), it would be useful to perform 5' rapid amplification of cDNA ends (5' RACE) (Schaefer 1995).

### 4.4 Characterisation of PlnKW30 of *L. plantarum* KW30

In order to carry out different analyses of PlnKW30 it was necessary to purify large amounts of the native bacteriocin, and many different conditions were tested to find the optimal conditions for PlnKW30 production (Section 3.4.1). As expected, the growth in a chemically defined medium (CDM) instead of a complex medium (MRS) resulted in lower amounts of PlnKW30 being secreted. The availability of oxygen (open versus closed tubes; figure 3.13) did not seem to have any effect on PlnKW30 production in MRS medium, which is not surprising in light of the fact that *L. plantarum* species are known to be aerotolerant (Archibald & Fridovich 1981). Contrary to this, in CDM, availability of oxygen seemed to have an adverse effect on PlnKW30 production (Figure 3.13). Possibly, the growth of *L. plantarum* KW30 is already restricted by the low amount of nutrients in CDM, resulting in the induction of a large number of general stress proteins (Hecker & Volker 2001) which may affect the production of some proteins, including PlnKW30. In *Bacillus subtilis* a specific phosphatase is required for sensing glucose and oxygen starvation (Hecker *et al.*
and if a similar sensing mechanism is present in L. plantarum it might lead to a stronger stress response under both nutrient starvation and presence of oxygen.

Differences in PlnKW30 production were observed after incubation at different temperatures in MRS media. The bacteriocin appeared to be produced in greater amounts at lower temperatures in MRS, as shown in figure 3.13. The optimal growth temperature of L. plantarum KW30 is 30 °C, but lower temperatures may generate a stress signal leading to increased bacteriocin production. The occurrence of higher bacteriocin concentrations at suboptimal growth temperatures has been described before (Krier et al. 1998; Moretro et al. 2000; Caldéron-Santoyo et al. 2001; Delgado et al. 2005). It is logical that bacteriocin production is not only stimulated by growth-associated mechanisms, but is also affected by several environmental conditions. For example, it has been reported that the occurrence of stress factors, such as low temperatures or high salt conditions, are able to stimulate bacteriocin production in other bacteria (Delgado et al. 2007).

The routine purification of native PlnKW30 involved growing L. plantarum KW30 in 8 L of MRS at room temperature (~25 °C) well into stationary phase (3 – 4 days). PlnKW30 was then purified from the growth medium as described in section 2.4.8. The yield from each purification was very variable, not only because of variability in culture growth and bacteriocin production, but also because of the state of the organism and the source and preparation of medium used. The efficiency of the chromatography resin used for the purification also affected the yield, as efficiency declined with each use. This variability in production was especially noticeable when the cell culture had already been transferred several times. Kelly et al. (1996) reported the appearance of bacteriocin negative cultures, although it was not a frequent occurrence. Furthermore, bacteriocin production is influenced by a large number of parameters, such as temperature and pH, carbon and nitrogen sources, Tween 80, ethanol and NaCl (Delgado et al. 2007). Although the growth temperature and the preparation of the growth medium was the same from preparation to preparation, some differences might have occurred because of variations in the complex medium MRS or the purity of the water used. In fact, tests of L. plantarum KW30 grown in MRS medium from different suppliers produced different amounts of bacteriocin. Paik et al (1998) found that the production of sublancin 168 varied markedly in different cultures for no apparent reason.
A chromatogram of the last step of the purification, RP-HPLC separation, is shown in figure 3.14. This step was very reproducible in all purifications, and the discolouration of some samples, especially when highly concentrated, was probably due to a putative residual compound from autoclaved MRS broth, which co-purifies with PInKW30. It was noticed that if the broth was dark brown, the colour was darker than when the broth was lighter (not autoclaved). The brown colour did not appear to have a negative effect on the activity of PInKW30, and there was no evidence either by mass spectrometry or NMR spectroscopy of other molecules being covalently bonded to PInKW30.

The minimum inhibitory concentration (MIC) for a bacteriocin is normally determined based on its activity towards a susceptible indicator strain. *L. plantarum* ATCC 8014 is the preferred indicator strain for PInKW30 (Kelly *et al.* 1996), although purified PInKW30 is also active against *L. plantarum* WCFS1 and NC8, and inactive against *L. plantarum* 965, the indicator strain used for *pln* locus bacteriocins (Dr. M.L. Patchett – personal communication). A simple plate assay is commonly used to determine the MIC of bacteriocins, but a culture test produces more accurate results. A 96-well plate format liquid culture assay was used to show that the MIC of PInKW30 is approximately 12.5 ng/mL (~2.5 nM) (Figure 3.16). The MIC is defined as the minimal concentration of the bacteriocin that inhibits the growth of the indicator strain by 50%. In the presence of 12.5 ng/mL PInKW30 an *L. plantarum* ATCC 8014 culture has an OD 600 close to 0.2 after 7 hours, whereas a control culture without PInKW30 has an OD 600 of about 0.4. Thus, the addition of 12.5 ng/mL PInKW30 reduces the turbidity of the culture by about 50% in comparison to the turbidity of the control culture. It is not possible to directly compare the activity of one bacteriocin to another, because different indicator strains are used. However, PInKW30 is active at nanomolar concentrations, which shows that it is a potent bacteriocin and comparable to nisin, a well characterised potent bacteriocin. Its potency is also comparable to *L. plantarum* C11 class II bacteriocins such as PlnJK and PlnEF, which have MICs of 0.1 nM and 7 nM, respectively, against the indicator strain *L. plantarum* 965 (Anderssen *et al.* 1998).

Experiments using a range of different fungi from various sources or habitats did not show any sensitivity to PInKW30. The cell walls and membranes of fungi have a very different composition to Gram-positive bacterial cells, and PInKW30 has been shown to be active
only against a small range of closely related lactic acid bacteria, mostly other *L. plantarum* strains (Kelly *et al.* 1996).

### 4.4.1 Antimicrobial Activity of PlnKW30 Against the Indicator Strain *L. plantarum ATCC 8014*

Growth curves of the producer strain *L. plantarum* KW30 and the indicator strain *L. plantarum* ATCC 8014 were measured and shown to be very similar. Using 10 % inoculum, cultures reached mid-log phase after about 6 hours growth with an OD$_{600}$ between 3.0 (*L. plantarum* ATCC 8014) and 4.0 (*L. plantarum* KW30), as shown in figure 3.17. This cell density was chosen as starting condition for all following experiments with *L. plantarum* ATCC 8014, because under normal conditions this is the cell density at which bacteriocin production starts in many lactobacilli species (Diep *et al.* 1995; Kelly *et al.* 1996; Eijsink *et al.* 2002). Production of bacteriocins is normally regulated by quorum sensing (Section 1.4) and therefore occurs after a certain threshold cell density is reached. Thus cultures at an OD$_{600}$ of 3.0 were used to reproduce the conditions under which bacteriocin “wars” begin. It had been observed that cells in log phase were more susceptible to PlnKW30 than cells in stationary phase, and needed lower concentrations of PlnKW30 to completely inhibit further growth. Cells that have reached stationary phase have a thicker cell wall with a different complement of membrane proteins than cells in log phase and appear to be more resistant to PlnKW30.

As expected, the addition of PlnKW30 to *L. plantarum* ATCC 8014 cultures stopped cell growth for 13 hours, after which cell density began to increase, leading to the conclusion that PlnKW30 has a predominantly bacteriostatic effect, rather than a bactericidal effect, on the cells of the indicator strain (Figure 3.18). Increased growth at 22 hours might be due to cells that have gained immunity to PlnKW30 over that time. Lower concentrations of PlnKW30 showed its effect to be concentration dependent, as cell lower PlnKW30 concentrations resulted in higher cell densities of *L. plantarum* ATCC 8014 at the final OD$_{600}$ at 22 hours.
The Live/Dead cell assay showed that PlnKW30 does in fact actively kill cells of *L. plantarum* ATCC 8014. The addition of PlnKW30 to cells of the indicator strain clearly shows the bactericidal effect increases over 22 hours after exposure to PlnKW30, killing about 26% of cells compared to 6% in the non-bacteriocin control (Figure 3.19). This was also reported by Kelly and co-workers (1996), who determined PlnKW30 to have a bactericidal rather than bacteriolytic mode of action. However, in their assays they used neutralised culture supernatant rather than purified PlnKW30. The Live/Dead cell assay, as already seen in the cell density assay, showed that bactericidal activity of PlnKW30 on the indicator strain was dependent on the concentration of the bacteriocin (Figure 3.20). Clear differences in percentage of dead cells were seen when 1 µg/mL and 250 ng/mL PlnKW30 was added to the cell cultures. These concentrations are much higher than those observed for the minimum inhibitory concentration. A comparison of these two experiments is not possible; because they were started at very different cell densities (MIC at 0.4; Live/Dead cell assay at 3) and the state of the cells used was different as well. For the MIC experiment, cells were diluted to an OD$_{600}$ of 0.4 from a stationary phase culture and used directly; in the Live/Dead cell assay they had already recovered from stationary phase, which takes about half a generation (about 1 hour). The method used to determine the MIC was chosen, because it is commonly used and so might allow comparisons with other bacteriocins in spite of the difficulties inherent in such comparisons (see above). The Live/Dead cell assays were also tested under comparable conditions starting at OD$_{600}$ of about 0.3. These experiments were not possible to carry out because the *L. plantarum* ATCC 8014 cells had for some reason an elongated morphology at this low cell density, which made it impossible to count them correctly.

The differentiation between the two modes of action, bacteriostatic versus bactericidal, is difficult, especially when only cell density measurements are performed, as normally has been the method of choice to investigate the antibacterial activity of bacteriocins. The use of the Live/Dead cell assay gives new insight into the way PlnKW30 works and could also be used to investigate the antimicrobial activity of other bacteriocins in more detail. Cell density measures all cells, dead or alive, present in the sample and so cannot distinguish between bacteriostatic and bactericidal effects. A decrease in cell density is only observed after cells begin to lyse, which starts in the case of *L. plantarum* ATCC 8014 at about...
22 hours of growth. Comparing the results from both experiments, it is apparent that PlnKW30 exhibits both bacteriostatic and bactericidal activities. The producer strain *L. plantarum* KW30 was, as expected, not affected by its own bacteriocin.

The effect of different sugars on the activity of PlnKW30 was tested, because work by Diep *et al.* (2007) showed that some class II bacteriocins use components of the mannose phosphotransferase system (man-PTS) of susceptible cells as receptors and because the unusual post-translational modifications of PlnKW30 led to speculation about their involvement in antimicrobial activity. PTS-type systems comprise up to half of the sugar transport systems found in low G + C Gram-positive lactic acid bacteria (*Lorca et al.* 2007), and *L. plantarum* WCFS1 contains 25 complete PTS sugar transport systems (*Kleerebezem et al.* 2003). Since PlnKW30 is modified by two N-acetylglucosamines (GlcNAc) (Section 3.5) it seemed possible that the mode of action of PlnKW30 involves a hexosamine transporter as a receptor. Interestingly, the addition of both GlcNAc and glucosamine to cultures challenged with PlnKW30 resulted in a decrease in the bactericidal activity of PlnKW30 as measured by the Live/Dead cell assay (Figures 3.21 and 3.23). The addition of mannose (Man) or N-acetylgalactosamine (GalNAc) did not have any protective effect (Figure 3.24). The man-PTS transporter has been shown to import both mannose and glucose (*Cochu et al.* 2003), indicating that this transporter is unlikely to be the receptor for PlnKW30. The results show however that either a GlcNAc or glucosamine transporter is likely to be involved in the antimicrobial mechanism of PlnKW30. The *L. plantarum* WCFS1 genes of two man-PTS transporters, transporting mannose or galactosamine, and a glc-PTS transporter that transports N-acetylglucosamine, are present in the genome of *L. plantarum* KW30 as near identical orthologues. *L. plantarum* KW30 and WCFS1 also have two further genes annotated as GlcNAc-specific PTS EIIICBA transporters (lp_2531 and lp_2969). However, *L. plantarum* KW30 contains an additional transporter IID subunit of the mannose/fructose/sorbose PTS family, which has also been identified in *L. plantarum* subsp. *plantarum* ATCC 14917 (GI: 227896999).

The interaction between a transporter and PlnKW30 might be initiated by the binding of either one or both GlcNAc molecules or by protein-protein interactions between transporter and PlnKW30 (Figure 4.8E). After the attachment of PlnKW30 to the transporter one or both of the GlcNAcs of PlnKW30 might mimic a substrate sugar and bind
to the transporter. Consequently, the transporter would be blocked, because it cannot transfer the PlnKW30 bound GlcNAc into the cell (Figure 4.8F). This would result in an inactive sugar transporter, which would compromise the cells ability to acquire GlcNAc. It might be possible to use other sugars as energy source, but if the transporter is blocked in an open state it will result in leakage of cellular components, followed by cell death (Diep et al. 2007). Another indication for PlnKW30 targeting a specific receptor is that its MIC is in the nanomolar range (Section 3.4.3), because micromolar concentrations of bacteriocins would be necessary to cause inhibitory effects, such as membrane damage, in pure lipid bilayers (Hechard & Sahl 2002).

*L. plantarum* ATCC 8014 cells in mid-log phase are more susceptible to PlnKW30 than those that have reached stationary phase as they are susceptible to much lower concentrations of bacteriocin as shown by Kelly et al (1996). Cells that have reached stationary phase might have a different composition of proteins in the membrane compared to still dividing cells. Sugar transport proteins are probably not as essential to survival for a non-dividing cell, because the cells stopped growing and these transporters therefore became less abundant. A decrease in the number of transporters together with the development of a thicker cell wall would explain why cells of the indicator strain in stationary phase are less susceptible than log phase ones.

The model for the mechanism of bactericidal action of PlnKW30, proposed in figure 4.8, can explain the concentration dependent, protective effect of GlcNAc and glucosamine (Figure 3.23). The model demonstrates the competitive manner in which GlcNAc or glucosamine could inhibit the interaction of PlnKW30 with a GlcNAc transporter. For a complete protection from 0.2 µM PlnKW30 the concentration of added GlcNAc or glucosamine has to be much higher (4.5 mM) than that of the added bacteriocin (Figure 3.23). This would be consistent with the hypothesis that protein-protein interactions between PlnKW30 and the receptor enhance the affinity of the PlnKW30 GlcNAc group(s) for the GlcNAc transporter. Both GlcNAc and glucosamine were also added to *L. plantarum* ATCC 8014 cells without any PlnKW30 to show that the excess of additional sugar did not have any influence on cell growth or death (Figure 3.22).
Figure 4.8: Schematic overview of the hypothetical interaction between PlnKW30 and a GlcNAc transporter.
A: hypothetical normal situation of GlcNAc transporter; B: PlnKW30 is not present, GlcNAc binds to transporter; C: GlcNAc is transferred across membrane into cell; D: GlcNAc transporter situation with PlnKW30 present; E: PlnKW30 interacts with transporter; F: PlnKW30 blocks function of transporter; possibly blocking it in open state, resulting in leakage of cellular components and cell death.

Also consistent with the model in figure 4.8, the addition of Man or GalNAc failed to provide the indicator strain with any protection from PlnKW30. Even at extremely high concentrations of 10 mg/mL (45 mM), which is much higher than the concentration of PlnKW30 (0.2 µM) and 10 x higher than the highest concentration of GlcNAc or glucosamine used (4.5 mM), there was no observed decrease of cell death (Figure 3.24). Although the error bars of the 10 mg/mL and 1 mg/mL Man and GalNAc are quite large, the
results are significant: no decrease in the number of dead cells was observed. These results support the theory that at least one of the GlcNAc modifications of PlnKW30, either that of Ser 18 or Cys 43, might be involved in its mechanism of action.

### 4.4.2 Future work

Future work could include further analysis of the antimicrobial effect of PlnKW30 on the indicator strain and other susceptible strains using the Live/Dead cell assay. These could include experiments using fully deglycosylated PlnKW30.

Further analysis of the putative interaction of PlnKW30 with the sugar transporter could be carried out via pull-down experiments (similar to those described by Diep et al. (2007)) and subsequent identification of the ‘pulled-down proteins’ by mass spectrometry. Similar experiments could be performed to identify possible immunity proteins and their interaction with PlnKW30 and the sugar transporter.

It is unlikely that the model for PlnKW30 proposed in figure 4.8 can fully explain the bacteriostatic activity of PlnKW30. Analysis of the interaction of PlnKW30 with other components of the cell membrane, and of the orientation, organisation and position of PlnKW30 in the membrane, might also give insight into its mode of action. NMR analysis of PlnKW30 could be performed in the presence of DPC (dodecylphosphocholine) or TFE (trifluoroethanol) micelles. These approaches have been informative for other antimicrobial peptides, such as magainin (Porcelli et al. 2006) and sakacin P (Uteng et al. 2003).

### 4.5 Analyses of N- and C-terminal fragments of PlnKW30

The tryptic digest of PlnKW30 was performed in order to obtain two fragments each containing one of the N-acetylglucosamine (GlcNAc) modifications (Section 3.5.1), so that they could be analysed separately (Section 2.6.1). Digestion with trypsin resulted a 32
amino acid and a 11 amino acid peptide (Figure 4.9), separating the two GlcNAs and potentially providing the opportunity to investigate the biological importance of each GlcNAc modification. The results obtained by circular dichroism (Section 3.7) showed that the N-terminal fragment retains its structure, while the C-terminal fragment is a random coil, as expected from the structure of the whole molecule (Figure 3.34; H. Venugopal – personal communication).

Two experiments, tricine SDS-PAGE overlaid with *L. plantarum* ATCC 8014 (Figure 3.27) and the Live/Dead cell assay (Figure 3.28), showed that there was a total loss of antimicrobial activity when PlnKW30 was hydrolysed into two separate peptides. Using only the N- and C-terminal fragments of PlnKW30 in the Live/Dead cell assay at molar concentrations comparable to full-length PlnKW30, resulted in percentages of dead cells similar to the control with no bacteriocin added (Figure 3.28). Addition of both fragments together resulted in a small increase of dead cells, but it was not significantly different from the addition of each fragment alone, proving that intact PlnKW30 is necessary for bactericidal activity.

![Figure 4.9: Schematic overview of N-terminal and C-terminal tryptic fragments of PlnKW30. Hexagons indicate GlcNAc modifications.](image)

Cells of the indicator strain in mid-log phase were not susceptible to the N-terminal fragment (Figure 3.27); however, some activity was visible when the cells were in early log phase (Figure 3.30). This suggests that the N-terminal fragment retains some of the
bacteriostatic activity of the intact PlnKW30, whereas bactericidal activity is a property of the intact molecule only.

### 4.5.1 Discovery of C-terminal glycan modification

The modification at the C-terminal cysteine was assumed to be a farnesyl group, because the original average mass for the C-terminal chymotryptic peptide was determined to be 462.2, which left a mass of 204 unaccounted for. The most likely match was to a farnesyl group. As farnesylation is a well characterised modification of C-terminal cysteine residues in eukaryotes, it was thought that it might represent the first C-terminal farnesylated peptide in prokaryotes. The other possibility was an N-acetylhexoseamine, but as the mass was out by one mass unit, and there were no convincing precedents for biological S-linked glycans (Section 1.5.3), this possibility was not pursued.

The monoisotopic mass spectrometry analysis of the C-terminal HC-X fragment of PlnKW30 (Figure 3.33) showed that the mass of the X was 203.0778 (N-acetylhexoseamine) and therefore only 0.1016 mass units less than a farnesyl group (203.1794 m/z). This small difference in mass units was not possible to detect without a highly accurate mass spectrometer capable of determining the exact monoisotopic masses of the peptides and fragments, and such an instrument was not available at the beginning of this thesis.

The specific interaction of the native PlnKW30 with wheat germ lectin strongly suggested that indeed PlnKW30 contained at least one N-acetylglucosamine residue. Attempts to pull down the N- and C-terminal fragments separately failed, probably because the peptide concentrations were too low to be visible on a tricine gel.

The glycosylation linkage types present in PlnKW30 were analysed using anti-O-GlcNAc antibodies. The results support the existence of only one O-linked GlcNAc, because only full length PlnKW30 and the N-terminal tryptic fragment, but not the C-terminal fragment, were recognised by the anti-O-GlcNAc antibodies (Figure 3.36). This shows the high specificity of the antibody, which was raised against a serine linked O-GlcNAc. As it was not
possible to visualize the C-terminal tryptic fragment on SDS-PAGE, ninhydrin was used to show the presence of peptide in all samples (Figure 3.37).

The clear advantage of producing an S-glycosidic modification is that these linkages are more stable than O-linkages (Zhu et al. 2004; Liang et al. 2009). Furthermore, target bacteria might secret O-glycosylhydrolases as a defence against O-glycosylated bacteriocins, such as the N-acetyl-β-D-glucosaminidase (GcnA) from Streptococcus gordonii that has been shown to O-deglycosylate PlnKW30 (Figure 3.44). Thioglucosidases (EC 3.2.1.147) have been identified mainly in plants and fungi, where they catalyse the hydrolysis of thioglycosides to a sugar and thiol, but no S-glycosylhydrolases that catalyse the deglycosylation of peptides or proteins have yet been identified.

4.6 ANALYSIS OF THE ROLE OF POST TRANSLATIONAL MODIFICATIONS IN PLNKW30 ACTIVITY

4.6.1 ANALYSIS OF THE DISULFIDE BRIDGES OF PLNKW30

The importance of the disulfide bridges for the antimicrobial activity of PlnKW30 was examined by reduction and alkylation. The purification of the reduced and alkylated sample resulted in a single peak eluting at approximately 10 min (30 % B), which is considerably earlier than the peak of the native PlnKW30 (16 min and 36 % B; figure 3.38). This could be due to a change of structure of the bacteriocin that occurs when the disulfide bonds are broken. Mass spectrometry analysis confirmed the reduction and alkylation of the disulfide bridges of the treated PlnKW30 sample (Figure 3.39), although three serine residues might also be acetylated. The effect of these serine acetylations on the antimicrobial activity of PlnKW30 cannot be determined. However, the results from the Live/Dead cell assay, where reduced but non-acetylated PlnKW30 was used, showed that the breakage of the disulfide bonds results in total loss of bactericidal activity (Figure 3.41).

The reduced PlnKW30 was subjected to the biological plate assay (Figure 3.40) and the Live/Dead cell assay (Figure 3.41). These assays show that the reduced PlnKW30 does not
Chapter 4: Discussion

4.6 Analysis of post-translational modifications

exhibit any bactericidal or bacteriostatic activity. In the Live/Dead cell assay an increase of dead cells is visible in the sample with both PlnKW30 and TCEP added, which is slightly higher than both the TCEP-only control and the control without added bacteriocin. However, this increase is not significantly higher and is not visible after 22 hours.

These results clearly illustrate the significance of the disulfide bonds in maintaining the proper folding, and consequently antimicrobial activity, of PlnKW30. These findings are consistent with results from other disulfide bonded bacteriocins, which also lose activity after destruction of their disulfide bonds (Section 1.3.2).

4.6.2 Analysis of the glycan modifications of PlnKW30

The analysis of the glycan modifications was difficult because it was only possible to remove the O-linked N-acetylglucosamine (GlcNAc) (Section 2.4.13) from the serine, but not the S-linked GlcNAc from the C-terminal cysteine. The C–S bond is much stronger than the C–O bond and is not susceptible to β-elimination under alkaline conditions. As soluble recombinant PlnKW30, which would not have been glycosylated, could not be produced (Appendix 6), it was not possible to analyse a completely deglycosylated PlnKW30.

The enzymatic O-deglycosylation of PlnKW30 using the N-acetyl-β-D-glucosaminidase of Streptococcus gordonii was monitored by mass spectrometry as shown in figure 3.44. The deglycosylated sample of PlnKW30 was purified by RP-HPLC (Figure 3.42) and eluted slightly later than the native bacteriocin, as expected because of the removal of the hydrophilic GlcNAc.

The question is whether the GlcNAc modification on serine 18, or that on cysteine 43, or both, are involved in the targeting of PlnKW30 to susceptible cells. Analysis of the partially deglycosylated disulfide bonded PlnKW30 showed the importance of the modification at serine 18, as the bactericidal activity of the deglycosylated PlnKW30 was reduced by two thirds using the Live/Dead cell assay (Figure 3.45). This result of the Live/Dead cell assay was supported by the biological plate assay, which also showed a reduction of the combined bacteriostatic and bactericidal activity due to the deglycosylation of the serine
residue of PlnKW30 (Figure 3.43). The fact that the O-deglycosylated PlnKW30 shows some antimicrobial activity indicates that its disulfide bonds have to be intact, as reduction of those results in total loss of activity. The mass spectrometry result also showed that the disulfide bonds of the deglycosylated PlnKW30 are intact.

The presence of two N-acetylglucosamines in PlnKW30 makes it the first proven class IV bacteriocin according to the Klaenhammer classification scheme (Section 1.2). Klaenhammer’s class IV comprises bacteriocins containing lipid and/or carbohydrate moieties, and despite several reports of bacteriocins that lose antimicrobial activity after incubation with amylase or lipase, so far no bacteriocin had been classed as such. Moreover in the latest classification schemes this class had been withdrawn for lack of evidence (Figure 1.1). Kelly and co-workers (1996) incubated culture supernatant containing PlnKW30 with amylase and lipase, which had no effect on activity. This emphasises the limited utility of such tests in detecting class IV bacteriocins.

However, a glycosylated antimicrobial peptide, microcin E492, from the Gram-negative Klebsiella pneumoniae had been described before (Lagos et al. 2001; Thomas et al. 2004; Nolan & Walsh 2009). It is post-translationally modified by a trimer of N-(2,3-dihydroxybenzoyl)-L-serine units linked via a C-glycosidic linkage to a β-D-glucose, that is itself O-glycosidically linked to the C-terminal serine via the carboxyl group.

**4.6.3 Future work**

The preparation and analysis of fully deglycosylated PlnKW30 could be carried out.

Future work could also include the preparation of the deglycosylated N-terminal tryptic fragment; to test its antimicrobial activity against cells of the indicator strain in early log phase using the biological plate assay.
4.7 Circular Dichroism

Circular dichroism (CD) spectroscopy can be used to quantitatively estimate secondary structure elements, which can be compared to NMR or X-ray crystallography structures. In contrast to the detailed structural information NMR and X-ray crystallography provide, CD spectra can only describe overall structural features. However, the advantage of CD is that it is a non-destructive and less experimentally demanding technique, and allows broad comparisons to be made between proteins using a “fingerprinting” technique. Thus CD can be used to identify structural changes, rather than to identify the actual structure (Kelly et al. 2005).

Measurements in the far UV (180 – 260 nm) arise predominantly from two transition states of the peptide bond at 220 nm and 190 nm. In the near UV the spectra are mainly due to the aromatic amino acid side chains with an absorption in the range 260 - 320 nm. The analysis of oligopeptides is in general unreliable, except where particular secondary structures are predominant, because structures of oligopeptides are not included in public databases (Kelly et al. 2005). In this case however, CD analysis was used mainly to compare the structural integrity of the modified samples to the native peptides, rather than to obtain structural information. The most accurate results are obtained for helical secondary structures, because they have regular angles, resulting in very similar spectra and they also produce intense CD signals. The CD spectra of β-sheets are more variable and less intense. The type “random coil” refers to secondary structures that adopt a wide range of angles dissimilar to those of helix and β-sheet, which are difficult to identify exactly.

Our measurements concentrated on the far UV range in order to compare the intactness of the structural features of PlnKW30 and its N-terminal fragment without interference from specific amino acids. We did not analyse the C-terminal fragment, because it is small and unstructured (NMR analysis; H. Venugopal – personal communication).

Several methods and databases are available to analyse the acquired CD data (Whitmore & Wallace 2008). We refrained from using any analysis software to deconvolute our data, because there are no suitable databases to accurately identify the structure of peptides. Peptides tend to have low spectral magnitudes and can be present in equilibrium of multiple conformations rather than a single structure. However, the structure of PlnKW30
Chapter 4: Discussion  4.7 Circular dichroism analysis of PlnKW30

is comparatively stable, because of its two nested disulfide bonds, with the C-terminal fragment being the most flexible part.

In order to show that the secondary structure of the modified PlnKW30 samples are still intact, we analysed the native full length PlnKW30, its reduced and O-deglycosylated form, and the native, reduced and O-deglycosylated form of the N-terminal tryptic fragment by CD. The spectra obtained for PlnKW30 and the N-terminal fragment are shown in figures 3.46, 3.47 and 3.48, respectively, together with their residuals. The latter shows if any distortion has occurred during the smoothing of the spectra. The noise should be randomly distributed around zero, as is visible for all analysed samples.

In comparison to the native PlnKW30 the CD spectra of the reduced peptide shows a loss in helicity (Figure 3.49). The change in the spectrum of the reduced PlnKW30 indicates that both disulfide bonds have been reduced, as the strong decrease in helicity indicates that the secondary structure has disappeared, resulting in a more random conformation. The presence of TCEP in the solution might be the cause of the uneven spectra between 180 to 190 nm of the reduced PlnKW30, although the baseline was measured on the same buffer with TCEP and subtracted from the spectra. The spectrum of the O-deglycosylated PlnKW30 shows helical characteristics similar to the native sample (Figure 3.49), which indicates that the enzymatic O-deglycosylation did not disrupt the secondary structure of the peptide.

The spectrum of the N-terminal tryptic fragment shows the typical helical bands and a peak at about +230 nm (Figure 3.50), which is probably due to aromatic residues, such as tryptophans, which have positive bands between 220 – 230 nm (Woody 1994). Calculations have shown that CD depends on both backbone and side chain conformations. Because the indole side chain of tryptophan has low symmetry, rotation of the side chain about the Cβ-Cγ bond can lead to large changes in the CD, often causing bands to change sign. This is especially likely in proteins of low helix content, and likely to be a fraction of this structure, which is a small peptide (43 amino acids), containing two tryptophans, four tyrosines and one phenylalanine. Thus 16.3 % of the residues of PlnKW30 are aromatic, a much larger than the normal proportion of about 11.6 % (Fukuchi & Nishikawa 2001). When the C-terminal peptide is removed, the percentage of aromatic residues in the even
smaller N-terminal peptide is 18.75%. Thus it is not surprising that the positive bands between 220 and 230 nm begin to dominate the spectrum. Removal of the C-terminal tail might also act to change the conformation of the tryptophan and one of the tyrosines which has been shown by modelling to be on the outside of the helices (Dr. G.E. Norris – personal communication).

Reduction of the N-terminal fragment results in total loss of structure, indicating that both disulfide bonds were reduced (Figure 3.50). The O-deglycosylated N-terminal fragment has a very similar helical spectrum compared to the native fragment. This shows that the enzymatic O-deglycosylation did not affect the disulfide bonds, leaving the secondary structure intact.

The comparison of the full length PlnKW30 and its N-terminal fragment was carried out by overlaying both spectra (Figure 3.51), and shows that both peptides have mainly helical secondary structures. The peak at ~230 nm in the spectrum of the N-terminal fragment is most likely due to the influence of aromatic residues that have a larger percentage in the N-terminal fragment (as discussed above). These results confirm that the structure of the N-terminal fragment was not disrupted by the enzymatic removal of the C-terminal fragment using trypsin.

The structural integrity of the modified peptides is important for the significance of the experiments carried out with these fragments, such as the analysis of bactericidal activity using the Live/Dead cell assay (Section 3.5).

### 4.7.1 Future work

Future work could include CD analysis of the fully deglycosylated form of PlnKW30, as it is important to show that the structure of the peptide was not modified by the deglycosylation.
4.8 CONCLUSIONS

The phylogenetic analysis showed that Lactobacillus plantarum KW30 is closely related to Lactobacillus plantarum WCFS1.

The core plnKW30 gene cluster comprises six open reading frames that are most likely to be essential for the production and maturation of the bacteriocin PlnKW30.

The plnKW30 gene cluster is probably transcribed as one large mRNA, but the plnKW30 gene gccA is also transcribed from its own promoter during bacteriocin production. The relatively abundant gccA transcript can be detected in early log phase cells and increases up to the beginning of stationary phase.

The minimum inhibitory concentration of PlnKW30 against the indicator strain Lactobacillus plantarum ATCC 8014 is about 2.5 nM. PlnKW30 exhibits bacteriostatic and bactericidal activity. The indicator strain can be protected from the bactericidal activity of PlnKW30 by the addition of N-acetylglucosamine or glucosamine to the growth medium, but not by the addition of mannose or N-acetylgalactosamine.

Digestion of PlnKW30 into a 32 amino acid and 11 amino acid fragment by trypsin virtually abolishes bactericidal activity. The 32 amino acid N-terminal fragment retains some bacteriostatic activity of PlnKW30. Reduction of the disulfide bonds of the full length PlnKW30 abolishes antimicrobial activity. A decrease in bactericidal activity of PlnKW30 by 2/3 is observed after the removal of the O-linked GlcNAc from the serine residue.

PlnKW30 has two N-acetylglucosamine modifications: one O-linked at serine 18 and another S-linked at the C-terminal cysteine 43. The modification of the cysteine 43 is a biologically unprecedented S-glycosylation and the first verified example of a naturally occurring peptide S-glycosidic bond. It was possible to deglycosylate serine 18, but not cysteine 43. PlnKW30 is the first verified class IV bacteriocin, containing carbohydrate moieties. A suggested name for this new class of glycopeptide bacteriocins is 'glycocins'.

The circular dichroism spectroscopy of the wild type PlnKW30 and its N-terminal fragment showed a similar, mostly helical secondary structure for both peptides. The helicity was decreased after reduction of both disulfide bonds in PlnKW30.
## 5 Appendix

### Appendix 1: The Genetic Code

The following translation table was used to generate predicted protein translations from DNA sequence data.

<table>
<thead>
<tr>
<th>First letter</th>
<th>Second letter</th>
<th>Third letter</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>UUU (Phe)</td>
<td>UCA (Ser)</td>
</tr>
<tr>
<td></td>
<td>UUC (Leu)</td>
<td>UCC (Ser)</td>
</tr>
<tr>
<td></td>
<td>UUA (Leu)</td>
<td>UCA (Ser)</td>
</tr>
<tr>
<td></td>
<td>UUG (Leu)</td>
<td>UCG (Ser)</td>
</tr>
<tr>
<td>C</td>
<td>CUU (Leu)</td>
<td>CCU (Pro)</td>
</tr>
<tr>
<td></td>
<td>CUC</td>
<td>CCC (Pro)</td>
</tr>
<tr>
<td></td>
<td>CUA</td>
<td>CCA (Pro)</td>
</tr>
<tr>
<td></td>
<td>CUG</td>
<td>CGG (Pro)</td>
</tr>
<tr>
<td>A</td>
<td>AUU (Ile)</td>
<td>ACU (Thr)</td>
</tr>
<tr>
<td></td>
<td>AUC (Ile)</td>
<td>ACC (Thr)</td>
</tr>
<tr>
<td></td>
<td>AUA (Ile)</td>
<td>ACA (Thr)</td>
</tr>
<tr>
<td></td>
<td>AUG (Met)</td>
<td>ACG (Thr)</td>
</tr>
<tr>
<td>G</td>
<td>GUU (Val)</td>
<td>GCU (Ala)</td>
</tr>
<tr>
<td></td>
<td>GUC</td>
<td>GCC (Ala)</td>
</tr>
<tr>
<td></td>
<td>GUA</td>
<td>GCA (Ala)</td>
</tr>
<tr>
<td></td>
<td>GUG</td>
<td>GCG (Ala)</td>
</tr>
</tbody>
</table>

The table uses the following codons for amino acids:

- **UUU, UUC**: Phenylalanine (Phe)
- **UUA, UUG**: Leucine (Leu)
- **UCU, UCC, UCA, UCG**: Serine (Ser)
- **UAU, UAC**: Tyrosine (Tyr)
- **UAA, UAG**: Stop
- **UGU, UGC, UGA, UGG**: Cysteine (Cys)
- **CUU, CUC, CUA, CUG**: Leucine (Leu)
- **CCU, CCC, CCA, CGG**: Proline (Pro)
- **CAU**: Histidine (His)
- **CGU, CGC, CGA, CGG**: Arginine (Arg)
- **AUU, AUC, AUA, AUG**: Isoleucine (Ile)
- **ACU, ACC, ACA, ACG**: Threonine (Thr)
- **AAU, AAC, AAG, AAG**: Asparagine (Asn)
- **AGU, AGC, AGA, AGG**: Arginine (Arg)
- **GUU, GUC, GUA, GUG**: Valine (Val)
- **GCU, GCC, GCA, GCG**: Alanine (Ala)
- **GAU, GAC**: Aspartic Acid (Asp)
- **GGU, GGC, GGA, GGG**: Glycine (Gly)

This table is a standard genetic code used in biology and biochemistry to translate DNA sequences into proteins.
### APPENDIX 2: STRUCTURES AND SYMBOLS OF THE STANDARD AMINO ACIDS

<table>
<thead>
<tr>
<th>Alanine (Ala / A)</th>
<th>Arginine (Arg / R)</th>
<th>Asparagine (Asn / N)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Alanine structure" /></td>
<td><img src="image" alt="Arginine structure" /></td>
<td><img src="image" alt="Asparagine structure" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Aspartic acid (Asp / D)</th>
<th>Cysteine (Cys / C)</th>
<th>Glutamic acid (Glu / E)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Aspartic acid structure" /></td>
<td><img src="image" alt="Cysteine structure" /></td>
<td><img src="image" alt="Glutamic acid structure" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Glutamine (Gln / Q)</th>
<th>Glycine (Gly / G)</th>
<th>Histidine (His / H)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Glutamine structure" /></td>
<td><img src="image" alt="Glycine structure" /></td>
<td><img src="image" alt="Histidine structure" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Isoleucine (Ile / I)</th>
<th>Leucine (Leu / L)</th>
<th>Lysine (Lys / K)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Isoleucine structure" /></td>
<td><img src="image" alt="Leucine structure" /></td>
<td><img src="image" alt="Lysine structure" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Methionine (Met / M)</th>
<th>Phenylalanine (Phe / F)</th>
<th>Proline (Pro / P)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Methionine structure" /></td>
<td><img src="image" alt="Phenylalanine structure" /></td>
<td><img src="image" alt="Proline structure" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Serine (Ser / S)</th>
<th>Threonine (Thr / T)</th>
<th>Tryptophan (Trp / W)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Serine structure" /></td>
<td><img src="image" alt="Threonine structure" /></td>
<td><img src="image" alt="Tryptophan structure" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tyrosine (Tyr / Y)</th>
<th>Valine (Val / V)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Tyrosine structure" /></td>
<td><img src="image" alt="Valine structure" /></td>
</tr>
</tbody>
</table>
# APPENDIX 3: LISTS OF STRAINS AND VECTORS USED IN THIS THESIS

## Table 5.1: List of strains used in this thesis

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant Genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> TOP10</td>
<td>F', araD139 Δ(ara, leu)7697, ΔlacX74, galU, galK, rpsL, deoR, Φ80dlacZΔM15, endA1, nupG, recA1, mcrA, Δ(mrr hsdRMS mcrBC)</td>
<td>Invitrogen</td>
</tr>
<tr>
<td><em>E. coli</em> XL1-Blue</td>
<td>recA1, endA1, gyrA96, thi-1, hsdR17 (rK, mK+), supE44, relA1, lac [F', proAB, lacIq, lacZΔM15, Tn10(Tc')]</td>
<td>(Bullock et al. 1987)</td>
</tr>
<tr>
<td><em>E. coli</em> DH5α</td>
<td>supE44 hsdR17 recA1 gyrA96 thi-1 relA1</td>
<td>(Hanahan 1983)</td>
</tr>
<tr>
<td><em>E. coli</em> EC1000</td>
<td>RepA1 MC1000, Km, carrying a single copy of the pWV01 repA gene in the glgB gene; host for pORI28-based plasmids</td>
<td>(Law et al. 1995)</td>
</tr>
<tr>
<td><em>E. coli</em> BL21 (DE3)</td>
<td>F', ompT hsdS(rB mB)gal dcm (DE3)</td>
<td>Novagen</td>
</tr>
<tr>
<td><em>E. coli</em> Origami (DE3)</td>
<td>Δara–leu7697 ΔlacX74 ΔphoA PvuII phoR araD139 ahpC galE galK rpsL F' [lac' lacIq pro] gor522 ::Tn10 (Tc') trxB::kan (DE3)</td>
<td>Novagen</td>
</tr>
<tr>
<td><em>E. coli</em> Rosetta (DE3)</td>
<td>F+ ompT hsdS8(rB mB) gal dcm lacY1 (DE3) pRARE (Cm')</td>
<td>Novagen</td>
</tr>
<tr>
<td><em>E. coli</em> Rosetta pLysS (DE3)</td>
<td>F+ ompT hsdS8(rB mB) gal dcm lacY1 (DE3) pLysSRARE (Cmr)</td>
<td>Novagen</td>
</tr>
<tr>
<td><em>E. coli</em> Rosetta Blue (DE3)</td>
<td>endA1 hsdR17(rK12 mK12') supE44 thi-1 recA1 gyr96 relA1 lac F' [proAB' lacIq] ZΔM15 ::Tn10 (Tc') pRARE (Cm')</td>
<td>Novagen</td>
</tr>
<tr>
<td><em>E. coli</em> Rosetta-gami (DE3)</td>
<td>Δara–leu7697 ΔlacX74 ΔphoA PvuII phoR araD139 ahpC galE galK rpsL F' [lac' (lacIq) pro] gor522 ::Tn10 (Tc') trxB::kan (DE3) pRARE (Cm')</td>
<td>Novagen</td>
</tr>
<tr>
<td><em>E. coli</em> TB1</td>
<td>F+ ara Δ(lac-proAB) [Φ80dlac Δ(lacZ)M15] rpsL(Str') thi hsdR</td>
<td>New England Biolabs</td>
</tr>
<tr>
<td><em>L. plantarum</em> KW30</td>
<td>PlnKW30 producer strain</td>
<td>(Kelly et al. 1996)</td>
</tr>
<tr>
<td><em>L. plantarum</em> ATCC 8014</td>
<td>PlnKW30 indicator strain</td>
<td>ATCC</td>
</tr>
<tr>
<td><em>L. plantarum</em> NC8</td>
<td>Host strain, silage isolate</td>
<td>(Aukrust &amp; Blom 1992)</td>
</tr>
<tr>
<td><em>L. sakei</em> Lb790</td>
<td>Host strain, meat isolate</td>
<td>(Schilling &amp; Lucke 1989)</td>
</tr>
<tr>
<td><em>L. lactis</em> MG1363</td>
<td>Derivative of <em>L. lactis</em> NCD0712</td>
<td>(Gasson 1983)</td>
</tr>
<tr>
<td><em>L. lactis</em> NZ9000</td>
<td>Model strain MG1363 with nisRK integrated in chromosome; host strain for nisin inducible vectors</td>
<td>(Kuipers et al. 1998)</td>
</tr>
</tbody>
</table>
**Chapter 5: Appendix**

**Appendix 3: Lists of strains and vectors**

*L. lactis NZ9800*  
NZ9700 DnisA, not able to produce nisin, but still immune; host strain for nisin inducible vectors (Kuipers et al. 1998)

*Streptococcus pyogenes*  
IMBS CC

*Oenococcus oeni*  
IMBS CC

*Saccharomyces cerevisiae*  
IMBS CC

---

a Cm\(^r\), Chloramphenicol resistance; Km\(^r\), Kanamycin resistance; Str\(^r\), Streptomycin resistance; Tc\(^r\), Tetracycline resistance.

b ATCC: American Type Culture Collection; IMBS CC: Institute of Molecular BioSciences culture collection, Massey University, NZ; Invitrogen, CA, USA; Novagen, Merck KGaA, Darmstadt, GER.

---

### Table 5.2: Vectors and plasmids used in this thesis

<table>
<thead>
<tr>
<th>Vector</th>
<th>Size [bp]</th>
<th>Selection and other features</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pT7-7</td>
<td>2469</td>
<td>Amp(^r)</td>
<td>(Tabor &amp; Richardson 1985)</td>
</tr>
<tr>
<td>pPROExHTb</td>
<td>4778</td>
<td>Amp(^r)</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>pMAL_p2G</td>
<td>6721</td>
<td>Amp(^r); lac(^I); P(<em>{lac}); malE; lacZ(</em>\alpha); M13 origin</td>
<td>Novagen</td>
</tr>
<tr>
<td>pMAL_c2G</td>
<td>6721</td>
<td>Amp(^r); lac(^I); P(<em>{lac}); malE (signal sequence deleted); lacZ(</em>\alpha); M13 origin</td>
<td>Novagen</td>
</tr>
<tr>
<td>pET32a(+)</td>
<td>5900</td>
<td>Amp(^r); T7 promoter; optional C-terminal or internal His-tag; internal S-tag; N-terminal Trx-tag</td>
<td>Novagen</td>
</tr>
<tr>
<td>pETDuet_TrxA_HST_Nco</td>
<td>5704</td>
<td>Amp(^r); two MCS preceded by P(_{T7}) &amp; RBS; pBR322 origin; lacI; His-tag &amp; rTEV site into site 1; TrxA inserted into site 2 (NdeI/XhoI)</td>
<td>Novagen</td>
</tr>
<tr>
<td>pETDuet_DsbC_HST_Nco</td>
<td>6025</td>
<td>Amp(^r); two MCS preceded by P(_{T7}) &amp; RBS; pBR322 origin; lacI; His-tag &amp; rTEV site into site 1; DsbC inserted into site 2 (NdeI/XhoI)</td>
<td>Novagen</td>
</tr>
<tr>
<td>pETDuet_MalE_HST_Nco</td>
<td>6478</td>
<td>Amp(^r); two MCS preceded by P(_{T7}) &amp; RBS; pBR322 origin; lacI; His-tag &amp; rTEV site into site 1; MalE inserted into site 2 (NdeI/XhoI)</td>
<td>Novagen</td>
</tr>
<tr>
<td>pETDuet_SUMO_HST_Nco</td>
<td>5705</td>
<td>Amp(^r); two MCS preceded by P(_{T7}) &amp; RBS; pBR322 origin; lacI; His-tag &amp; rTEV site into site 1; SUMO inserted into site 2 (NdeI/XhoI)</td>
<td>Novagen</td>
</tr>
<tr>
<td>Vector Name</td>
<td>Size</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>pORI28</td>
<td>1597</td>
<td>Em\textsuperscript{r}, ori (pWV01), replicates only with repA provided \textit{in trans} (Law \textit{et al.} 1995)</td>
<td></td>
</tr>
<tr>
<td>pTRK669</td>
<td>1462</td>
<td>ori (pWV01), Cm\textsuperscript{r}, RepA1 (Russell &amp; Klaenhammer 2001)</td>
<td></td>
</tr>
<tr>
<td>pSIP409</td>
<td>5500</td>
<td>spp-based expression vector, Em\textsuperscript{r}; SppKR expression driven by ermL read-through; with P\textsubscript{ori} :: \textit{gusA} (Sorvig \textit{et al.} 2005)</td>
<td></td>
</tr>
<tr>
<td>pSIP412</td>
<td>5500</td>
<td>spp-based expression vector, Em\textsuperscript{r}; SppKR expression driven by ermL read-through; with SH71rep and P\textsubscript{ori} :: \textit{pepN} (Sorvig \textit{et al.} 2005)</td>
<td></td>
</tr>
<tr>
<td>pNZ8112</td>
<td>3478</td>
<td>Broad host range vector with \textit{nisA}-promoter and mcs; terminator sequence behind the mcs; \textit{usp45} signal sequence translationally fused to the \textit{nisA}-promoter and \textit{NaeI} site for gene fusion; possibility for C-terminal His-tag; Cm\textsuperscript{r} NIZO</td>
<td></td>
</tr>
<tr>
<td>pNZ8113</td>
<td>3368</td>
<td>Broad host range vector with \textit{nisA}-promoter and mcs; terminator sequence behind the mcs; possibility for C-terminal His-tag; Cm\textsuperscript{r} NIZO</td>
<td></td>
</tr>
<tr>
<td>pNZ8148</td>
<td>3165</td>
<td>Broad host range vector with \textit{nisA}-promoter and mcs; terminator sequence behind the mcs; with a deletion of 184 bp to remove a small \textit{B. subtilis} DNA fragment; Cm\textsuperscript{r} NIZO</td>
<td></td>
</tr>
<tr>
<td>pNZ5319</td>
<td>7028</td>
<td>Low copy vector pIL252 with \textit{nisR} and \textit{nisK} genes; compatible with pNZ8*** vectors in other hosts than \textit{Lactococcus lactis}; Em\textsuperscript{r} NIZO</td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a} Amp\textsuperscript{r}, Ampicillin resistance; Cm\textsuperscript{r}, Chloramphenicol resistance; Em\textsuperscript{r}, Erythromycin resistance; Gm\textsuperscript{r}, Gentamycin resistance; Km\textsuperscript{r}, Kanamycin resistance; MCS, multiple cloning site; RBS, ribosome binding site; rTEV.

\textsuperscript{b} Invitrogen, CA, USA; Novagen, Merck KGaA, Darmstadt, GER; NIZO food research, The Netherlands.
## APPENDIX 4: PRIMER SEQUENCES

### Table 5.3: Primers used in sequencing and PCR

<table>
<thead>
<tr>
<th>No.</th>
<th>Primer name</th>
<th>Length (nt)</th>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
</table>

#### Classification of *L. plantarum KW30*

| 1   | Lab16                | 19          | CCTTTCCCTCACGGTACTG                                                             |
| 2   | Lacto1               | 21          | GAGTTTGATCCTGGCTCAGGA                                                          |
| 3   | Lacto2               | 21          | AACGGCGGTCTTGTTCTGTA                                                           |

**plnKW30 cloning**

| 4   | PreBacExF (Ndel)     | 33          | GGTTGTCATGGGTAAACCTGATGGTTGTTGAT                                               |
| 5   | MatBacEx/F (Ndel)    | 34          | AAAACACATGAACTGATGGTGGTGTTGAT                                                 |
| 6   | BacExR (BamHI)       | 30          | TCGGATCTTTAAACTGATAGCTACTAC                                                 |

**NICE system**

| 7   | Pln_Nael_F           | 39          | GGTTGTCATGGGTAAACCTGATGGTTGTTGAT                                               |
| 8   | Pln_HindIII_R        | 33          | GGTTGTCATGGGTAAACCTGATGGTTGTTGAT                                               |
| 9   | Pln XbaI_R           | 31          | GTCTAGACTACATAGCTACTGTTAGCAAGAGGA                                               |
| 10  | MatBacDPF            | 22          | CGAAACACATGATGGTGGTGTTGTTGAT                                                  |
| 11  | PreBacDPF            | 24          | CGAAACACATGATGGTGGTGTTGTTGAT                                                  |

**Modification of bacteriocin**

| 12  | plnMatModNcoI        | 35          | GGTTGTCATGGGTAAACCTGATGGTTGTTGAT                                               |
| 13  | PlnMatModHindIII (Met)| 31         | GGTGTCATGGGTAAACCTGATGGTTGTTGAT                                               |
| 14  | plnModSer_HindIII (Ser)| 31        | GGTGTCATGGGTAAACCTGATGGTTGTTGAT                                               |

**Cloning of RR (former putative FTase)**

| 15  | FTase.RR HindIII_R  | 36          | GGTTGTCATGGGTAAACCTGATGGTTGTTGAT                                               |
| 16  | FTase.RRNcoI        | 32          | GGTTGTCATGGGTAAACCTGATGGTTGTTGAT                                               |
| 17  | FTase.RR BglR       | 37          | GGTTGTCATGGGTAAACCTGATGGTTGTTGAT                                               |
| 18  | FTase.RNdeF         | 33          | GGTTGTCATGGGTAAACCTGATGGTTGTTGAT                                               |
| 19  | RR R                | 27          | GTCTAGACTACATAGCTACTGTTAGCAAGAGGA                                               |
| 20  | RR int R            | 22          | CGAAACACATGATGGTGGTGTTGTTGAT                                                  |
| 21  | RR int F            | 20          | TAGTTCACAGACTAGAA                                                                |

**NICE system**

| 22  | RR Nael_F           | 36          | GGTTGTCATGGGTAAACCTGATGGTTGTTGAT                                               |
| 23  | RR XbaI_R           | 35          | GTCTAGACTACATAGCTACTGTTAGCAAGAGGA                                               |

**Yeast system**

| 24  | RNdeF_Yeast         | 37          | GGTTGTCATGGGTAAACCTGATGGTTGTTGAT                                               |
| 25  | RRBamR_Yeast        | 37          | ATGTTCAAGACTAGAA                                                                |

**Cloning of new putative FTase target (lp_1715; PTase; *L. plantarum* WCFS1)**

| 26  | Lp_1715N.col        | 37          | GGTTGTCATGGGTAAACCTGATGGTTGTTGAT                                               |
| 27  | Lp_1715 HindIII     | 24          | GGTTGTCATGGGTAAACCTGATGGTTGTTGAT                                               |

**Cloning of GTase**

| 28  | GlycoT KW30_BamHI_F | 39          | ATATCGGATCATGAAAAATAGACAAAGTAAAATGAGTTGACA                                       |
| 29  | GlycoT KW30 HindIII_R| 57         | TCGGATCTTTAAACTGATAGCTACTAC                                                   |
| 30  | GT R                | 27          | TTAATTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTA |
### NICE system

<table>
<thead>
<tr>
<th>Primer sequence</th>
<th>GC content</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GTase_NcoI_F</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>GTase_HindIII_R</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>GTase_NaeI_F</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>GTase_XbaI_R</td>
<td>31</td>
<td></td>
</tr>
</tbody>
</table>

### Yeast system

<table>
<thead>
<tr>
<th>Primer sequence</th>
<th>GC content</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GNdeF_Yeast</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>GTBamR_Yeast</td>
<td>38</td>
<td></td>
</tr>
</tbody>
</table>

### Other target genes of PlnKW30 gene cluster

<table>
<thead>
<tr>
<th>Primer sequence</th>
<th>GC content</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC1KW30_Bam_F</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>ABC1KW30_Not_R</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>ABC2KW30_Ncol_F</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Trx1KW30_Ncol_F</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Trx1KW30_XhoI_R</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Trx2KW30_Bam_F</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>Trx2KW30_HindIII_R</td>
<td>34</td>
<td></td>
</tr>
</tbody>
</table>

### Modification of pET vector

<table>
<thead>
<tr>
<th>Primer sequence</th>
<th>GC content</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEVNcolEcoRIHindIII</td>
<td>47</td>
<td></td>
</tr>
</tbody>
</table>

### Gene disruption using Russell & Klaenhammer 2001

<table>
<thead>
<tr>
<th>Primer sequence</th>
<th>GC content</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KW30FTaseKO_BamF</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>KW30FTaseKO_EcoR</td>
<td>29</td>
<td></td>
</tr>
</tbody>
</table>

### Gene disruption using cre-lox system (Lambert et al. 2007)

<table>
<thead>
<tr>
<th>Primer sequence</th>
<th>GC content</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pln_KO_1F</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Pln_KO_1R</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Pln_KO_2F</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Pln_KO_2R</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Pln_KO_1F_XhoI</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Pln_KO_2R_BglII</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>GTase_KO_1F</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>GTase_KO_1R</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>GTase_KO_2F</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>GTase_KO_2R</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>GTase_KO_1F_XhoI</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>GTase_KO_2R_BglII</td>
<td>34</td>
<td></td>
</tr>
</tbody>
</table>

### Cre-lox verifying primer

<table>
<thead>
<tr>
<th>Primer sequence</th>
<th>GC content</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>85_65</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>85_67</td>
<td>22</td>
<td></td>
</tr>
</tbody>
</table>

### Housekeeping genes

<table>
<thead>
<tr>
<th>Primer sequence</th>
<th>GC content</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>gap_F</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>gap_R</td>
<td>20</td>
<td></td>
</tr>
</tbody>
</table>

### Sequencing of plnKW30 gene cluster

<table>
<thead>
<tr>
<th>Primer sequence</th>
<th>GC content</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BacLin/F (degenerate primer for first sequencing)</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Primer Sequence</td>
<td>Length (bp)</td>
</tr>
<tr>
<td>---</td>
<td>------------------------------------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>67</td>
<td>BacCon/R (degenerate primer for first sequencing)</td>
<td>24</td>
</tr>
<tr>
<td>68</td>
<td>F6ii</td>
<td>26</td>
</tr>
<tr>
<td>69</td>
<td>F6(CT)</td>
<td>22</td>
</tr>
<tr>
<td>70</td>
<td>LplgDNA F5ii</td>
<td>28</td>
</tr>
<tr>
<td>71</td>
<td>LplgDNA F5</td>
<td>28</td>
</tr>
<tr>
<td>72</td>
<td>LplgDNA F4</td>
<td>28</td>
</tr>
<tr>
<td>73</td>
<td>LplgDNA F3</td>
<td>29</td>
</tr>
<tr>
<td>74</td>
<td>LplgDNA F2</td>
<td>30</td>
</tr>
<tr>
<td>75</td>
<td>LplgDNA F</td>
<td>28</td>
</tr>
<tr>
<td>76</td>
<td>LplgDNA R</td>
<td>28</td>
</tr>
<tr>
<td>77</td>
<td>LplgDNA Rgap1</td>
<td>28</td>
</tr>
<tr>
<td>78</td>
<td>LplgDNA Rgap2</td>
<td>28</td>
</tr>
<tr>
<td>79</td>
<td>LplgDNA Rgap2ii</td>
<td>33</td>
</tr>
<tr>
<td>80</td>
<td>LplgDNA R2</td>
<td>30</td>
</tr>
<tr>
<td>81</td>
<td>LplgDNA R3</td>
<td>30</td>
</tr>
<tr>
<td>82</td>
<td>LplgDNA R3ii</td>
<td>29</td>
</tr>
<tr>
<td>83</td>
<td>R4</td>
<td>25</td>
</tr>
<tr>
<td>84</td>
<td>R4i(CT)</td>
<td>22</td>
</tr>
<tr>
<td>85</td>
<td>R4ii(CT)</td>
<td>22</td>
</tr>
<tr>
<td>86</td>
<td>R5(CT)</td>
<td>25</td>
</tr>
<tr>
<td>87</td>
<td>R5ii(CT)</td>
<td>25</td>
</tr>
<tr>
<td>88</td>
<td>R6</td>
<td>26</td>
</tr>
<tr>
<td>89</td>
<td>R7</td>
<td>26</td>
</tr>
<tr>
<td>90</td>
<td>R8</td>
<td>23</td>
</tr>
<tr>
<td>91</td>
<td>R8a</td>
<td>21</td>
</tr>
<tr>
<td>92</td>
<td>R9</td>
<td>23</td>
</tr>
<tr>
<td>93</td>
<td>R10</td>
<td>26</td>
</tr>
<tr>
<td>94</td>
<td>R11</td>
<td>26</td>
</tr>
<tr>
<td>95</td>
<td>R12</td>
<td>29</td>
</tr>
<tr>
<td>96</td>
<td>R13</td>
<td>25</td>
</tr>
<tr>
<td>97</td>
<td>R14</td>
<td>22</td>
</tr>
<tr>
<td>98</td>
<td>FR9</td>
<td>27</td>
</tr>
<tr>
<td>99</td>
<td>FR10</td>
<td>22</td>
</tr>
<tr>
<td>100</td>
<td>FR11</td>
<td>21</td>
</tr>
<tr>
<td>101</td>
<td>FR12</td>
<td>24</td>
</tr>
<tr>
<td>102</td>
<td>FR15</td>
<td>24</td>
</tr>
<tr>
<td>103</td>
<td>20103_R</td>
<td>21</td>
</tr>
<tr>
<td>104</td>
<td>Rsub13</td>
<td>23</td>
</tr>
<tr>
<td>105</td>
<td>Ex18R1</td>
<td>23</td>
</tr>
</tbody>
</table>
APPENDIX 5: RESULTS OF EARLY EXPERIMENTS

This section describes the experiments that were done and results achieved before it was discovered that the C-terminal modification of PInKW30 was actually an N-acetylglucosamine rather than a farnesyl group.

5.1 INTRODUCTION INTO FARNESYLATION

Prenylation is a lipid modification where the C1 of farnesyl (15-carbon) or geranylgeranyl (20-carbon) isoprenoids is bound via a thioether linkage to conserved cysteine residues at or near the C-terminus of proteins (Zhang & Casey 1996). There are three known enzymes which catalyze this reaction: farnesyltransferase (FTase), geranylgeranyltransferase I (GGTase I) and geranylgeranyltransferase II (GGTase II). FTase and GGTase I recognize a CaaX (C, cysteine; a, typically an aliphatic amino acid; X, C-terminal amino acid) motif at the C-terminus of proteins as the site for prenylation, whereas GGTase II recognizes C-terminal CC or CxC motifs. In general, the C-terminal residue of the CaaX motif determines whether farnesyl or geranylgeranyl will be transferred to a protein. Serine, methionine, glutamine or alanine as “X” results in modification by FTase, while a leucine at this position is recognized by GGTase I (Casey et al. 1991; Yokoyama et al. 1991). Following the prenylation step, CaaX-containing proteins are typically further modified in two steps: (1) the three C-terminal residues (i.e. the “aaX”) are cleaved by an endoprotease (Ashby 1998); (2) the prenylated cysteine is carboxylmethylated by a specific methyltransferase (Stimmel et al. 1990).

It is estimated that about 2% of eukaryotic proteins are prenylated. These prenylated proteins include fungal mating factors, nuclear lamins, Ras and Ras-related G proteins, subunits of heterotrimeric G proteins and protein kinases. The group of CC- or CxC-containing proteins is not so diverse and consists almost exclusively of members of the Rab family of small GTP-binding proteins that participate in intracellular membrane trafficking (Novick & Brennwald 1993).

The hydrophobic prenyl groups support membrane interactions, but the farnesyl group alone makes only a small contribution to membrane association. Methylation of prenylated
cysteines abolishes the negative charge at the carboxyl terminus and thereby increases hydrophobicity. Prenylation also promotes specific protein-protein interactions (Sinensky 2000). For example, the attachment of a farnesyl group to Ras2 increases its affinity for adenylate cyclase about 100-fold (Kuroda et al. 1993).

To date, no bacterial FTases have been described, although a recent publication shows that the structure of the lanthionine cyclase NisC produced by *L. lactis* is similar to the β subunit of mammalian FTase. NisC is involved in the biosynthesis of the five cyclic thioethers of the lantibiotic nisin (Li et al. 2006).

5.2 EXPERIMENTAL RESULTS

5.2.1 WESTERN BLOTTING USING ANTI-FARNESYL ANTIBODIES

The work on this thesis started with the assumption that the C-terminal cysteine of PlnKW30 was modified by a thioether linked farnesyl group. That the C-terminal residue was a modified cysteine was first indicated by N-terminal Edman degradation of proteolytic fragments and the presence of a C-terminal sequence modification was confirmed by tandem mass spectrometry (SCIEX 300 triple quadrupole instrument with collision induced fragmentation (CID)). The additional molecular mass of 204 mass units and identification of the C-terminal residue as cysteine, strongly suggested the modification was a farnesyl group. To confirm the identity of the modification PlnKW30 was analysed by Western blot using anti-farnesyl antibodies.

In this work, the Western blots were repeated with various controls in order to verify that one of the post-translational modifications of PlnKW30 was a farnesyl group. The samples included purified PlnKW30, recombinant (unmodified) bacteriocin produced in *E. coli*, whole cell extracts of *L. plantarum* KW30, whole cell extracts of *Streptococcus pyogenes* (negative control) and a human cell line whole cell extract (positive control). All samples were separated by tricine SDS-PAGE (Section 2.4.2; figure 5.1A) and then subjected to Western blotting using anti-farnesyl antibodies (batch 2) (Section 2.4.6; figure 5.1B).
The native PInKW30 (Figure 5.1B, lane 2) but not the recombinant sample (lane 1) was visible on the blot. In the whole cell extracts of *L. plantarum* KW30 three bands of about 16, 17 and 35 kDa (indicated by arrows) were reproducibly detected with the anti-farnesyl antibodies. The bands on the blot are not the most intense bands in comparison to the tricine SDS-PAGE, which shows the high specificity of the antibodies. As expected, in the positive control sample (human cell extract) several proteins showed up on the Western blot (lane 4). The most prominent band is 16 - 17 kDa, which probably is Ras protein (human Ras is 21 kDa) and the other bands might be less abundant small GTP-binding proteins (20 – 25 kDa) and nuclear lamins (~66 kDa). No bands are visible in the negative control sample (lane 7).
5.2.2 **TWO-DIMENSIONAL GEL ELECTROPHORESIS**

To identify the three proteins detected by the anti-farnesyl antibodies (16, 17 and 35 kDa; figure 5.1B) in the whole cell extracts of *L. plantarum* KW30 proteomic methods were used. First the cell extracts were subjected to duplicate two-dimensional (2D) gel electrophoresis (Section 2.4.4). One gel was stained with Coomassie blue G250, while the second was subjected to Western blotting using anti-farnesyl antibodies (Section 2.4.6; figure 5.2).

![2D-electrophoresis gel of whole cell extract of L. plantarum KW30.](image)

The circles indicate the spots identified by Western blotting with anti-farnesyl antibodies, which were excised and subjected to tryptic digest. M = protein marker in kDa.

Spots on the stained gel that coincided with those that appeared on the Western blot (indicated in figure 5.2 by circles) were excised, then subjected to tryptic digestion (Section 2.4.5). The extracts were analysed by mass spectrometry (Section 2.5.2), but unfortunately, no definitive results were obtained, because the masses could not be matched to *L. plantarum* proteins. The detected bands at 16 and 17 kDa in figure 5.1 were not detected in figure 5.2, but several spots are visible at 46, 51 and 64 kDa. The experiments were not repeated because the true identity of the C-terminal modification of PlnKW30 was determined.
5.2.3 Farnesyltransferase Activity Assay

In order to detect the putative farnesyltransferase (FTase) activity in different \textit{L. plantarum} KW30 cell extracts, an \textit{in vitro} FTase activity assay was developed. Firstly, the assay was developed using crude \textit{Saccharomyces cerevisiae} (yeast hereafter) extract or purified recombinant yeast FTase (Sigma) as positive control and a yeast-specific peptide (GCVLS), as reported in the literature (Hightower \textit{et al.} 2001). Later, \textit{L. plantarum} KW30 cell extracts were used as a source of putative FTase and the C-terminal synthetic peptide of PlnKW30 (FGIKHHSSGSSSYHC) was used as the acceptor peptide. All assays were carried out using the appropriate controls (Section 2.4.15).

The FTase reaction mixes and controls were separated by thin layer chromatography (TLC) and reaction products containing tritiated farnesyl pyrophosphate ($^{3}$H-FPP) were visualised by exposure to film (Section 2.4.15). In figure 5.3A spots are visible in the two lanes with yeast extract (lanes 3 and 4), which are not present in the negative control reactions (lanes 5 - 7). Although this could have been radio labelled farnesylated peptide, there were no corresponding spots visible in lanes 1 and 2, with the purified recombinant FTase. The two possible explanations of this anomaly were (1) that the pure yeast enzyme was not active because of some event, such as freeze thawing during delivery from the USA, and (2) that the spots were not farnesylated peptide, but some other metabolite which had reacted with the $^{3}$H-FPP.

This assay was repeated using the same samples, except for the whole cell yeast extract, which was prepared fresh each time (Figure 5.3B). However, the spots visible in lanes 3 and 4 of figure 5.3B with freshly prepared yeast extract were not in the same place as those in the first assay (Figure 5.3A), despite the assay and TLC conditions being identical. When these results were compared to those obtained by Vogt \textit{et al.} (1995), it was possible to assign identities to spots in figures 5.3A and B.

The TLC plates shown in figure 5.3 were sprayed with ninhydrin (0.2 % in ethanol) and heated at 110 °C for 5 – 10 minutes. All spots seen on the autoradiography films were also visible after ninhydrin treatment.
Chapter 5: Appendix

Appendix 5: Results of early experiments

Figure 5.3: Autoradiography films of FTase activity assays.
The results of two different experiments are shown (a and b). The farnesylation of the peptide GCVLS was detected by thin layer chromatography followed by autoradiography. 1: Pure FTase; 1µl ³H-FPP; 2: Pure FTase; 3µl ³H-FPP; 3: Yeast extract; 1µl ³H-FPP; 4: Yeast extract; 3µl ³H-FPP; 5: No FTase; 6: No peptide; 7: ³H-FPP only, 1 µL.

Because of the lack of consistency with TLC, another assay was developed using RP-HPLC. Whole cell extracts of *L. plantarum* KW30 were used as enzyme source with an appropriate acceptor peptide (FGIKHHSSGSSSYHC). The assay was performed as described in section 2.4.15 and subjected to RP-HPLC as described in section 2.5.1. The chromatogram (black) is shown in figure 5.4 overlaid with a chromatogram of the assay performed without the cell lysate (red), and the assay performed without peptide (blue). The separate peaks at 3.5 minutes, 4.2 minutes, 4.5 minutes and 5.2 minutes (indicated by arrows) were collected and analysed by mass spectrometry, but again the results did not correspond to the mass of the peptide + farnesyl group. The peak eluting at 4.5 minutes was not present in both the control samples (Figure 5.4), but it is unlikely to represent the peptide covalently modified by farnesyl pyrophosphate. The current knowledge (Section 3.5) shows that it is unlikely that there is a peptide FTase in *L. plantarum* KW30.
Figure 5.4: FTase activity assay separated by RP-HPLC.
A Jupiter C18 column was used with solutions A: H₂O, 0.1 % TFA and B: Acetonitrile, 0.08 % TFA, conditions used as shown in chromatogram. 1 (black): FTase assay with peptide FGIKHHSSGSSSYHC; 2 (red): FTase activity assay without KW30 cell lysate; 3 (blue): FTase activity assay without peptide FGIKHHSSGSSSYHC.

5.2.4 Removal of Farnesyl Group by Methyl iodide Treatment

PlnKW30 was treated with methyl iodide (CH₃I) to remove the putative farnesyl group from its C-terminal cysteine using the method of Casey et al. (1989) (Section 2.4.14). After the reaction was complete, the sample was purified using RP-HPLC as shown in figure 5.5. The main peak in the CH₃I-treated sample had a lower retention time than PlnKW30, eluting at 28 - 29 % B compared to 36 - 37 % B, respectively.
Chapter 5: Appendix

Appendix 5: Results of early experiments

5.3 DISCUSSION

A lot of work was carried out to optimise the conditions for the Western blot with anti-farnesyl antibodies as it was difficult to obtain consistent results. This included using different membranes, different sample preparations, and different blotting methods. It was also found that different batches of the same commercial (Sigma) anti-farnesyl antibodies gave different results. One batch (batch 2) of antibodies resulted in the pattern of bands shown in figure 5.1B, where PlnKW30 is visible, as well as the bands in the human cell extract and L. plantarum KW30 cell extracts. In contrast, another batch (batch 1) of
antibodies that were used in earlier experiments only reacted with proteins of the human cell extract and *L. plantarum* KW30 cell extracts (result not shown).

These results can be explained by a variation in specificity of the antibodies. The anti-farnesyl antibodies are raised against *N*-acetyl-5-farnesyl-L-cysteine conjugated to keyhole limpet hemocyanin (KLH) as immunogen, and can recognise farnesyl-cysteine-BSA and cross-react with KLH or geranylgeranyl-cysteine. A possible explanation for the recognition of PlnKW30 by these antibodies is that batch 2 of the antibodies recognised the cysteine thiol–carbon bond and not the farnesyl group itself.

Attempts to identify other farnesylated proteins in the *L. plantarum* KW30 proteome using peptide mapping also failed, despite the well separated and relatively intense spots (Figure 5.2). One of the reasons for this was contamination of the gel with keratin. Efforts to repeat these experiments were curtailed by the identification of the C-terminal modification as *N*-acetylglucosamine.

It is hardly surprising that all attempts to develop an assay for the putative farnesytransferase of *L. plantarum* KW30 failed. Initially it was thought that one ORF: gccF in the *plnKW30* gene cluster (now: putative response regulator) could be a farnesytransferase, although this gene had no sequence homology to other farnesytransferases in public data bases. C-terminal farnesylation had never been reported in bacteria and it could have been a new kind of farnesyltransferase, quite different to its eukaryotic counterparts.

The FTase activity assays were developed using as a ‘positive control’ yeast whole cell extracts and purified yeast farnesyltransferase (Sigma). The purified yeast FTase did not result in any change from the control samples, thus it is possible that it was not active. In contrast, reactions containing yeast whole cell extract had additional visible spots that differed from those seen in the negative controls (Figure 5.3).

The assay was not reproducible (Figure 5.3A versus B) despite great care being taken to ensure that the running conditions of the TLC plates were the same each time. Vogt *et al.* (1995) assigned possible identities to their TLC spots, which were used to identify the spots on the plates. However, these identifications are unconfirmed because of the
irreproducibility. In order to test where the peptide was on the plate, it was sprayed with ninhydrin. This was not successful as all spots were visualised and it was not possible to distinguish the peptide. It should have been possible to separate a farnesylated peptide from a non-farnesylated one on the basis of hydrophobicity. Therefore the products of the reaction using KW30 cell extract were subjected to RP-HPLC. The results obtained did not make sense, and although a "product peak" could be produced repeatedly, mass spectrometry showed it was not the peptide substrate, the farnesylated peptide, nor the farnesyl pyrophosphosphate. These results are not surprising when it is understood that *L. plantarum* KW30 most likely does not contain a peptide FTase (Section 3.5).

### 5.3.1 CONCLUSIONS

The discovery that the modification at the C-terminal cysteine was an *N*-acetylglucosamine (Section 3.5.1) considerably changed the direction of the thesis and explains some of the disappointing results of early experiments.

However, the apparent specificity of the anti-farnesyl antibodies still remains a mystery, and it would be interesting to identify the epitope responsible for the apparently specific interaction. Future work may be done to clarify this, once the 2D gels are repeated using anti-\(\text{O-GlcNAc}\) antibodies. Nevertheless, there were reproducible changes that occurred in the assays which, together with the presumed specificity of the Western blot, suggested that the modification was a farnesyl group.

Furthermore, the mass results from the methyl iodide treatment of PlnKW30 could be explained by the reduction of the disulfide bonds plus modification of the four histidines. As PlnKW30 is inactivated by reduction of the disulfides the results were inconclusive as to whether a CH\(_3\)I-susceptible modification is required for activity.
APPENDIX 6: PROTEIN EXPRESSION

5.4 PROTEIN EXPRESSION

In order to express soluble protein for crystallization and structure determination the full-length *plnKW30 (gccA)*, response regulator (RR, *gccF*), PTase (prenyltransferase; possible FTase equivalent; homologous to lp_1715 of *L. plantarum* WCFS1) and glycosyltransferase (*gccB*) genes were cloned into several vectors as listed in table 5.5. The numbers of the primers used refer to the list in appendix 4. Some of these vectors included solubility tags such as the maltose binding protein or thioredoxin protein, which sometimes help to improve the solubility of the recombinant protein. All expression constructs were sequenced (Allan Wilson Centre Sequencing Service, Palmerston North) to confirm that the cloned genes were free of PCR errors and, in the case of expression constructs with N-terminal tags, in frame.

Expression trials were carried out using various *E. coli* strains (Table 5.4), different lactobacilli (pSIP system; Sorvig *et al.* 2005) and lactococci (Nisin-controlled expression (NICE) system; NIZO food research, The Netherlands) as expression hosts (Sections 2.4.7(b) and (c)). All constructs and hosts used in these expression trials are listed in table 5.5.

**Table 5.4: List of *E. coli* expression strains and their descriptions.**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> BL21 (DE3)</td>
<td>general expression host</td>
</tr>
<tr>
<td><em>E. coli</em> BL21 (DE3) Rosetta</td>
<td>provides rare codon tRNAs</td>
</tr>
<tr>
<td><em>E. coli</em> BL21 (DE3) Rosetta-gami</td>
<td>enhances disulfide bond formation in <em>E. coli</em> cytoplasm; provides rare codon tRNAs</td>
</tr>
<tr>
<td><em>E. coli</em> BL21 (DE3) Origami</td>
<td>enhances disulfide bond formation in <em>E. coli</em> cytoplasm</td>
</tr>
<tr>
<td><em>E. coli</em> BL21 (DE3) Origami pLysS</td>
<td>High stringency expression host; enhance disulfide bond formation in <em>E. coli</em> cytoplasm</td>
</tr>
</tbody>
</table>

Different growth temperatures and IPTG concentrations (or the respective induction peptide) were examined for each expression strain. One construct (pMAL_C2G_HST::RR) produced a small amount of soluble protein in *E. coli* TB1. However, cleavage of the maltose
binding protein from the RR and purification of the cleaved RR resulted in loss of the sample. It was possible to express a reasonable, but insoluble, amount of protein with the pETDuet_TrxA_RR construct. Despite much effort to produce soluble protein, none was produced in any other expression system or plasmid trialled.

5.4.1 **CELL-FREE expression**

Cell-free expression (Roche RTS 100 *E. coli* HY kit, Roche, GER) was carried out for the full-length *plnKW30_C43M*, response regulator and putative prenyltransferase genes using the plasmids pET32a(+)::PlnMet; pETDuet_HST_TrxA::PlnMet; pT7-7::RR; pET32a(+)::RR; pETDuet_HST_TrxA::RR and pETDuet_HST_TrxA::PTase. A GFP fusion vector was supplied and served as positive control.

It was not possible to detect any over-expression of the PlnMet, RR, PTase nor the GFP on SDS-PAGE. Expression of GFP from a positive control plasmid could be detected after Western blotting using anti-GFP antibodies as shown in figure 5.6.

![Western blot using anti-GFP antibodies](image)

**Figure 5.6: Western blot using anti-GFP antibodies.**
1: yeast whole cell extract with several GFP-labelled proteins; 2: molecular weight marker; 3: GFP positive control, Roche RTS 100 cell-free expression system.

The blot shows in lane 1 a yeast whole cell extract where several GFP-labelled proteins were detected with anti-GFP antibodies. The GFP positive control in lane 3 also gave a strong positive response. The streaky and indistinct appearance of the band is due to overexposure of the blot.
Because recombinant protein was required in concentrations large enough for structure determination using x-ray diffraction methods this method of producing recombinant protein was not pursued. The GFP result suggests that only antibody methods would be sensitive enough to detect expression with this cell-free method.

Table 5.5: Expression constructs and strains used for expression trials of \textit{plnKW30}, RR and GT in this thesis

<table>
<thead>
<tr>
<th>Construct</th>
<th>Restriction sites used</th>
<th>Modifications in vector</th>
<th>Expression host tried</th>
<th>Expression(^a)</th>
<th>Primer No.(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pT7-7::PlnKW30</td>
<td>NdeI – BamHI</td>
<td>-</td>
<td>\textit{E. coli} BL21 (DE3)</td>
<td>-</td>
<td>5, 6</td>
</tr>
<tr>
<td>pT7-7::RR</td>
<td>NdeI – BglII</td>
<td>-</td>
<td>\textit{E. coli} BL21 (DE3)</td>
<td>-</td>
<td>17, 18</td>
</tr>
<tr>
<td>pT7-7::PlnKW30</td>
<td>NdeI – BamHI</td>
<td>-</td>
<td>\textit{E. coli} BL21 (DE3) Rosetta</td>
<td>-</td>
<td>5, 6</td>
</tr>
<tr>
<td>pT7-7::RR</td>
<td>NdeI – BglII</td>
<td>-</td>
<td>\textit{E. coli} BL21 (DE3) Rosetta</td>
<td>-</td>
<td>17, 18</td>
</tr>
<tr>
<td>pT7-7::RR</td>
<td>NdeI – BglII</td>
<td>-</td>
<td>\textit{E. coli} BL21 (DE3) Origami pLysS</td>
<td>-</td>
<td>17, 18</td>
</tr>
<tr>
<td>pET32a(+)_HST::RR</td>
<td>Ncol – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>\textit{E. coli} BL21 (DE3)</td>
<td>-</td>
<td>15, 16</td>
</tr>
<tr>
<td>pET32a(+)_HST::RR</td>
<td>Ncol – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>\textit{E. coli} BL21 (DE3) Origami</td>
<td>-</td>
<td>15, 16</td>
</tr>
<tr>
<td>pET32a(+)_HST::PlnKW30</td>
<td>Ncol – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>\textit{E. coli} BL21 (DE3) Origami</td>
<td>-</td>
<td>8, 12</td>
</tr>
<tr>
<td>pMAL.C2G.HST::RR</td>
<td>Ncol – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>\textit{E. coli} BL21 (DE3)</td>
<td>-</td>
<td>15, 16</td>
</tr>
<tr>
<td>pMAL.C2G.HST::RR</td>
<td>Ncol – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>\textit{E. coli} TB1</td>
<td>+</td>
<td>15, 16</td>
</tr>
<tr>
<td>pMAL.C2G.HST</td>
<td>Ncol – HindIII</td>
<td>Addition of His-tag, spacer, rTEV</td>
<td>\textit{E. coli} BL21</td>
<td>-</td>
<td>15, 16</td>
</tr>
<tr>
<td>Construct</td>
<td>Vectors</td>
<td>E. coli</td>
<td>Comment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pMAL_C2G_HST::RR</td>
<td>NcoI – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>E. coli BL21 (DE3) Rosetta-gami</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pMAL_P2G_HST::RR</td>
<td>NcoI – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>E. coli BL21 (DE3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pMAL_P2G_HST::RR</td>
<td>NcoI – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>E. coli TB1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pETDuet_HST_TrxA::PlnKW30</td>
<td>NcoI – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>E. coli BL21 (DE3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pETDuet_HST_TrxA::RR</td>
<td>NcoI – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>E. coli BL21 (DE3) ++, IB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pETDuet_HST_MaE::PlnKW30</td>
<td>NcoI – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>E. coli BL21 (DE3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pETDuet_HST_MaE::PlnKW30</td>
<td>NcoI – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>E. coli BL21 (DE3) +, IB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pETDuet_HST_MaE::RR</td>
<td>NcoI – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>E. coli BL21 (DE3) Rosetta</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pETDuet_HST_MaE::RR</td>
<td>NcoI – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>E. coli BL21 (DE3) +</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pETDuet_HST_DsbC::PlnKW30</td>
<td>NcoI – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>E. coli BL21 (DE3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pETDuet_HST_DsbC::RR</td>
<td>NcoI – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>E. coli BL21 (DE3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pETDuet_HST_SUMO::PlnKW30</td>
<td>NcoI – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>E. coli BL21 (DE3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pETDuet_HST_SUMO::RR</td>
<td>NcoI – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>E. coli BL21 (DE3) +, IB V</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pETDuet_HST_SUMO::RR</td>
<td>NcoI – HindIII</td>
<td>Addition of His-tag, spacer, rTEV cleavage site</td>
<td>E. coli BL21 (DE3) Rosetta ++, IB</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

206
5.5 DISCUSSION

The majority of expression trials carried out used different *E. coli* strains, which included strains supporting the usage of rare codons and those that enhance the formation of disulfide bonds in the *E. coli* cytoplasm (Table 5.4). The codon usage of *L. plantarum* differs from that of *E. coli* especially in 15 codons. In order to facilitate expression in *E. coli* the sequences of the cloned genes could be adapted to the codon usage of *E. coli*.

After having no success in *E. coli*, a *Lactobacillus* expression system was trialled. The use of an expression strain from the same species could be advantageous in successfully expressing genes of the *plnKW30* gene cluster of *L. plantarum* KW30. The vectors for the pSIP system (Sorvig et al. 2005) were constructed and confirmed by sequencing, but it was not possible to express any of the target genes. Both pSIP vectors pSIP409 and pSIP412 were supplied to us containing the reporter genes *gusA* and *pepN*, respectively. At the time
it was not possible to demonstrate the expression of these reporter genes, although recently pepN has been expressed. It seems that conditions for the transformation and selection of expression hosts, and the expression conditions themselves, had to be extensively optimized and the efforts undertaken at the time were not sufficient.

The NICE system (NIZO food research, The Netherlands) has been used within the last ten years for expression of a number of genes originating from different strains, including *Lactococcus*, *Lactobacillus* (Christensson *et al.* 2002) and *Bacillus* species. Codon usage is the greatest problem, but genes of closely related Gram-positive bacteria, such as *Lactobacillus*, are mostly expressed successfully (Mierau & Kleerebezem 2005). The codon usage of *Lactobacillus plantarum* and *Lactococcus lactis* is quite similar, and only differs at nine codons. It was therefore disappointing that it was not possible to obtain any soluble protein for genes of the *plnKW30* gene cluster using this system.

DNA sequencing of the expression constructs showed them to be correct so there must be another reason why no soluble protein could be expressed. It is possible that a high concentration of some of the gene products, such as *plnKW30*, could be toxic to the host cells, although PlnKW30 has a very narrow spectrum of activity (Kelly *et al.* 1996). Another reason could be that the products of some of the *gcc* genes associate to form a multi-protein complex, with specific protein-protein interactions being required for folding and/or solubility. Associations between the proteins might also cause problems when they are expressed separately. Recently, work has been carried out to clone the complete *plnKW30* gene cluster (GTase, ABC-transporter, TRX1 & 2, RR and *plnKW30*), but the large size (~5.4 kbp) and an inconveniently located restriction site in the middle of the ABC-transporter gene make it difficult and progress is slow.

It has been shown that LytTR response regulators (Risoen *et al.* 1998; Sidote *et al.* 2008) and glycosyltransferases (pdb 3bcv) can be expressed successfully in *E. coli* expression systems. In the case of the response regulator PnnC from *L. plantarum* C11, expression in *E. coli* BL21 (DE3) pLysS yielded little soluble protein (Risoen *et al.* 1998), but expression was improved using the pSIP expression system (Straume *et al.* 2006). Expression of several class II bacteriocins in *E. coli* BL21 (DE3) was shown by Ingham *et al.* (2005) using secretable fusion constructs. Their constructs contained a C-terminal intein tag for
purification, whereas our expression constructs had no tags or were N-terminal fusions. In
the case of PlnKW30 it might have been advantageous to have a C-terminal fusion that
protects the unmodified C-terminal cysteine and thus promotes the formation of the native
disulfide bonding pattern from the remaining four cysteines. However, we concentrated on
the use of different solubility tags and trialling different *E. coli* expression strains to
promote expression of soluble protein. Direction of the expressed protein into the
periplasm was trialled using the pMAL_p2G vector, but no better results were obtained.
Very recently, a soluble (but presumably inactive) truncation of the PlnKW30
glycosyltransferase (*gccB*) (T.S. Loo & Dr. G.E. Norris - personal communication) was
expressed in *E. coli*. A truncated form of TRX2 (*gccE*) has also been expressed in a soluble
form in *E. coli* (T.S. Loo & Dr. G.E. Norris - personal communication).

Refolding was not pursued because of the complexity of the disulfide bonds, such as the
nested disulfide bonds of PlnKW30. It was possible to express wild type PlnKW30 as
inclusion bodies, but refolding was not tried because of the uneven number of cysteines.
Expression trials were undertaken to express modified PlnKW30 that had the C-terminal
cysteine mutated to a methionine or serine in order to simplify proper refolding of the two
nested disulfide bonds, but, unlike the wild type PlnKW30, the C43M mutant was not
overexpressed in the *E. coli* hosts tested.

The use of other expression systems, such as yeast expression systems or the baculovirus
expression system, was also considered. Preliminary efforts to make constructs for
expression in *Schizosaccharomyces pombe* (Craven *et al.* 1998) were stopped because of
time limitations. Time constraints also prevented the use of a baculovirus expression
system (Invitrogen). It uses slow-growing insect cells, resulting in time consuming
preparations of phage stock, subsequent transfection of the cells and finally expression.
5.5.1 CONCLUSIONS

Although all expression constructs were verified by sequencing and several expression systems and strains were used, none of the expression trials produced soluble protein. A plausible explanation for these difficulties is that several gcc gene products function in a membrane-associated multi-protein complex.

Further work possibly includes the use of truncated versions of some gcc genes to overcome this expression problem.
APPENDIX 7: TARGETED GENE DISRUPTION OF THE **plnKW30** GENE CLUSTER

Targeted gene disruption was used to analyse the functions of the genes in the **plnKW30** gene cluster. The disruption of the **plnKW30** gene should result in a non-producer strain, which is still immune to the bacteriocin. Disrupting **gccB** or **gccD** and **gccE** might result in non-native **plnKW30** that is inactive, because it lacks post-translational modifications essential for activity.

5.6 TARGETED GENE DISRUPTION USING A TEMPERATURE SENSITIVE PLASMID

The targeted gene disruption method described by Russell & Klaenhammer (2001) works via homologous recombination and uses two plasmids (Section 2.2.16(a)). The helper plasmid (**pTRK669**) is temperature sensitive and unstable at temperatures over 42 °C. This helper plasmid supplies the **repA** gene product necessary for the replication of **pORI28**, which contains a portion of the gene targeted for disruption.

The temperature sensitivity of **L. plantarum** KW30 was tested by growing it at 43 °C and as control at 30 °C (Figure 5.7). The results show that **L. plantarum** KW30 can grow at elevated temperatures (43 °C), although the OD_{600} reaches only ~2.6 compared to an OD_{600} of ~4 at 30 °C. These cultures were started with a 10 % inoculum from an overnight culture.

In order to disrupt the putative farnesyltransferase **gccF** (now putative response regulator; Section 3.2) an internal fragment of this gene was cloned into **pORI28**, resulting in the FTase mutagenesis vector **pORI28::FTaseKO**. The preparation of competent cells and the electroporation conditions were carried out using the method of Kaneko *et al.* (2000) (Section 2.2.14(f)). First, **L. plantarum** KW30 was transformed with the helper plasmid **pTRK669**; a plasmid harbouring colony was isolated and then transformed with **pORI28::FTaseKO**.

It was not possible to isolate the plasmids from the **L. plantarum** KW30 colony that possibly harboured both plasmids. However, as untransformed cells were not able to grow on the selection media, it was presumed that this colony harbours both plasmids.
A colony harbouring both plasmids was grown overnight at 30 °C in MRS broth with 10 µg/mL erythromycin (Em) and 5 µg/mL chloramphenicol (Cm). The culture was then transferred into MRS broth with 5 µg/mL Em and incubated at 43 °C. However, no growth was detectable after 24, 48 or 72 hours of incubation. This experiment was repeated and a L. plantarum KW30 colony harbouring both plasmids was grown at 30 °C and then transferred to growth cabinets at 40, 41 and 43 °C (Table 5.6).

After the transfers (the number of transfers is listed in Table 5.6) a sample of each culture was plated onto MRS and MRS with 5 µg/mL Cm plates. However, colonies grew on all plates, indicating that none of the cells had lost the helper plasmid pTRK669.
Table 5.6: Results of disruption attempts with *L. plantarum* KW30 (pTRK669, pORI28::FTaseKO) at 40, 41 and 43 °C.

1. test

<table>
<thead>
<tr>
<th>Temperature tested</th>
<th>30 °C</th>
<th>43 °C</th>
<th>43 °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of transfers</td>
<td>3</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>OD&lt;sub&gt;600&lt;/sub&gt;</td>
<td>3.857</td>
<td>Only minimal growth after two days</td>
<td>Only minimal growth after two days</td>
</tr>
</tbody>
</table>

2. test

<table>
<thead>
<tr>
<th>Temperature tested</th>
<th>30 °C</th>
<th>40 °C</th>
<th>40 °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of transfers</td>
<td>3</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>OD&lt;sub&gt;600&lt;/sub&gt;</td>
<td>3.791</td>
<td>3.370</td>
<td>3.378</td>
</tr>
</tbody>
</table>

3. test

<table>
<thead>
<tr>
<th>Temperature tested</th>
<th>30 °C</th>
<th>41 °C</th>
<th>41 °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of transfers</td>
<td>11</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>OD&lt;sub&gt;600&lt;/sub&gt;</td>
<td>4.0</td>
<td>2.919</td>
<td>2.928</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>15</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>OD&lt;sub&gt;600&lt;/sub&gt;</td>
<td>3.842</td>
<td>2.912</td>
<td>2.930</td>
</tr>
</tbody>
</table>

5.7 TARGETED GENE DISRUPTION USING A SUICIDE VECTOR

This gene deletion system (Lambert *et al.* 2007) is also based on homologous recombination, but uses a non-replicating vector (Section 2.2.16(b)). With the cre-lox system multiple gene deletions can be carried out using the same mutagenesis vector (pNZ5319), because the antibiotic resistance-cassette can be excised.

The *plnKW30* mutagenesis vector was constructed by cloning a ~1 kbp upstream fragment upstream of the *gccA* (*plnKW30*) gene into the *Pmel* restriction site and a ~1 kbp fragment downstream of the *gccA* (*plnKW30*) gene into the *Ecl*136II restriction site. The *Pmel*
fragment contained the first five codons and the Ecl136II-fragment contained the last five codons of the *plnKW30* gene. Both restriction sites are blunt end, so that the fragment orientation in the resulting plasmids had to be verified by sequencing.

Attempts were made to transform the verified *plnKW30* mutagenesis vector (pNZ5319::plnKO 13) into *L. plantarum* KW30 by electroporation (Kaneko *et al.*, 2000; section 2.2.14(f)). In order to improve transformation efficiency into *L. plantarum* KW30 protoplasts were prepared and transformed (Morelli *et al.*, 1987; section 2.2.14(g) and (h)), but these also did not result in any colonies. Despite many attempts with different concentrations and several plasmid preparations to transform this vector into *L. plantarum* KW30, no colonies grew on plates.

### 5.8 Discussion

Several attempts were made to disrupt genes of the *plnKW30* gene cluster using two different mutagenesis systems. The first system was based on a temperature sensitive vector (Russell & Klaenhammer *et al.*, 2001), which becomes unstable at increased temperatures. However, the increased temperature (43 °C) combined with the *L. plantarum* KW30 cells harbouring two additional vectors (they naturally harbour five plasmids already (Kelly *et al.* 1996)), stopped their growth. The helper plasmid is only unstable and finally lost after at least three transfers at 43 °C, which cannot be achieved when the cells stop growing. At lower temperatures, where *L. plantarum* KW30 was able to grow, even up to 15 transfers did not lead to the loss of the helper plasmid (Table 5.6). This system obviously is not suitable for *L. plantarum* KW30.

The second gene disruption system used a non-replicating vector (Lambert *et al.*, 2007). The plasmid for the *plnKW30* disruption was constructed and verified by sequencing. This mutagenesis system probably did not work in *L. plantarum* KW30 because of the fact that it already harbours five plasmids, which would restrict the introduction of another one. While it was possible to introduce a different vector into *L. plantarum* KW30 cells (Klaenhammer mutagenesis system), the fact that this vector is non-replicating may make a difference. Once transformation has occurred, a double-crossover is necessary to obtain
colonies. If the transformation efficiency is low, and only a fraction of plasmids are actually transformed into the cells, it will be difficult to obtain a double-crossover. The transformation efficiency should have been improved by transforming protoplasts, but also did not result in colonies. The recovery of protoplasts after electroporation is harder and takes longer, than for normal cells. Protoplasts are more susceptible to environmental changes, but great care had been taken to ensure that the recommended conditions were provided. The Lambert et al. gene disruption system would be very useful, if the transformation problems for \textit{L. plantarum} KW30 could be solved, as this system allows multiple gene disruptions with the same antibiotic resistance-cassette.

5.8.1 \textit{CONCLUSIONS}

Despite several attempts to disrupt the \textit{plnKW30} (\textit{gccA}) or putative FTase (now putative response regulator; \textit{gccF}) genes using the two different systems, no gene disruption was observed.
Chapter 5: Appendix

Appendix 8: 16-23S rDNA sequence of L. plantarum KW30

APPENDIX 8: 16-23SRDNA SEQUENCE OF L. PLANTARUM KW30
ATTGGTGCTTGCATCATGATTTAAGCTACTAAGGTGGCGAACTGGTGAGTAACACGTGGGAAACCTCCTTAAAAGCG
GGGGATAACACCTGGAAACAGATGCTAATACCGCATAACAACTTGGACCGCATGGTCCGAGTTTGAAAGATGGCTTC
GGCTATCACTTTTGGATGGTCCCGCGGCGTATTAGCTAGATGGTGGGGTAACGGCTCACCATGGCAATGATACGTAGC
CGACCTGAGAGGGTAATCGGCCACATTGGGACTGAGACACGGCCCAAACTCCTACGGGAGGCAGCAGTAGGGAATCTT
CCACAATGGACGAAAGTCTGATGGAGCAACGCCGCGTGAGTGAAGAAGGGTTTCGGCTCGTAAAACTCTGTTGTTAA
AGAAGAACATATCTGAGAGTAACTGTTCAGGTATTGACGGTATTTAACCAGAAAGCCACGGCTAACTACGTGCCAGC
AGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTTTTAAGTC
TGATGTGAAAGCCTTCGGCTCAACCGAAGAAGTGCATCGGAAACTGGGAAACTTGAGTGCAGAAGAGGACAGTGGAA
CTCCATGTGTAGCGGTGAAATGCGTAGATATATGGAAGAACACCAGTGGCGAAGGCGGCTGTCTGGTCTGTAACTGA
CGCTGAGGCTCGAAAGTATGGGTAGCAAACAGGATTAGATACCCTGGTAGTCCATACCGTAAACGATGAATGCTAAG
TGTTGGAGGGTTTCCGCCCTTCAGTGCTGCAGCTAACGCATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGCTGA
AACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCTACGCGAAGAACCTTAC
CAGGTCTTGACATACTATGCAAATCTAAGAGATTAGACGTTCCCTTCGGGGACATGGATACAGGTGGTGCATGGTTG
TCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTATTATCAGTTGCCAGCATTAAGT
TGGGCACTCTGGTGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGACC
TGGGCTACACACGTGCTACAATGGATGGTACAACGAGTTGCGAACTCGCGAGAGTAAGCTAATCTCTTAAAGCCATTC
TCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAAGTCGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGA
ATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGAGAGTTTGTAACACCCAAAGTCGGTGGGGTAACCTTT
TAGGAACCAGCCGCCTAAGGTGGGACAGATGATTAGGGTGAAGTCGTAACAAGGTAGCCGTAGGAGAACCTGCGGCT
GGATCACCTCCTTTCTAAGGAATATTACGGAAACCTACACACTCGTCGAAACTTTGTTTAGTTTTGAGAGATTTAACT
CTCAAAACTTGTTCTTTGAAAACTAGATAATATCAAATATATTTTTTCATAATGAAACCGAGAACACCGCGTTTTTT
GAGTTTTTTATTGAAGTTTAATTATCGCTAAACTCATTAATCGCATTTACCGTTAGGTAAATGAGGTTAAGTTAACA
AGGGCGCATGGTGAATGCCTTGGCACTAGGAGCCGATGAAGGACGGGACTAACACCGATATGCTTCGGGGAGCTGTA
CGTAAGCTATGATCCGGAGATTTCCGAATGGGGCAACCCAGCAGTTTTAATCAACTGTTACCACTAGATGAATTCATA
GTCTAGTTGGAGGTAAACGCTGTGAACTGAAACATCTCATTAGCAGCAGGAATATAAAGAAATTTCGATTCCCTAAG
TAGCGGCGAGCGAACGGGGAACAGCCCAAACCAAAGTGCTTGCACTTTGGGGTTGTAGGACTGAACATTTGAGTTAC
CAAAGAACTTGATAGTCGAAGGATTTGGGAAAATCCGCCATAGATGGTGATAGCCCAGTAGATTAAATCAAATTCTC
TCAGTTCAGGATCCAGTCTAATTGCCGGAACACGTGAAATTCCGTCGGAATCCGGGAGGACCATCT

216


APPENDIX 9: NUCLEOTIDE AND AMINO ACID SEQUENCES

Nucleotide sequence and conceptual translation of *plnKW30* gene cluster of L. plantarum KW30.

**GTase nucleotide sequence**, 1269 bp (4118 – 5386 bp in appendix 10):

atgaaaatagcaaaatatggacagtattttttgtgcaaatatttattttgtggaatatcagttttttttattttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
TRX 1 nucleotide sequence, 420 bp (7478 – 7897 bp in appendix 10):

tagaaaaaccaattctactattttgctatacttttccatataatagggatatatctttgagcaagcaattcataaataagcgcctatcagcataaatatgaagatgttatgctactattttccgcttagtttatcataataaatattttacacttttcaaaattataaaccagaagtcatagcatttatttttttmcttttttttcgcttttactttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
GTase amino acid sequence: 422aa
MKRNQRNEDSYNLHTPLRVHKSDFGNLTVIDQFRHIHYYSIVICKNSQATIERCVCNSAQMNMEGNDELIVLDTGSTDETVHLLVKNMQPASKISVTNWKNDSEVRNKALASKVDWHEYFDSDEWDVDGQKLLKFKVQAANFKFVINTFSDHSGQIYQTVGRIFPKSSHYPHYAKHEVERKQDKGLGVDVRHFACDDILYLHDGDEKVLRRDKDILRINRLLQEMTC
EEOQNAWMDFLADDGFDLQDKLQVLRKTLDDLVADSGLEQKDYSFKAKKLLRLGRILLRREGKIIQAVLFLDVLQIQTGGE
EDSDAIYYIEFSKINIEIAAEKSIIEVKMLRYNLKHKMIDVNSDIGNNYHIAQVILEDIIISANYHFLPIPEIKNSGDIKS
SVASKVLYESLQGDSKNENN

ABC-transporter amino acid sequence, 706aa:
MTAKMRIIKQIDQNCGPAVATITGMISISKIDWQRQIITTTNQYGTNFIISISLGVQVESVKCVKNDPVFDEIEEFPVLVQMQNGYLHFLVTLKCSGKLYWADPGSKGIESIDKAKFMHWTPLLITPRTSQNFQLTKVESLKVLPLWFSE
RKHACLLLLVYVLMIIIFATTLLMFLATWMSAYFNRIVPNSFIAIPITTLTVQVSAIFLIEVINFIALIAYGTGNISIKELYDI
FKAFPPLKKRLVIVLDFEDEGEIIIFNAYIQAIAGQRLVILGDLPIDIITMGITLFLLARINPFLLILFIPMLLMLYLSSDAIL
KKRSTNLFKEQETFQDIIEQIKNFSTIKTFHAGSINSKTKMSEKYTEANARSGFLYDRLQIARARSGVQFLSYLFGHGV
FLVQIKVGNLTLLSFYATGVLNPVFKITMTQVNLQSQGKVALERYDLIDSPEDNNKSKLTLNSIEIRKVAFSDYG
VPIKIDMGVFSKFPATIIGNSSGKISIAKLMAGFYPAEGLSNGLPPYIAIGAESLGSNITVQEVPSIQSFDTVMNITLG
RKRTNVKSNRIASSGDECLNHLPNGYTQLNGSSYLSGGQKQLNNRSMVIPSRIIFDEITNGLDINTKLKVENYLLS
QNPKLIFITHDLTILARIDIDITVKNGEMHIETKNSTESLRALL

TRX 1 amino acid sequence, 139aa:
MKKTIITIALVSVIIGYVSQRLTLQVPELYSIHKTQLQADIKNKDDVVFYFQGCEGECHESTPILNRFIRNKRTVK
AIDINADPSKDFILGTLGSSSTPILIFIIHGGKIVQRNQVFGYEDELTYKGYTD

TRX 2 amino acid sequence, 177 aa:
MNLLKRRKLVLILLILLILLVIGLFLSVIITKRSNYEHENVRALAEIESKNTMMYKELISITPTFEKEISSKRDVLVYIRG
PTCSCDNCFLDPILVNELKEDMSTNVSFLNVAVERQKWTSVQFQKGYFKQTPAIIHHYHHNGKVLISIQWGNKGIS
KGDHLWLJIKQKEI

LytTR amino acid sequence, 258 aa:
MAYQAGEELVQPRTFRVLVTQKPYIQNLHNHLQMLNLNSNHIFKCHPKLSSSENFQFICSLSILNLDDEEIDLKKIS
IAYKRLPSNYFIYIGSSHRFENLIFHMYLDAIMTKTFQKIDDNLIALADICNLIHJNSIPLNSSNTIQISVRFQKFL
RLDDFLVTIDNQHPKHVFHLCYDEYQRMLTKQHHQCRNLTLACHEGAVLNKNITFNSRNLYFENGTECVR
RYTKEKISDLIPMKTF

PlnKW30 precursor protein, 64 aa:
MSKLVKTLTISEISKQAQNNGGGPAWCWYTAMCGAGYDSGTCDYMISHCFIKHHSSGSSSYHC
APPENDIX 10: COMPLETE SEQUENCE OF PLNKW30 GENE CLUSTER (14149 BP)

1  ATGTGC CGAA  TTCTCGGTGTT  TTCCAGAGCT  CAGTATTATC  GTTATCGATC
51  CCCCAA AAAACCT  TCAAACG GCCG  GGGCCGAAGA  TGCGGACTTG  AAACAACGGA
101  TTCTCGCGGAT  CTTTGCGGAA  TTTAAGCAGC  GATACGGTGT  TATGAAGATC
151  CACCATGAAT  TGAATCTGGA  ACTTCAACCA  CTGCAGCTTC  GGTGCAGTCC
201  ACGACGGATT  TCCCGGCTCA  TGAAGAAACT  GGATATCCAC  TCCGTTACCG
251  TCAATAAGTG  GAAAGC GGCT  TCGGCTTCCA  AAACCAAGGA  TGAAACAGGT
301  CCCAACTTGC  TTAAGCAGGA  TTTCTCGACC  ACTGGTTTAA  ATCAAAAATG
351  GACCGCTGAT  ATGACCTATA  TTCAAACGAA  GCGTAATGCT  GGTGTTACTT
401  ATCAACCATC  ATGGACCTGC  ACTCACGACG  GATTATCGGC  TATTCGTTCT
451  CAAAAAAGAT  GGCTACTGAT  TTAGTCTTAA  AGACCCTTGA  AAGCGGGTT
501  AAAAATTGAA  CCATTACTGG  GGACCTGATT  ATCCATAGCG  ATTTGAGATC
551  ACAGTATACC  AGCGATGATT  ACAATCAACG  TTTAACTGAG  CTACATATCC
601  GCCACTCATA  CAGCTGTAAG  GGTTGTCCGT  ATGATAATGC  GCCAATGGAA
651  TCCTTTCACG  CTTCCCTCAA  AAAGGAATGT  GTTTATCCAG  TGCCGGTCTT
701  TGAAGATTAT  GAAACTGCCG  CTGCCGTCCT  TTTTGAATAT  GTGCATGCTT
751  TCTACA TTAATGAACT  TGCAAAACCT  TTGAATTTA GCTTCTCTTA
801  TTTCCAGGGT  CTTTGCGGAA  TTTAAGCAGC  GATACGGTGT  TATGAAGATC
851  CACCATGAAT  TGAATCTGGA  ACTTCAACCA  CTGCAGCTTC  GGTGCAGTCC
901  TCTACA TTAATGAACT  TGCAAAACCT  TTGAATTTA GCTTCTCTTA
951  TTTGACTTCAA  TCCATATCCT  TTGAATTTA GCTTCTCTTA
1001  TTTTAAACTCT  TTTCATTTTT  AGGTAGTCA  AAACATATT  TATGTTCTAT
1051  CTCAAAACAA CTGATTAACT  TAAATTTGAT  AAAGGAAGAG  CAGAATAAGTG
1101  ACCTGAAAGC  ACTTAGAACT  TAGGAATTAT  GGGCCACAC  CATATTGCTCC
1151  AGATTTGAATC  GTGATATTTG  TAAATTTGAT  GAGGACACTG  ACATTTGCTCC
1201  GGACGGGTGA  TCATTTTCAA  GTCACTTTGA  TGGCAATTCC  AGCTGGTGGT
1251  GGAGATATCG  GAATGGAGAT  TCACCATGGT  AATGACCAGT  TCATTTATTT
1301  GGTAGACGGT  GTTGGTCACG  TTCAAATGGG  CAAAGATAAA  AATAAATGTTGC
1351  CAATGACGTT  CTTTGCGGAA  TTTAAGCAGC  GATACGGTGT  TATGAAGATC
1401  ACTTGCAGTTG  AATTTAATTT  AGTCTTAAGT  GAAAGAAGGT  TGCCAAGTTT
1451  ATGATGCGAT  TCAGCAAGAA  GGCCCACTAG  AGGGCACTGG  TGAATAAAGTG
1501  CATGGAATTT  TTTTTCAGCA  TGCTAATTTA  GCTTCTCTTA  ACCTGATTTA
1551  CTGCAAAACAA CTGATTAACT  TAAATTTGAT  AAAGGAAGAG  CAGAATAAGTG
1601  CTGCAAAACAA CTGATTAACT  TAAATTTGAT  AAAGGAAGAG  CAGAATAAGTG
1651  TGCTAATTTA  AGAAAGAAGGT  TGCCAAGTTT  TTGACAAGCA  TCAACGTGGA
1701  GTGATATTTG  TAAATTTGAT  GAGGACACTG  ACATTTGCTCC
1751  AGATTTGAATC  GTGATATTTG  TAAATTTGAT  GAGGACACTG  ACATTTGCTCC
1801  TTACTAAATG  CCAAAATATT  ACCCATCAAA  CACATGCATC  ACGCGTTTGA
1851  GGTAGACGGT  GTTGGTCACG  TTCAAATGGG  CAAAGATAAA  AATAAATGTTGC
1901  CAATGACGTT  CTTTGCGGAA  TTTAAGCAGC  GATACGGTGT  TATGAAGATC
1951  CTGCAAAACAA CTGATTAACT  TAAATTTGAT  AAAGGAAGAG  CAGAATAAGTG
2001  ATCTCAAAAGT  CGCGCACTGG  TTGTCATAC  GGGCCAGGCC  ATTCAGTTG
2051  GGTAGACGGT  GTTGGTCACG  TTCAAATGGG  CAAAGATAAA  AATAAATGTTGC
2101  CAATGACGTT  CTTTGCGGAA  TTTAAGCAGC  GATACGGTGT  TATGAAGATC
2151  TTACTAAATG  CCAAAATATT  ACCCATCAAA  CACATGCATC  ACGCGTTTGA
2201  ATCTCAAAAGT  CGCGCACTGG  TTGTCATAC  GGGCCAGGCC  ATTCAGTTG
2251  GGTAGACGGT  GTTGGTCACG  TTCAAATGGG  CAAAGATAAA  AATAAATGTTGC
2301  CATGGAATTT  TTTTTCAGCA  TGCTAATTTA  GCTTCTCTTA  ACCTGATTTA
2351  TAAGATCTGGG  AGATGATATGAA  ACCCATCGCC  TTCTTGCTCA  TGGTCTACG
2401  ATCTCAAAAGT  CGCGCACTGG  TTGTCATAC  GGGCCAGGCC  ATTCAGTTG
2451  CAATGACGTT  CTTTGCGGAA  TTTAAGCAGC  GATACGGTGT  TATGAAGATC
2501  TTACTAAATG  CCAAAATATT  ACCCATCAAA  CACATGCATC  ACGCGTTTGA
2551  TGCAATTTTG  TTAAGATTTT  GAAAGAAGGT  TGCCAAGTTT  TTGACAAGCA
2601  ATCTTAATTTG  TTAAGATTTT  GAAAGAAGGT  TGCCAAGTTT  TTGACAAGCA
2651  CAATGACGTT  CTTTGCGGAA  TTTAAGCAGC  GATACGGTGT  TATGAAGATC
2701  ATCTCAAAAGT  CGCGCACTGG  TTGTCATAC  GGGCCAGGCC  ATTCAGTTG
2751  ATGACTTCAACC  GCTAAGCAGAC  AGCTGATATGAA  ACCCGGGAGATGCTG

220
2801 ACCCAGCAAA ACCCAGCACT CAGTACGCTA AGTCAACGGT TACTAAAGTA
2851 CTTAGAGCTG CCCCACCGAG TCAATCAGCA AAATATTCGT TAGTATGAGT
2901 GGCTTAGTGTG AAGCTTATAG TCACGATGCT GATGATGCAA CTTCCGATGG
2951 TTACGATGCT CAACTCTGTT TAGGATGCAA CAGGAGAAA CTTGGCAATT
3001 AAATTTTTTG AAATGAGATG AGTATATGTT AACTAATATG
3051 TGTTTTTTTT GTACACCATC CTTTTTTTTA TTTTTTTTTT
3101 GTGGTTTTTT TTATTTTTTT GTTTTTTTTT ATTTTTTTTT
3151 ATTTTTTTTT GTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
3201 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
3251 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
3301 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
3351 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
3401 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
3451 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
3501 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
3551 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
3601 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
3651 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
3701 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
3751 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
3801 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
3851 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
3901 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
3951 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4001 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4051 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4101 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4151 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4201 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4251 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4301 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4351 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4401 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4451 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4501 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4551 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4601 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4651 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4701 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4751 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4801 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4851 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4901 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
4951 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
5001 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
5051 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
5101 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
5151 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
5201 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
5251 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
5301 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
5351 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
5401 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
5451 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
5501 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
5551 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
5601 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
5651 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
5701 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT

221
5751  CCGCTTACTCT TAAGCATTAC TAGATCACAA AATT TTCAA TA TCTAA AAC
5801  AAAGTTGAA AGCTTAAGAG TCCCTATT GTG GGTTTAGT TTTAGGAAAC
5851  ACGCGACG AGCTTTTAAT TAATTTGAGT TTTATTTGCT
5901  ACACAATCTT TTCTCAAAAG TATGTTAACG TTCTATTTA ACCGACATCT
5951  TCCCAAATAG TGTTATTTAC GATACCTTAC TATAACTTCT GTATAATT
6001  CTGTTCCCTTT CTTGAGAATT TACAATTTAC AATTTTTAAT TTCTAGTAAT
6051  GCTTTATATG GAAAAAGATG GAAGATATGA GATATTTTAA
6101  AACATTAT TTTATTTT GAGTGTG TTTAGTCTT ATGTTACCTT
6151  AAATATTATC TTTGAAAG TGTAAAAATA TTTGTTCAAA CTACCTTTA
6201  GTGTCTTCTAT CTTGACTTAC AATTGCAATT ATTACATTTG GTAATTACCT
6251  TTTTCTTCTTT GCTTGAATT TTTCACTATT AAAAATTTTT CATGGAG
6301  GCTACATCAA TTCCCAAATA CAAAAGAATA TAAAAG GATATACTGCA
6351  5A00 CAGGAGTAGT TGTGCTTATA TATGATGTTA AGATTTTGT GAAGAAGTTG
6401  TTTTGTTTTG TATAGATTTG TTTTGGTATC AGTTTTTTTT AATTGATTTG
6451  AAACGACGCTT TTTTTTCAAA GAAAGAAAGA TGTTAGTAGT AGATTTTATT
6501  GTTTTCTTAA CTTTTGAAAG TTTGAAAGAA TTATTTTATT CTCGGTGGG
6551  GCTTAAATGT GGTTAGTAGT AGATTTTATT CTCGGTGGG
6601  TATATTATAT CTTTTTCTAT TTTTGGATTT AGATTTTATT
6651  GCTACATCAA TTCCCAAATA CAAAAGAATA TAAAAG GATATACTGCA
6701  5A00 CAGGAGTAGT TGTGCTTATA TATGATGTTA AGATTTTGT GAAGAAGTTG
6751  TTTTGTTTTG TATAGATTTG TTTTGGTATC AGTTTTTTTT AATTGATTTG
6801  AAACGACGCTT TTTTTTCAAA GAAAGAAAGA TGTTAGTAGT AGATTTTATT
6851  GTTTTCTTAA CTTTTGAAAG TTTGAAAGAA TTATTTTATT CTCGGTGGG
6901  GCTACATCAA TTCCCAAATA CAAAAGAATA TAAAAG GATATACTGCA
6951  5A00 CAGGAGTAGT TGTGCTTATA TATGATGTTA AGATTTTGT GAAGAAGTTG
7001  TTTTGTTTTG TATAGATTTG TTTTGGTATC AGTTTTTTTT AATTGATTTG
7051  AAACGACGCTT TTTTTTCAAA GAAAGAAAGA TGTTAGTAGT AGATTTTATT
7101  GTTTTCTTAA CTTTTGAAAG TTTGAAAGAA TTATTTTATT CTCGGTGGG
7151  GCTACATCAA TTCCCAAATA CAAAAGAATA TAAAAG GATATACTGCA
7201  5A00 CAGGAGTAGT TGTGCTTATA TATGATGTTA AGATTTTGT GAAGAAGTTG
7251  TTTTGTTTTG TATAGATTTG TTTTGGTATC AGTTTTTTTT AATTGATTTG
7301  AAACGACGCTT TTTTTTCAAA GAAAGAAAGA TGTTAGTAGT AGATTTTATT
7351  GTTTTCTTAA CTTTTGAAAG TTTGAAAGAA TTATTTTATT CTCGGTGGG
7401  GCTACATCAA TTCCCAAATA CAAAAGAATA TAAAAG GATATACTGCA
7451  5A00 CAGGAGTAGT TGTGCTTATA TATGATGTTA AGATTTTGT GAAGAAGTTG
7501  TTTTGTTTTG TATAGATTTG TTTTGGTATC AGTTTTTTTT AATTGATTTG
7551  AAACGACGCTT TTTTTTCAAA GAAAGAAAGA TGTTAGTAGT AGATTTTATT
7601  GCTACATCAA TTCCCAAATA CAAAAGAATA TAAAAG GATATACTGCA
7651  5A00 CAGGAGTAGT TGTGCTTATA TATGATGTTA AGATTTTGT GAAGAAGTTG
7701  TTTTGTTTTG TATAGATTTG TTTTGGTATC AGTTTTTTTT AATTGATTTG
7751  AAACGACGCTT TTTTTTCAAA GAAAGAAAGA TGTTAGTAGT AGATTTTATT
7801  GCTACATCAA TTCCCAAATA CAAAAGAATA TAAAAG GATATACTGCA
7851  5A00 CAGGAGTAGT TGTGCTTATA TATGATGTTA AGATTTTGT GAAGAAGTTG
7901  TTTTGTTTTG TATAGATTTG TTTTGGTATC AGTTTTTTTT AATTGATTTG
7951  AAACGACGCTT TTTTTTCAAA GAAAGAAAGA TGTTAGTAGT AGATTTTATT
8001  GCTACATCAA TTCCCAAATA CAAAAGAATA TAAAAG GATATACTGCA
8051  5A00 CAGGAGTAGT TGTGCTTATA TATGATGTTA AGATTTTGT GAAGAAGTTG
8101  GCTACATCAA TTCCCAAATA CAAAAGAATA TAAAAG GATATACTGCA
8151  5A00 CAGGAGTAGT TGTGCTTATA TATGATGTTA AGATTTTGT GAAGAAGTTG
8201  GCTACATCAA TTCCCAAATA CAAAAGAATA TAAAAG GATATACTGCA
8251  5A00 CAGGAGTAGT TGTGCTTATA TATGATGTTA AGATTTTGT GAAGAAGTTG
8301  CAAAACAGCTT TTTTCTTACA TAATAAATCA TAAATAAATCA TTTCTCTTCT
8351  GCTTTTCAAA CTTTTTCAAA TTTTCTTCTA TTTTCTTCTC TTTTCTTCTC
8401  GCTTTTCAAA CTTTTTCAAA TTTTCTTCTA TTTTCTTCTC TTTTCTTCTC
8451  GCTTTTCAAA CTTTTTCAAA TTTTCTTCTA TTTTCTTCTC TTTTCTTCTC
8501  GCTTTTCAAA CTTTTTCAAA TTTTCTTCTA TTTTCTTCTC TTTTCTTCTC
8551  GCTTTTCAAA CTTTTTCAAA TTTTCTTCTA TTTTCTTCTC TTTTCTTCTC
8601  GCTTTTCAAA CTTTTTCAAA TTTTCTTCTA TTTTCTTCTC TTTTCTTCTC
8651  GCTTTTCAAA CTTTTTCAAA TTTTCTTCTA TTTTCTTCTC TTTTCTTCTC
Chapter 5: Appendix  Appendix 10: Complete sequence of plnKW30 gene cluster

8701 TCACCTAGAC ATGTTTTTGA AAACCTTAAAT TCTCAATCATT ATATGCCTTTA
8751 TGATGCCAATA AGCACCAAAA CTTTTCAAAA AGATGATGGG AAAATATGGA
8801 TCAGCTGATATT TGTTAATATT CAATACTAACT ATCTATCTCT CTTAATAATG
8851 TCACGAAACCA TACAACTTTC TACAAATTTT TCACAGATTT GCATGGATAT
8901 GCAGATATATA TTTATTTGTA TAACTGACAA TCAACCGCAT AATGCTATAT
8951 TTATTTGAGTA TATATTTGAAT TATATGTGCT TTATATGTGCT TAAACAGGTT
9001 CATCTGCAATT CTGTCATATCT TACTGTCATT GATAGAGAG CCTTATTAAG
9051 CCTAAGAATG ATACCATTTA TTTAACAGAG CAAATATTCA AAAAACATTG
9101 AAAAAATGACG AGCATAATCA TATACATGTT GCTTCAAGAA ATTGCACCAT
9151 GATTATATCC CGATGAGAGC ATTTTTGTTTA CCAAGTAAAG TATCTGCTCT
9201 CCAATATAGT TCTTTTCTTT TTTAATTATT TTTAATTATT AAAACAGGTT
9251 AGGATGTTAG ATATGACATT TATAATTCTA ATATGACATT GATTTCTTAA
9301 CTCTAACAAG CCGTTGAAAA CTTGTGATGT TTTGTTTATT TTTGAAGATT
9351 TGGTGGCTGTA GAGCTAGTAG GACATGAAAG TATTTTCTCT TTTTCCTTCT
9401 TTATTTGATAA AAGGAGGATA TATACATGTT GATGTTGCAC AACATCTTAA
9451 TTATTTGATAA AAGGAGGATA TATACATGTT GATGTTGCAC AACATCTTAA
9501 AAAATAGTGA TATATTTGAA TAACTGACAA TCAACCGCAT AATGCTATAT
9551 GGATGATATA TTATTTGTAA TAACTGACAA TCAACCGCAT AATGCTATAT
9601 TTCATTTGAA ATATGATGAA TATCAAGTCC GTATGACTTT AAAACAAGTT
9651 CATCATCAAT GTCGTAATCT TACTGCTTGC CATGAAGGAG CCTTAGTTAA
9701 CCTAAAAAAC ATCACTAAAT TTAACAAGAG CAATAGGATC CTATATTTTG
9751 AAAATGGTGA TACTTGCGAG GTTTCAAGAA GATATACAAA AAAGATTAGT
9801 GATTTAATCC CGATGAGAGC ATTTTTGTTTA CCAAGTAAAG TATCTGCTCT
9851 CCAATATAGT TCTTTTCTTT TTTAATTATT TTTAATTATT AAAACAGGTT
9901 AATTATGAGT AAGGAGGATA TATACATGTT GATGTTGCAC AACATCTTAA
9951 TTATTTGATAA AAGGAGGATA TATACATGTT GATGTTGCAC AACATCTTAA
10001 TCTTATTCTTA ATGATATCCT AATTTTTGAA TAACTGACAA TCAACCGCAT
10051 AGCATCAGTT GTGTTTATTTA GCAAGGAAAA GACACATGAA TTATTTTCAA
10101 CAGATTGAGT ATATATCTGA AATCAGGAAA AACACACTATT TTTAATTATT
10151 GCATATTAGT TCTTTTCTTT TTTAATTATT TTTAATTATT AAAACAGGTT
10201 TGGATTAGAT AGGCCTGAGT ATACAAAAA ATAGAGGTTA TGATGTTGATG
10251 ATACGTTGTT GAGCTAGTAG GACATGAAAG TATTTTCTCT TTTTCCTTCT
10301 CCCAGGAAG AATAAGGCAT CCAACAAATT TTATCTTTAT ATGACACATC
10351 ACTAAATACT TGGAGACCTT CAAATACCTAA AAAATATGAA AGCAGTATTTA
10401 ATTTTTATTTA ATGATATCCT AATTTTTGAA TAACTGACAA TCAACCGCAT
10451 GATAATGATGG TAAATAATAT AAAGCAGTCC GAAACACAGCA GTGCTATTCA
10501 GCTACCTGTT TCTTTTCTTT TTTAATTATT TTTAATTATT AAAACAGGTT
10551 ATTTAGTTGT AGGCAAACAA TTCATTTCTG TCTTACTATT TTTATATTTT
10601 CAGTTGATGT AGGCAAACAA TTCATTTCTG TCTTACTATT TTTATATTTT
10651 GCATATTAGT TCTTTTCTTT TTTAATTATT TTTAATTATT AAAACAGGTT
10701 ACCCCATAAA AATCAATTC TTTAAGCGCA CAGGGAAAA AAGGATAGTT
10751 GATTATTTCTA ATGATACCTT GAGTTCTGAA CTCAAGAGCT GAAGTTTTCA
10801 TGGTGCAGGT AATTCACATC TACTCTAAA GAAATTTGTT GGATGTTCCA
10851 CTTGAAATAA ATGTTTGTAA TTATTTGAAA TAACTGACAA TCAACCGCAT
10901 CTTTTTTCTT TTTTGTATTT GTATATTCTA ATGATACCTT GAGTTCTGAA
10951 ATTTAGTTGT AGGCAAACAA TTCATTTCTG TCTTACTATT TTTATATTTT
11001 CATCTGATTT ACAAATTAAA GCACATGAAC TAGAAACATA TCTTTTTAAA
11051 GAAGATTTTGA AAACCTTAAAT TCTCAATCATT ATATGCCTTTA
11101 AACAAATTCAA AAACCTTAAAT TCTCAATCATT ATATGCCTTTA
11151 ACCCATAAA AATCAATTC TTTAAGCGCA CAGGGAAAA AAGGATAGTT
11201 ATGATTCTCA ATGATGCTCT TCTAACGACA GAAAACACGT GTAAGCTTTA
11251 GATTTTTTTT TTTTATTTTA AATTCACATC TACTCTAAA GAAATTTGTT GGATGTTCCA
11301 AATCTCCGCTT AAGCAGCTTT CAGAGATAC TAACTGACAA TCAACCGCAT
11351 GATTTTTTTT TTTTATTTTA AATTCACATC TACTCTAAA GAAATTTGTT GGATGTTCCA
11401 CAGATTGAGT ATATATGTTT CATATATTTC TTTAATTATT TTTAATTATT
11451 CTTGTCAGAG ATGTTTGTAA TTATTTGAAA TAACTGACAA TCAACCGCAT
11501 ATTTAGTTGT AGGCAAACAA TTCATTTCTG TCTTACTATT TTTATATTTT
11551 CTTTATGAGA TGTATTTGACT GCTATGATTG TTAATGTTTT AATATCTATT
11601 CTTTATGAGA TGTATTTGACT GCTATGATTG TTAATGTTTT AATATCTATT

223
11651  CCGCTCAATGA TTTTGGATTGA AAAAAACACC TTAGCCAAAAA AAGAAGACGC
11701  TACGGCATCC CTAAATGATT TAAAAATCAAG GAATTAAAAG GACAGCAAGC
11751  AATTAAGCCAAG TAGTATAGGA TCAATTAAAT GATATTCTCTGA
11801  TTTTCAACCCG AGAAGATTTTT CAGGGAAGCA ATACAAGATCG TTTTGAGGCT
11851  TTATAGGCCTG GATTTTTCAGT AAAAAAAGCTG CAAACTGACAT AATAGGAGTG
11901  GTTGCAGCAG CAACTTGAAG CTTTCTTCAC CCAATTTGCA AGTATTAGAA
11951  GATTTAAATCC AGATTTACAC TTTCAAGATT AGCAAGGAGG ATATTTCAAG
12001  AATACTTTGTGA TTTTCTGAGA AAAGTTTTAAT CTTTAAATCT
12051  TATTTTATGTT TTTTAGGGGT CAGGTTTATG TTTTTTAGCA AGATTTGCTC
12101  TCAATTGGTTA TTTTTTTTAT CGTTTTGAAA AAAAAATTTAAA CTTTAAACAT
12151  CTTTGCACTG CAAATCTATGG CTTTCTGATA CCAATTCACC AGTTTCAAGAT
12201  GAGCAGCTCT CAGCAGTTCA AAAGTTTTAT TTTTTTATGT TTTTTCTGAT
12251  ATTTTTTATT GAGGTTTATA TCCAAACTTA AGATTTTTTT TTTTTTATGA
12301  GTTTTCGCTTG CCGCTCCTAC TAAGTAGGTA TGACTGAGGT GATTTTCTCT
12351  GATTTTTTAT CCAATTGTTA TTTTTTTTTT CTGACCCTGG CTTATTTTCT
12401  GATTACCCAA ATTAATTTTTT GGAGTACAGA GTTACGCAAA AATGTTTATT
12451  TTACACATTGA CGTATACGAT CAGCTGAGGT GATTTTATCT TTTTTTATCA
12501  TTATTTTTTT TTGGGATTCA GGACTTTTGT TTTTTTATCA ATTATTTTGA
12551  GATTTTTTAT CCAATTGTTA TTTTTTTTAT CGTTTTGAAA AAAAAATTTAAA
12601  CTTTGCACTG CAAATCTATGG CTTTCTGATA CCAATTCACC AGTTTCAAGAT
12651  GAGCAGCTCT CAGCAGTTCA AAAGTTTTAT TTTTTTATGT TTTTTCTGAT
12701  ATTTTTTATT GAGGTTTATA TCCAAACTTA AGATTTTTTT TTTTTTATGA
12751  GTTTTCGCTTG CCGCTCCTAC TAAGTAGGTA TGACTGAGGT GATTTTCTCT
12801  GATTTTTTAT CCAATTGTTA TTTTTTTTTT CTGACCCTGG CTTATTTTCT
12851  GATTACCCAA ATTAATTTTTT GGAGTACAGA GTTACGCAAA AATGTTTATT
12901  TTACACATTGA CGTATACGAT CAGCTGAGGT GATTTTATCT TTTTTTATCA
12951  GATTTTTTAT CCAATTGTTA TTTTTTTTAT CGTTTTGAAA AAAAAATTTAAA
13001  CTTTGCACTG CAAATCTATGG CTTTCTGATA CCAATTCACC AGTTTCAAGAT
13051  GAGCAGCTCT CAGCAGTTCA AAAGTTTTAT TTTTTTATGT TTTTTCTGAT
13101  ATTTTTTATT GAGGTTTATA TCCAAACTTA AGATTTTTTT TTTTTTATGA
13151  CTTTGCACTG CAAATCTATGG CTTTCTGATA CCAATTCACC AGTTTCAAGAT
13201  GATTTTTTAT CCAATTGTTA TTTTTTTTAT CGTTTTGAAA AAAAAATTTAAA
13251  CTTTGCACTG CAAATCTATGG CTTTCTGATA CCAATTCACC AGTTTCAAGAT
13301  GAGCAGCTCT CAGCAGTTCA AAAGTTTTAT TTTTTTATGT TTTTTCTGAT
13351  ATTTTTTATT GAGGTTTATA TCCAAACTTA AGATTTTTTT TTTTTTATGA
13401  CTTTGCACTG CAAATCTATGG CTTTCTGATA CCAATTCACC AGTTTCAAGAT
13451  GAGCAGCTCT CAGCAGTTCA AAAGTTTTAT TTTTTTATGT TTTTTCTGAT
13501  ATTTTTTATT GAGGTTTATA TCCAAACTTA AGATTTTTTT TTTTTTATGA
13551  CTTTGCACTG CAAATCTATGG CTTTCTGATA CCAATTCACC AGTTTCAAGAT
13601  GATTTTTTAT CCAATTGTTA TTTTTTTTAT CGTTTTGAAA AAAAAATTTAAA
13651  CTTTGCACTG CAAATCTATGG CTTTCTGATA CCAATTCACC AGTTTCAAGAT
13701  GAGCAGCTCT CAGCAGTTCA AAAGTTTTAT TTTTTTATGT TTTTTCTGAT
13751  ATTTTTTATT GAGGTTTATA TCCAAACTTA AGATTTTTTT TTTTTTATGA
13801  CTTTGCACTG CAAATCTATGG CTTTCTGATA CCAATTCACC AGTTTCAAGAT
13851  GAGCAGCTCT CAGCAGTTCA AAAGTTTTAT TTTTTTATGT TTTTTCTGAT
13901  CTTTGCACTG CAAATCTATGG CTTTCTGATA CCAATTCACC AGTTTCAAGAT
13951  GAGCAGCTCT CAGCAGTTCA AAAGTTTTAT TTTTTTATGT TTTTTCTGAT
14001  CTTTGCACTG CAAATCTATGG CTTTCTGATA CCAATTCACC AGTTTCAAGAT
14051  GAGCAGCTCT CAGCAGTTCA AAAGTTTTAT TTTTTTATGT TTTTTCTGAT
14101  GAGCAGCTCT CAGCAGTTCA AAAGTTTTAT TTTTTTATGT TTTTTCTGAT

Chapter 5: Appendix
Appendix 10: Complete sequence of plnKW30 gene cluster
### APPENDIX 11: SEQUENCE ALIGNMENTS

| 3bcv (3)  | lip------------------------kVSIVpIynve |
| GT        | MKNRQNEIDSYLNLRLFVHKSFDFGKLNIDQFRHIIYYSIVICKNSQ |
|           | bbbbbbbb |

| 3bcv (18) | kyLdqCVqalAQtLsdIEIIILIDDeSpdnCpkiCddyaqynIkVihk |
| GT        | ATIERCVNSIAQNMENDELIVLDGTGDETHVKKKMFQ-AKISVNTW |
|           | aaaaaaaaa | bbbbbbbb | aaaaaaaaaaa | bbbbb |

| 3bcv (68) | knaglGACnsGLdvAtGeYVAfCSDdyvdsY----tynvAqkytCD |
| GT        | KNDFSEVRNKALKLASKDWVYSDWEVLDDGQLKIKFKVQAKNFK |
|           | aaaaaaaa | bbbbb | aaaaaaaa | bbb |

| 3bcv (116) | AVFtfkly----------------knknkneIhtlLkdIlAsdpYaReerAi |
| GT        | FVINPTFSHDGQIYQTVGRIPPKSSFYKYAHIEVRQKEDQKLYGVR |
|           | bbb | bbb | aaaaaaaa | bbb |

| 3bcv (170) | qVsakvLYrrnLIekkhLrFvseriLpSedLIFNVDLanSniVCvlp |
| GT        | HFACDDIILYHDGYDKEVLR----------DKDKKRNIRLLQEMT--CEEPQ |
|           | aa | bbb | aaaaaaaa | bbb |

| 3bcv (4)  | ipkVSVIVpIynveklDqCVqalAQtLsdIEIIILIDDeSpdnCpkiCd |
| GT        | NARWPFLARGDGPFDLPQDKLQKVRLNT-----LVA---SDSIQEYK |
|           | bbbbb | bbbbb | aaaaaaaa | bbbbb |

| 3bcv (54) | dyaaqynIkVih---kknaGlGACnsGLdvAtGeYVAfCSDydvdsY |
| GT        | PFAKLLGRII'REEGKTTQAVLSFKDVLQITGGE-----DSAIYIESF |
|           | aaaa | bbb | aaaaaaa | bbb |

| 3bcv (104) | tynvAqkytCDAVFtfklyknkneIhtlLkdIlAsdpYaReerAiqVsak |
| GT        | KINEIIEAASKIYMLRNYLNKGMISGNYHHIQAVIDCIDI |
|           | aaaa | bbb | aaaaaaaa | bbb |

| 3bcv (175) | vVLyr---rnllkekklrFvseriLpSedLIFNVDLanSniVCvlp |
| GT        | SANYSHFPFILSEIKPNFSDIKSSVAKLYKQGDSKKENN- |
|           | bbb | aaaa | aaaaaaaa | bbb |

---

Figure 5.8: Amino acid alignment of the PlnKW30 GTase and a putative GTase from *Bacteroides fragilis* (pdb 3bcv) using FUGUE (Shi *et al*. 2001) (*a*: alpha helix; *b*: beta strand; *3*: 3_10 helix; lower case: solvent accessible; UPPER CASE: solvent inaccessible).
Figure 5.9: Amino acid alignment of the plnKW30 GTase and SpsA from *Bacillus subtilis* (pdb 1qg8) using FUGUE (Shi et al. 2001) (a: alpha helix; b: beta strand; 3:310 helix; lower case: solvent accessible; UPPER CASE: solvent inaccessible).
## APPENDIX 12: MS/MS FRAGMENTATION OF PlnKW30

### Table 5.7: MS/MS fragmentation of PlnKW30

**PlnKW30:** KPAWCWYTLAMCGAGYDSGTCDYMYSHCFGIKHHSSGSSSYHC

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th># Res:</th>
<th># x</th>
<th>y&quot;</th>
<th>z&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>101.108</td>
<td>129.103</td>
<td>146.129</td>
<td>1</td>
<td>Lys</td>
<td>43</td>
<td>-</td>
</tr>
<tr>
<td>198.161</td>
<td>226.156</td>
<td>243.182</td>
<td>2</td>
<td>Pro</td>
<td>42</td>
<td>5099.955 5075.991 5058.965</td>
</tr>
<tr>
<td>269.198</td>
<td>297.193</td>
<td>314.219</td>
<td>3</td>
<td>Ala</td>
<td>41</td>
<td>5002.902 4978.939 4961.912</td>
</tr>
<tr>
<td>455.277</td>
<td>483.272</td>
<td>500.299</td>
<td>4</td>
<td>Trp</td>
<td>40</td>
<td>4931.865 4907.901 4890.875</td>
</tr>
<tr>
<td>558.286</td>
<td>586.281</td>
<td>603.308</td>
<td>5</td>
<td>Cys</td>
<td>39</td>
<td>4745.786 4721.822 4704.796</td>
</tr>
<tr>
<td>744.366</td>
<td>772.36</td>
<td>789.387</td>
<td>6</td>
<td>Trp</td>
<td>38</td>
<td>4642.776 4618.813 4601.786</td>
</tr>
<tr>
<td>907.429</td>
<td>935.424</td>
<td>952.45</td>
<td>7</td>
<td>Tyr</td>
<td>37</td>
<td>4456.697 4432.734 4415.707</td>
</tr>
<tr>
<td>1008.477</td>
<td>1036.472</td>
<td>1053.498</td>
<td>8</td>
<td>Thr</td>
<td>36</td>
<td>4293.634 4269.667 4252.644</td>
</tr>
<tr>
<td>1121.561</td>
<td>1149.556</td>
<td>1166.582</td>
<td>9</td>
<td>Leu</td>
<td>35</td>
<td>4192.586 4168.623 4151.596</td>
</tr>
<tr>
<td>1192.598</td>
<td>1220.593</td>
<td>1237.619</td>
<td>10</td>
<td>Ala</td>
<td>34</td>
<td>4079.502 4055.539 4038.512</td>
</tr>
<tr>
<td>1323.638</td>
<td>1351.633</td>
<td>1368.66</td>
<td>11</td>
<td>Met</td>
<td>33</td>
<td>3984.501 3967.475 3950.45</td>
</tr>
<tr>
<td>1426.647</td>
<td>1454.642</td>
<td>1471.669</td>
<td>12</td>
<td>Cys</td>
<td>32</td>
<td>3774.706 3750.452 3733.425</td>
</tr>
<tr>
<td>1483.669</td>
<td>1511.664</td>
<td>1528.69</td>
<td>13</td>
<td>Gly</td>
<td>31</td>
<td>3556.034 3530.781 3514.538</td>
</tr>
<tr>
<td>1554.706</td>
<td>1582.701</td>
<td>1599.727</td>
<td>14</td>
<td>Ala</td>
<td>30</td>
<td>3340.302 3314.056 3297.812</td>
</tr>
<tr>
<td>1611.727</td>
<td>1639.722</td>
<td>1656.749</td>
<td>15</td>
<td>Gly</td>
<td>29</td>
<td>3129.357 3093.113 3066.969</td>
</tr>
<tr>
<td>1774.791</td>
<td>1802.786</td>
<td>1819.812</td>
<td>16</td>
<td>Tyr</td>
<td>28</td>
<td>2989.357 2963.113 2936.969</td>
</tr>
<tr>
<td>1899.818</td>
<td>1927.813</td>
<td>1944.839</td>
<td>17</td>
<td>Asp</td>
<td>27</td>
<td>2774.357 2748.113 2721.969</td>
</tr>
<tr>
<td>2179.929</td>
<td>2207.924</td>
<td>2224.951</td>
<td>18</td>
<td>Ser</td>
<td>26</td>
<td>2556.357 2529.113 2502.969</td>
</tr>
<tr>
<td>2236.951</td>
<td>2264.946</td>
<td>2281.972</td>
<td>19</td>
<td>Gly</td>
<td>25</td>
<td>2337.357 2311.113 2284.969</td>
</tr>
<tr>
<td>2337.998</td>
<td>2365.993</td>
<td>2383.02</td>
<td>20</td>
<td>Thr</td>
<td>24</td>
<td>2115.357 2089.113 2062.969</td>
</tr>
<tr>
<td>2441.007</td>
<td>2469.002</td>
<td>2486.029</td>
<td>21</td>
<td>Cys</td>
<td>23</td>
<td>1995.357 1969.113 1942.969</td>
</tr>
<tr>
<td>2556.034</td>
<td>2584.029</td>
<td>2601.056</td>
<td>22</td>
<td>Asp</td>
<td>22</td>
<td>1875.357 1849.113 1822.969</td>
</tr>
<tr>
<td>2719.098</td>
<td>2747.093</td>
<td>2764.119</td>
<td>23</td>
<td>Tyr</td>
<td>21</td>
<td>1655.357 1629.113 1602.969</td>
</tr>
<tr>
<td>2890.138</td>
<td>2918.133</td>
<td>2935.156</td>
<td>24</td>
<td>Met</td>
<td>20</td>
<td>1435.357 1409.113 1382.969</td>
</tr>
<tr>
<td>3013.202</td>
<td>3041.196</td>
<td>3058.223</td>
<td>25</td>
<td>Tyr</td>
<td>19</td>
<td>1215.357 1189.113 1162.969</td>
</tr>
<tr>
<td>3100.234</td>
<td>3128.228</td>
<td>3145.255</td>
<td>26</td>
<td>Ser</td>
<td>18</td>
<td>1005.357 979.113 952.969</td>
</tr>
<tr>
<td>3237.292</td>
<td>3265.287</td>
<td>3282.314</td>
<td>27</td>
<td>His</td>
<td>17</td>
<td>800.357 774.113 747.969</td>
</tr>
<tr>
<td>3340.302</td>
<td>3368.297</td>
<td>3385.323</td>
<td>28</td>
<td>Cys</td>
<td>16</td>
<td>600.357 574.113 547.969</td>
</tr>
<tr>
<td>3487.37</td>
<td>3515.365</td>
<td>3532.392</td>
<td>29</td>
<td>Phe</td>
<td>15</td>
<td>400.357 374.113 347.969</td>
</tr>
<tr>
<td>3543.392</td>
<td>3572.386</td>
<td>3589.413</td>
<td>30</td>
<td>Gly</td>
<td>14</td>
<td>200.357 174.113 147.969</td>
</tr>
<tr>
<td>3657.476</td>
<td>3685.471</td>
<td>3702.497</td>
<td>31</td>
<td>Ile</td>
<td>13</td>
<td>100.357 734.113 607.969</td>
</tr>
<tr>
<td>3795.571</td>
<td>3813.565</td>
<td>3830.592</td>
<td>32</td>
<td>Lys</td>
<td>12</td>
<td>90.357 634.113 497.969</td>
</tr>
<tr>
<td>3922.629</td>
<td>3950.624</td>
<td>3967.651</td>
<td>33</td>
<td>His</td>
<td>11</td>
<td>80.357 534.113 397.969</td>
</tr>
<tr>
<td>4059.688</td>
<td>4087.683</td>
<td>4104.71</td>
<td>34</td>
<td>His</td>
<td>10</td>
<td>70.357 434.113 297.969</td>
</tr>
<tr>
<td>4146.72</td>
<td>4174.715</td>
<td>4191.742</td>
<td>35</td>
<td>Ser</td>
<td>9</td>
<td>60.357 304.113 167.969</td>
</tr>
<tr>
<td>4233.752</td>
<td>4261.747</td>
<td>4278.774</td>
<td>36</td>
<td>Ser</td>
<td>8</td>
<td>50.357 234.113 134.969</td>
</tr>
<tr>
<td>4290.774</td>
<td>4318.769</td>
<td>4335.795</td>
<td>37</td>
<td>Gly</td>
<td>7</td>
<td>40.357 164.113 94.969</td>
</tr>
<tr>
<td>4377.806</td>
<td>4405.801</td>
<td>4422.827</td>
<td>38</td>
<td>Ser</td>
<td>6</td>
<td>30.357 84.113 24.969</td>
</tr>
<tr>
<td>4464.838</td>
<td>4492.833</td>
<td>4509.859</td>
<td>39</td>
<td>Ser</td>
<td>5</td>
<td>20.357 74.113 14.969</td>
</tr>
<tr>
<td>4551.87</td>
<td>4579.865</td>
<td>4596.891</td>
<td>40</td>
<td>Ser</td>
<td>4</td>
<td>10.357 64.113 3.969</td>
</tr>
<tr>
<td>4719.933</td>
<td>4747.928</td>
<td>4759.955</td>
<td>41</td>
<td>Tyr</td>
<td>3</td>
<td>10.357 44.113 2.969</td>
</tr>
<tr>
<td>4851.992</td>
<td>4879.987</td>
<td>4907.014</td>
<td>42</td>
<td>His</td>
<td>2</td>
<td>10.357 24.113 1.969</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>43</td>
<td>Cys</td>
<td>1</td>
<td>10.357 8.113 0.969</td>
</tr>
</tbody>
</table>
# APPENDIX 13: LISTS OF ENZYMES AND OTHER MATERIALS

## Table 5.8: List of polymerases and restriction enzymes

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>KOD Hot Start DNA Polymerase</td>
<td>Novagen, Merck KGaA, Darmstadt, GER</td>
</tr>
<tr>
<td>Platinum <em>Taq</em> DNA Polymerase</td>
<td>Invitrogen, CA, USA</td>
</tr>
<tr>
<td><em>Taq</em> DNA Polymerase</td>
<td>Roche, Mannheim, GER</td>
</tr>
<tr>
<td>T4 DNA Ligase</td>
<td>Invitrogen, CA, USA</td>
</tr>
<tr>
<td>Tth DNA Polymerase</td>
<td>Roche, Mannheim, GER</td>
</tr>
<tr>
<td><em>Bam</em>HI</td>
<td>Invitrogen, CA, USA</td>
</tr>
<tr>
<td><em>Bgl</em>II</td>
<td>Roche, Mannheim, GER</td>
</tr>
<tr>
<td><em>Ecl</em>136II</td>
<td>Fermentas</td>
</tr>
<tr>
<td><em>Eco</em>RI</td>
<td>Roche, Mannheim, GER</td>
</tr>
<tr>
<td><em>Eco</em>RV</td>
<td>Invitrogen, CA, USA</td>
</tr>
<tr>
<td><em>Hind</em>II</td>
<td>Roche, Mannheim, GER</td>
</tr>
<tr>
<td><em>Kpn</em>I</td>
<td>Roche, Mannheim, GER</td>
</tr>
<tr>
<td><em>Msc</em>I</td>
<td>New England Biolabs, MA, USA</td>
</tr>
<tr>
<td><em>Nae</em>I</td>
<td>Roche, Mannheim, GER</td>
</tr>
<tr>
<td><em>Nco</em>I</td>
<td>Invitrogen, CA, USA</td>
</tr>
<tr>
<td><em>Nde</em>I</td>
<td>Roche, Mannheim, GER</td>
</tr>
<tr>
<td><em>Pme</em>I</td>
<td>New England Biolabs, MA, USA</td>
</tr>
<tr>
<td><em>Sacl</em></td>
<td>Roche, Mannheim, GER</td>
</tr>
<tr>
<td><em>Sal</em>I</td>
<td>Invitrogen, CA, USA</td>
</tr>
<tr>
<td><em>Sma</em>I</td>
<td>Invitrogen, CA, USA</td>
</tr>
<tr>
<td><em>Xba</em>I</td>
<td>Roche, Mannheim, GER</td>
</tr>
<tr>
<td><em>Xho</em>I</td>
<td>Roche, Mannheim, GER</td>
</tr>
</tbody>
</table>
### Table 5.9: List of commercial kits

<table>
<thead>
<tr>
<th>Kit</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM Chemiluminescence Blotting Substrate (POD)</td>
<td>Roche, Mannheim, GER</td>
</tr>
<tr>
<td>RTS 100 <em>E. coli</em> HY cell free expression kit</td>
<td>Roche, Mannheim, GER</td>
</tr>
<tr>
<td>Complete, Mini, EDTA-free</td>
<td>Roche, Mannheim, GER</td>
</tr>
<tr>
<td>High pure PCR product purification kit</td>
<td>Roche, Mannheim, GER</td>
</tr>
<tr>
<td>High pure plasmid isolation kit</td>
<td>Roche, Mannheim, GER</td>
</tr>
<tr>
<td>Illustra RNAspin mini RNA isolation kit</td>
<td>GE Healthcare, UK</td>
</tr>
<tr>
<td>LIVE/DEAD BacLight Bacterial Viability Kit</td>
<td>Invitrogen, CA, USA</td>
</tr>
<tr>
<td>Perfectprep gel cleanup</td>
<td>Eppendorf, Hamburg, GER</td>
</tr>
<tr>
<td>ProbeQuant™ G-50 micro columns</td>
<td>Pharmacia Biotech, Uppsala, Sweden</td>
</tr>
<tr>
<td>Random primers DNA labelling system</td>
<td>Invitrogen, CA, USA</td>
</tr>
<tr>
<td>TRIzol® Reagent</td>
<td>Invitrogen, CA, USA</td>
</tr>
<tr>
<td>TURBO DNA-free™ Kit</td>
<td>Ambion, Applied Biosystems, CA, USA</td>
</tr>
<tr>
<td>Wizard Genomic DNA purification kit</td>
<td>Promega, Wisconsin, USA</td>
</tr>
</tbody>
</table>
### Table 5.10: List of consumables

<table>
<thead>
<tr>
<th>Consumable</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filter membranes</td>
<td>Millipore, MA, USA</td>
</tr>
<tr>
<td>Chelating Sepharose Fast Flow</td>
<td>GE Healthcare, SE</td>
</tr>
<tr>
<td>Gloves, latex</td>
<td>Ansell, Muenchen, GER</td>
</tr>
<tr>
<td>Gloves, nitrile</td>
<td>LabServ, Biolab Ltd., Auckland, NZ</td>
</tr>
<tr>
<td>Immobiline Dry Strips</td>
<td>GE Healthcare, UK</td>
</tr>
<tr>
<td>Immobiline Dry Strip Cover Fluid</td>
<td>GE Healthcare, UK</td>
</tr>
<tr>
<td>Immobilon-P PVDF membrane, 0.45 µm</td>
<td>Millipore, MA, USA</td>
</tr>
<tr>
<td>Maximum Recovery tubes</td>
<td>Axygen, CA, USA</td>
</tr>
<tr>
<td>Nitrocellulose membrane</td>
<td>Sigma-Aldrich, MO, USA</td>
</tr>
<tr>
<td>Parafilm</td>
<td>Pechiney Plastic Packaging, USA</td>
</tr>
<tr>
<td>Petri dishes</td>
<td>LabServ, Biolab Ltd., Auckland, NZ</td>
</tr>
<tr>
<td>PP-Test tubes, 15 mL</td>
<td>Greiner bio-one, Frickenhausen, GER</td>
</tr>
<tr>
<td>PP-Test tubes, 50 mL</td>
<td>Greiner bio-one, Frickenhausen, GER</td>
</tr>
<tr>
<td>Quartz SUPRASIL® precision cell, 0.1 mm</td>
<td>Hellma GmbH &amp; Co. KG, Müllheim, GER</td>
</tr>
<tr>
<td>path length</td>
<td></td>
</tr>
<tr>
<td>Quartz SUPRASIL® precision cell holder</td>
<td>Hellma GmbH &amp; Co. KG, Müllheim, GER</td>
</tr>
<tr>
<td>Screw cap tubes, sterile</td>
<td>NUNC™ CryoTubes</td>
</tr>
<tr>
<td>Syringes, sterile</td>
<td>Terumo Corporation, Tokyo, Japan</td>
</tr>
<tr>
<td>Syringe filters</td>
<td>Millipore, MA, USA</td>
</tr>
<tr>
<td>Tips, 5 mL</td>
<td>Eppendorf, Hamburg, GER</td>
</tr>
<tr>
<td>Tips, 1 mL</td>
<td>Axygen, CA, USA</td>
</tr>
<tr>
<td>Tips, 200 µL</td>
<td>Axygen, CA, USA</td>
</tr>
<tr>
<td>Tips, 10 µL</td>
<td>Axygen, CA, USA</td>
</tr>
<tr>
<td>Tubes, 1.5 mL</td>
<td>Axygen, CA, USA</td>
</tr>
<tr>
<td>Tubes, 0.7 mL</td>
<td>Axygen, CA, USA</td>
</tr>
<tr>
<td>Whatman 3 mm paper</td>
<td>Whatman International Ltd., Kent, UK</td>
</tr>
<tr>
<td>ZipTip&lt;sub&gt;C18&lt;/sub&gt; pipette tips</td>
<td>Millipore, MA, USA</td>
</tr>
</tbody>
</table>
### Table 5.11: List of equipment used

<table>
<thead>
<tr>
<th>Equipment</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABI Prism 377-64 sequencer</td>
<td>Perkin Elmer, MA, USA</td>
</tr>
<tr>
<td>ABI Prism 3730 capillary sequencer</td>
<td>Perkin Elmer, MA, USA</td>
</tr>
<tr>
<td>Bead mill</td>
<td>Fast prep cell disruptor, Thermo Savant, Qbiogene, Carlsbad, USA</td>
</tr>
<tr>
<td>Biofuge fresco centrifuge</td>
<td>Heraeus, Thermo Scientific, MA, USA</td>
</tr>
<tr>
<td>Chirascan Circular Dichroism spectrometer</td>
<td>Applied Photophysics, UK</td>
</tr>
<tr>
<td>CoverWell, 20-mm diameter, 0.5 mm deep</td>
<td>Molecular Probes, Invitrogen, CA, USA</td>
</tr>
<tr>
<td>ESI-MS, micromass ZMD</td>
<td>Micromass, Manchester, UK</td>
</tr>
<tr>
<td>Etan IPGphor II Isoelectric Focusing Unit</td>
<td>GE Healthcare, UK</td>
</tr>
<tr>
<td>GASTIGHT syringe</td>
<td>Hamilton CO., NV, USA</td>
</tr>
<tr>
<td>Hamamatsu ORCA-ER C4742-80 digital charge-coupled device camera</td>
<td>Hamamatsu Corporation, Japan</td>
</tr>
<tr>
<td>Intelligent Dark Box II, LAS-1000</td>
<td>Fujifilm Holdings Corporation, Tokyo, Japan</td>
</tr>
<tr>
<td>Jupiter 5u C4 300A columns, 250 x 10 mm</td>
<td>Phenomenex, CA, USA</td>
</tr>
<tr>
<td>Jupiter 5u C4 300A columns, 250 x 4.6 mm</td>
<td>Phenomenex, CA, USA</td>
</tr>
<tr>
<td>Liquid scintillation &amp; luminescence counter</td>
<td>Perkin Elmer, MA, USA</td>
</tr>
<tr>
<td>Mini-Protean II system</td>
<td>BioRad Laboratories, Milan, I</td>
</tr>
<tr>
<td>MiniSpin plus centrifuge</td>
<td>Eppendorf, Hamburg, GER</td>
</tr>
<tr>
<td>Mini Trans-blot electrophoresis transfer cell</td>
<td>BioRad Laboratories, Milan, I</td>
</tr>
<tr>
<td>Nano drop ND-1000 Spectrophotometer</td>
<td>Thermo Scientific, MA, USA</td>
</tr>
<tr>
<td>NANOpure II filtration system</td>
<td>Sybron/Barnstead, Maryland, USA</td>
</tr>
<tr>
<td>Olympus IX 71 microscope</td>
<td>Olympus Europa Holding GmbH, Hamburg, GER</td>
</tr>
<tr>
<td>PCR cycler T gradient</td>
<td>Biometra, Goettingen, GER</td>
</tr>
<tr>
<td>Rotation shaker Model G25</td>
<td>New Brunswick Scientific, New Jersey, USA</td>
</tr>
<tr>
<td>RP-HPLC Waters 2790</td>
<td>Waters Corporation, MA, USA</td>
</tr>
<tr>
<td>RTS ProteoMaster</td>
<td>Roche, Mannheim, GER</td>
</tr>
<tr>
<td>Equipment</td>
<td>Manufacturer/Location</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>Smart Spec™ Plus Spectrophotometer</td>
<td>BioRad Laboratories, Milan, I</td>
</tr>
<tr>
<td>Sonicator</td>
<td>Soniclean Pty. Ltd., Adelaide, AUS</td>
</tr>
<tr>
<td>Sorvall Evolution RC</td>
<td>Thermo Fisher Scientific, MA, USA</td>
</tr>
<tr>
<td>Sorvall RT7</td>
<td>Thermo Fisher Scientific, MA, USA</td>
</tr>
<tr>
<td>SpeedVac Concentrator</td>
<td>Savant, Global Medical Instrumentation, Inc., Minnesota, USA</td>
</tr>
<tr>
<td>Summit HPLC</td>
<td>Dionex, Corporation, CA, USA</td>
</tr>
<tr>
<td>UltiMate 3000 HPLC</td>
<td>Dionex, Corporation, CA, USA</td>
</tr>
<tr>
<td>UV-Trans-Illuminator, Bio Rad Gel Doc</td>
<td>BioRad Laboratories, Milan, I</td>
</tr>
</tbody>
</table>
APPENDIX 14: DNA, RNA AND PROTEIN MARKERS

1kb Plus DNA Ladder (Invitrogen) on 0.9% agarose gel stained with ethidium bromide.

0.5-10 kb RNA Ladder (Invitrogen) on 1.2% formaldehyde agarose gel stained with ethidium bromide.

Pre-stained SeeBlue Plus2 protein marker (Invitrogen)

<table>
<thead>
<tr>
<th>Protein</th>
<th>Approximate Molecular Weights (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tris-Glycine</td>
</tr>
<tr>
<td>Myosin</td>
<td>250</td>
</tr>
<tr>
<td>Phosphorylase</td>
<td>148</td>
</tr>
<tr>
<td>BSA</td>
<td>98</td>
</tr>
<tr>
<td>Chymotrypsin Dehydrogenase</td>
<td>64</td>
</tr>
<tr>
<td>Alcohol Dehydrogenase</td>
<td>50</td>
</tr>
<tr>
<td>Carbonic Anhydrase</td>
<td>36</td>
</tr>
<tr>
<td>Myoglobin Red</td>
<td>22</td>
</tr>
<tr>
<td>Lysozyme</td>
<td>16</td>
</tr>
<tr>
<td>Aprotinin</td>
<td>6</td>
</tr>
<tr>
<td>Insulin, B Chain</td>
<td>4</td>
</tr>
</tbody>
</table>

NuPAGE® Novex® Bis-Bis 4-12% Gel
6 REFERENCES


References


References


References


References


References


References


247
References


References


References


References


References


References


References


