

**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.



## ACKNOWLEDGEMENTS

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 immer währende 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
<b>1 Chapter 1: Introduction .....</b>	<b>1</b>
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) <i>Anti-listerial pediocin-like bacteriocins</i> .....	7
(b) <i>Two-peptide bacteriocins</i> .....	8
(c) <i>Cyclic bacteriocins</i> .....	8
(d) <i>Genetic organisation and production of class II bacteriocins</i> .....	9
(e) <i>Regulation of bacteriocin production</i> .....	9
(f) <i>Processing and secretion of bacteriocins</i> .....	10
(g) <i>Mode of action</i> .....	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) <i>Protein architecture of two-component systems</i> .....	16
(b) <i>Quorum sensing in Lactobacillus plantarum C11</i> .....	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	<i>Lactobacillus plantarum</i> .....	34
1.6.1	Plantaricin KW30.....	34
<b>2</b>	<b>Chapter 2: Material and methods .....</b>	<b>37</b>
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)	<i>Luria-Bertani (LB) media</i> .....	38
(b)	<i>De Man, Rogosa and Sharpe (MRS) media</i> .....	38
(c)	<i>M17 media with 0.5 % glucose</i> .....	38
(d)	<i>Yeast media (YM)</i> .....	38
(e)	<i>Chemically defined medium (CDM)</i> .....	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 <i>L. plantarum</i> KW30 .....	41
2.2.3	Isolation of plasmid DNA from <i>E. coli</i> .....	42
2.2.4	Isolation of plasmid DNA from <i>L. plantarum</i> .....	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)	<i>Phenol/Chloroform extraction</i> .....	44
(b)	<i>Precipitation of DNA</i> .....	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 .....	48
(a)	<i>Preparation of competent E. coli cells</i> .....	48
(b)	<i>Transformation and selection of E. coli cells</i> .....	49
(c)	<i>Preparation of electrocompetent E. coli cells</i> .....	49
(d)	<i>Electroporation of E. coli cells</i> .....	50
(e)	<i>Production of electrocompetent cells of Lactobacillus and Lactococcus species</i> .....	51
(f)	<i>Electroporation of Lactobacillus and Lactococcus species</i> .....	51
(g)	<i>Protoplast formation of L. plantarum KW30</i> .....	52
(h)	<i>Protoplast transformation</i> .....	52
2.2.15	Polymerase chain reaction (PCR) .....	53
(a)	<i>Polymerase chain reaction (PCR)</i> .....	53
(b)	<i>Colony PCR</i> .....	53
2.2.16	Targeted gene disruption .....	54
(a)	<i>Targeted gene disruption via homologous recombination in lactobacilli</i> .....	54
(b)	<i>Targeted gene disruption via cre-lox-based homologous recombination system in lactobacilli</i> .....	55
2.3	Methods for ribonucleic acid (RNA) work .....	56
2.3.1	Isolation of RNA .....	56
2.3.2	Agarose/formaldehyde gel electrophoresis of RNA .....	56
2.3.3	Northern blot .....	57
(a)	<i>Preparation of <sup>32</sup>P-labelled probe</i> .....	58
(b)	<i>Northern blotting and hybridisation</i> .....	58
2.3.4	Reverse Transcriptase (RT)-PCR .....	58
2.4	Protein biochemical methods .....	59
2.4.1	Determination of protein concentration .....	59
(a)	<i>Bradford method for determination of protein concentration</i> .....	59
(b)	<i>Concentration determination by UV absorption</i> .....	59
2.4.2	SDS-polyacrylamide gel electrophoresis .....	60
2.4.3	Detection of proteins in SDS-polyacrylamide gels .....	60
(a)	<i>Coomassie blue staining</i> .....	61
(b)	<i>Silver staining</i> .....	61
2.4.4	2D-electrophoresis .....	62
2.4.5	In-gel tryptic digest for protein identification by mass spectrometry .....	66
2.4.6	Western blot .....	68

## 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 <i>E. coli</i> .....	72
(b)	Expression of recombinant proteins in lactobacilli.....	73
(c)	Expression of recombinant proteins in <i>Lactococcus</i> .....	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

3	Chapter 3: Experimental results .....	88
3.1	Phylogenetic classification of <i>L. plantarum</i> KW30.....	89
3.2	Analysis of the <i>plnKW30</i> gene cluster.....	91
3.2.1	Bioinformatic analyses of the proteins of the <i>plnKW30</i> gene cluster .....	94
(a)	<i>Analysis of the putative GTase (GccB)</i> .....	95
(b)	<i>Analysis of the putative ABC-transporter (GccC)</i> .....	96
(c)	<i>Analysis of the putative TRX1 (GccD)</i> .....	99
(d)	<i>Analysis of the putative TRX2 (GccE)</i> .....	100
(e)	<i>Analysis of the putative response regulator (GccF)</i> .....	102
(f)	<i>Analysis of the bacteriocin PlnKW30 (GccA)</i> .....	103
3.2.2	Brief summary of bioinformatics analysis .....	104
3.3	Transcriptional analysis .....	105
3.3.1	Transcriptional analysis by RT-PCR.....	105
3.3.2	Northern blot analysis .....	108
3.3.3	Growth-phase-dependent gene expression of <i>plnKW30</i> .....	109
3.4	Characterisation of the bacteriocin PlnKW30 from <i>L. plantarum</i> KW30.....	110
3.4.1	Influence of growth conditions on PlnKW30 production .....	110
3.4.2	Purification of native PlnKW30 .....	111
3.4.3	Minimum inhibitory concentration of PlnKW30 .....	113
3.4.4	Antimicrobial activity of PlnKW30 against the indicator strain <i>L. plantarum</i> ATCC 8014.....	114
3.4.5	Analysis of antimicrobial activity of PlnKW30 .....	117
3.5	Analysis of N- and C-terminal fragments of PlnKW30 .....	124
3.5.1	Identification of the C-terminal glycan modification.....	130
(a)	<i>Mass spectrometry of N- and C-terminal fragments</i> .....	130
(b)	<i>Wheat germ agglutinin pull-down experiments</i> .....	134
(c)	<i>Dot blot using anti-O-GlcNAc antibodies</i> .....	135
3.6	Analysis of the post-translational modifications of PlnKW30 .....	136
3.6.1	Analysis of the disulfide bridges of PlnKW30 .....	136
3.6.2	Analysis of the glycan modifications of PlnKW30.....	139
3.7	Circular dichroism analysis of PlnKW30.....	144
3.7.1	Comparison of native, reduced and deglycosylated PlnKW30.....	148
3.7.2	Comparison of native, reduced and deglycosylated N-terminal fragment.....	149
3.7.3	Comparison of PlnKW30 with the N-terminal fragment.....	150

<b>4</b>	<b>Chapter 4: Discussion .....</b>	<b>151</b>
4.1	Phylogenetic classification of <i>L. plantarum</i> KW30 .....	151
4.2	Analysis of the <i>plnKW30</i> gene cluster .....	152
4.2.1	Bioinformatic analyses of the proteins of the <i>plnKW30</i> 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 <i>plnKW30</i> .....	166
4.3.3	Future work.....	166
4.4	Characterisation of <i>PlnKW30</i> of <i>L. plantarum</i> KW30.....	167
4.4.1	Antimicrobial activity of <i>PlnKW30</i> against the indicator strain <i>L. plantarum</i> ATCC 8014 .....	170
4.4.2	Future work.....	175
4.5	Analyses of N- and C-terminal fragments of <i>PlnKW30</i> .....	175
4.5.1	Discovery of C-terminal glycan modification .....	177
4.6	Analysis of the role of post translational modifications in <i>PlnKW30</i> activity.....	178
4.6.1	Analysis of the disulfide bridges of <i>PlnKW30</i> .....	178
4.6.2	Analysis of the glycan modifications of <i>PlnKW30</i> .....	179
4.6.3	Future work.....	180
4.7	Circular Dichroism.....	181
4.7.1	Future work.....	183
4.8	Conclusions .....	184
<b>5</b>	<b>Appendix .....</b>	<b>185</b>
	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 methyl iodide treatment.....	199

## Table of Contents

5.3	Discussion.....	200
5.3.1	Conclusions .....	202
	Appendix 6: Protein expression.....	203
5.4	Protein expression.....	203
5.4.1	Cell-free expression .....	204
5.5	Discussion.....	207
5.5.1	Conclusions .....	210
	Appendix 7: Targeted gene disruption of the <i>plnKW30</i> gene cluster .....	211
5.6	Targeted gene disruption using a temperature sensitive plasmid.....	211
5.7	Targeted gene disruption using a suicide vector .....	213
5.8	Discussion.....	214
5.8.1	Conclusions .....	215
	Appendix 8: 16-23SrDNA sequence of <i>L. plantarum</i> KW30 .....	216
	Appendix 9: Nucleotide and amino acid sequences.....	217
	Appendix 10: Complete sequence of <i>plnKW30</i> gene cluster (14149 bp) .....	220
	Appendix 11: Sequence alignments .....	225
	Appendix 12: MS/MS fragmentation of PlnKW30.....	227
	Appendix 13: Lists of enzymes and other materials .....	228
	Appendix 14: DNA, RNA and Protein Markers .....	233
6	References .....	234



## 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 <sub>C</sub> -DNA complex.....	18
Figure 1.8: Genetic map of <i>pln</i> locus of different <i>L. plantarum</i> strains.....	19
Figure 1.9: Glycan linkages to peptide chains.....	21
Figure 1.10: Pentasaccharide inner-core of <i>N</i> -linked glycans.....	21
Figure 1.11: Model for the biosynthesis of <i>N</i> - and <i>O</i> -linked glycoproteins in eukaryotes. ....	23
Figure 1.12: Model for the biosynthesis of <i>N</i> - and <i>O</i> -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 <i>et al.</i> (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 <i>Lactobacillus</i> species.....	90
Figure 3.2: Schematic bacteriocin gene cluster with putative functions of ORFs identified <i>in silico</i> ..	91
Figure 3.3: Possible regulatory sites of <i>gccC</i> (putativeABC-transporter).....	93
Figure 3.4: Nucleotide sequence and deduced protein of <i>gccA</i> ( <i>plnKW30</i> ).....	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 <i>plnKW30</i> gene cluster..	106
Figure 3.11: Northern blot analysis using a <i>plnKW30</i> 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 <i>L. plantarum</i> KW30 (producer strain) and <i>L. plantarum</i> ATCC 8014 (indicator strain). .....	115
Figure 3.18: Growth of <i>L. plantarum</i> ATCC 8014 after addition of different concentrations of PlnKW30. ....	116
Figure 3.19: Antimicrobial effect of PlnKW30 on <i>L. plantarum</i> ATCC 8014 cells.....	118
Figure 3.20: Concentration-dependent antimicrobial activity of PlnKW30 against <i>L. plantarum</i> ATCC 8014.....	119
Figure 3.21: Effect of different sugars on antimicrobial activity of PlnKW30.....	120
Figure 3.22: Growth of <i>L. plantarum</i> ATCC 8014 after addition of 1mg/mL <i>N</i> -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 <i>L. plantarum</i> ATCC 8014.....	126
Figure 3.28: Antimicrobial effect of N- & C-terminal fragments of PlnKW30.....	128
Figure 3.29: Growth of <i>L. plantarum</i> ATCC 8014 after addition of the N-terminal fragment of PlnKW30. ....	129

## List of Figures

Figure 3.30: Tricine SDS-PAGE of the N-terminal fragment overlaid with <i>L. plantarum</i> ATCC 8014 at OD <sub>600</sub> of 0.3.....	130
Figure 3.31: Monoisotopic mass spectrum of the N-terminal fragment of PlnKW30.....	131
Figure 3.32: Monoisotopic mass spectrum of the C-terminal fragment of PlnKW30.....	131
Figure 3.33: Monoisotopic mass spectrometry result for the C-terminal HC-X fragment of PlnKW30 .....	132
Figure 3.34: Cartoon representation of the NMR structure of PlnKW30.....	133
Figure 3.35: Tricine SDS-PAGE of lectin pull-downs of PlnKW30.....	134
Figure 3.36: Dot blot using anti- <i>O</i> -GlcNAc antibodies to detect PlnKW30 and its N- & C-terminal tryptic fragments.....	135
Figure 3.37: Visualisation of N- and C-terminal fragments using ninhydrin.....	135
Figure 3.38: RP-HPLC chromatogram of native and reduced/alkylated PlnKW30.....	136
Figure 3.39: Monoisotopic mass spectrometry result of the reduced and alkylated PlnKW30.....	137
Figure 3.40: Activity of native and reduced PlnKW30 in a bioassay overlay.....	138
Figure 3.41: Antimicrobial activity of reduced PlnKW30.....	139
Figure 3.42: RP-HPLC chromatogram of native and deglycosylated PlnKW30.....	140
Figure 3.43: Monoisotopic mass spectrometry result of the deglycosylated PlnKW30.....	141
Figure 3.44: Activity of native and deglycosylated PlnKW30 in a bioassay overlay.....	142
Figure 3.45: Antimicrobial effect of deglycosylated PlnKW30.....	143
Figure 3.46: Smoothed, base-line corrected spectra of native PlnKW30 (A), deglycosylated PlnKW30 (B) and their corresponding residuals.....	145
Figure 3.47: Smoothed, base-line corrected spectra of reduced and alkylated PlnKW30 (A), reduced PlnKW30 (B) and their corresponding residuals.....	146
Figure 3.48: Smoothed, base-line corrected spectra of the N-terminal fragment (A), reduced N-terminal fragment (B) and their corresponding residuals.....	147
Figure 3.49: Circular dichroism spectra of native, reduced and deglycosylated PlnKW30.....	148
Figure 3.50: Circular dichroism spectra of native, reduced and deglycosylated N-terminal fragment. .....	149
Figure 3.51: Circular dichroism spectra of native PlnKW30 and native N-terminal fragment.....	150
Figure 4.1: Comparison of the <i>plnKW30</i> gene cluster with similar gene clusters.....	154

## 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 <i>L. plantarum</i> 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 CH <sub>3</sub> I-treated PlnKW30.....	200
Figure 5.6: Western blot using anti-GFP antibodies. ....	204
Figure 5.7: Growth curves of <i>L. plantarum</i> KW30 at 30 °C and 43 °C.....	212
Figure 5.8: Amino acid alignment of the PlnKW30 GTase and a putative GTase from <i>Bacteroides fragilis</i> (pdb 3bcv) using FUGUE (Shi <i>et al.</i> 2001) (a: alpha helix; b: beta strand; 3:3 <sub>10</sub> helix; lower case: solvent accessible; UPPER CASE: solvent inaccessible). ....	225
Figure 5.9: Amino acid alignment of the plnKW30 GTase and SpsA from <i>Bacillus subtilis</i> (pdb 1qg8) using FUGUE (Shi <i>et al.</i> 2001) (a: alpha helix; b: beta strand; 3:3 <sub>10</sub> 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 <i>L. plantarum</i> KW30.....	89
Table 3.2: Characteristics of the genes and gene products of the <i>plnKW30</i> cluster of <i>L. plantarum</i> 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 <i>plnKW30</i> 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 <i>E. coli</i> expression strains and their descriptions.....	203
Table 5.5: Expression constructs and strains used for expression trials of <i>plnKW30</i> , RR and GT in this thesis.....	205
Table 5.6: Results of disruption attempts with <i>L. plantarum</i> 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

## Abbreviations

Å	Ångström ( $10^{-10}\text{m}$ )
aa	Amino acid
ABC-transporter	ATP-binding cassette transporter
ACN	acetonitrile
AHL	acyl-homoserine lactone
Amp	Ampicillin
APS	Ammonium persulfate
ATCC	American Type Culture Collection
ATP	Adenosine-5'-triphosphate
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
BSA	Bovine serum albumin
°C	Degrees Celsius
C	Carbon
ca.	Circa
CAZy	Carbohydrate-Active enZymes
CBM	Carbohydrate-binding module
Cds	Conserved domains in sequences
Cm	Chloramphenicol
CTP	Cytidine-5'-triphosphate
ddH <sub>2</sub> O	Double distilled water
Dha	Dehydroalanine
Dhb	Dehydrobutyrine
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxyribose nucleotide triphosphate
DTT	Dithiothreitol
EDTA	Ethylene diamine tetraacetic acid

## Abbreviations

Em	Erythromycin
<i>et al.</i>	<i>et alteri</i> (and others)
ER	Endoplasmic reticulum
EtBr	Ethidium bromide
ETD	Electron Transfer Dissociation
EtOH	Ethanol
ExPASy	Expert Protein Analysis System
FLP	FNR-like regulatory proteins
FNR	fumarate-nitrate reduction
FPLC	Fast Protein Liquid Chromatography
FTase	Farnesyltransferase
g	Gramm; standard gravity (9.81 m/s <sup>2</sup> )
GalNAc $\alpha$ 1-Ser/Thr	<i>N</i> -acetyl-D-galactosamine and L-serine or L-threonine
gDNA	Genomic DNA
GDP	Guanidine-5'-diphosphate
G+C %	Percentage of guanine and cytosine
Gal	Galactose
GalNAc	<i>N</i> -acetylgalactosamine
GlcNAc	<i>N</i> -acetylglucosamine
GlcNAc $\beta$ 1-Asn	<i>N</i> -acetylglucosaminyl-asparagine
G protein	GTP binding protein
GT	Glycosyltransferase
GTP	Guanosine-5'-triphosphate
hrs	Hours
HEPES	n-(2-hydroxyethyl)piperazine-n-(2-ethanesulfonic acid)
HPK	Histidine Protein Kinase
HPLC	High Pressure Liquid Chromatography
HTH	Helix-turn-helix
IEX	Ion Exchange Chromatography
IM	Inner membrane
IMAC	Immobilised Metal Affinity Chromatography

## Abbreviations

IPTG	Isopropyl- $\beta$ -D-Thiogalactopyranoside
IR	Infrared radiation
IUBMB	International Union of Biochemistry and Molecular Biology
Kan	Kanamycin
Kbp	Kilo basepairs
kDa	kilo Daltons
L	Liter
$\lambda$	Wavelength
LAB	Lactic Acid Bacteria
LB	Luria Bertani broth
M	Molar (mol/L)
m	Meter
Man	Mannose
man-PTS	mannose phosphotransferase system
MCS	Multiple Cloning Site
min	Minute
mol	$6.023 \times 10^{23}$ molecules
MRS	De Man, Rogosa and Sharpe broth
MW	Molecular mass
m/z	Mass-to-charge ratio
n/a	Not applicable
NaCl	Sodium chloride
NBD	Nucleotide binding domain
NCBI	National Centre for Biotechnology Information
ND	Not determined
NDP	Nucleotide diphosphate
NEB	New England Biolabs
NH <sub>2</sub>	Amine
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	Ammonium sulphate
NMR	Nuclear magnetic resonance
No.	Number

## Abbreviations

O	Oxygen
OD	Optical Density
OM	Outer membrane
o/n	overnight
ORF	Open Reading Frame
p.a.	<i>pro analysis</i>
Pa	Pascal (= $10^{-5}$ bar = $145.04 \times 10^{-6}$ psi)
PCR	Polymerase Chain Reaction
pH	Negative decadal logarithm of the proton concentration
PlnKW30	plantarin KW30
ppm	Parts per million
PTase	Prenyltransferase
PTM	Post-translational modification
PVDF	Polyvinylidene difluorine
QS	Quorum sensing
RBS	Ribosome Binding Sequence
RNA	Ribose nucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
RT	Room temperature
RP-HPLC	Reverse-Phase High Pressure Liquid Chromatography
RR	Response regulator (originally thought to be FTase)
SAM	S-adenosyl-methionine
s/sec	Second
S/D	Shine-Delgarno-Sequence
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
sec pathway	Translocase general secretion pathway
S-layer	surface-layer
Sm	Streptomycin
SOB	Super optimal broth

## Abbreviations

SOC	SOB with catabolite repression, indicative of the presence of glucose
Sp.	Species
SSC	Saline sodium citrate buffer
Subsp.	Subspecies
T	Temperature
t	time
TAE	Tris-Acetate-EDTA
TCEP	tris(2-carboxyethyl)phosphine
TCS	Two-Component System
Tc	Tetracycline
TE	Tris-EDTA
TLC	Thin layer chromatography
TMD	Transmembrane domain
TMH	Transmembrane helix
TRIS	Tris (hydroxymethyl) aminomethane
TTP	Thymidine-5'-triphosphate
TRX	Thioredoxin
UDP	Uridine diphosphate
UDP-Glc	Uridine diphosphate- $\alpha$ -D-glucose
USA	United States of America
UV	Ultraviolet light
V	Volts
v/v	Volume per volume
W	Watt
w/v	Weight per volume
w/w	Weight per weight

## Abbreviations of Nucleic Acids

<b>One letter code</b>	<b>Base(s) represented</b>
A	Adenine
C	Cytosine
G	Guanine
T	Thymine
U	Uracil
R	GA
Y	TC
K	GT
M	AC
S	GC
W	AT
B	GTC
D	GAT
H	ACT
V	G or C or A
N	ANY

## Abbreviations of Amino Acids

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



# 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  $\beta$ -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.

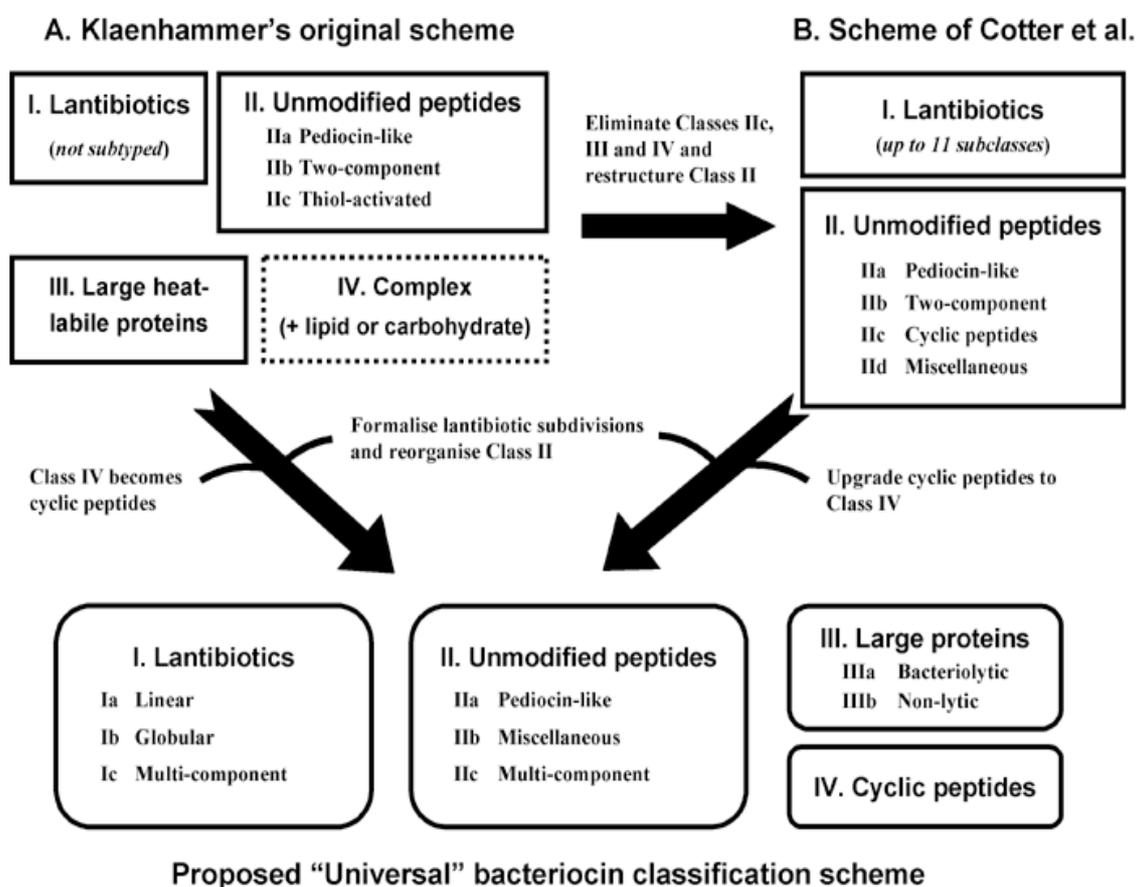
Class II bacteriocins are small (< 10 kDa), heat-stable, unmodified peptides, typically with a diglycine processing site in the bacteriocin prepeptide. Class II can be subdivided into three groups. Class IIa are the *Listeria*-active bacteriocins (e.g. pediocin PA-1/AcH), containing an N-terminal sequence 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.



**Figure 1.1: Proposed classification scheme for bacteriocins.**  
(from correspondence to Cotter *et al.* 2005; by Heng & Tagg 2006)

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  $\beta$ -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).

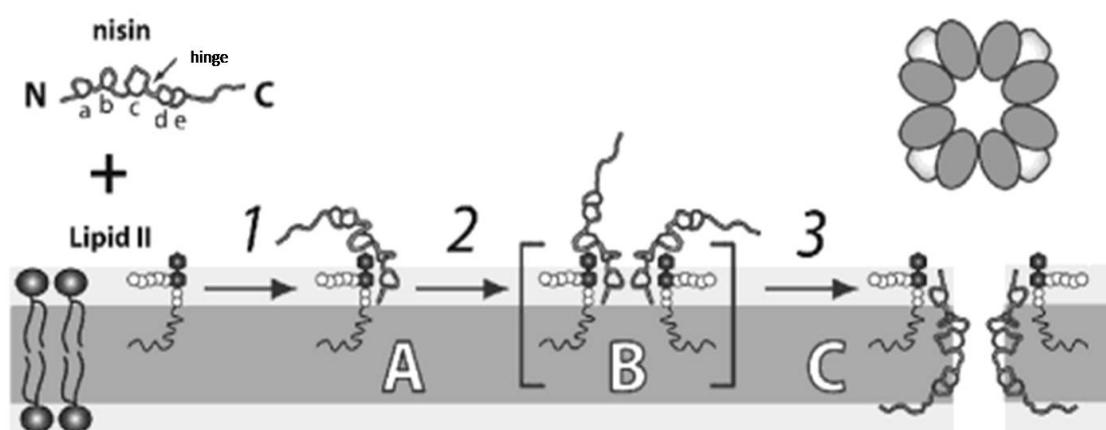
**Table 1.1: Open reading frames and functions of the nisin gene cluster.**

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

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  $\beta$ -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).



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).



**Figure 1.4: Model for the formation of nisin-lipid II pores in lipid bilayers.**

The polypeptide antibiotic nisin is presented here in a very simplified way, showing only the main structural properties such as the five thioether rings (a-e) and the position of the hinge region (top left). The formation of the nisin-lipid II pore (C) is thought to occur in 3 steps; via the formation of two intermediate structures (A and B). The top view of complex C (top right) shows the proposed arrangement of the molecules in a pore complex composed of 8 nisin (grey ellipse) and 4 lipid II (white triangle) molecules (from 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).

### 1.3.2 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.

	1	10	20	30	40
Pediocin PA-1/AcH	KYY <u>YGNGV</u> TCGKHSCSVDWGKATTCIINNGAMAWATGGHQGNIKC				
Mesentericin Y105	KYY <u>YGNGV</u> HCTKSGCSVNWGEAASAGIHRLANGGNGFW				
Carnobacteriocin B2	VNY <u>YGNGV</u> SCSKTKCSVNWQAFQERYTAGINSFVSGVASGAGSIGRRP				
Carnobacteriocin BM1	ISY <u>YGNGV</u> YCNKEKCWVNKAENKQAITGIVIGGWASSLAGMGH				
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  $\alpha$  and  $\beta$  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 as 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 (QX<sub>4</sub>D/ECX<sub>2</sub>AX<sub>3</sub>MX<sub>4</sub>Y/FGX<sub>4</sub>I/L) and a histidine motif (HY/FY/VVX<sub>10</sub>I/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)**

Bacteriocin	Mode of action	$\Delta pH$	PMF	ATP	Efflux	References
Mesentericin Y105	Pore formation	ND	dissipation	ND	aa	(Hechard <i>et al.</i> 1992; Maftah <i>et al.</i> 1993; Fleury <i>et al.</i> 1996; Robichon <i>et al.</i> 1997; Dalet <i>et al.</i> 2000; Dalet <i>et al.</i> 2001)
Pediocin PA-1/AcH	Pore formation	dissipation	dissipation	Depletion of intracellular ATP	aa, K <sup>+</sup>	(Bhunja <i>et al.</i> 1991; Christensen & Hutkins 1992; Chikindas <i>et al.</i> 1993; Chen <i>et al.</i> 1997a; Chen <i>et al.</i> 1997b; Chen <i>et al.</i> 1998; Fimland <i>et al.</i> 1998; Waite <i>et al.</i> 1998)
Lactacin F	Pore formation	ND	dissipation	Depletion of intracellular ATP	K <sup>+</sup> and P <sub>i</sub>	(Abee <i>et al.</i> 1994)
Lactococcin G	Cation pores	No effect	dissipation	Depletion of intracellular ATP	aa, Na <sup>+</sup> , K <sup>+</sup> , Li <sup>+</sup> , Cs <sup>+</sup> , Rb <sup>+</sup>	(Nissen-Meyer <i>et al.</i> 1992; Moll <i>et al.</i> 1996; Moll <i>et al.</i> 1998)
Plantaricin EF	Cation pores	dissipation	dissipation	ND	Cations	(Moll <i>et al.</i> 1999)
Plantaricin JK	Anion pores	dissipation	dissipation	ND	Anions	(Moll <i>et al.</i> 1999)

aa, amino acids; ND, not determined; P<sub>i</sub>, 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 (Dridger *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).

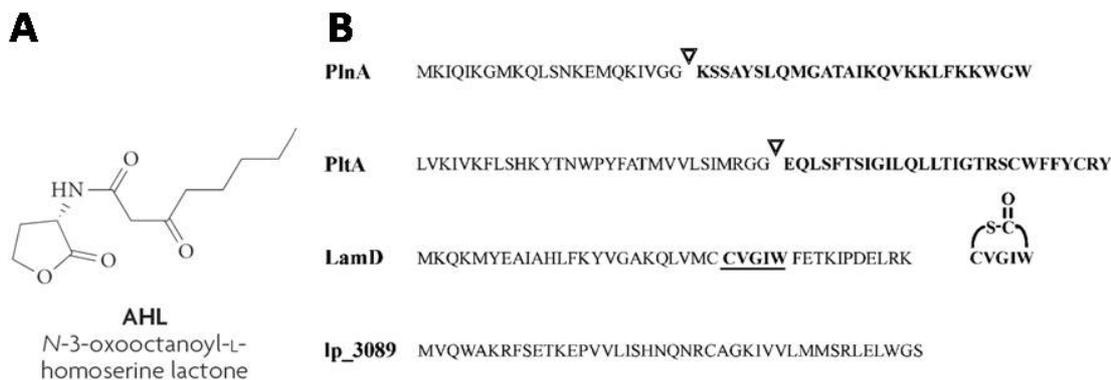
There are several proposed modes of action for immunity proteins: they can interact with the bacteriocin-induced pore and physically block it. Alternatively, they can interact with the bacteriocin receptor, thereby altering the receptor conformation so it can no longer bind bacteriocins and/or masking 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 wide spread 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.**

**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; Uden *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

production of bacteriocins in LAB (Nes *et al.* 1996; Dunny & Leonard 1997; Kleerebezem *et al.* 1997; Nes & Eijsink 1999).

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.

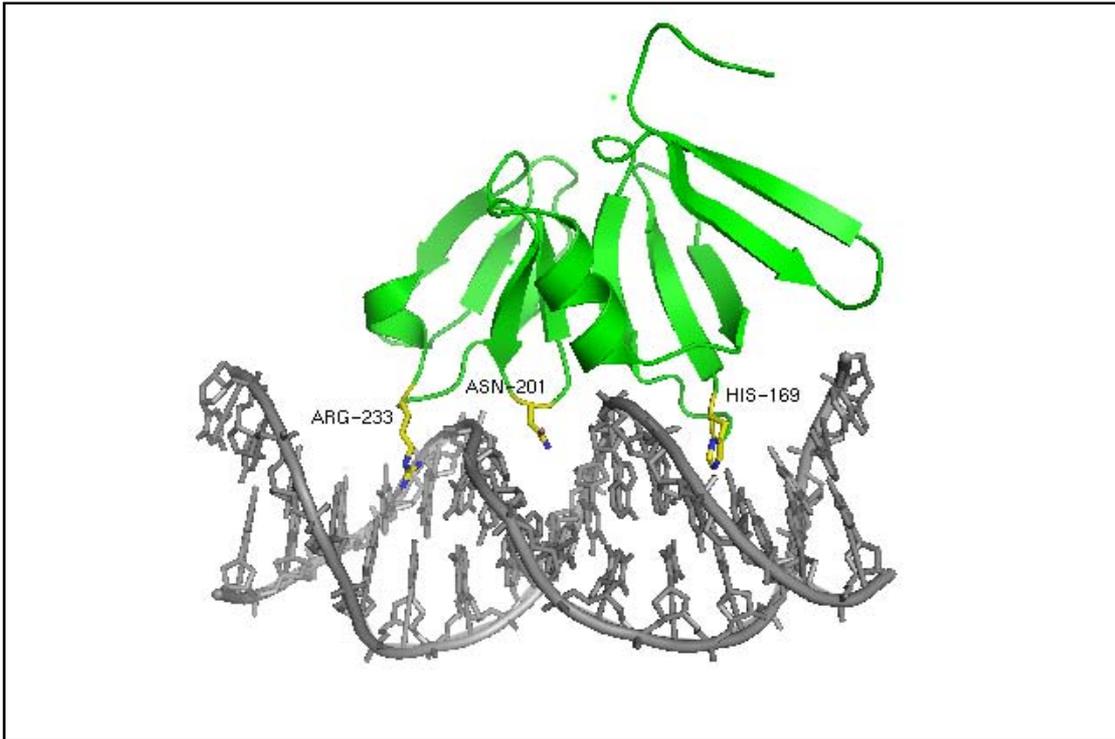
#### *(a) Protein architecture of two-component systems*

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 HPK<sub>10</sub> 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<sub>10</sub> subfamily belong to the R<sub>D</sub>/ComE subfamily (Sturme *et al.* 2007). The majority of response regulators of the R<sub>D</sub>/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 *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 *et al.* 2008). The C-terminal DNA-binding domain of *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  $\beta$ -strands organized in three antiparallel  $\beta$ -sheets and a two-turn  $\alpha$ -helix as shown in figure 1.7 (Sidote *et al.* 2008).

Base-specific contacts are made between residues located in the loops between the  $\beta$ -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).



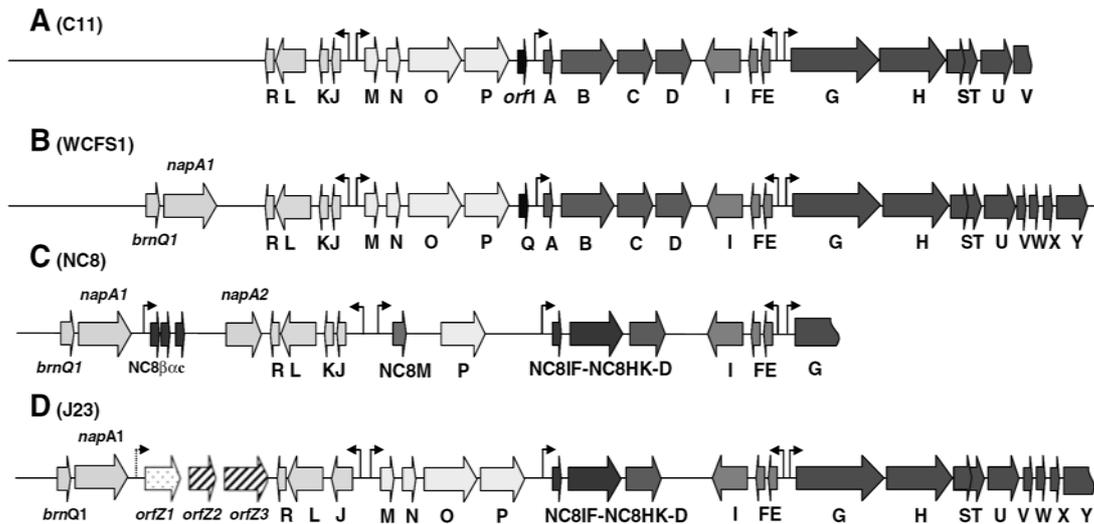
**Figure 1.7: Cartoon representation of the structure of the AgrAc-DNA complex.** Residues that make base-specific contacts with the DNA are shown in stick representation. The DNA is coloured in gray (PyMol; DeLano 2002).

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).



**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. AL935253); **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).

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) and J51 (Navarro *et al.* 2008). The *pln* locus in these *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 (*plnGH*), bacteriocin maturation (*plnSXY*) and variable regions related to bacteriocin regulation, production and immunity (*plnABCD*; *plnIEF*; *plnJKL*).

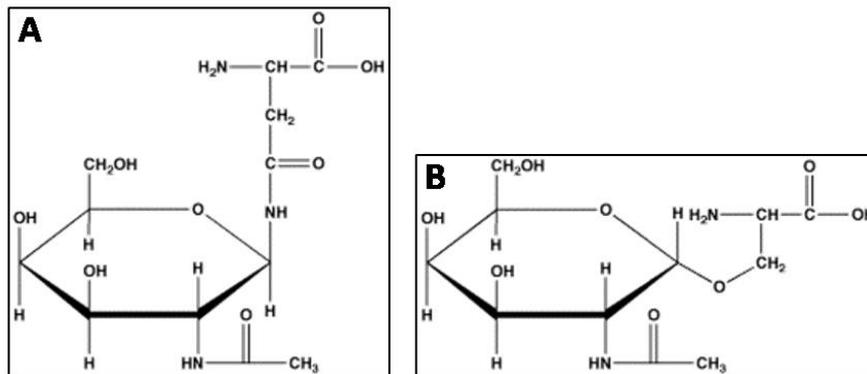
## 1.5 GLYCOSYLATION

### 1.5.1 GLYCOSYLATION 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. *N*-glycosyl and *O*-glycosyl linkages occur in all three kingdoms, whereas *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 physico chemical 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: *N*-glycosylproteins, *O*-glycosylproteins and *N,O*-glycosylproteins. In *N*-glycosylated proteins the sugar chains are attached to the protein via the side chain amide nitrogen of an asparagine residue, while in *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).

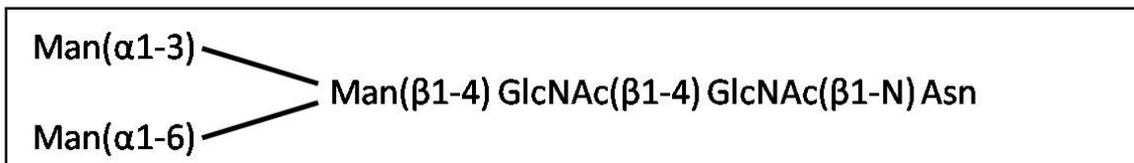


**Figure 1.9: Glycan linkages to peptide chains.**

**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 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<sub>3</sub>-GlcNAc<sub>2</sub> (Figure 1.10), which is followed by three different types: the high-mannose-type (or oligomannosidic-type), the complex-type, and the hybrid-type.



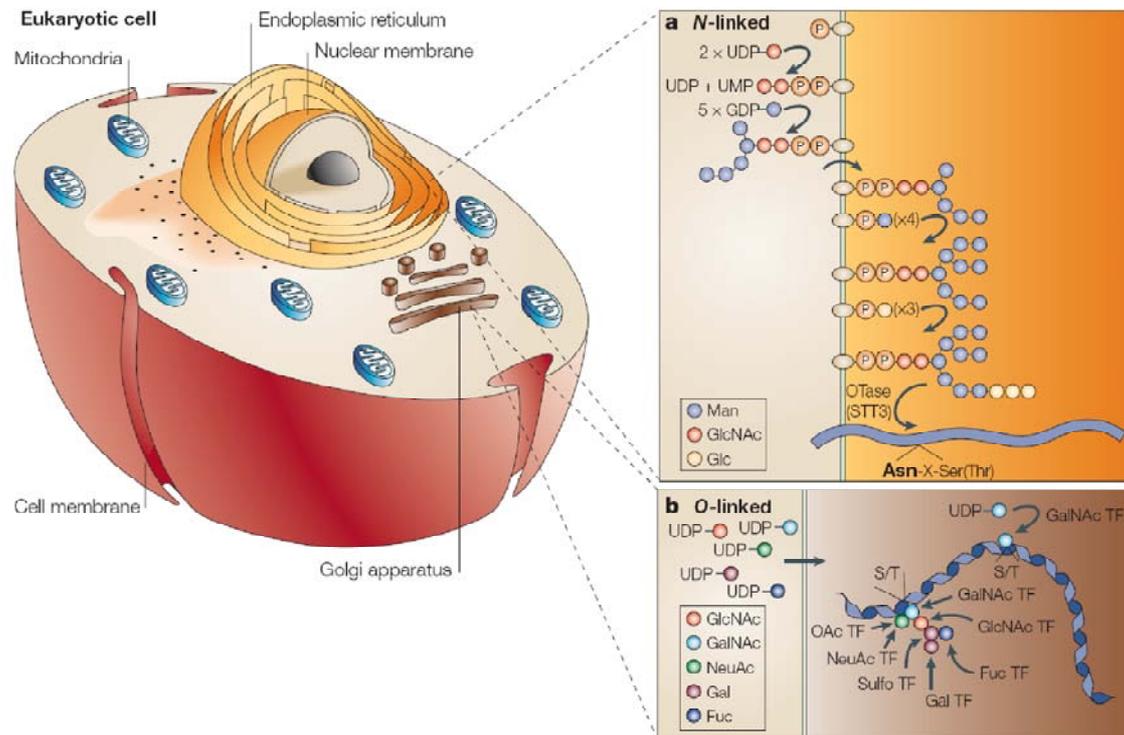
**Figure 1.10: Pentasaccharide inner-core of *N*-linked glycans.**

Asn, asparagines; GlcNAc, *N*-acetylglucosamine; Man, mannose.

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.

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)



**Figure 1.11: Model for the biosynthesis of *N*- and *O*-linked glycoproteins in eukaryotes.**

**a:** *N*-linked glycosylation; see text for details. **b:** *O*-linked glycosylation: see text for details. Fuc, fucose; Gal, galactose; GalNAc, *N*-acetylgalactosamine; Glc, glucose; GlcNAc, *N*-acetylglucosamine; NeuAc, *N*-acetylneuraminic acid; OTase, oligosaccharyltransferase; TF, transferase (from Szymanski & Wren 2005).

Modification of peptides and proteins with a single *O*-linked  $\beta$ -*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  $\beta$ -*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 physico chemical properties (Upreti *et al.* 2003).

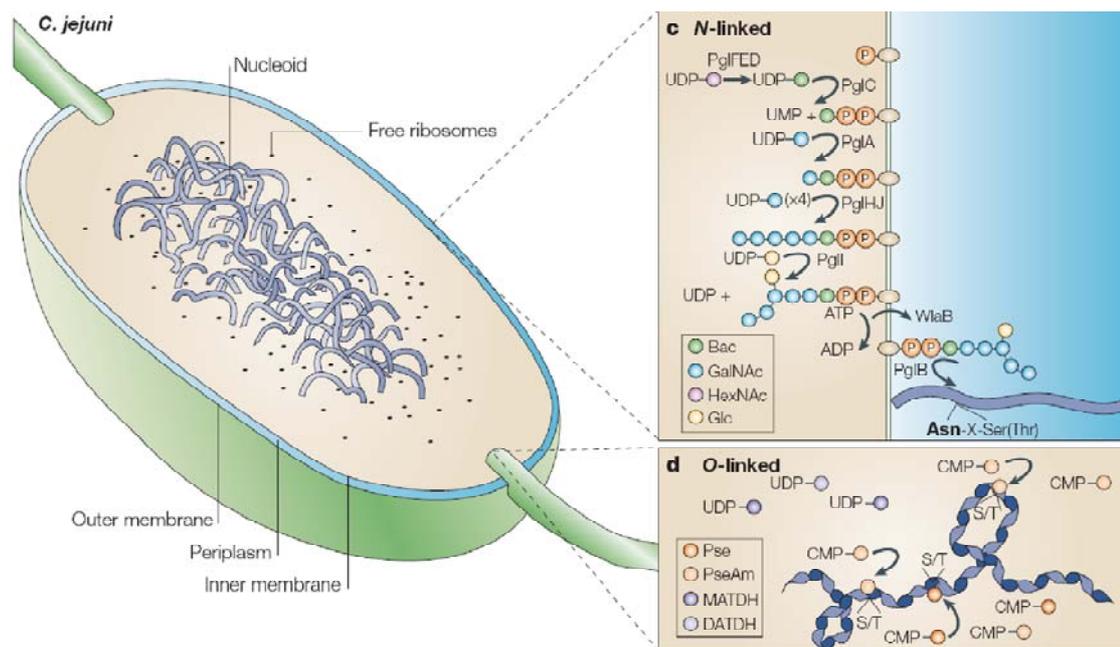
Only N- and O-glycosylation have so far been found in prokaryotes, but the available data shows that the structures of prokaryotic glycoproteins are far more diverse than the structures found in eukaryotes, although they share some characteristics (Moens &

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.



**Figure 1.12: Model for the biosynthesis of *N*- and *O*-linked glycoproteins in bacteria.**

**c:** *N*-linked glycosylation; see text for details. **d:** *O*-linked glycosylation: see text for details. Bac, 2,4-diacetamido-2,4,6- trideoxyglucose; DATDH, diacetamido-trideoxyhexose; Fuc, fucose; Gal, galactose; GalNAc, *N*-acetylgalactosamine; Glc, glucose; GlcNAc, *N*-acetylglucosamine; HexNAc, *N*-acetylhexosamine; MATDH, monoacetamido-trideoxyhexose; Pse, 5,7-diamino-3,5,7,9-tetraoxynon-2-ulosonic acid (Pse5Ac7Ac); PseAm, 5-acetamidino analogue of Pse (Pse5Am7Ac) (from Szymanski & Wren 2005).

### 1.5.3 *S*-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.

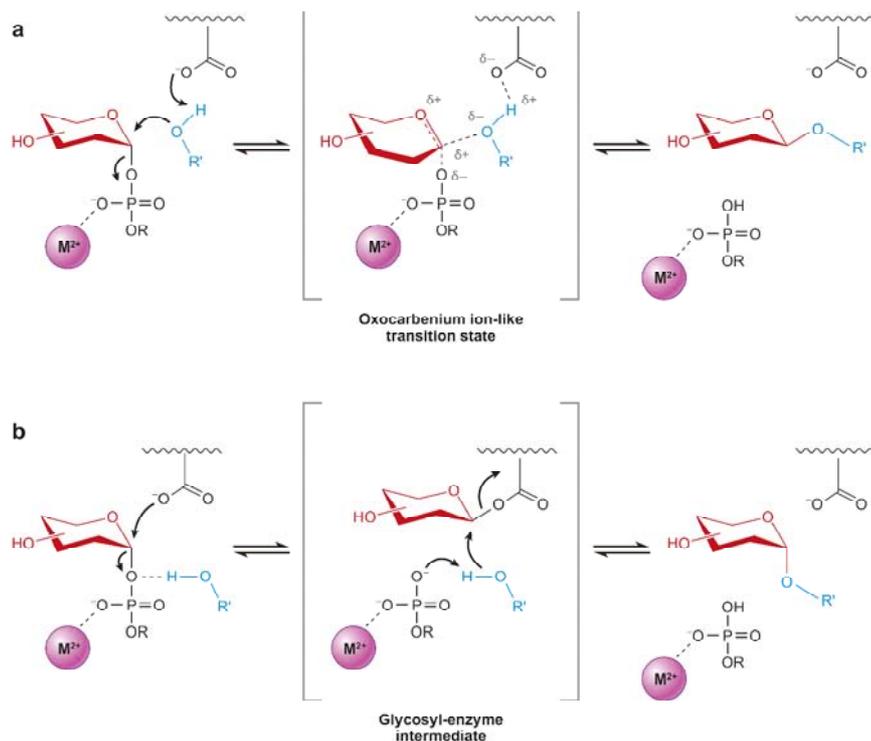
The glycosylation of a cysteine residue in a polypeptide chain as we found in PlnKW30 (see section 3.5.1), is totally new in biology and no natural examples of protein or peptide *S*-glycosylation exist (Taylor 1998; Thibodeaux *et al.* 2007).

#### 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  $S_N2$  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).



**Figure 1.13: Mechanism of inverting and retaining glycosyltransferases.**

**a:** Inverting glycosyltransferases utilize a direct-displacement  $S_N2$ -like reaction that results in an inverted anomeric configuration via a single oxocarbenium ion-like transition state. **b:** A proposed double-displacement mechanism for retaining glycosyltransferases involves the formation of a covalently bound glycosyl-enzyme intermediate. Abbreviations: R, a nucleoside, a nucleoside monophosphate, a lipid phosphate, or phosphate (phosphorylases classified as glycosyltransferases); and R'OH, an acceptor group (e.g., another sugar, a protein, or an antibiotic) (from Lairson *et al.* 2008).

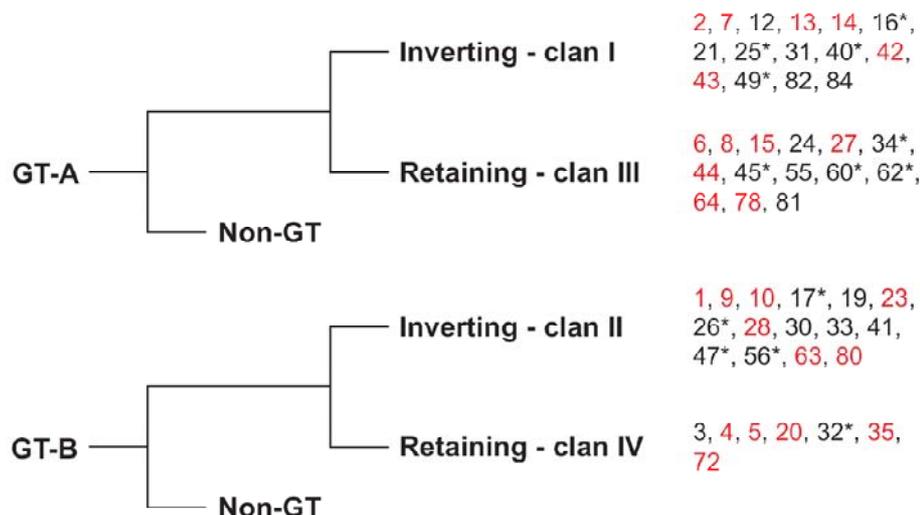
Several studies (Boix *et al.* 2001; Persson *et al.* 2001; Boix *et al.* 2002; Ly *et al.* 2002; Pedersen *et al.* 2003) of retaining glycosyltransferases could not identify the presence of an active-site nucleophile involved in the formation of the enzyme-sugar adduct. However, a mutant  $\alpha$  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](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 furiosus* (Igura *et al.* 2007).

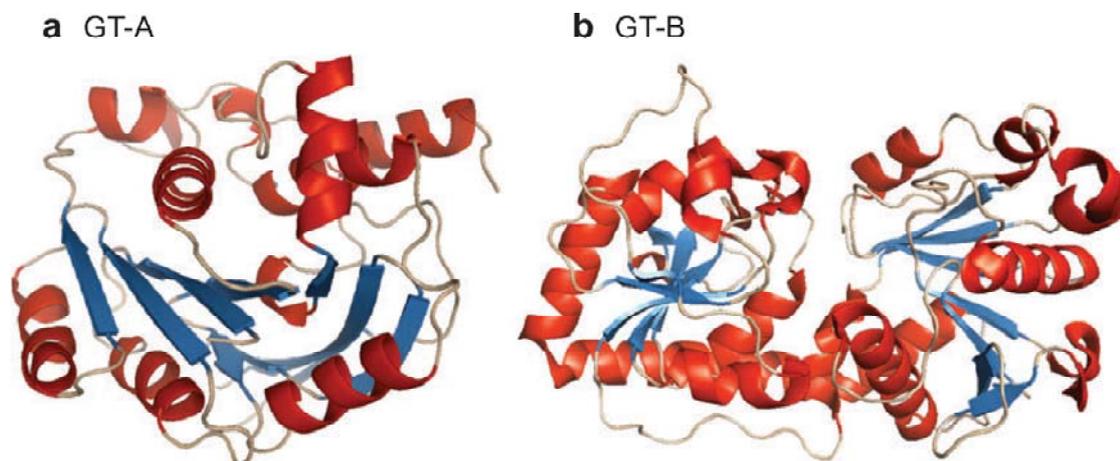
However, Lovering and co-workers (2007) described the structure of a new family of glycosyltransferases that adopts a different fold from both the GT-A and GT-B families, but was not a predicted GT-C member. This bifunctional glycosyltransferase/transpeptidase enzyme is involved in peptidoglycan biosynthesis in *Staphylococcus aureus* and its glycosyltransferase domain (GT51 family) shows a primarily  $\alpha$ -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  $\beta$ -glucosyltransferase (BGT; Vrieling *et al.* 1994) (Figure 1.15b). The BGT catalyses the transfer of glucose from uridine diphosphate- $\alpha$ -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  $\beta/\alpha/\beta$  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  $\beta$ -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  $\beta/\alpha/\beta$  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  $\alpha/\beta/\alpha$  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.**

**a:** GT-A fold represented by SpsA from *B. subtilis*; resolution 1.5 Å, Protein Data Bank (pdb) 1qgq; **b:** GT-B fold represented by the DNA modifying bacteriophage T4  $\beta$ -glucosyltransferase, pdb 1jg7 (from Lairson *et al.* 2008).

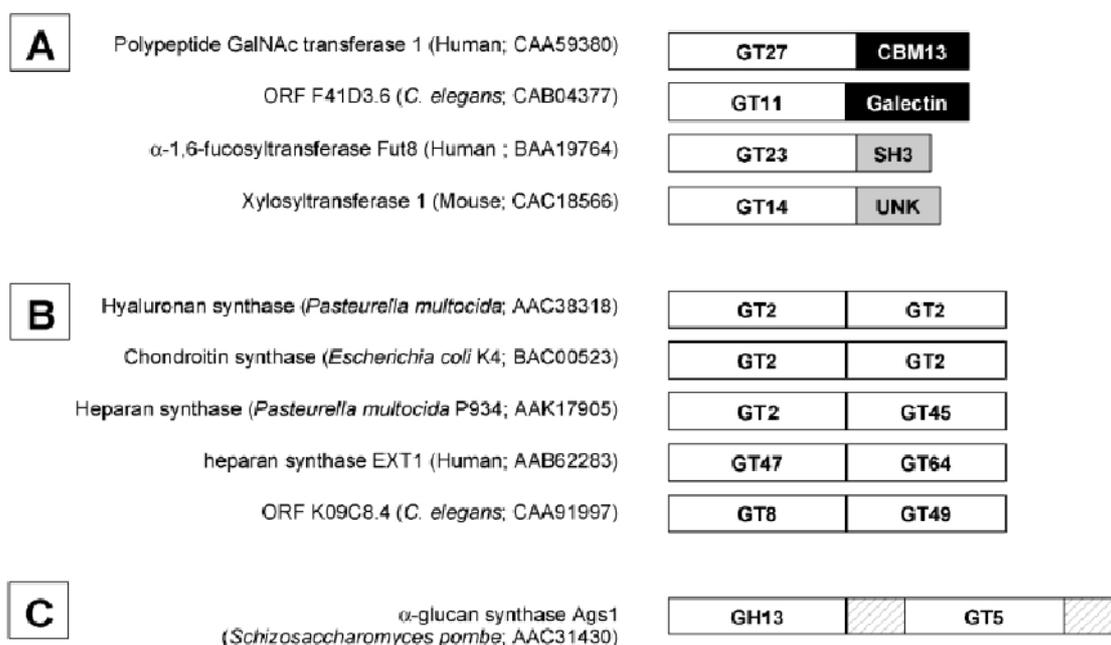
The GT-B fold has also been found in the catalytic domain of *N*-acetylglucosamine (*O*-GlcNAc) transferase of family GT41 from *Xanthomonas campestris*. This enzyme catalyses the intracellular modification of serine and threonine residues with *O*-GlcNAc, which interferes with protein phosphorylation, 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 tetratricopeptide (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  $\alpha$ -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.**

**A:** Examples of glycosyltransferases with appended non-catalytic modules; **B:** examples of tandems of two glycosyltransferases on a single polypeptide; **C:** a glycosyltransferase with an appended (trans)glycosidase domain. CBM13, carbohydrate-binding module of family 13 (a family classification of CBMs is available from the CAZy database at: <http://afmb.cnrs-mrs.fr/CAZY/CBM.html>); galectin, module displaying sequence similarity to galectins; SH3, Src homology 3 domain; UNK, module of unknown function; GH13, transglycosidase module from glycoside hydrolase family GH13. A family classification of glycoside hydrolases is available from the CAZy database at: <http://afmb.cnrs-mrs.fr/CAZY/GH.html> (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  $\alpha$ -amylase, lipase A or lysozyme (Kelly *et al.* 1996). PlnKW30 was shown to be stable over a

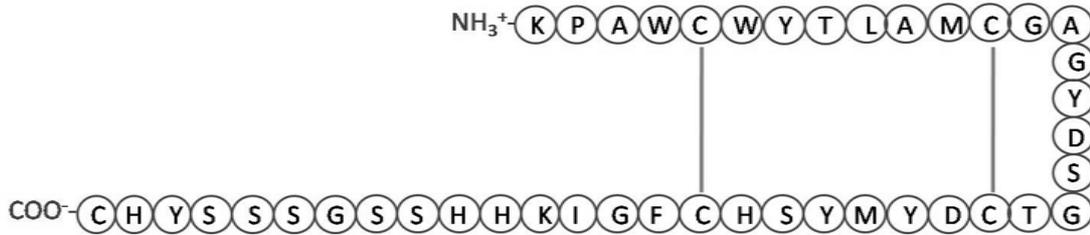
pH range from 2 - 10 at 28 °C for 16 hours and at pH 3.5 at 100 °C for 60 minutes, however, all activity was lost after autoclaving. Bacteriocin production was shown to begin and rapidly increase when cell cultures reached stationary phase. Kelly *et al.* (1996) suggested that the mode of action of PlnKW30 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) led 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<sub>0</sub>, 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*-acetylated 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 CHAPTER 2: MATERIAL AND METHODS

A list of strains and plasmids used can be found in appendix 3 table 5.1 and table 5.2, respectively, and all expression plasmids constructed for this work are summarised in table 5.5.

Appendix 10 lists polymerases and restriction enzymes (Table 5.8), commercial kits (Table 5.9), consumables (Table 5.10) and equipment (Table 5.11) used.

### 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 H<sub>2</sub>O 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 10<sup>5</sup> 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<sub>2</sub>O made up to 1 L

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

MRS media                    52.2    g

Pure H<sub>2</sub>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<sub>2</sub>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<sub>2</sub>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)**

Chemical	g / L	Chemical	g / L
glucose	10.0	DL-Alanine	0.20
potassium acetate	10.0	L-Arginine	0.30
potassium dihydrogen orthophosphate	2.0	L-Asparagine	0.20
sodium thioglycollate	0.5	L-Aspartic acid	0.20
magnesium sulphate 7H <sub>2</sub> O	0.15	L-Cysteine-HCl	0.20
manganese sulfate 4H <sub>2</sub> O	0.02	L-Glutamic acid	0.15
ferrous sulphate 7H <sub>2</sub> O	0.01	L-Glycine	0.30
Tween 80	0.001	L-Histidine-Hal	0.20
adenine	0.05	L-Isoleucine	0.20
cytidylic acid	0.05	L-Leucine	0.30
deoxyguanosine	0.05	L-Lysine-HCl	0.30
guanine HCl	0.05	L-Methionine	0.20
p-amineobenzoic acid	0.01	L-Phenylalanine	0.20
Vitamins B12	0.001	L-Proline	0.30
Calcium pantothenate	0.001	L-Serine	0.30
D-biotin	0.01	L-Threonine	0.20
Folic acid	0.001	L-Tryptophan	0.20
Niacin	0.001	L-Tyrosine	0.30
Piridoxal ethyl acetal	0.001	L-Valine	0.30
HCl			
Riboflavin	0.001		
Thiamine HCl,	0.001		

**2.1.4 ANTIBIOTIC STOCK SOLUTIONS**

The antibiotic stock solutions were prepared in pure H<sub>2</sub>O 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**

Antibiotic	Stock solution (mg/mL)	Final concentration ( $\mu\text{g/mL}$ )		
		<i>E. coli</i>	<i>Lb. plantarum</i>	<i>L. lactis</i>
Ampicillin (Na-salt)	100 in pure H <sub>2</sub> O	100	-	-
Tetracycline (-Hydrochloride)	12.5 in ethanol (70 %)	12.5	-	-
Kanamycin (-Sulfate)	40 in pure H <sub>2</sub> O	15	-	-
Chloramphenicol	34 in pure H <sub>2</sub> O	30	5 or 10	5 or 10
Erythromycin	50 in ethanol (absolute)	200	5 or 10	5 or 10
Streptomycin	500 in pure H <sub>2</sub> O	-	500	-

### 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™ *CryoTubes*) 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<sub>600</sub> 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  $2 \times 10^5$  Pa to deactivate any DNases present. Non autoclavable equipment was treated with 70 % (v/v) ethanol and then rinsed with sterile pure H<sub>2</sub>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<sub>600</sub> 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<sub>3</sub>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-harboring *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 H<sub>2</sub>O 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.

<u>Alkaline SDS solution:</u>	3.0 %	SDS
	0.2 N	NaOH
<u>TE-RNase:</u>	10.0 mM	Tris/HCl
	1.0 mM	EDTA
	0.1 mg/mL	RNase
		pH 8.5

### *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 H<sub>2</sub>O. DNA was illuminated by exposure to UV-light ( $\lambda = 254$  nm) and recorded using an UV-Trans-Illuminator (BioRad Gel Doc, BioRad Laboratories, Milan, I).

<u>TAE buffer (50x):</u>	2.0	M	Tris/HCl
	5.0	mM	EDTA
	57.1	mL	Glacial acetic acid
			pH 8,0
	Pure H <sub>2</sub> O		up to 1 L

<u>Stop-mix (6x):</u>	60.0	mM	Tris-HCl
	60.0	mM	EDTA
	0.2	% (w/v)	Orange G
	0.05	% (w/v)	Xylene Cyanol FF
	60.0	% (v/v)	Glycerol

### *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.

<u>Ligation conditions:</u> 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 OD<sub>600</sub> 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.

<u>RF1 solution:</u>	100.0 mM	RbCl
	50.0 mM	MnCl <sub>2</sub>
	30.0 mM	Potassium acetate
	10.0 mM	CaCl <sub>2</sub>
		pH 5.8 (conc. Acetic Acid)

<u>RF2 solution:</u>	10.0 mM	RbCl
	10.0 mM	MOPS
	75.0 mM	CaCl <sub>2</sub> x 6 H <sub>2</sub> 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<sub>600</sub> 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	MgCl <sub>2</sub>

1. Dissolve tryptone, yeast extract and NaCl in 950 mL H<sub>2</sub>O.
2. Make 250 mM KCl solution by dissolving 1.86 g of KCl in 100 mL H<sub>2</sub>O. 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 H<sub>2</sub>O.
4. Autoclave, cool to ~55 °C and add 10 mL of sterile 1 M MgCl<sub>2</sub>.

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 OD<sub>600</sub> of about 0.05. The cells were grown at 30 °C for 4 hours until an OD<sub>600</sub> 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 H<sub>2</sub>O, 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 x *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 x *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.

<u>Protoplast buffer (PB):</u>	10 mM	Tris/HCl (pH 8)
	20 mM	CaCl <sub>2</sub>
	0.5 M	Sucrose

*(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 MgCl<sub>2</sub>. After the incubation, the cells were centrifuged again and resuspended in 50 µL protoplast buffer and then plated on RM agar.

<u>Regeneration medium (RM):</u>		MRS
	0.5 M	Sucrose
	20 mM	MgCl <sub>2</sub>
	2.5 %	gelatine
	0.5 %	heat-inactivated BSA

### 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), Cm<sup>r</sup>, provides *repA in trans*) is temperature sensitive and becomes unstable above 43 °C. The mutagenesis vector pORI28 (Em<sup>r</sup>, 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 *Bam*HI/*Eco*RI 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 *lox66-P<sub>32</sub>-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 *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 *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 *PmeI* and *Ecl136II* blunt-end restriction sites of pNZ5319. The primers were designed so that the 5'- and 3'-flanking regions of the *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 *plnKW30*-mutagenesis vector was transferred into *L. plantarum* KW30 by electroporation as described in section 2.2.14(f). Chloramphenicol-resistant (Cm<sup>r</sup>) transformants were selected and replica plated to check for an erythromycin-sensitive (Em<sup>s</sup>) 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<sup>r</sup> Em<sup>s</sup>) by PCR amplification using chloramphenicol (*cat*) and erythromycin (*ery*) specific gene primers. This would have been followed by confirming the correct integration

of the *lox66-P<sub>32</sub>-cat-lox71* cassette into the genome by PCR analysis using primers annealing to genomic sequences flanking the *lox66-P<sub>32</sub>-cat-lox71* cassette.

## 2.3 METHODS FOR RIBONUCLEIC ACID (RNA) WORK

### 2.3.1 ISOLATION OF RNA

*L. 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. 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 TRIzol® 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-free™ 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 *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<sub>2</sub>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.

<u>10x MOPS buffer:</u>	MOPS	0.4 M, pH 7.0
	Sodium acetate	0.1 M
	EDTA	0.01 M
	Store up to 3 months at 4°C	

<u>Sample loading buffer:</u>	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 <sup>32</sup>P-labelled probe*

The preparation of probes was performed following the instructions of the 'Random primers DNA labelling system' (Invitrogen, USA) using  $\alpha^{32}\text{P}$ -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  $\mu\text{L}$  sample of different dilutions of each probe was counted for 1 minute (cpm/ $\mu\text{L}$ ).

*(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/ $\mu$ L. 5  $\mu$ L of each standard and dilutions of the unknown sample were added to separate wells in a 96-well plate. 100  $\mu$ L 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.

<u>Bradford reagent (5x):</u>	Coomassie brilliant blue G-250	100 mg
	Ethanol (95 %)	50 mL
	Phosphoric acid (concentrated)	100 mL
	Pure H <sub>2</sub> O	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  $\pm 10$  %). 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:

$$E_{205}^{1\text{mg mL}^{-1}} = 27.0 + 120 \times \frac{A_{280}}{A_{205}}$$

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.**

Amount	Staining solution	Destaining solution
0.1 %	Coomassie blue R-250 (G-250)	-
40 %	Methanol	Methanol
10 %	Acetic acid	Acetic acid

*(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 H<sub>2</sub>O) 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 H<sub>2</sub>O. 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 H<sub>2</sub>O. 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 H<sub>2</sub>O mixture.

<u>Fixing solution:</u>	EtOH (40%)	100.0 mL
	Glacial acetic acid (10%)	25.0 mL
	Pure H <sub>2</sub> O	made up to 200.0 mL

<u>Sensitizing solution:</u>	EtOH (40%)	75.0 mL
	Sodium thiosulphate	0.5 g
	Sodium acetate	17.0 g
	Pure H <sub>2</sub> O	made up to 250.0 mL

<u>Silver nitrate solution:</u>	Silver nitrate	0.4 g
	Formaldehyde (37% w/v)	16.0 µL
	Pure H <sub>2</sub> O	made up to 40.0 mL

Developing solution (6% stock solution):

Sodium carbonate x 10 H <sub>2</sub> O	6.25 g
Formaldehyde (37% w/v)	100.0 µL
Pure H <sub>2</sub> O	made up to 250.0 mL

<u>Stopping solution:</u>	EDTA-Na <sub>2</sub> x 2 H <sub>2</sub> O	3.65 g
	Pure H <sub>2</sub> O	made up to 250.0 mL

#### 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.

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  $\mu$ L of urea rehydration solution with freshly added DTT and 50  $\mu$ L 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	<u>30 minutes</u>
		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  $\mu$ L of marker with 5  $\mu$ L 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:

	<b>Final concentration</b>	<b>Amount</b>
Urea (FW 60.06)	8 M	12 g
CHAPS	4% (w/v)	1.0 g
IPG Buffer§	2% (v/v)	500 $\mu$ L
DTT (FW 154.2)	40 mM	154 mg
Double-distilled water	—	to 25 mL (16 mL required)

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

Rehydration stock solution:

	<b>Final concentration</b>	<b>Amount</b>
Urea (FW 60.06)	8 M	12 g
CHAPS	2% (w/v)	0.5 g
IPG Buffer (same range as the IPG strip)	0.5 % (v/v)	125 $\mu$ L
1% Bromophenol blue stock solution	0.002%	50 $\mu$ L
Double-distilled water	—	to 25 mL (16 mL required)

\* 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.

SDS equilibration buffer:

	<b>Final concentration</b>	<b>Amount</b>
Urea (FW 60.06)	6 M	72.1 g
Tris-HCl, pH 8.8	75 mM	10.0 mL
Glycerol (87% w/w)	29.3% (v/v)	69 mL (84.2 g)
SDS (FW 288.38)	2% (w/v)	4.0 g
1% Bromophenol blue stock solution	0.002%	400 $\mu$ L
Double-distilled water	—	to 200 mL

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

1 % bromophenol blue stock solution:

	<b>Final concentration</b>	<b>Amount</b>
Bromophenol blue	1%	100 mg
Tris-base	50 mM	60 mg
Double-distilled water	—	to 10 mL

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

	<b>Final concentration</b>	<b>Amount</b>
Laemmli SDS electrophoresis buffer		100 mL
Agarose	0.5%	0.5 g
1% Bromophenol blue stock solution	0.002% (w/v)	200 $\mu$ L

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

5 % Coomassie Blue G-250 stock:

(modified from Neuhoff *et al.* 1988)

0.5 g      Coomassie Blue G-250  
to 10.0 mL    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 g      Ammonium sulphate (FW 132.1)  
6 mL    Phosphoric acid 85% (w/w)  
10 mL    5 % Coomassie Blue G-250 stock  
to 500 mL    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 mL    Colloidal Coomassie Blue G-250 dye stock solution  
100 mL    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  $\text{NH}_4\text{HCO}_3$  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  $-20^\circ\text{C}$  until required for further processing.



### 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 H<sub>2</sub>O 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.



Whatman 3 mm paper. Lastly, the pre-wetted membrane was placed on top of the wet filter paper and 5  $\mu$ L 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 peroxidase (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

<u>Blocking solution:</u>	PBS or TBS (PBST or TBST without Tween-20) 5 % (w/v) skim milk or 1 % (w/v) BSA 0.05 % Tween-20 (optional)
<u>Primary antibody solution:</u>	Blocking solution Antibody dilution 1:1,000 (anti-GFP antibody, Sigma)
<u>Secondary antibody solution:</u>	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 PlnKW30 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  $\mu$ l 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- $\beta$ -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<sub>600</sub> of ~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<sub>600</sub> of ~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<sub>600</sub> of ~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 *Nco*I and *Hind*III restriction sites of pNZ8148 and the RR and GTase genes were cloned into the *Nae*I and *Xba*I 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<sub>600</sub> 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 OD<sub>600</sub> of 0.6 and induced by addition of IPTG to a final concentration of 1.5 mM. After 6 hours growth to an OD<sub>600</sub> 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 %  $\text{NH}_4\text{HCO}_3$ , then packed into a 1.5 L glass column and eluted with 2 L of 2 %  $\text{NH}_4\text{HCO}_3$  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  $\text{C}_4$  300 Å column (10 x 250 mm, Phenomenex) using the following solutions to form a linear gradient: A:  $\text{H}_2\text{O}$ , 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  $\mu\text{L}$  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 OD<sub>595</sub> 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 2.4: Dilutions of PlnKW30 used determine its MIC.**

PlnKW30 samples were diluted in pure H<sub>2</sub>O.

Sample number	PlnKW30 dilutions used for Indicator plate	Sample number	PlnKW30 dilutions used for 96-well plate
1	0.1 mg/mL	1	1.0 µg/mL
2	0.05 mg/mL	2	0.1 µg/mL
3	0.025 mg/mL	3	50.0 ng/mL
4	0.0125 mg/mL	4	25.0 ng/mL
5	6.25 µg/mL	5	12.5 ng/mL
6	3.1 µg/mL	6	6.0 ng/mL
7	1.6 µg/mL	7	3.0 ng/mL
8	0.8 µg/mL	8	1.5 ng/mL
9	0.4 µg/mL	9	No culture (negative control)
10	0.2 µg/mL	10	No PlnKW30 (positive control)
11	0.1 µg/mL		
12	Water only (negative control)		

#### 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  $\mu$ g 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  $\mu$ L of 2 mg/mL activated chymotrypsin in 1 % ammonium bicarbonate at room temperature. Prior to the chymotryptic digest the  $\alpha$ -chymotrypsin had to be activated by addition of 20  $\mu$ L of 25 ng/ $\mu$ L 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  $\text{NH}_4\text{HCO}_3$  buffer and was incubated for 1 hour at 56 °C. After cooling the sample to room temperature, 55 mM iodoacetamide in 100 mM  $\text{NH}_4\text{HCO}_3$  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 5 $\mu$ m C<sub>18</sub> 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  $\beta$ -elimination under alkaline conditions (Greis *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- $\beta$ -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 OD<sub>600</sub> 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 CH<sub>3</sub>I (16.7 % per volume) for 48 hours in the dark. This chemical reaction was stopped by addition of 18 % NH<sub>4</sub>HCO<sub>3</sub> (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 H<sub>2</sub>O. Further purification of the CH<sub>3</sub>I-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<sub>600</sub> 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:	Hepes	50.0	mM
	MgCl <sub>2</sub>	0.1	mM
	EGTA	0.1	mM
	TCEP	5.0	mM
			pH 7.5

#### *(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  $\mu\text{L}$ : reaction buffer (5 x; 250 mM Hepes, pH 7.5; 50 mM  $\text{MgCl}_2$ )
  - 1  $\mu\text{L}$ : TCEP (150 mM)
  - 1 or 3  $\mu\text{L}$ :  $^3\text{H}$ -FPP (3.7  $\mu\text{M}$  in 50 % EtOH)
  - 1.5  $\mu\text{L}$ : peptide (GCVLS: specific substrate for yeast FTase, containing the typical CaaX motif; 33.3  $\mu\text{M}$ )
  - 2  $\mu\text{L}$ : purified recombinant yeast FTase (55  $\mu\text{g}/\text{mL}$ )
- or**
- 6  $\mu\text{L}$ : whole cell extract of yeast
- $\text{H}_2\text{O}$  to a total volume of 15  $\mu\text{L}$

The composition of the reaction buffer was changed to minimise the amount of ethanol added, although some ethanol derives from the  $^3\text{H}$ -FPP which is dissolved in 50 % EtOH and 50 % water. The controls included reaction mixes without peptide, without  $^3\text{H}$ -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/ $\text{NH}_4\text{OH}/\text{H}_2\text{O}$ ) as mobile phase and again dried overnight. The TLC plate was sprayed three times with EN<sup>3</sup>HANCE® 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.



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<sub>600</sub> 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<sub>2</sub> - C<sub>18</sub>; stationary phase) are covalently linked to the silica. The routine purification of PlnKW30 was performed using Jupiter 5µm C<sub>4</sub> 300 Å columns (250 x 10 mm, Phenomenex). Other peptide work was done with Jupiter 5µm C<sub>18</sub> 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).

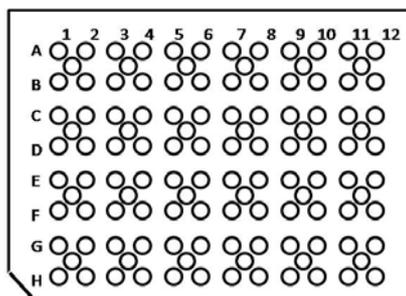
#### *(a) ESI-MS (Electrospray ionisation mass spectrometry)*

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  $\alpha$ -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.**

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  $3.90 \times 10^{-6}$  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*

ZipTip<sub>C18</sub> 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 C<sub>18</sub>, 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 H<sub>2</sub>O, 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 H<sub>2</sub>O the sample was eluted with 1 CV 80 % ACN, 0.1 % TFA, this being repeated several times, up to a total of 400  $\mu$ L. 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 KH<sub>2</sub>PO<sub>4</sub>, 10 % acetonitrile to a concentration of about 0.2 mg/mL. Approximately 300  $\mu$ L of each sample was pipetted into a Quartz SUPRASIL<sup>®</sup> 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 KH<sub>2</sub>PO<sub>4</sub>, 10 % acetonitrile, the baseline scans were carried out 10 x on 2 mM KH<sub>2</sub>PO<sub>4</sub>, 10 % acetonitrile without sample added, then averaged and subtracted from the sample scans. The reduced samples were resuspended in 2 mM KH<sub>2</sub>PO<sub>4</sub>, 10 % acetonitrile containing 10 mM TCEP and measured at 60 °C. 2 mM KH<sub>2</sub>PO<sub>4</sub>, 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**

Parameters	Condition
Path length	1 mm
Wavelength range	180 nm – 260 nm
Time per point	0.25 s
Bandwidth	1 nm
Step Size	1 nm
Repeats in set	20

**Table 2.6: Sample specific parameters for CD scans of native, reduced and deglycosylated PInKW30 and its N-terminal fragment.**

Parameter	PInKW30			N-terminal fragment		
	native	reduced	<i>O</i> -deglycosylated	native	reduced	<i>O</i> -deglycosylated
Molecular weight [Da]	5199.0561	5203.0561	4996.9827	3827.55043	3831.55043	3624.4726
Number of amino acids	43	43	43	32	32	32

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),
- Methyl iodide 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.

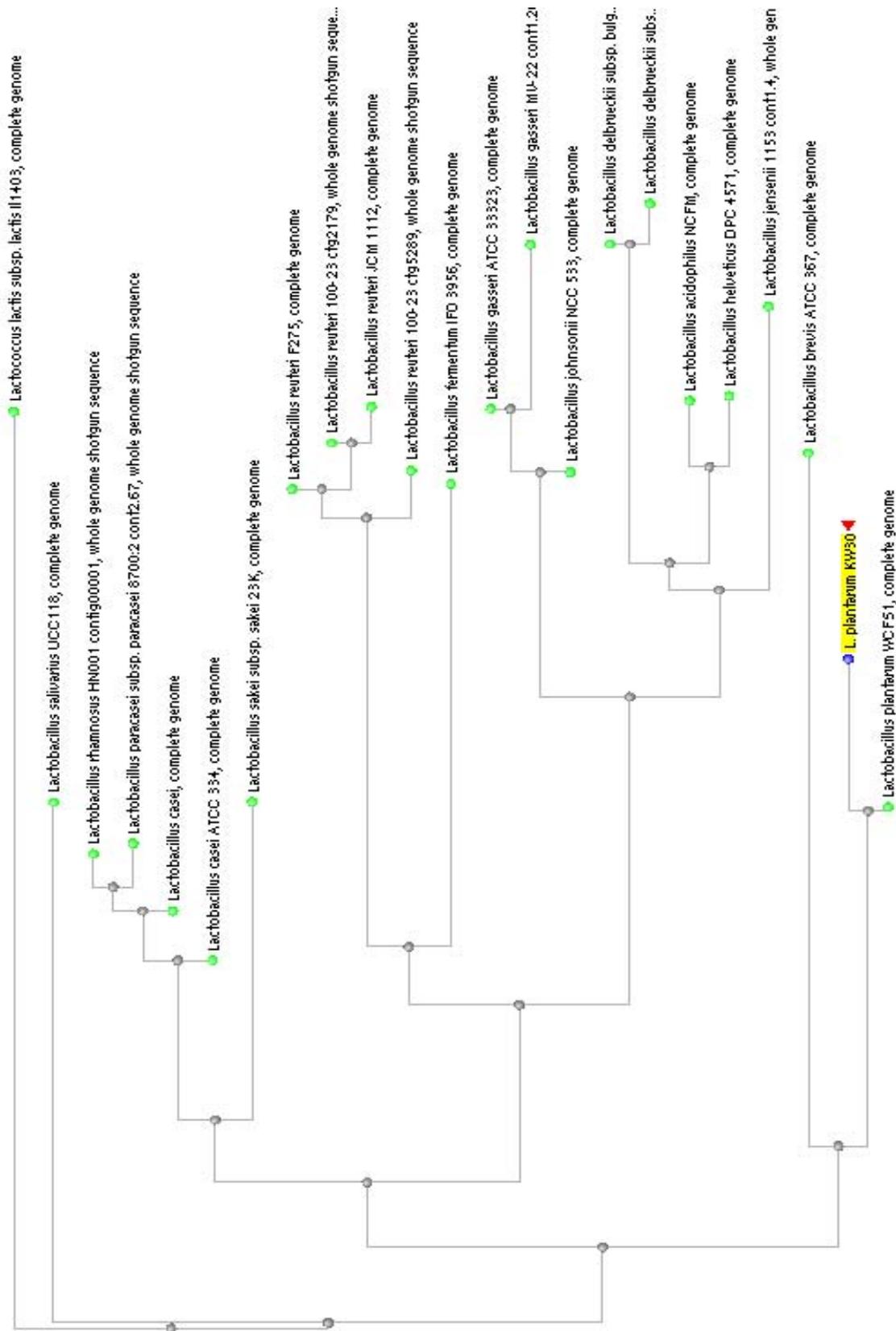
### 3.1 PHYLOGENETIC CLASSIFICATION OF *L. PLANTARUM* KW30

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 3.1: Results from nucleotide BLAST (NCBI) of 16S - 23S rDNA sequence of *L. plantarum* KW30.**

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

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* Il1403 was included in the genomic BLAST search as an outgroup for the alignment tree.



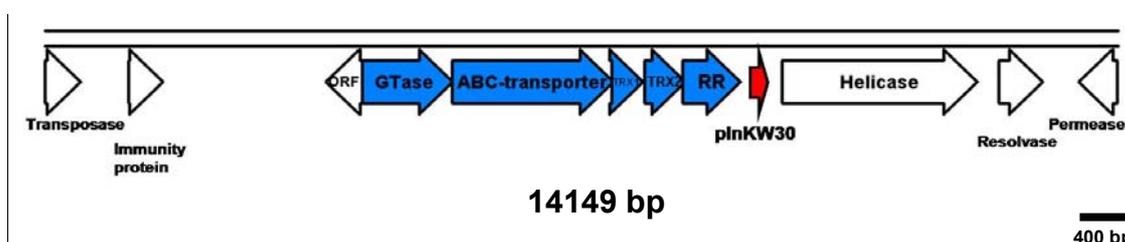
**Figure 3.1: Phylogenetic tree of *Lactobacillus* species.**  
 Classification of *L. plantarum* KW30 showing it to be closely related to *L. plantarum* WCFS1.

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*.**

ABC-transporter: ATP binding cassette transporter; GTase: glycosyltransferase; ORF: open reading frame; RR: response regulator; TRX: thioredoxin.

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 PlnKW30 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**

Gene name	ORF (bp)	Protein (aa)	Theoretical molecular weight (Da) <sup>a</sup>	Theoretical isoelectric point (pI) <sup>a</sup>	Amino acid sequence identity to the highest scoring orthologue from other species (%) <sup>b</sup>
<i>gccB</i>	1269	422	48776.65	7.65	32 (4e-29; Glycosyltransferase, GT2; <i>Enterococcus faecalis</i> V583; GI: 29377896)
<i>gccC</i>	2121	706	79086.62	9.29	29 (7e-61; ABC-type bacteriocin/ lantibiotic transporter; <i>Lactococcus lactis</i> subsp. <i>cremoris</i> SK11)
<i>gccD</i>	420	139	15774.37	9.34	30 (7e-10; hypothetical protein; thioredoxin-family; <i>Enterococcus faecalis</i> V583, plasmid pTEF2)
<i>gccE</i>	534	177	20674.47	9.60	31 (8e-13; putative secreted thioredoxin, TRX-family; <i>Listeria welshimeri</i> serovar 6b strain SLCC5334)
<i>gccF</i>	777	258	30139.02	9.09	34 (1e-09; putative response regulator of the competence regulon ComE1; <i>Streptococcus bovis</i> )
<i>gccA</i>	195	64	-	-	Pre-PlnKW30
		mature: 43	mature: 4800.36 actual: 5199.0561	mature: 7.04	Bacteriocin PlnKW30; no identifiable similarities to any other proteins

<sup>a</sup> Theoretical molecular weight and isoelectric point (pI) was predicted using 'Compute pI/MW' (ExpASY).

<sup>b</sup> 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.

```

TCAGGTTATTTT TAGAATGTGACATAATAAGTGCCAACTACAGCCATCTTTTCCCTTTAATATCAGAAATT
      -35                -10                +1(TS)   RBS
CCTAAAAACTTTTCAGGCGATATAAAAAAGTAGTGTCAAATCTGCAGTGAAACTGTATTCAAATTTACAAG
→ABC      traditional start codon
GTGACAGCAAAAAATGAGAAATAATTAAACAAATTGACCAAAAATGATTGTGGACCAGCAGCAGTAGCAACAA

```

**Figure 3.3: Possible regulatory sites of *gccC* (putative ABC-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 promoter 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).

```

AAGAAGATATACAAAAAAGATTAGTGATTTAAATCCCGATGAAGACATTTTAGTTACCGAATAAGATAGTT
-35 -10 +1 (TS) RBS →plnKW30
GCTCTCCAATATGTTTGTGTTTTTGGTTTATAATTTTCTTATTAAAAGGAGGGTATTAGGAATGAGTAAATT
GGTTAAGACACTTACTATAAGTGAAATTTCTAAGGCTCAAAACAACGGTGGAAAACCTGCATGGTGTGG
TATACTTTAGCAATGTGTGGTGTCTGGTTATGATTCGGGAACCTGTGATTATATGTATTCGCATTGTTTTG
GTATAAAGCATCATAGTAGTGGTAGTAGCAGTTATCATTTAGTTTTGTGAATGTTTTAGACTTATTAA
stop (pln)
GTTACGTTGATAGAGTATTTATTCTTTGATTTTGTAATTTCCACAATAAAGAGTTATTTGGAAAATTTAT
GACGCAAATAAGGAAGACAATAACATAATAATTATTGTCTTCCTTTAAATTTCTCTAAGTGAATAATTAG
CAGTTAGCGCTATAAAAAGAGTGTCTTGATCTTGACTATTTAGTATTAAATTC AAGATAGAAAGGTACA
GGATACATTTATGGACTAACATTAATGTGAACAATGGCTTATCTTGATGCTGTTTTTAATAGAACCCAAA
→Helicase
ATCACAAGGAGATGTTATTTATG

```

**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

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- $\alpha$ -D-sugars to generate  $\beta$ -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.**

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

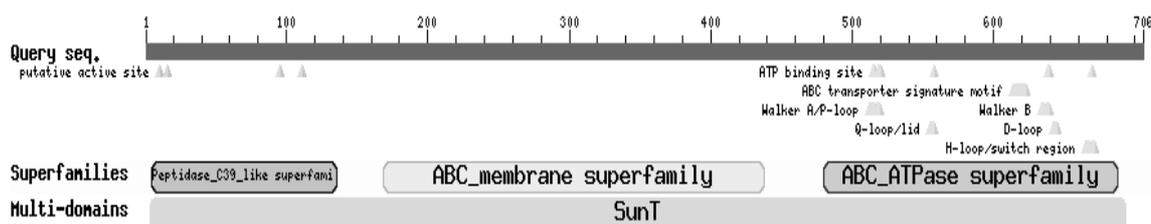
Conserved domains in GccC were predicted using cds BLAST, SMART, InterProScan and Pfam sequence searches and the predicted domains are listed in table 3.4.

**Table 3.4: Predicted conserved domains of GccC**

Prediction program	SunT*	Peptidase C39 domain*	ABC-membrane domain*	ABC transporter-like*	ATPase domain*
cds BLAST	4 - 693 (1e-72)	5 - 136 (cd02418; 2e-14)	169 - 438 (PF00664; 1e-12)	-	480 - 688 (cd03254; 3e-28)
SMART	-	8 - 141 (PF03412; 7.80e-10)	169 - 438 (PF00664; 7.30e-06)	-	504 - 687 (SM00382; 1.22e-11)
InterProScan	-	8 - 141 (PF03412; 6.9e-10)	169 - 438 (PF00664; 3.4e-07)	508 - 686 (PF00005; 4e-33)	504 - 687 (SM00382; 5.5e-11)
Pfam sequence search	-	10 - 136 (PS50990; 13.254)	8 - 141 (PF03412; 8.7e-10)	480 - 704 (PS50893; 12.945)	508 - 686 (PF00005; 1.9e-35)

\*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)*

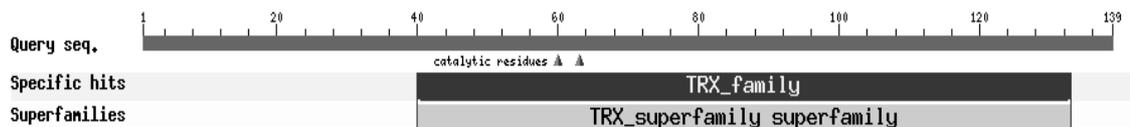
The prediction for GccD made by SignalP 3.0 is that it is a non-secreted protein, even although there is a probability of 0.465 (0.0 is low and 1.0 is high) that it contains a signal peptide. The most likely cleavage site for this signal peptide is between amino acids 21 and 22 (cleavage site probability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

Prediction program	TMH1	TMH2
TMHMM	5 - 22	
TopPred	3 - 23	93 - 113
SCAMPI-seq	4 - 24	
SCAMPI-msa	4 - 24	
PRODIV	3 - 23	
PRO	3 - 23	
OCTOPUS	2 - 22	
<b>TOPCONS</b>	<b>3 - 23</b>	

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.



**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.

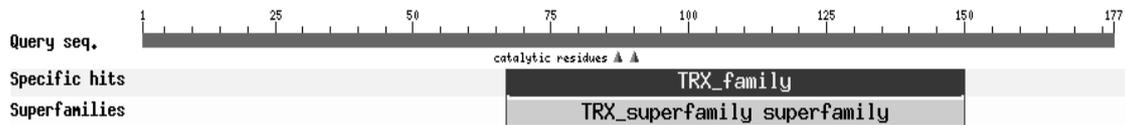
#### *(d) Analysis of the putative TRX2 (GccE)*

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).



**Figure 3.8: Schematic view of GccE with TRX and TRX-superfamily domains (from cds BLAST).**

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).

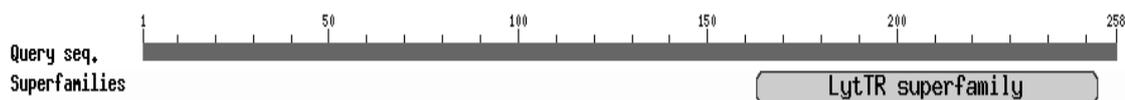
The best structural homologue as determined by GenTHREADER is a bacteriocin transport accessory protein from *Streptococcus pneumoniae* (pdb 1zma) which has a p-value of 0.003. The second best match is the structure of the thioredoxin mutant R82E from *Alicyclobacillus acidocaldarius* (pdb 1nw2; Bartolucci *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).**

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.

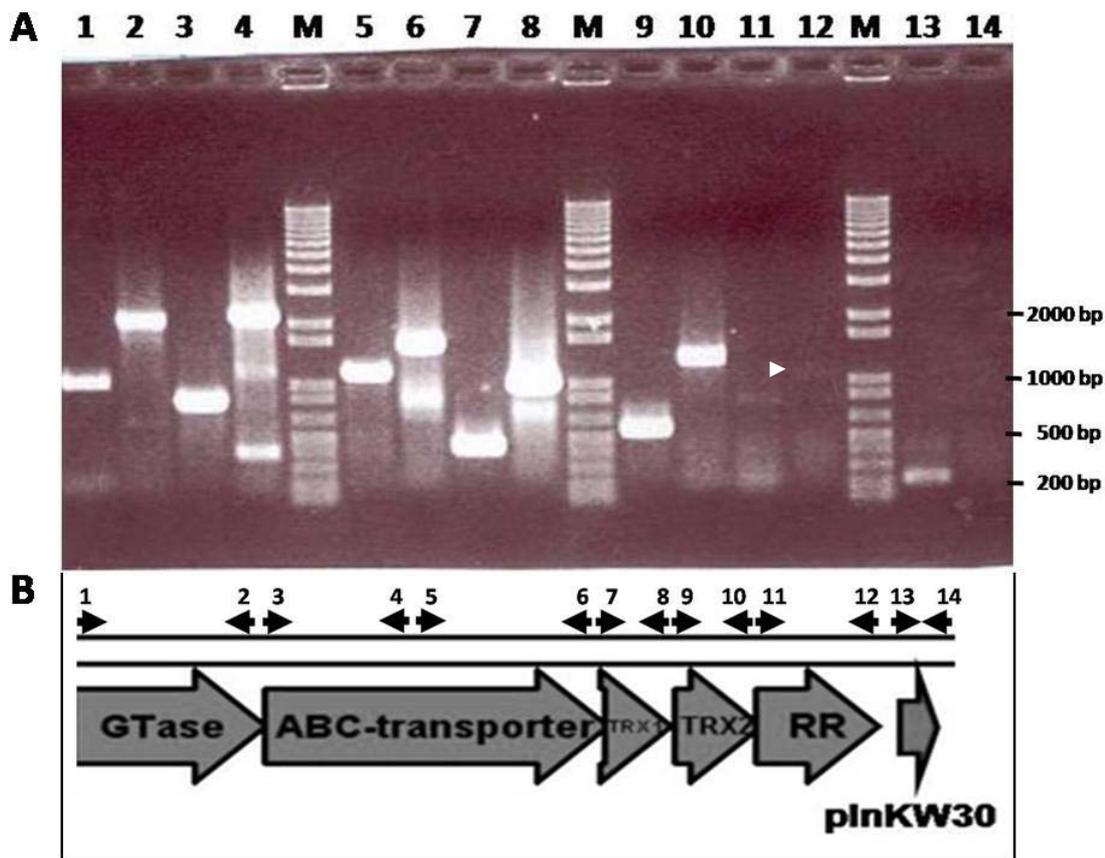
*GccA* contains the typical diglycine cleavage site for leader peptides in class II bacteriocins at positions 20 and 21. It does not contain any transmembrane helices or conserved domains, but has some structural similarity to a carbohydrate-binding wheat germ agglutinin in complex with *N*-acetylglucosamine, 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.**

**A:** 1: GTase; 2: GTase-ABC1; 3: ABC1; 4: ABC1-ABC2; 5: ABC2; 6: ABC2-TRX1; 7: TRX1; 8: TRX1-TRX2; 9: TRX2; 10: TRX2-RR; 11: RR; 12: RR-PlnKW30; 13: PlnKW30; 14: negative control using GTase primers; **M:** 1kb Plus DNA ladder. **B:** Positions of primers used in RT-PCR, for details of primers see table 3.6 and appendix 4.

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 *plnKW30* gene cluster.**

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

<sup>a</sup> Numbers refer to figure 3.10B and appendix 4

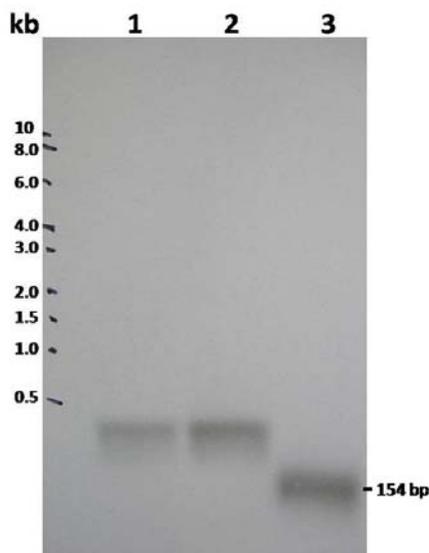
RT-PCR products were obtained for all intergenic regions, as shown in figure 3.10 and table 3.6, although the band for the RR and PlnKW30 was very faint (arrow in lane 12). The results show that possibly the whole *plnKW30* gene cluster is transcribed on one mRNA, although *gccA* (*plnKW30*) may be transcribed 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  $\mu\text{g}$  of *L. plantarum* KW30 total RNA, blotted and developed using *plnKW30*-specific  $^{32}\text{P}$ -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  $\mu\text{g}$  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).



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

**1:** 20  $\mu\text{g}$  of total RNA of *L. plantarum* KW30; **2:** 40  $\mu\text{g}$  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).

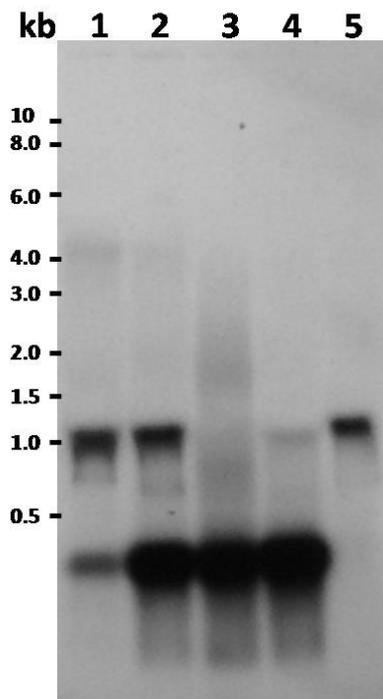
Efforts to establish the transcript size of the *plnKW30* gene cluster using probes for the response regulator and glycosyltransferase failed due to lack of specificity of the probes.

### 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* KW30 $\Delta$ *gcc*) 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* KW30Δ*gcc*.

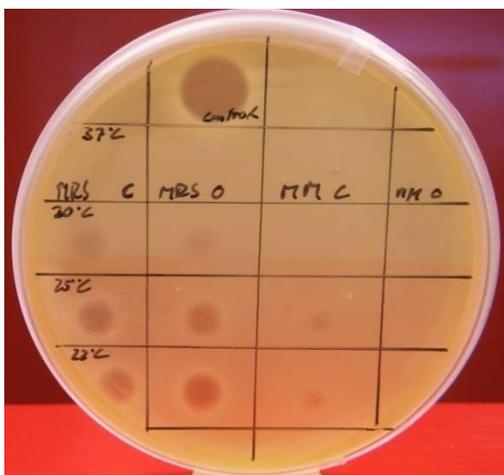
This would suggest that the intensity of the *gap* transcripts is not a good loading control and that the increase in *plnKW30* transcript is real. Consequently, the production of *plnKW30* mRNA increases with time until it peaks at the end of log phase.

### 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  $\mu$ L of each supernatant was spotted on an agar plate containing the indicator strain (Section 2.4.9). 2  $\mu$ L 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.**

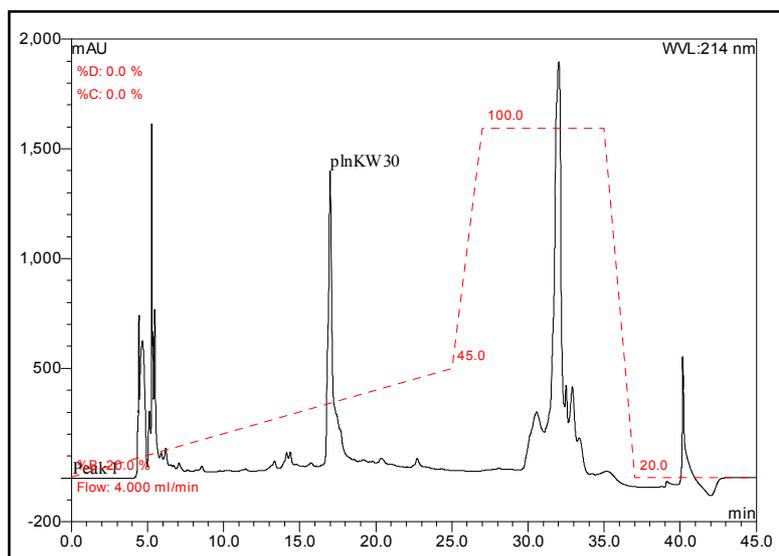
*L. plantarum* KW30 was grown in MRS or chemically defined medium (MM) at 22, 25, 30 and 37 °C in closed (C) or open (O) tubes. 2  $\mu$ L samples were spotted on indicator plate and incubated overnight at 25 °C. Control: purified PlnKW30.

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).

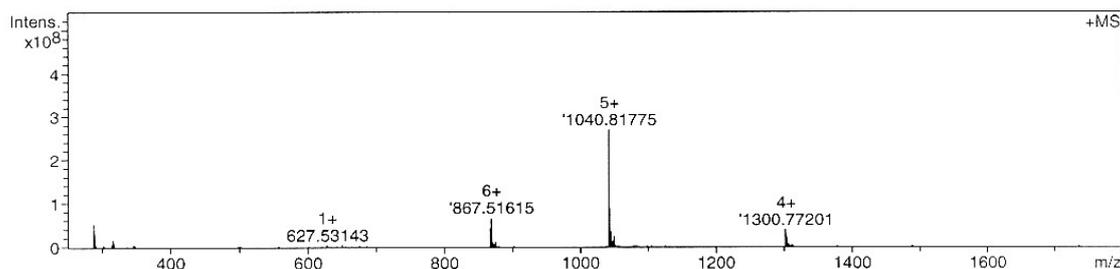


**Figure 3.14: RP-HPLC chromatogram of PlnKW30.**

A Jupiter C14 column was used with the mobile phase A: H<sub>2</sub>O, 0.1 % TFA and B: Acetonitrile, 0.08 % TFA, conditions used as shown in chromatogram. PlnKW30 was eluted at 17 minutes between 36 - 37 % B.

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.

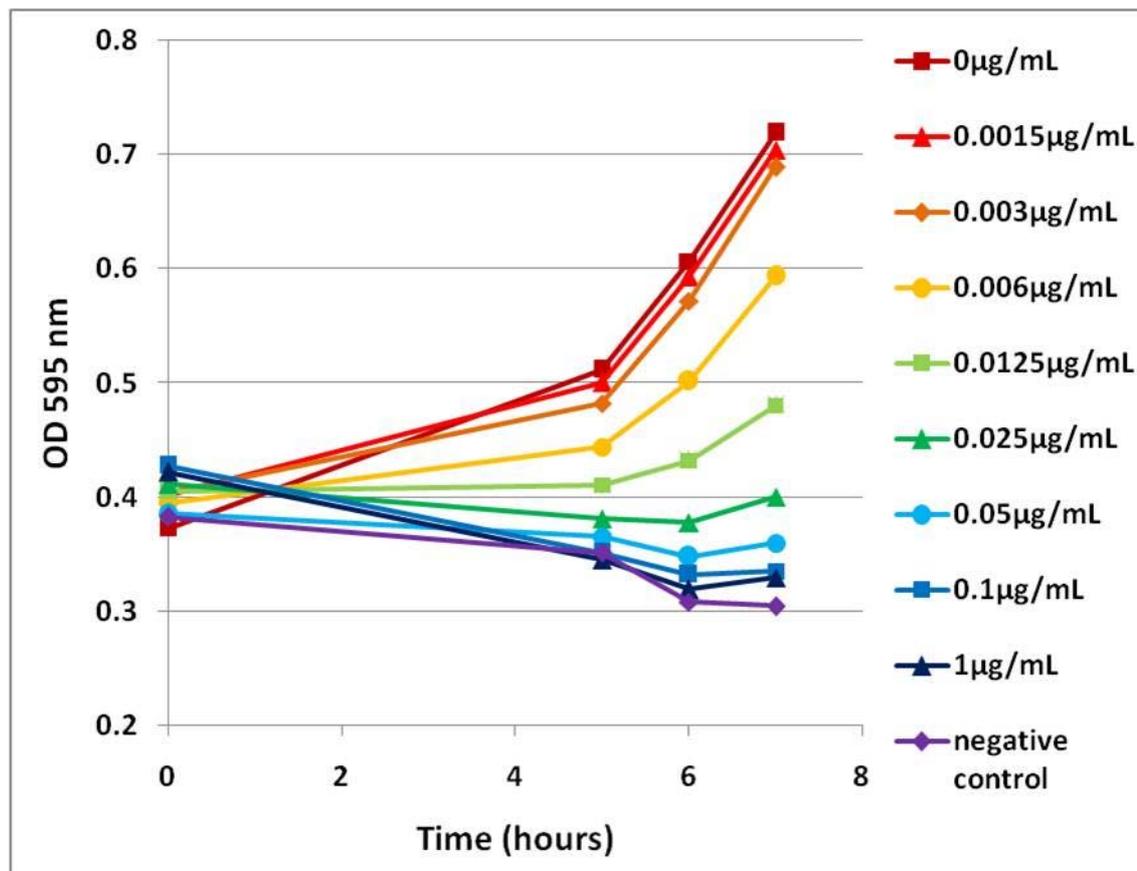


**Figure 3.15: Monoisotopic mass spectrometry result of the native PlnKW30.**

The bacteriocin PlnKW30 ( $m + H$ ) has a molecular mass of 5200.06050 Da (ESI, positive mode, Bruker Apex-Q-FTMS (9.4T Dual Source) Laboratory of Molecular Structure Characterization, Academy of Sciences of the Czech Republic).

### 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  $\mu$ L fresh MRS media and different concentrations of PlnKW30 were added. The cultures were grown at 30  $^{\circ}$ 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 ( $\sim$ 2.5 nM). A minimum inhibitory concentration in the nanomolar range makes PlnKW30 comparable to the well known potent bacteriocin nisin.

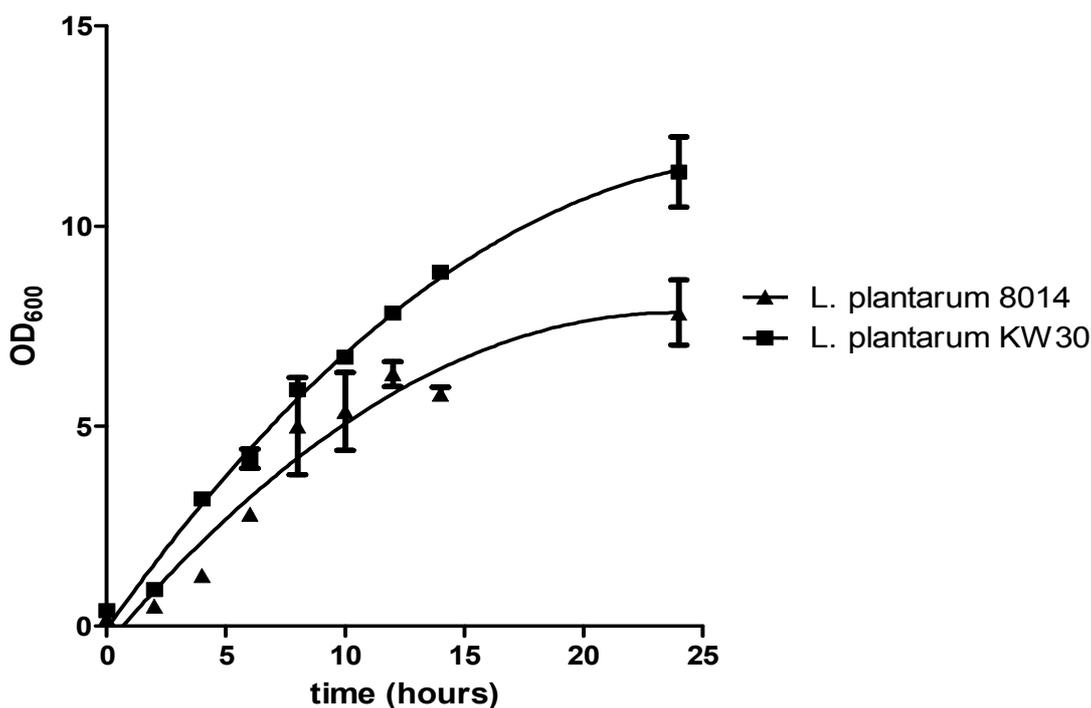


**Figure 3.16: Minimum inhibitory concentration assay for PlnKW30.**

Different concentrations of PlnKW30 were added to 200 µL of the indicator strain *L. plantarum* ATCC 8014 at about OD<sub>595</sub> 0.4. Cell growth was then monitored over 7 hours and the results plotted as shown. Negative control: MRS broth without indicator strain.

#### 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.

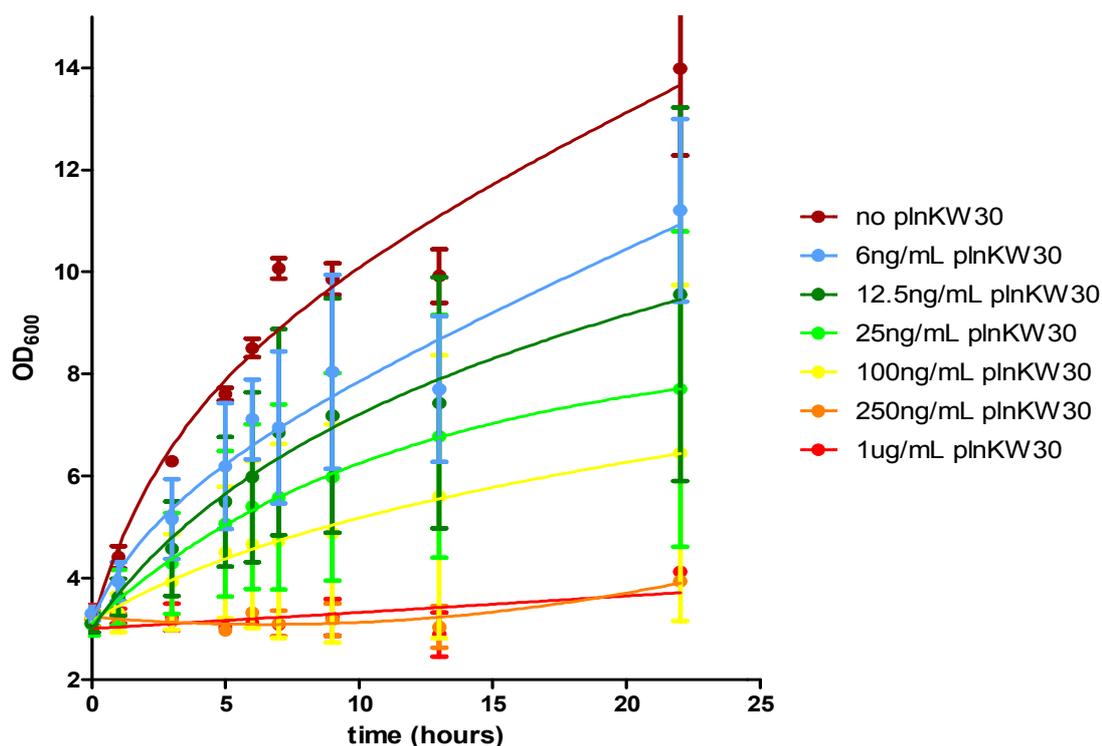


**Figure 3.17: Growth curve of *L. plantarum* KW30 (producer strain) and *L. plantarum* ATCC 8014 (indicator strain).**

Each data point represents the average of measurements from three separate cultures with standard errors shown.

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<sub>600</sub> of around 3.0) of the indicator strain *L. plantarum* ATCC 8014, whereas the MIC experiments were performed at OD<sub>600</sub> 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).



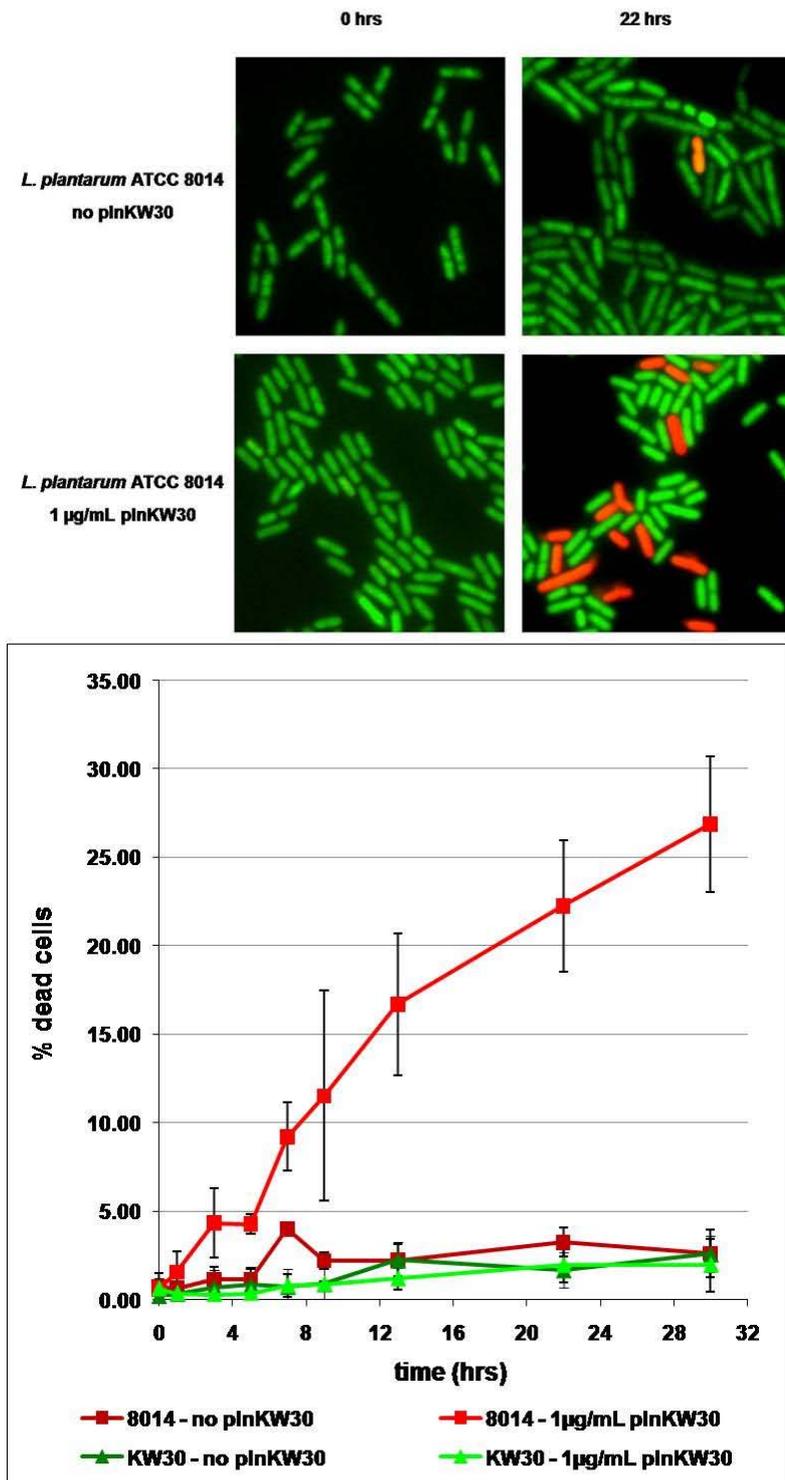
**Figure 3.18: Growth of *L. plantarum* ATCC 8014 after addition of different concentrations of PlnKW30.** PlnKW30 was added to mid-log phase *L. plantarum* ATCC 8014 cells. Each data point represents the mean of measurements from at least two separate cultures with standard errors shown.

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  $\mu\text{g}/\text{mL}$  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  $\text{ng}/\text{mL}$  and 1  $\mu\text{g}/\text{mL}$  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 OD<sub>600</sub> ~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.

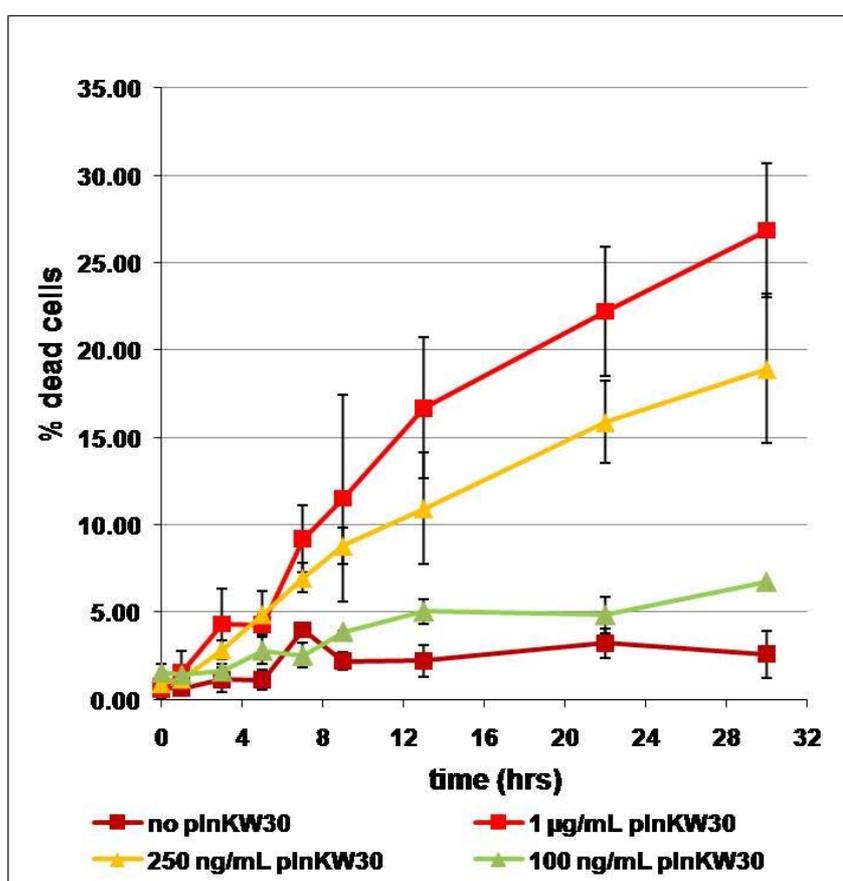


**Figure 3.19: Antimicrobial effect of PlnKW30 on *L. plantarum* ATCC 8014 cells.**

**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 ( $OD_{600} \sim 3.0$ ). The results showed that increased PlnKW30 concentration led to an increased percentage of dead *L. plantarum* ATCC 8014 cells (Figure 3.20).



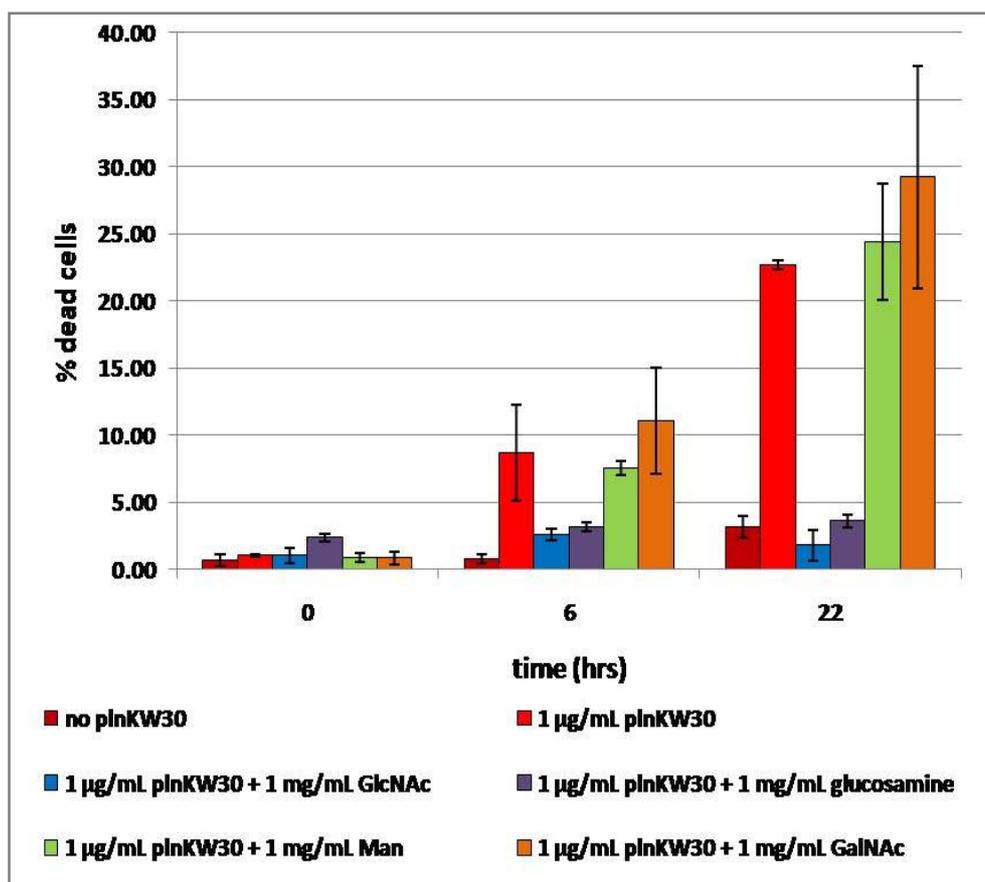
**Figure 3.20: Concentration-dependent antimicrobial activity of PlnKW30 against *L. plantarum* ATCC 8014.**

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).



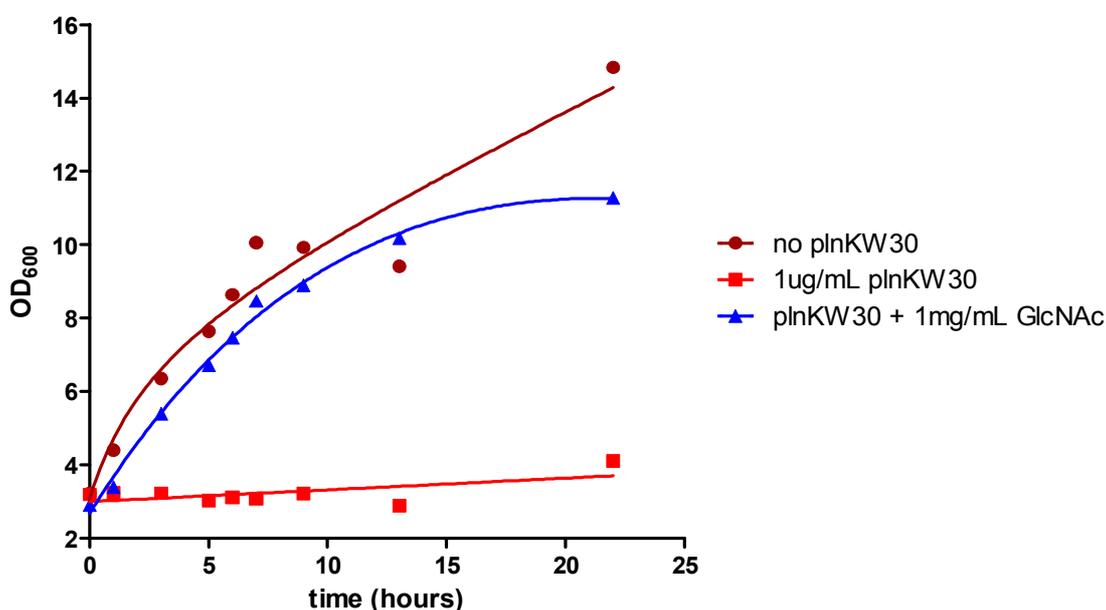
**Figure 3.21: Effect of different sugars on antimicrobial activity of PlnKW30.**

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 PlnKW30 (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 1 µg/mL PlnKW30 and 1 mg/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.

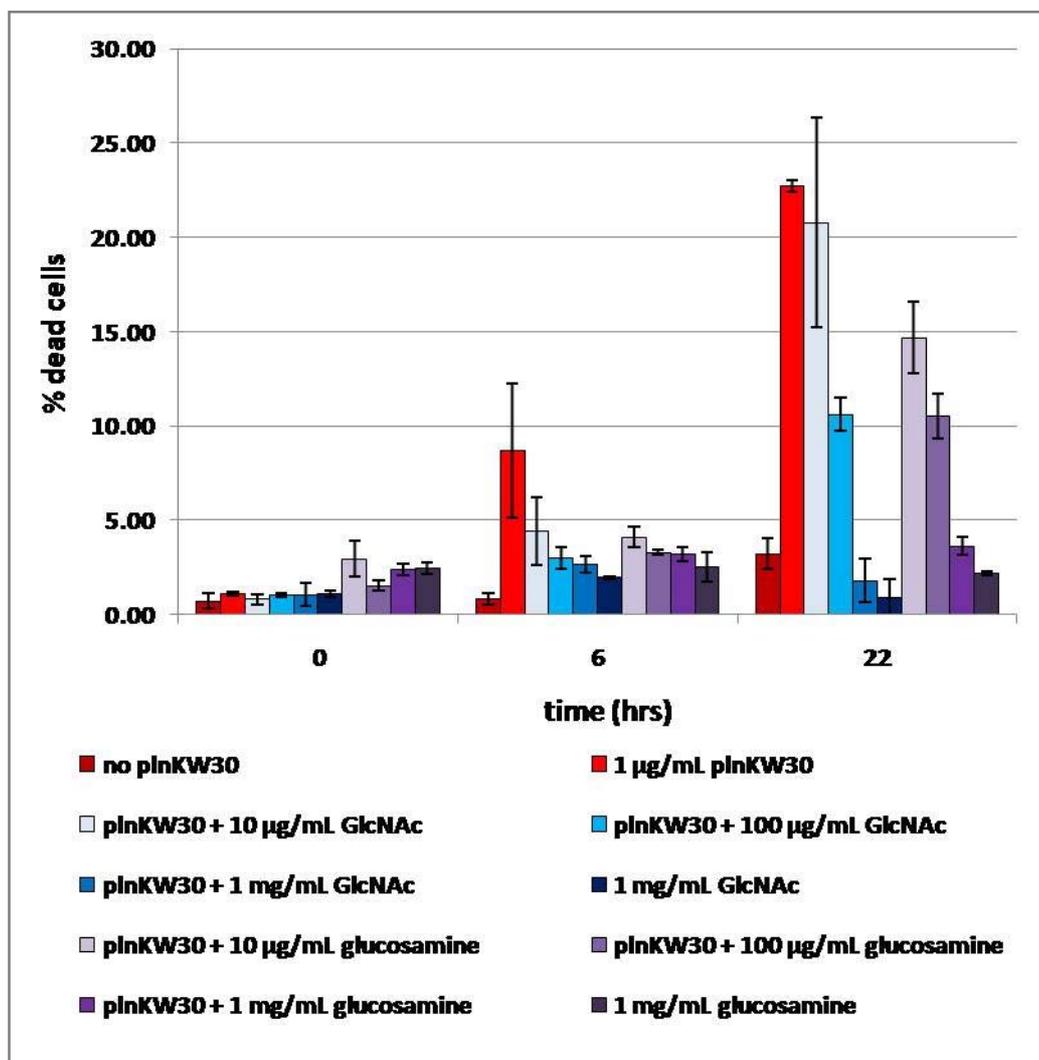
This protective effect of GlcNAc was also seen in the cell density test, where the addition of 1 mg/mL GlcNAc resulted in almost complete protection of the cells from the effects of PlnKW30 (Figure 3.22). The controls in this figure are the same as shown in figure 3.18.



**Figure 3.22: Growth of *L. plantarum* ATCC 8014 after addition of 1 mg/mL *N*-acetylglucosamine and PlnKW30.**

PlnKW30 was added to mid-log phase *L. plantarum* ATCC 8014 cells. Each data point represents the average of two measurements of separate cultures.

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.**

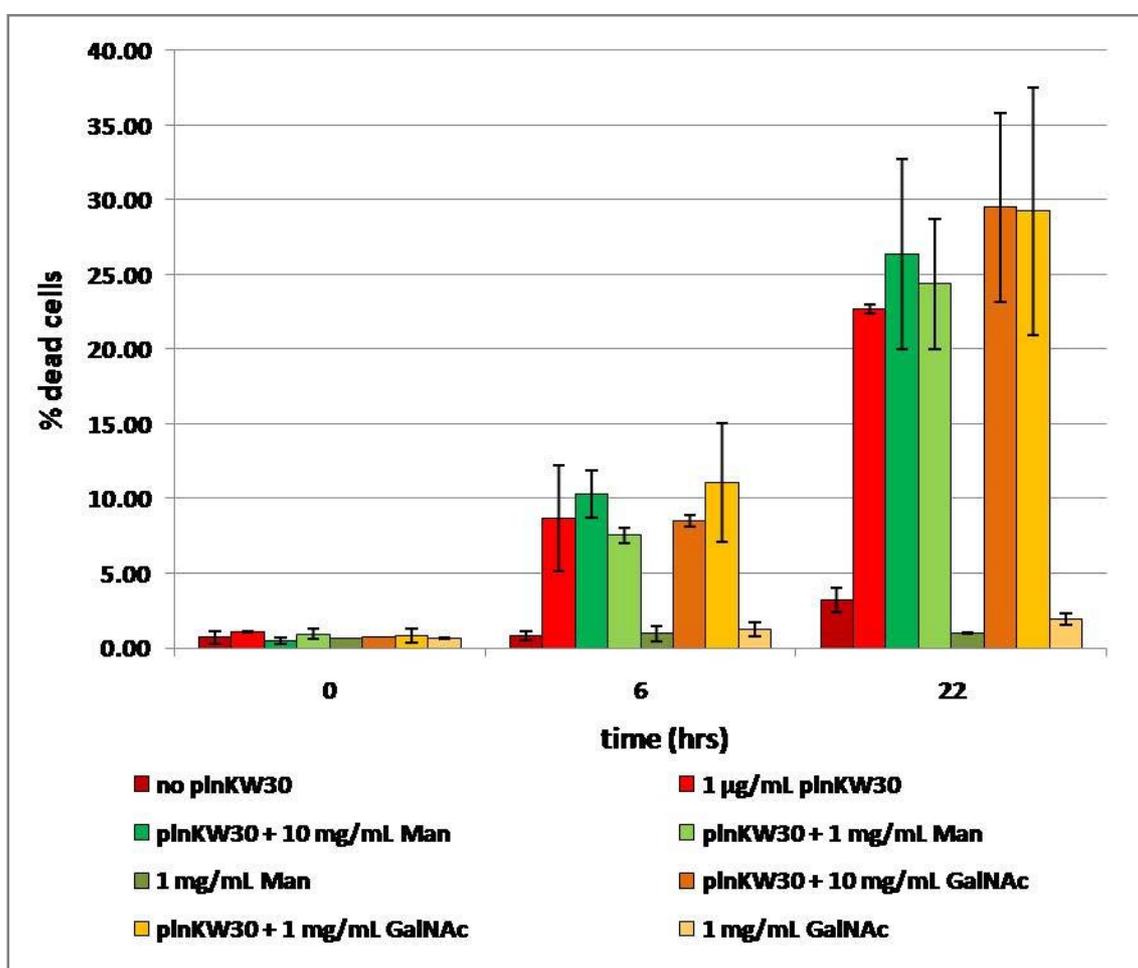
The percentage of dead cells of *L. plantarum* ATCC 8014 at times 0, 6 and 22 hours in the presence of 1 µg/mL PlnKW30 and different concentrations of N-acetylglucosamine or glucosamine is shown.

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.**

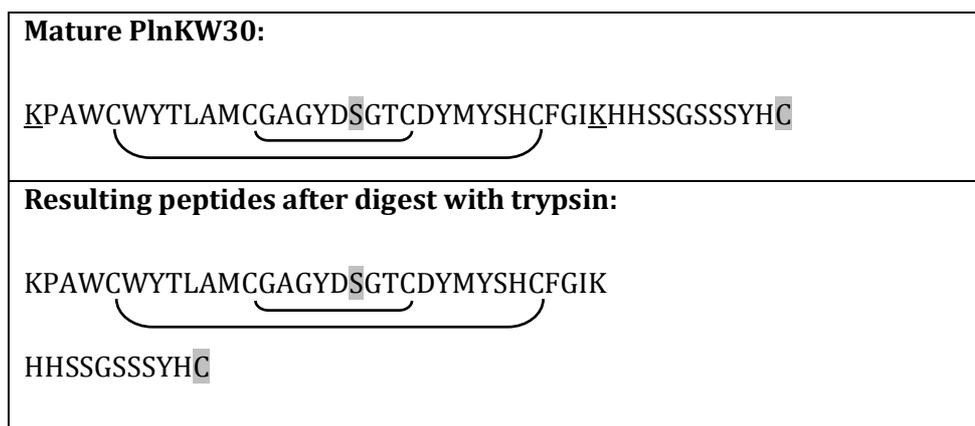
The percentage of dead cells of *L. plantarum* ATCC 8014 at time 0, 6 hours and 22 hours in the presence of 1 µg/mL PlnKW30 and different concentrations of mannose and N-acetylgalactosamine is shown.

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.

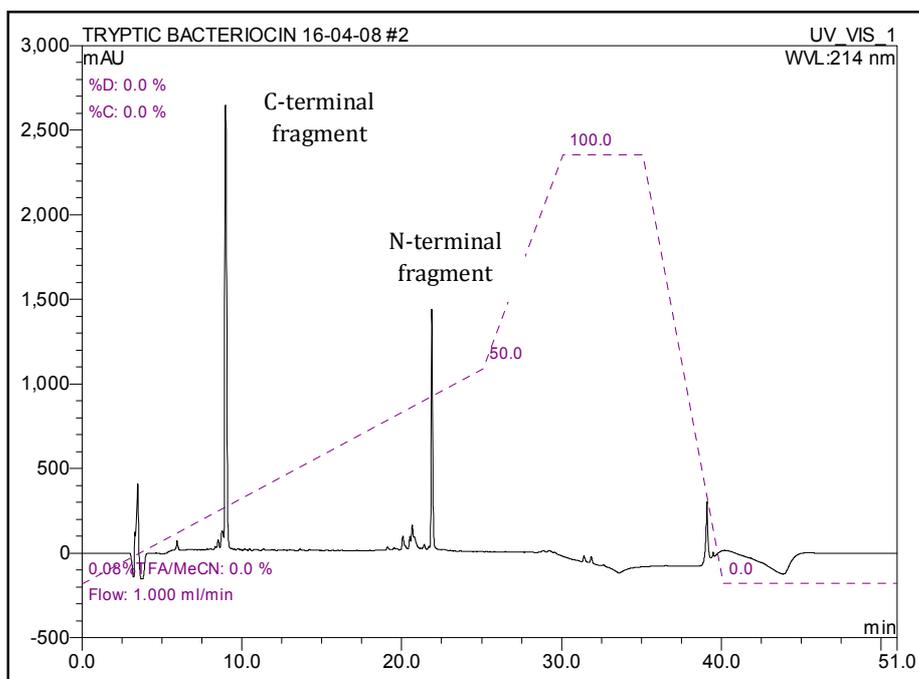


**Figure 3.25: Mature PlnKW30 sequence with trypsin cleavage sites and resulting peptides.**

Trypsin cleavage sites are underlined; post-translationally modified Ser 18 and Cys 43 are shadowed and the two nested disulfide bonds are indicated by brackets.

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 (KPAWCWYTLAMCGAGYDSGTCDYMYSHCFGIK) 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).

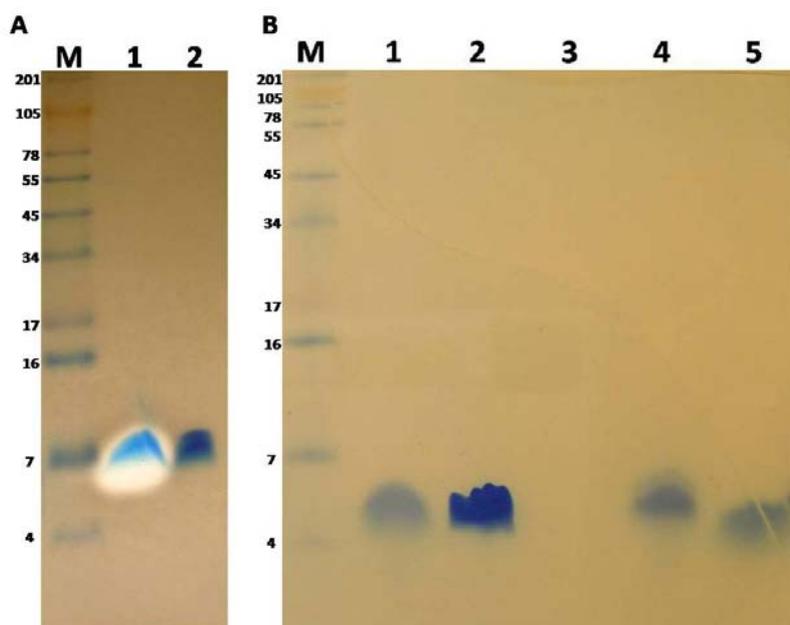
The C-terminal peptide is very hydrophilic and elutes earlier than the relatively hydrophobic N-terminal peptide. In comparison, the full length PlnKW30 elutes on a C18 column between both fragments at 40 % B, which is to be expected as it consists of both peptides.



**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: H<sub>2</sub>O, 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 (FGIKHHSSGSSSYHC 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 OD<sub>600</sub> of 3.

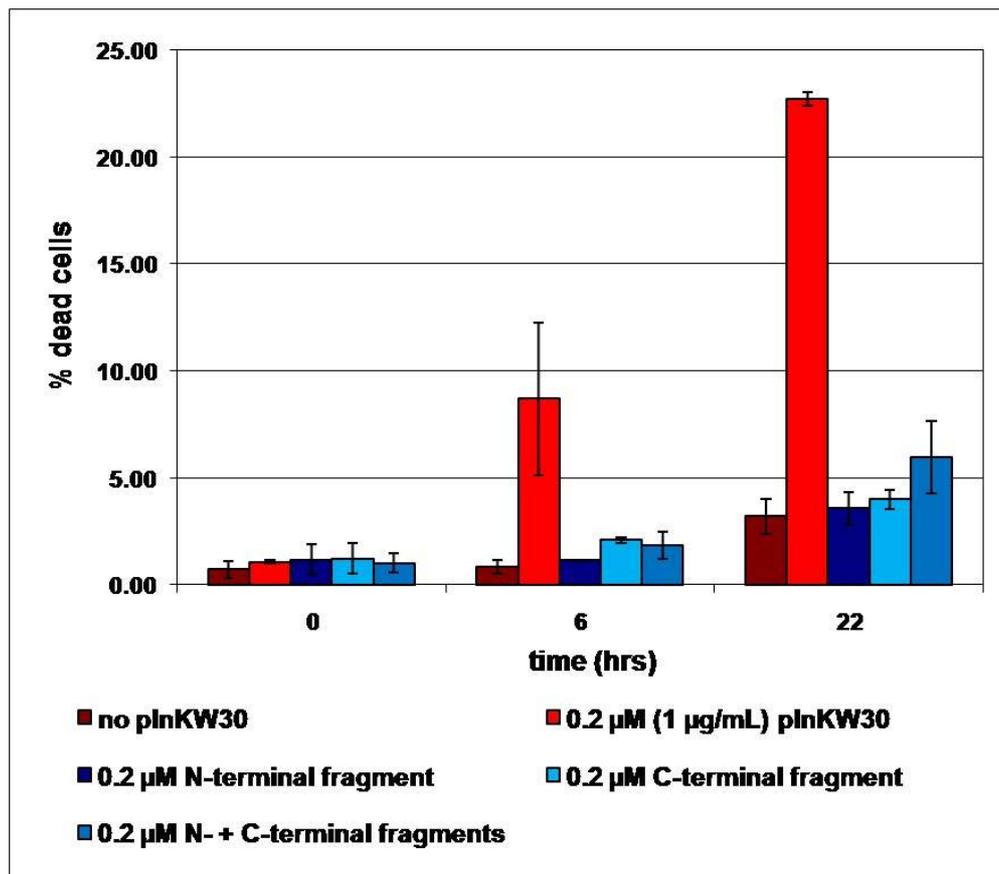
**A:** 1: PlnKW30 in non-reducing buffer; 2: PlnKW30 in reducing buffer;

**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 ( $2 \times 1633$  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  $\beta$  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 ( $0.2 \mu\text{M}$ ), as well as together to the indicator strain *L. plantarum* ATCC 8014 (Figure 3.28).

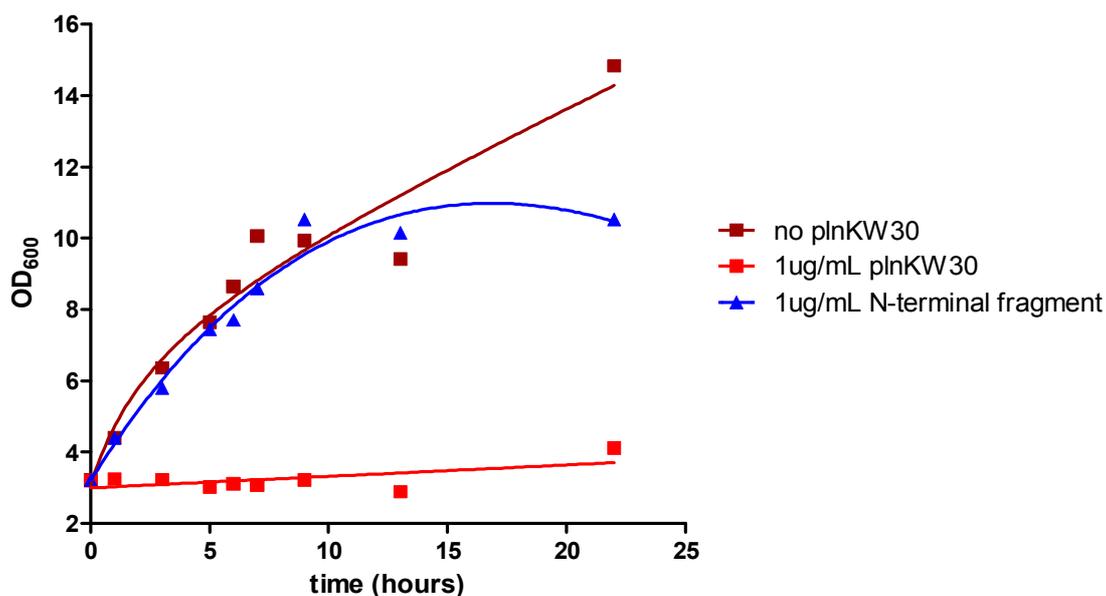


**Figure 3.28: Antimicrobial effect of N- & C-terminal fragments of PlnKW30.**

Percentage of dead cells of *L. plantarum* ATCC 8014 at time 0, 6 and 22 hours with addition of  $\sim 0.2 \mu\text{M}$  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).

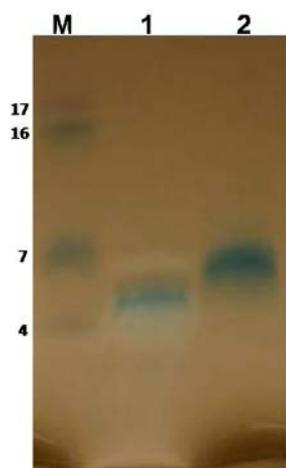
The effect of the N-terminal fragment on the growth of the indicator strain was also measured by OD (Figure 3.29). The controls in this figure are the same as shown in figure 3.18.



**Figure 3.29: Growth of *L. plantarum* ATCC 8014 after addition of the N-terminal fragment of PlnKW30.** PlnKW30 or N-terminal fragment was added to mid-log phase *L. plantarum* ATCC 8014 cells. The curves represent the mean of two independent measurements.

Using this method it was confirmed that the N-terminal fragment of PlnKW30 does not have an antimicrobial effect on *L. plantarum* ATCC 8014, as the OD<sub>600</sub> of this culture reached a similar density to that of a culture to which no PlnKW30 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<sub>600</sub> of about 0.3 (Figure 3.30). This antimicrobial activity is less effective than that of the full length PlnKW30, 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.**

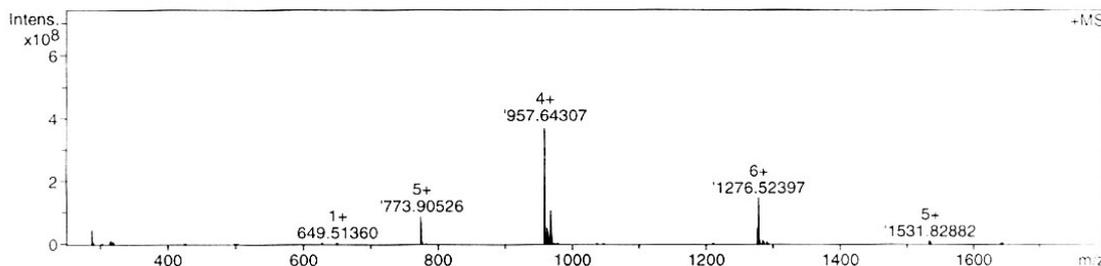
Samples of the N-terminal fragment were prepared in reducing and non-reducing sample buffer and separated on tricine SDS-PAGE (16 %, 3 % bisacrylamide) and then overlaid with the indicator strain *L. plantarum* ATCC 8014 OD<sub>600</sub> of 0.3. **1:** N-terminal fragment in non-reducing buffer; **2:**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

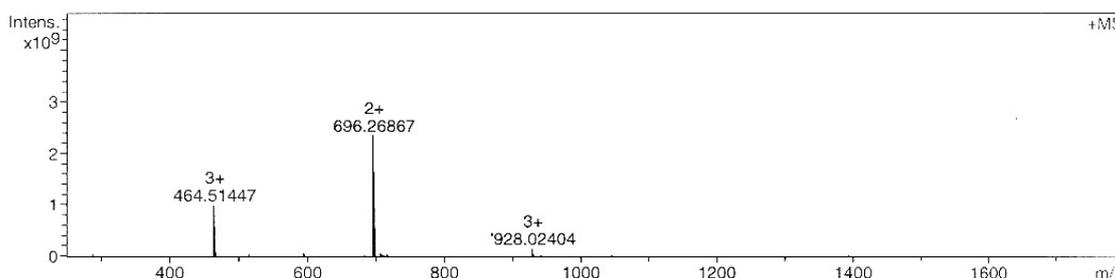
#### (a) Mass spectrometry of N- and C-terminal fragments

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.



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

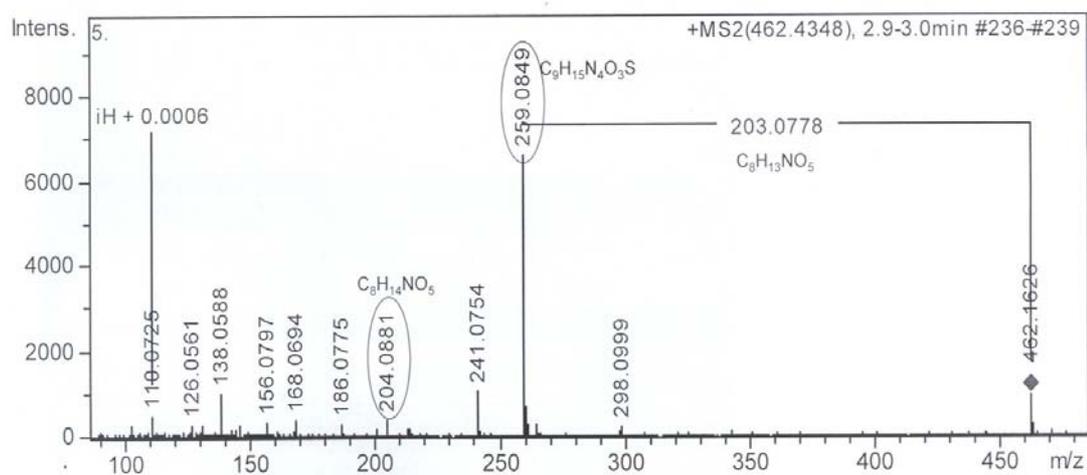
The molecular mass of the N-terminal fragment ( $m + H$ ) of PlnKW30 is 3827.55043 Da (ESI, positive mode, Bruker Apex-Q-FTMS (9.4T Dual Source), Laboratory of Molecular Structure Characterization, Academy of Sciences of the Czech Republic).



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

The molecular mass of the C-terminal fragment ( $m + H$ ) of PlnKW30 is 1391.52972 Da (ESI, positive mode, Bruker Apex-Q-FTMS (9.4T Dual Source), Laboratory of Molecular Structure Characterization, Academy of Sciences of the Czech Republic).

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 micrOTOF<sub>Q</sub> (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<sub>15</sub>H<sub>23</sub>.

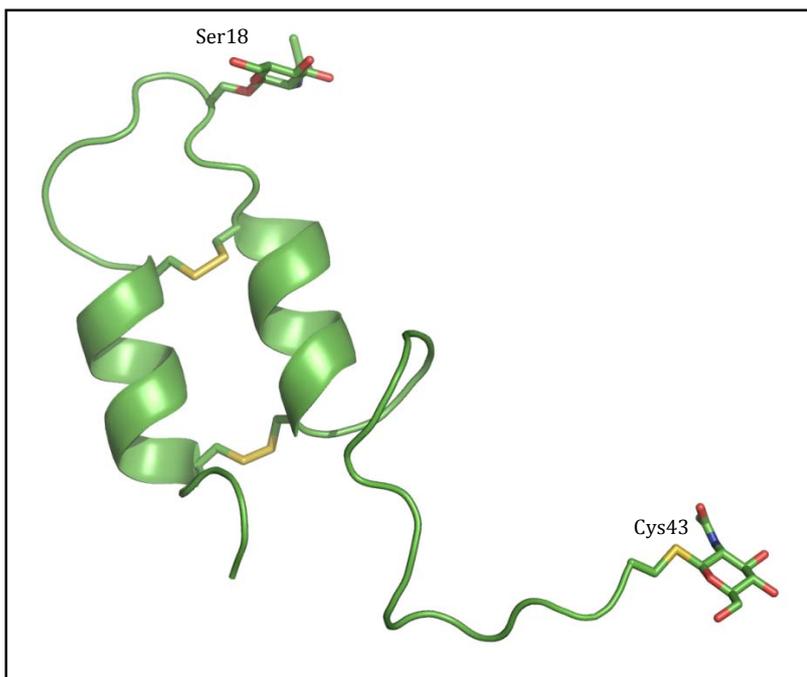


**Figure 3.33: Monoisotopic mass spectrometry result for the C-terminal HC-X fragment of PlnKW30** (Sample preparation: Trevor Loo, Massey University. Mass spectrometry analysis on Bruker ESI micrOTOFQ; Matthias Pelzig, Bruker, Australia).

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  $^1H$ -NMR spectroscopy (Figure 3.34; H. Venugopal – personal communication).



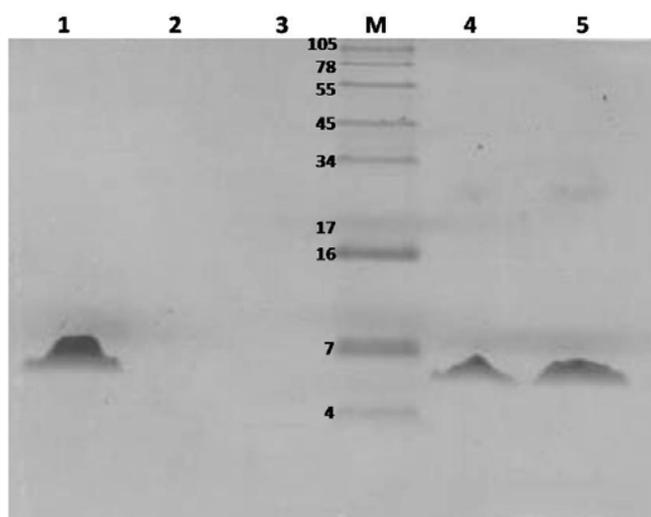
**Figure 3.34: Cartoon representation of the NMR structure of PlnKW30.**

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  $\beta$ 1-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  $\alpha$ - and  $3_{10}$ -helices. The secondary structural characteristics of this helix seem to be shifting dynamically between those of an  $\alpha$ -helix and a  $3_{10}$ -helix. A  $3_{10}$ -helix contains three amino acids within a  $360^\circ$  turn, stabilised by hydrogen bonds between residues  $i$  and  $i + 3$ , in contrast to the more usual  $\alpha$ -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  $\alpha$ -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  $\beta$ -*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  $\mu$ L 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.**

**1:** 20  $\mu$ g purified PlnKW30 control; **2:** supernatant of pull-down sample, 10  $\mu$ L; **3:** supernatant of pull-down sample, 20  $\mu$ L; **M:** protein standard (kDa); **4:** incubated beads, 10  $\mu$ L; **5:** incubated beads, 20  $\mu$ L.

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).

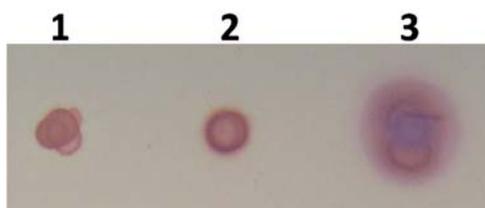


**Figure 3.36: Dot blot using anti-*O*-GlcNAc antibodies to detect PlnKW30 and its N- & C-terminal tryptic fragments.**

**1:** 4  $\mu$ g 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.**

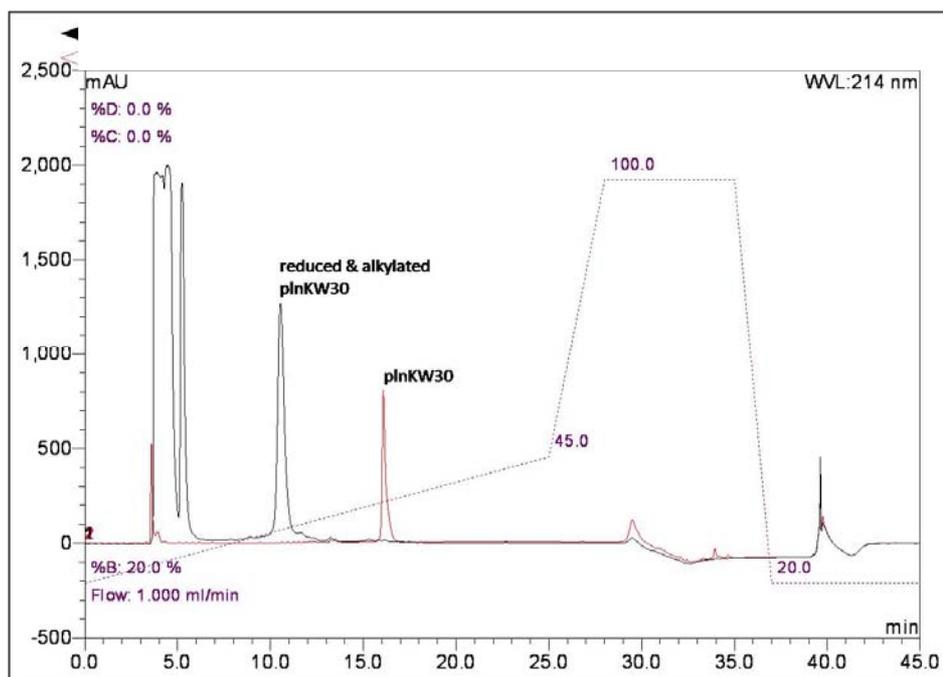
The peptide samples were spotted on TLC plates and sprayed with ninhydrin which reacts with peptides resulting in a purple colour. **1:** N-terminal fragment; **2:** C-terminal fragment; **3:** synthetic C15 C-terminal fragment.

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 H<sub>2</sub>O 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).

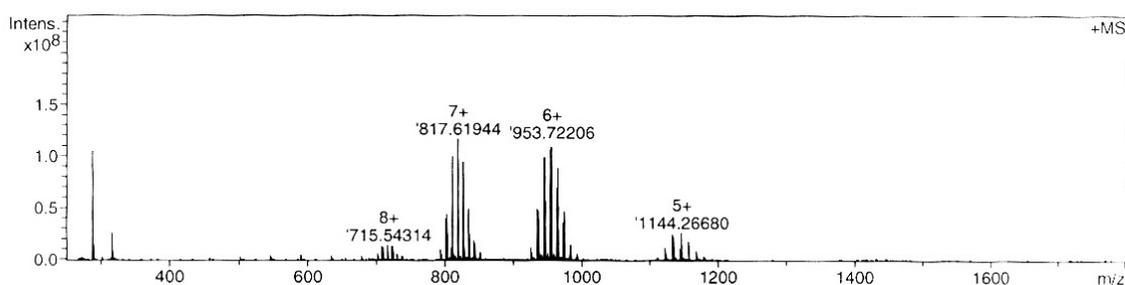


**Figure 3.38: RP-HPLC chromatogram of native and reduced/alkylated PlnKW30.**

A Jupiter C4 column was used with the solutions A: H<sub>2</sub>O, 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).

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.

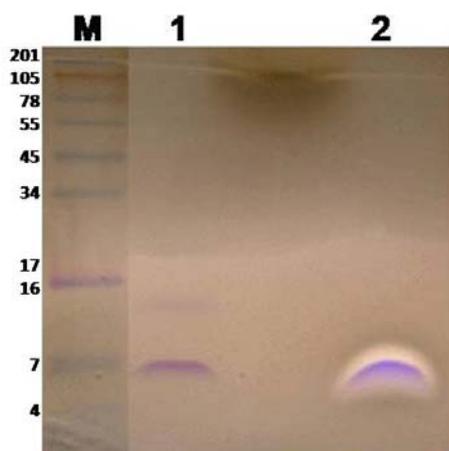
Mass spectrometry results revealed the addition of seven ethylacetate groups (7 x 57 Da). As iodacetamide is not specific for cysteines, and under some conditions acetylates serine residues, it is most likely that in addition to the four cysteines, three of the five serines in the C-terminal tail are 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.**

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).



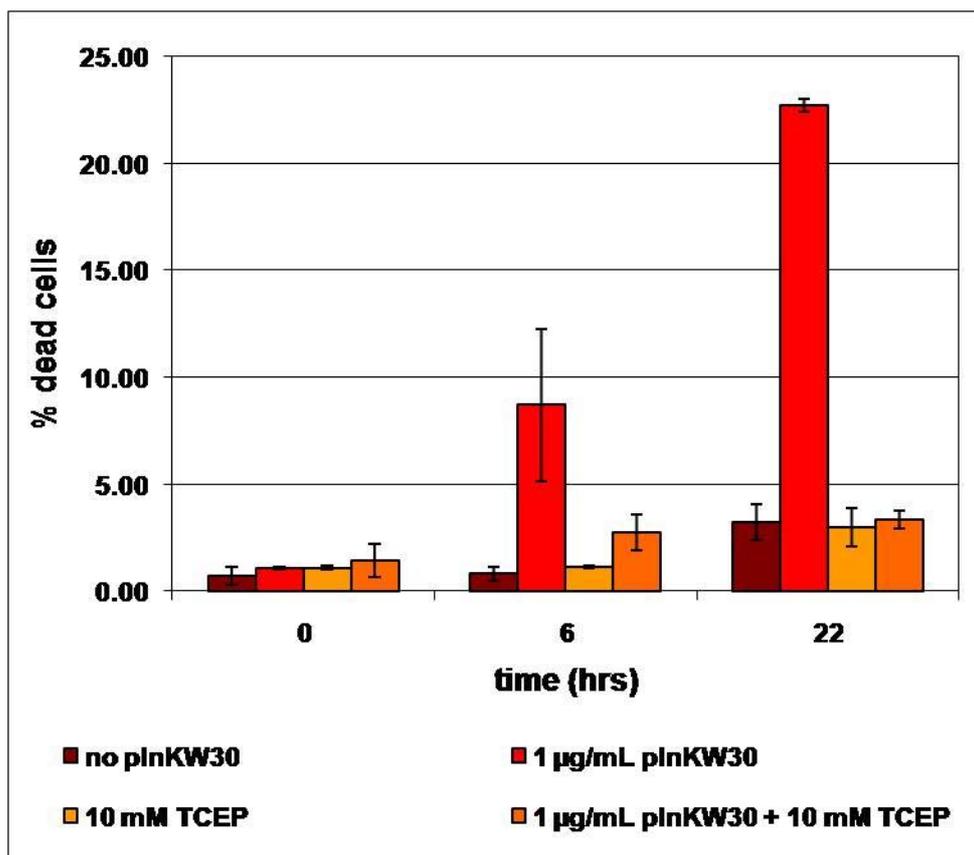
**Figure 3.40: Activity of native and reduced PlnKW30 in a bioassay overlay.**

About 1  $\mu\text{g}$  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 PInKW30.**

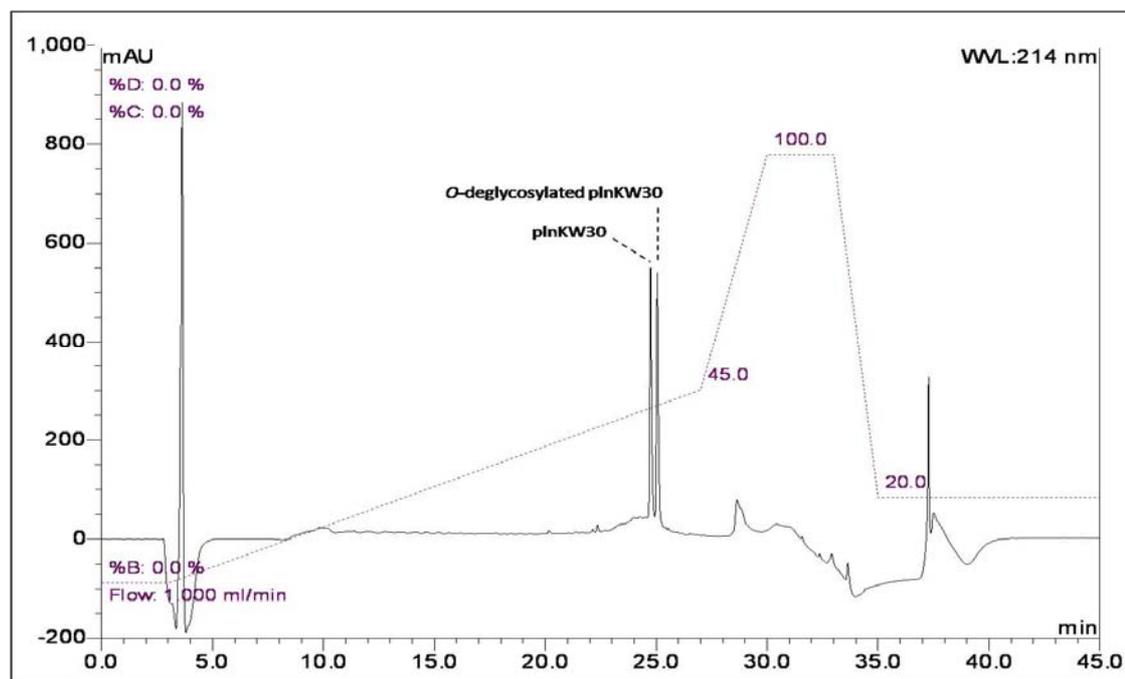
Percentage of dead cells of *L. plantarum* ATCC 8014 at time 0, 6 and 22 hours with addition of 1 µg/mL reduced PInKW30; each data point represents the mean of at least two measurements of separate cultures.

### 3.6.2 ANALYSIS OF THE GLYCAN MODIFICATIONS OF PLNKW30

Characterisation of both the N- and C-terminal fragments of PInKW30 is discussed in section 3.5, including the identification of the post-translational modification of both fragments.

To further characterise the antimicrobial activity of PInKW30, 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 PlnKW30 (Figure 3.42).

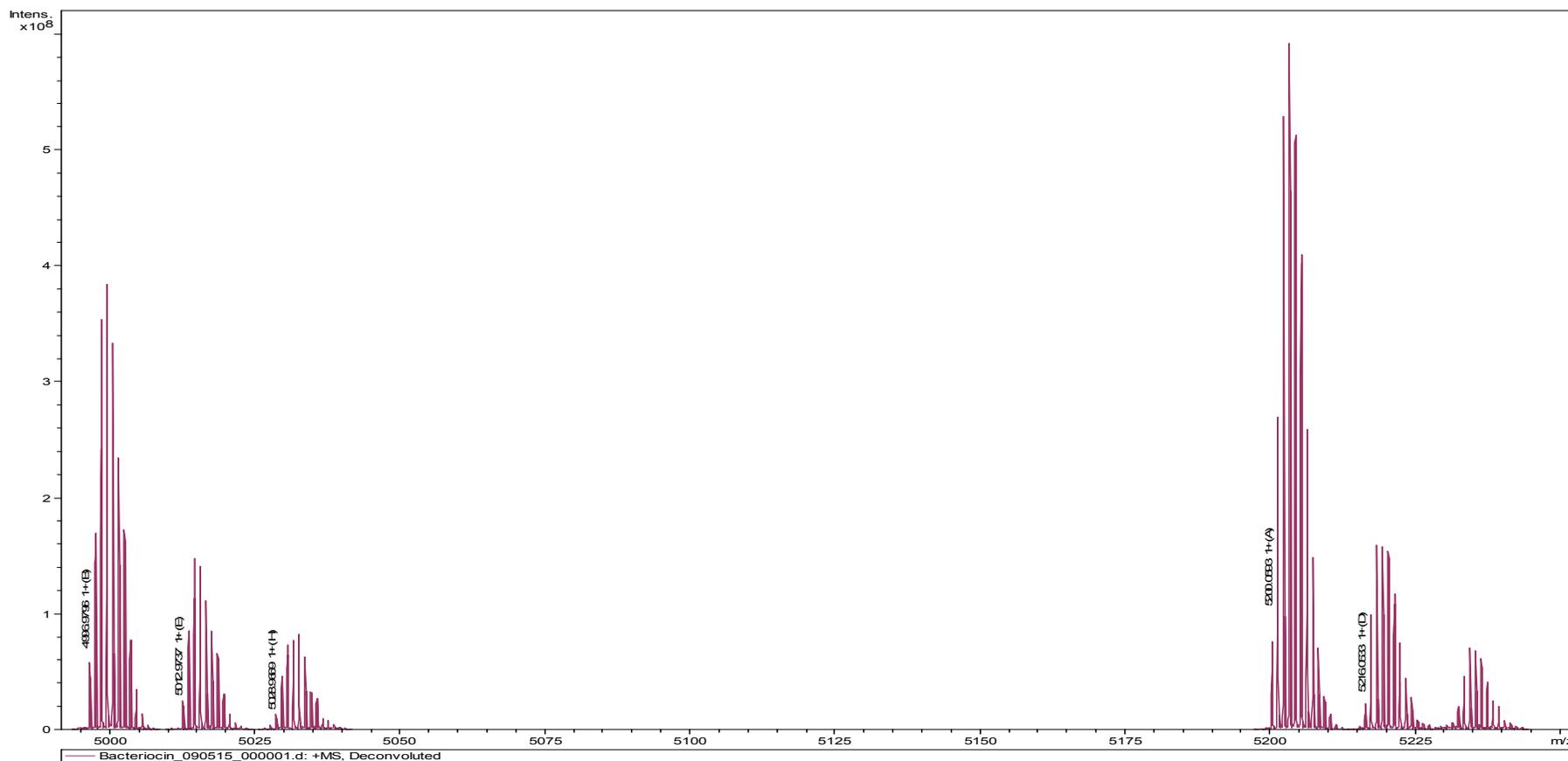


**Figure 3.42: RP-HPLC chromatogram of native and deglycosylated PlnKW30.**

RP-HPLC, Jupiter C18 column, solutions A: H<sub>2</sub>O, 0.1 % TFA and B: Acetonitrile, 0.08 % TFA, conditions used as shown in chromatogram. Native PlnKW30 was eluted at 24 min, 40.3 % B. and the deglycosylated PlnKW30 was eluted at 25 min, 41.7 % B.

Analysis by mass spectrometry confirmed that only one *N*-acetylglucosamine (GlcNAc) had been removed from PlnKW30 by the enzyme-treatment and that the disulfide bonds remained intact (Figure 3.43). The resulting deglycosylated PlnKW30 had a mass of 4996.9827 m/z, which is 5200.0605 m/z minus 203.0778 m/z of the GlcNAc.

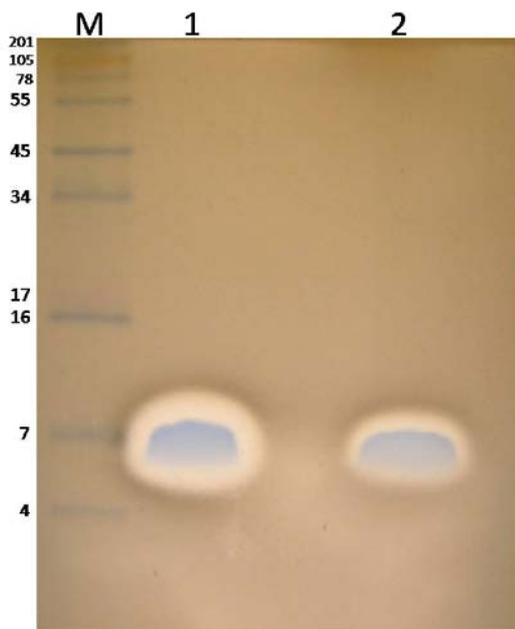
The deglycosylated PlnKW30 was subjected to collision induced fragmentation (CID) and no fragment ions corresponding to serine 18 modified with an *O*-linked GlcNAc could be detected. In contrast, only ions with a modified C-terminal cysteine residue could be detected (Dr. P. Man – personal communication).



**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)

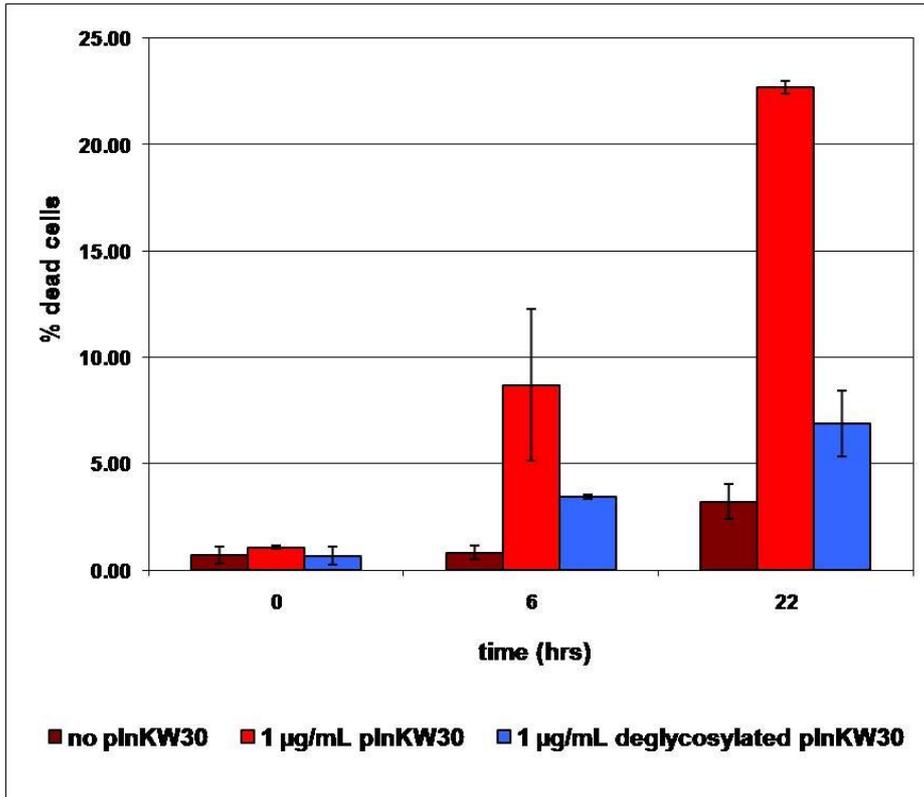
The deglycosylated PlnKW30 was subjected to the biological plate assay (Section 2.4.9) as well as to the Live/Dead cell assay (Section 2.4.17). The tricine SDS-PAGE overlaid with the indicator strain (Figure 3.44) consistently showed a decrease in the antimicrobial activity of the deglycosylated PlnKW30 compared to the native bacteriocin.



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

About 1  $\mu\text{g}$  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:  $\alpha$ -helices, parallel and antiparallel  $\beta$ -sheets,  $\beta$ -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 H<sub>2</sub>O 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.

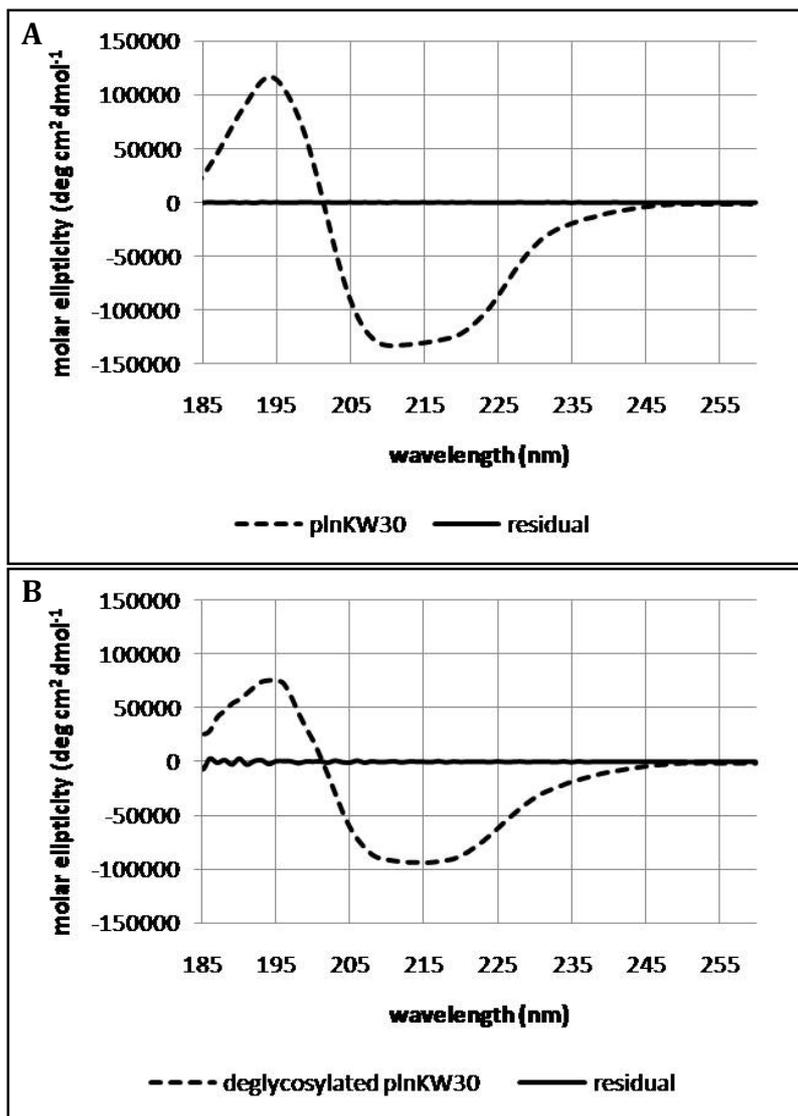


Figure 3.46: Smoothed, base-line corrected spectra of native PlnKW30 (A), deglycosylated PlnKW30 (B) and their corresponding residuals.

The spectra of the reduced and alkylated and reduced (using TCEP) PlnKW30 and their residuals are shown in figure 3.47.

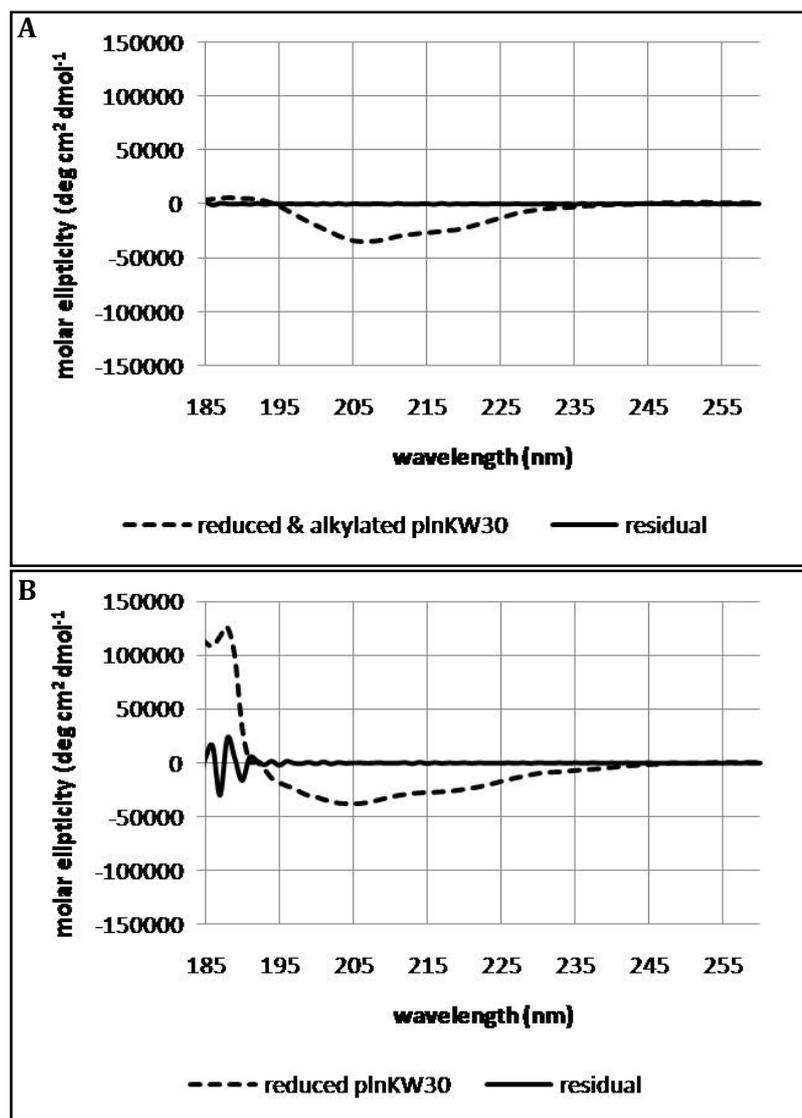


Figure 3.47: Smoothed, base-line corrected spectra of reduced and alkylated PlnKW30 (A), reduced PlnKW30 (B) and their corresponding residuals.

The spectra of the reduced and alkylated PlnKW30 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 PlnKW30 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.

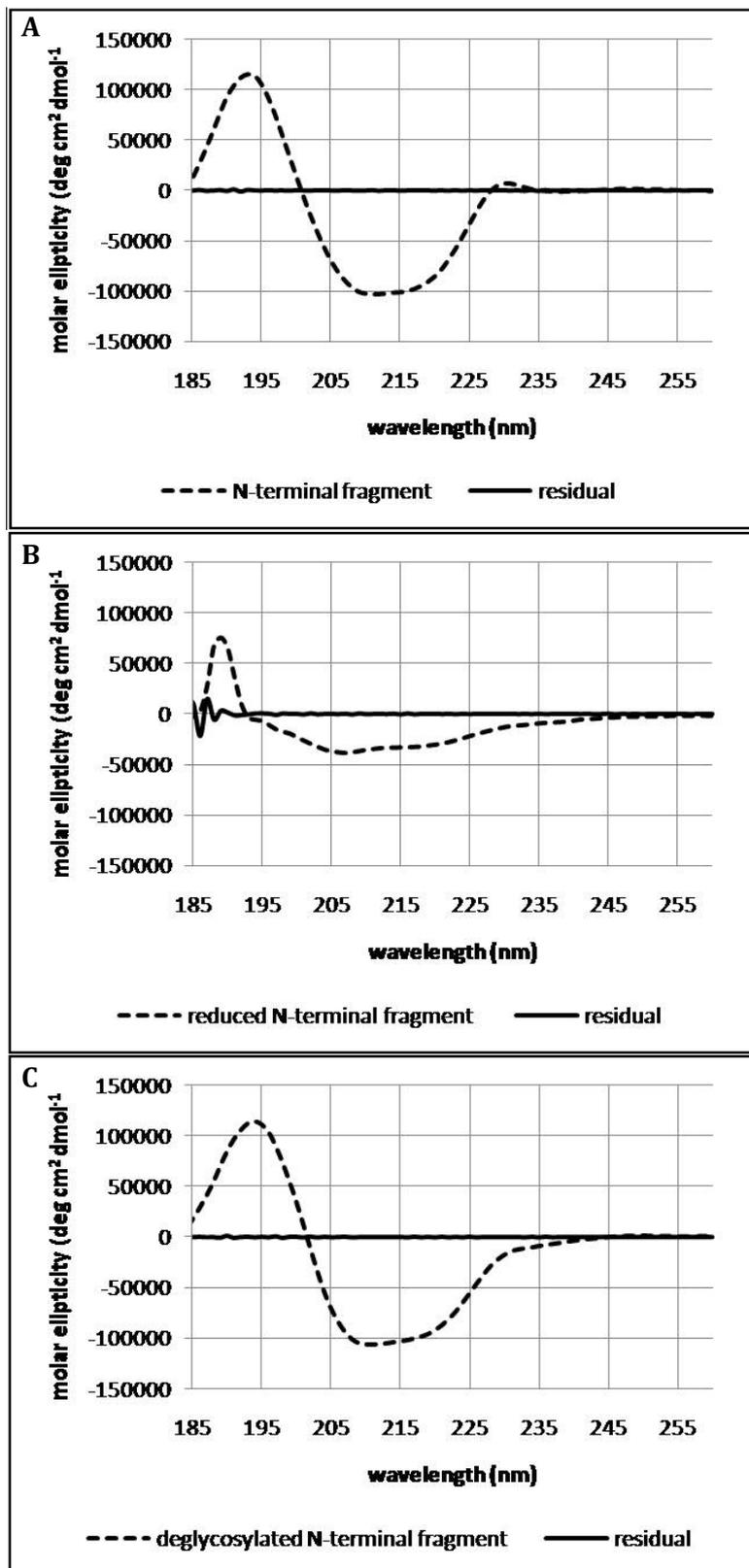


Figure 3.48: Smoothed, base-line corrected spectra of the N-terminal fragment (A), reduced N-terminal fragment (B) and their corresponding residuals.

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).

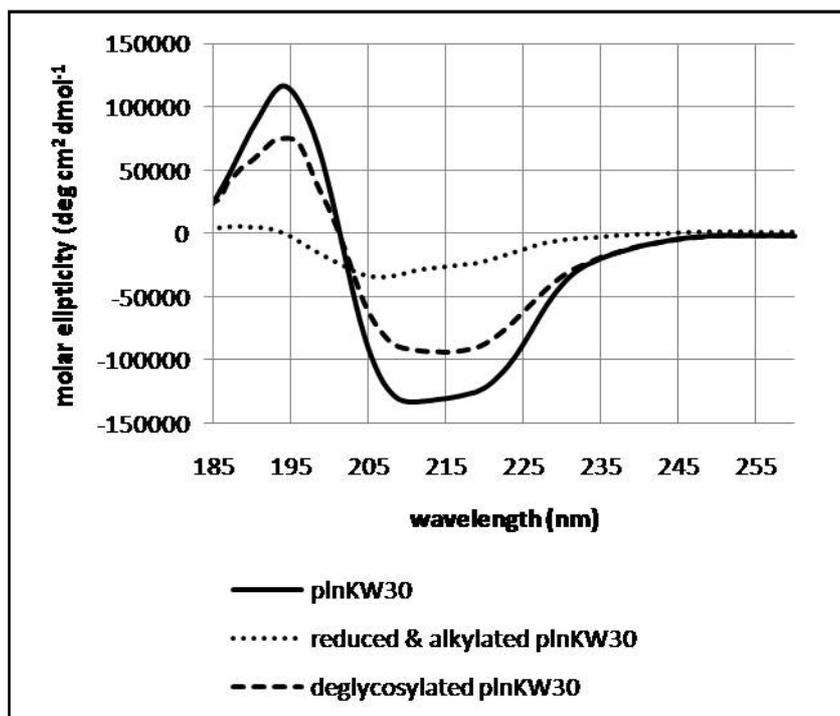


Figure 3.49: Circular dichroism spectra of native, reduced and deglycosylated PlnKW30.

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).

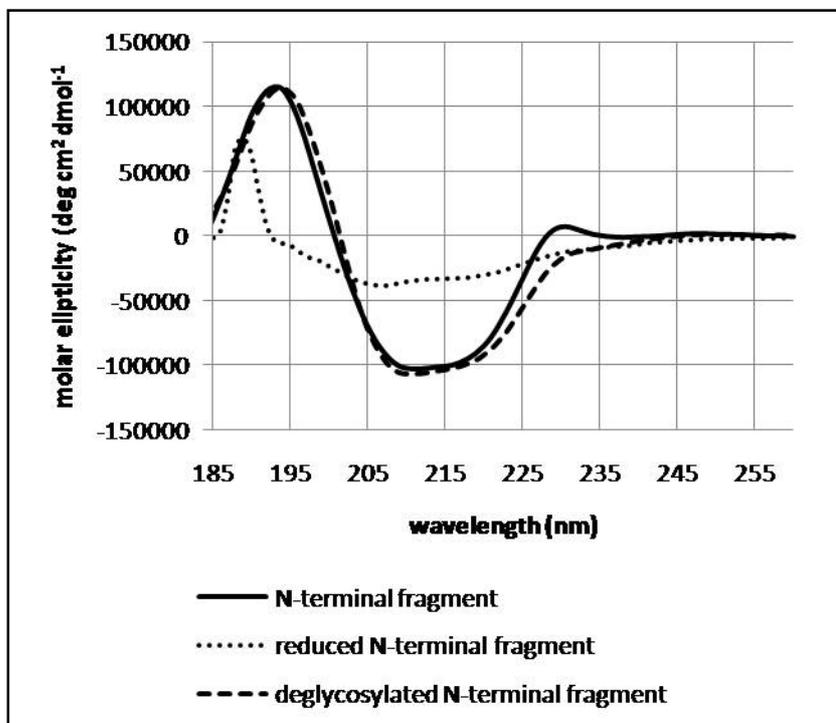


Figure 3.50: Circular dichroism spectra of native, reduced and deglycosylated N-terminal fragment.

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.

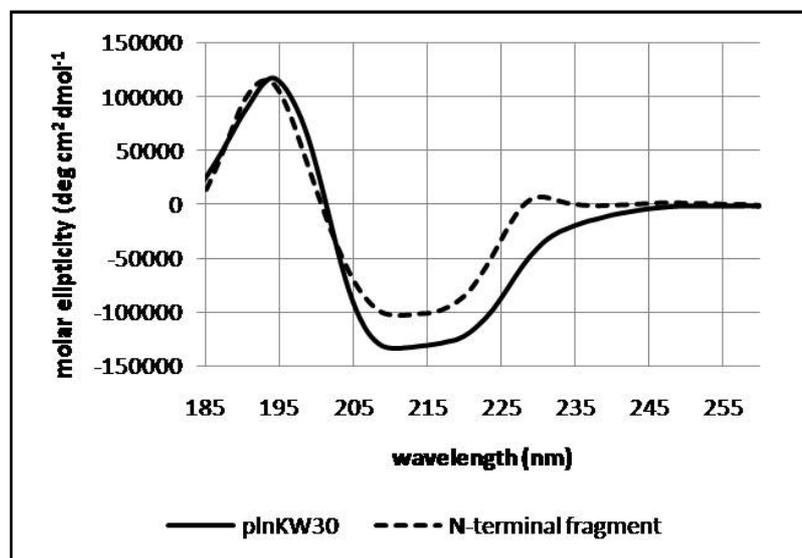


Figure 3.51: Circular dichroism spectra of native PlnKW30 and native N-terminal fragment.

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* Il1403 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.

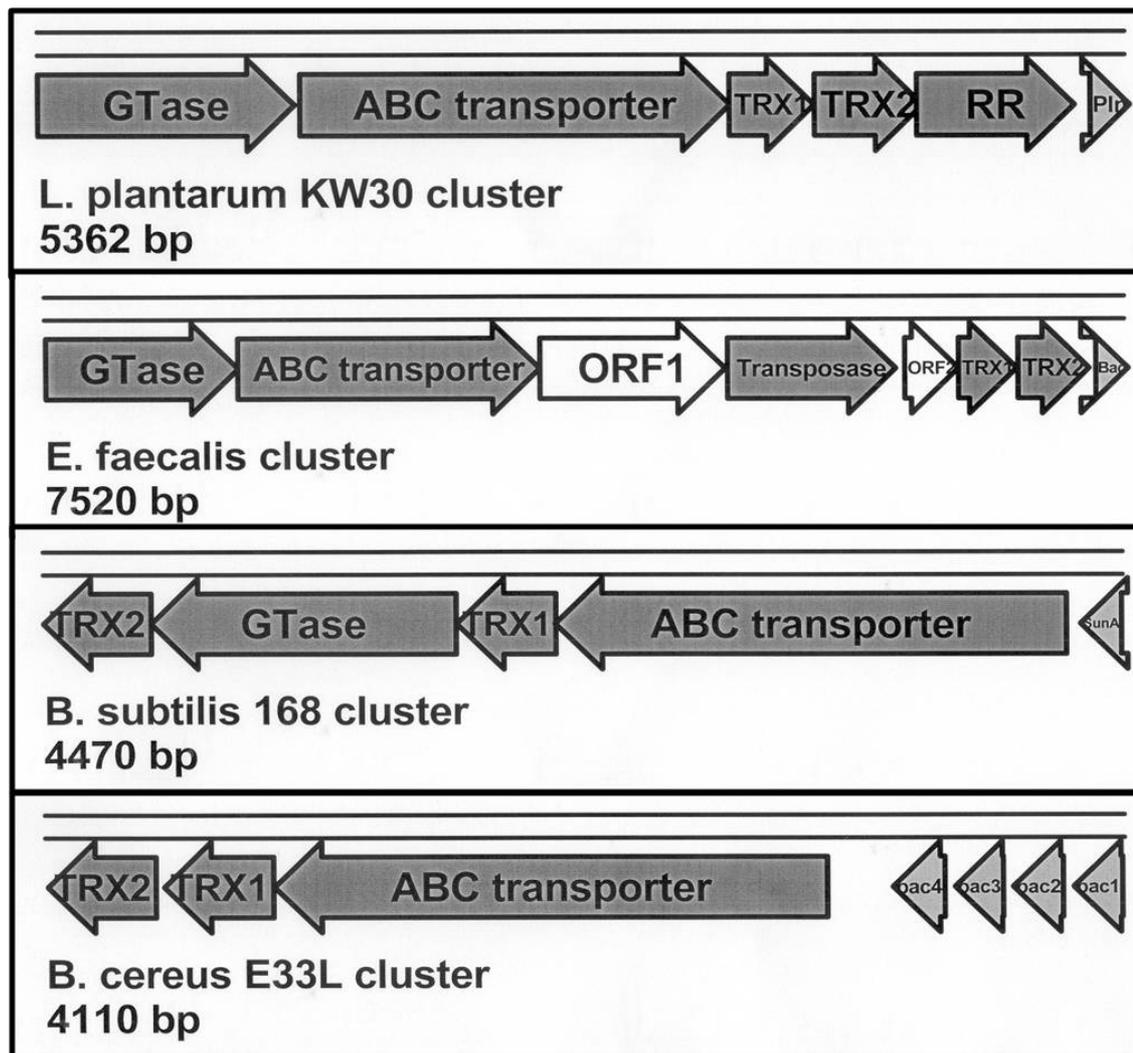
## 4.2 ANALYSIS OF THE *PLNKW30* GENE CLUSTER

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 *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 *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 *et al.* 1990; Rauch & Devos 1992). The presence of transposase, helicase and resolvase ORFs flanking the *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 *et al.* 2006). Horizontal gene transfer between bacteria usually occurs by natural competence or bacteriophage infection, and although *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 *B. subtilis* (Dubnau & Lovett 2002). These findings show that there is a reasonable possibility that the *plnKW30* gene cluster was acquired via horizontal gene transfer, especially when the *plnKW30* gene cluster is compared to clusters from other bacteria that have similar gene arrangements (Figure 4.1). The *plnKW30* gene cluster may have arrived as a functional block of genes or *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 *plnKW30*-like clusters in figure 4.1 suggests that these clusters are not evolutionarily related, but have resulted from convergent evolution.



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

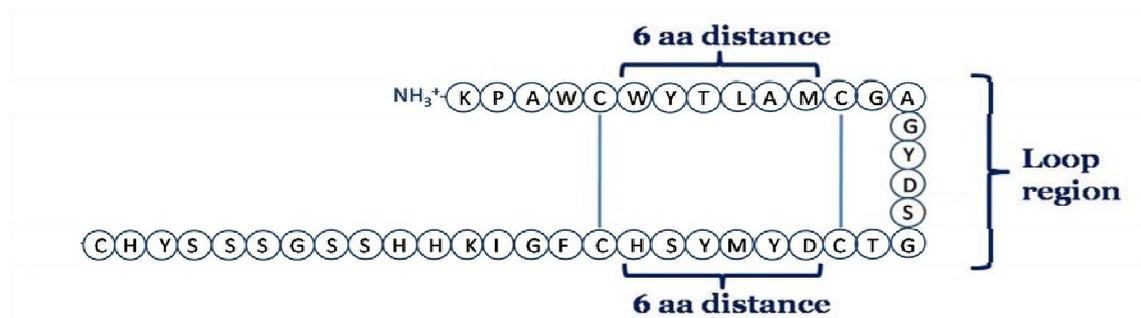
*PlnKW30* gene cluster from *L. plantarum* KW30; Gene cluster on pTEF2 plasmid from *Enterococcus faecalis* V583; Sublancin 168 gene cluster from *Bacillus subtilis* 168; Gene cluster on pE33L466 plasmid from *Bacillus cereus* E33L. ABC-transporter: ATP binding cassette transporter; bac: bacteriocin; 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.**

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..

PlnKW30	MSKLVKTLTISEISKAQNN <u>GG</u> KPAW----- <u>C</u> WYTLAM <u>C</u> GAGYDSGT----- <u>C</u> DYMYSH <u>C</u> FGIKHHSSGSSSYH <u>C</u>
EF_bac	MLNKKLLENGVVNAVITIDELDAQF <u>GG</u> MSKRD----- <u>C</u> NLMKAC <u>C</u> AGQAVTYAIHSLNRLGGDSSDPAG <u>C</u> NDIVRKY <u>C</u> K
SunA	MEKLFKEVKLEELENQK <u>GS</u> GLGKAQ----- <u>C</u> AALWLO <u>C</u> ASGGTIG <u>C</u> GGGAVA----- <u>C</u> ONYRQF <u>C</u> R
BC_1	MNKFIRELKTEELTKHV <u>GG</u> IDPKNETVIHDGGGAGPNC <u>C</u> KVLFVGC <u>C</u> RAKDWEL----- <u>C</u> QLWYDF <u>C</u> R
BC_2	MNKFIRELKTEELTKHV <u>GG</u> IDPKNETVIHDGGGAGPNC <u>C</u> KMLLLGC <u>C</u> QMKYISQ----- <u>C</u> NLWTQF <u>C</u> RS
BC_3	MNKFIRELKTEELTKHV <u>GG</u> IDPKNETVIHDGGGAGPNC <u>C</u> RRLAQS <u>C</u> ILKNFKS----- <u>C</u> GLFDQF <u>C</u>
BC_4	MNKFIRELKAEELTKHV <u>GG</u> VDSKNETVIHDGGGAGPNC <u>C</u> SLLNQC <u>C</u> ILKNFKA----- <u>C</u> GLFDQY <u>C</u> YREWL
SM_bac	MQKATDFKELSKELLATYE <u>GG</u> ASTA----- <u>C</u> KWATLT <u>C</u> AHDIALAGYGTSDPGSA----- <u>C</u> AYMRRV <u>C</u> KK
SS_bac	MTQMEKFNKISKTVLNEIR <u>GG</u> ASESAN----- <u>C</u> KWAIVT <u>C</u> AHDIILAGLGTSDPGSS----- <u>C</u> AYMRRV <u>C</u> KKN

**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  $\beta$ -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.

```

EF_GT      MYSENFITNDWFNVEVFNKNKYTLTNQENKDVTELWLQILKGLKFPNELKETVSYSKNLK 60
BS_GT      -----MKLSDIYLELKKGYADSLLYSDLVLLVNIEMEYKDDIDVMSIQSLVA 46
LP_GT      -----MKNRQNEIDSYLNHLHLP-----VHKSFDFGNLTN 30
           :.: . * : . . .

EF_GT      ELSLKTHAEVSVCIIAKNEQDSIRKINSIYEFSD--EIIFIDTGSIDLTKKIVKEIAS 117
BS_GT      GYEKSDTPTITCGIIVYNESKRIKCLNSVKDDFN---EIIVLDSYSTDDTVDIICKDFP 103
LP_GT      IDQFRHHIYVSYIVICKNSQATIERCVNSIAQNMENGDDELIVLDTGSTDETIVHLVKKNMP 90
           . : : * *.. *.:*:*: : : *:*:*: * * * .:.* .

EF_GT      EKVKIFDYTWQDDFSDARNYSIQKASKEWILIIDADEYVSSDELTKLRLLDMLDRFKFK 177
BS_GT      D-VEIKYEKWKNDFSYARNKII EYATSEWIYFIDADNLYSKENKGIKAKVARVLEFFSID 162
LP_GT      Q-AKISVTNWKNDFSEVRNKALKLASKDWFYVDSDEWLDVDDGAQLKKILFKVQAKNFK 149
           : .:* .*:*** .* : : *.:* :*:* . : : : : : : .:

EF_GT      DSLRVSCAIYQLDNVITHGQSRLFRNKKIKIYGLIHEELRNKNG---LDPIFNVESEIT 234
BS_GT      --CVVSPYIEEYTGHLYSDDRMRFLNGKVKFHKVHEEPMNYN---HSLPFNFIVNLK 216
LP_GT      --FVINPTFSDHSGQIYQTVGRIFPKKSSFHYAKIHEEVRKEDQKLGVDVRHFACDDII 207
           :. : : . : *:* :.:*:* :*:* : . . . : :

EF_GT      FFHDGYKEILR--KEKCERNIRLLAKMLEKEPDNVRWAYLYCRDSFSINANIDFEKILLP 292
BS_GT      VYHNGYNPSENNIKSKTRRNINLTEMLRLEPENPKWLFFFGRELHLLDK---DEEAD 272
LP_GT      LYHDGYDKEVLRDKDKIKRNIRLLQEMTCEEPQARWPFLLARDGFVLP---QDKLKQ 263
           .:*:* .*. *.*.* * :* *:*:* * : : * : . : :

EF_GT      FLIKNMDESISYEN--ILLTNYTHLILFLITKKYIIDGKSSLASKCIK-VLEKMLPNSSD 349
BS_GT      YLKKSINNYKFNDRHFIDALVLLCTLLLRNNYVD---LTLYLD-ILETEYPRCVD 326
LP_GT      LVKRTLDLVASDS---LQEKYSPFAKLLGRILLREGKTTQAVLSFKDVLQITGGEDSD 319
           : :.: . . : : * : : : : : : . : *

EF_GT      VTFYKFLNKQHSLYEQQFEFLKEVIQFRKNNE--YDQYSQIGCNLLHYDLLISGLLFDVK 407
BS_GT      VDYFR---SAILLVDMQNKLTSLSNMIDEALT--DERYSAINTTKDHFKRILISLNIQLE 381
LP_GT      AIYYIESFKINEIIAEAKSIEVKMLRYLNKHKGMIDVNSDISGNYHIAQVILECDIISA 379
           . : : . : . : : : * * . * : : :

EF_GT      SYDYSYQYFLKLDLANYPSELEIPDEYKMLINKYRENES----- 446
BS_GT      NWERVKEISGEIKNDNMKEIKQYLANSLHNI EHVLRKGLIEV--- 422
LP_GT      NYSHLFPLISEIP-KNFSGDIKSSVKSVAKLYSKLQGDGSKNENN 422
           :. . : : * : : : : .

```

**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 regions 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).

```

BS_SunT      MNKKKKYVHTKQFNSHDCGLACISSILKFHN-LNYGIDFLLDLIGD-KEGYSLRDLIVIF 58
BC_ABC      ----MKYVHTKQEEYDCGIACVASIFKYH-LHHGINYL RDQVTY-KNGYDLKDLLALF 54
EF_ABC      ---MIKRPHVMQDSIQDCGVACIEMICKFYN-INIDRRYIQEETGYGMIGISLKAMEKFF 56
LP_ABC      --VTAKMRI IKQIDQNDCGPAAVATIIGMISKIKIDDWQIRQIITTTNYGTNFISIINGL 58
              *      * . *** *.: *      :. . : :      * .: : :

BS_SunT      KKMG----IKTRPLELQENKTFEALKQIKLPCIALLEGEE----YGHYITIYEIRNNYLL 110
BC_ABC      SQVNH---FSCKAVEINKEDIEEALKHIEGPCIALINKQVNESYEGHYIIIIYKRKKNKLV 111
EF_ABC      SKVDANPEIVNISKINRLNKENREMINNSLPAIVFLEEEE---IINHVVVIWHIGKKRIL 113
LP_ABC      KSVQ-----VESVVEKCVKNPVDVFEIEFPVLTQIMQNG----YLHFVVLTKCSGSKLY 108
              ..:      . . . : . . * :. : :      *:: : . . :

BS_SunT      VSDPDKDKITKIKKEDFESKFTNFIEIDKESIP--EKEKDQKKHSYFFKDILFRNKLV 168
BC_ABC      ISDPGNDKISTITIDEFKKHTTGIFLLIESKNS---NSDKKTSYHQKFFRELKNNKLSI 168
EF_ABC      VSDPTHTKKEWINNKHFEKRAISYLFVEKPKNIFLNRSPKKVRFYGFYIKRNLKIFSLVM 173
LP_ABC      WADPGSGKIESIDKAKFMMHWTPILITIPRSQNFQILKTKVESLKVLPILWFSFRKHWACL 168
              :** * * . * : : : .. . * : : : :

```

**Figure 4.5: Comparison of the N-terminal amino acid sequences 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. The conserved cysteine and histidine regions that are part of the active site of the C39 peptidase domain are shadowed. Identities are marked \*; highly conserved :; conserved . (Clustal W; Larkin *et al.*, 2007).

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).

```

BS_SunT      LSKLYKVPDKSIYLNGLDINRYDHL SIRKRIVYIDENPFLFKGTIKENLCMG-EIFDQNE 587
BC_ABC      LNKLYQANASQIYINGVDINEINDNSLRKELIYLNENPFLFKNTIKENICMG-EYFTEDE 584
EF_ABC      LVRLDDDDYRGQILINNIDIKKINLDCLRSKLVFVEPNPKFLEGTIRDNLLLG-HKVPNSI 591
LP_ABC      MAGFYPAEG--LSLGNIPYYAIGAESLGNLTITTYVEQSPQIFSDTVMNNTLGRKNITVKS 582
           :  :           : ...: . .: . : :: . * :...*: **: :* . . .

BS_SunT      IENACIMSQCHEFICNLDKQYSYKLSNGSNLSTGQKQRLALARAILHQPOVLILDESLS 647
BC_ABC      IIKACQIAHIYDVITSLPKGFNFVINDSNSNLSTGQKQRLCLARAILHKPSVLILDESLS 644
EF_ABC      FNKLIRDFEINKILDDPLGINFPGAAIKCLSSGQKQKLALFRGILKRPDILLVDEGFS 651
LP_ABC      INRIASSIGFDECLNHLPNGIYTLGSSGYSLSGGQKQLLNIIIRSMVIPSRFIIFDEITN 642
           : .           . : *           ** **** * : *.: : . .:.* * .

BS_SunT      NIDPDNTKLIYETLHRMD-CLIIILITHNDPSNFKYNKKLVFRNNRIESSYSENKEYSI- 705
BC_ABC      NVDPDNFMKIYDDLDDLN-TIIIFI THNPEYITKYDKKFI FREKSI FEMQSEFEKQLN-- 701
EF_ABC      NMDKEYLDRILPKFDSWG-IKLIVIDHSN----RVTKDIDYITMENYDIKNSWM----- 700
LP_ABC      GLDINTKLKVENYLLSQNKPKLIFITHDLTLAIRCDDIFTVKNEMHQIETNKNSTESLR 702
           .:* : : : . :*. * * . : . : : . .

```

**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

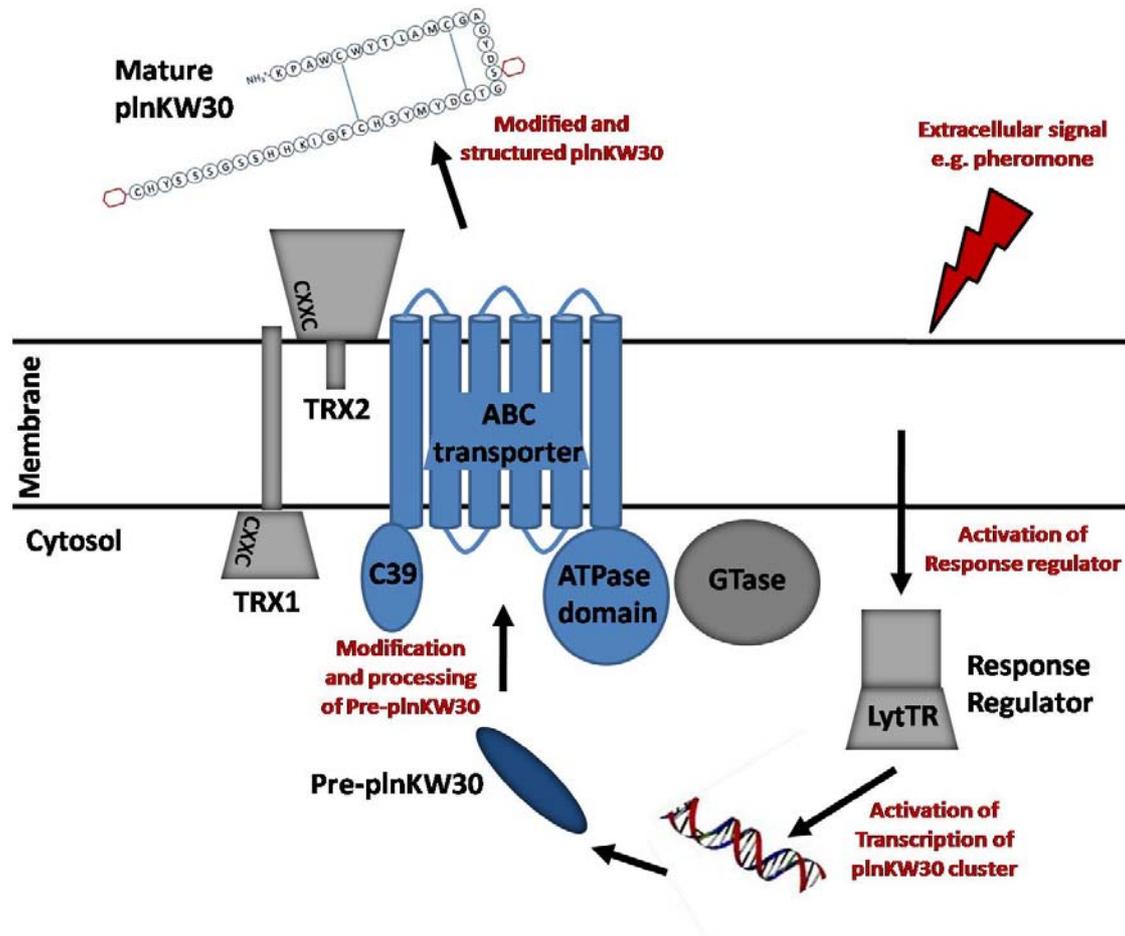
Bioinformatic analyses were used to predict the function of the proteins encoded by the *plnKW30* gene cluster. Such analyses are based on the knowledge that proteins with similar amino acid sequences usually carry out similar functions.

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  $\beta$ -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 *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 *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 *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-*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 *E. faecalis* and *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 *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 *plnKW30* specific probe is very clear: only one band is visible with the size of the pre-*plnKW30* mRNA product (Figure 3.11). This includes the *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 *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 *plnKW30* gene is transcribed in two forms: a short, abundant *gccA* transcript and a longer very low abundance transcript comprising *gccBCDEFA*. Although a longer transcript was not visible on the *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 *gccE-gccF*.

#### 4.3.2 GROWTH-PHASE-DEPENDENT GENE EXPRESSION OF *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 *plnKW30* expression. The blot (Figure 3.12) showed that *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 *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 *plnKW30* transcript increases strongly from 4 to 24 hours and thus probably reduces the relative amount of *gap* transcript in the total RNA sample. This could explain the lower concentrations of *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.*

2007) 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 PlnKW30. 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 PlnKW30, and there was no evidence either by mass spectrometry or NMR spectroscopy of other molecules being covalently bonded to PlnKW30.

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 PlnKW30 (Kelly *et al.* 1996), although purified PlnKW30 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 PlnKW30 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 PlnKW30 an *L. plantarum* ATCC 8014 culture has an OD<sub>600</sub> close to 0.2 after 7 hours, whereas a control culture without PlnKW30 has an OD<sub>600</sub> of about 0.4. Thus, the addition of 12.5 ng/mL PlnKW30 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, PlnKW30 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 PlnKW30. The cell walls and membranes of fungi have a very different composition to Gram-positive bacterial cells, and PlnKW30 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<sub>600</sub> 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<sub>600</sub> 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<sub>600</sub> 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<sub>600</sub> 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<sub>600</sub> 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.

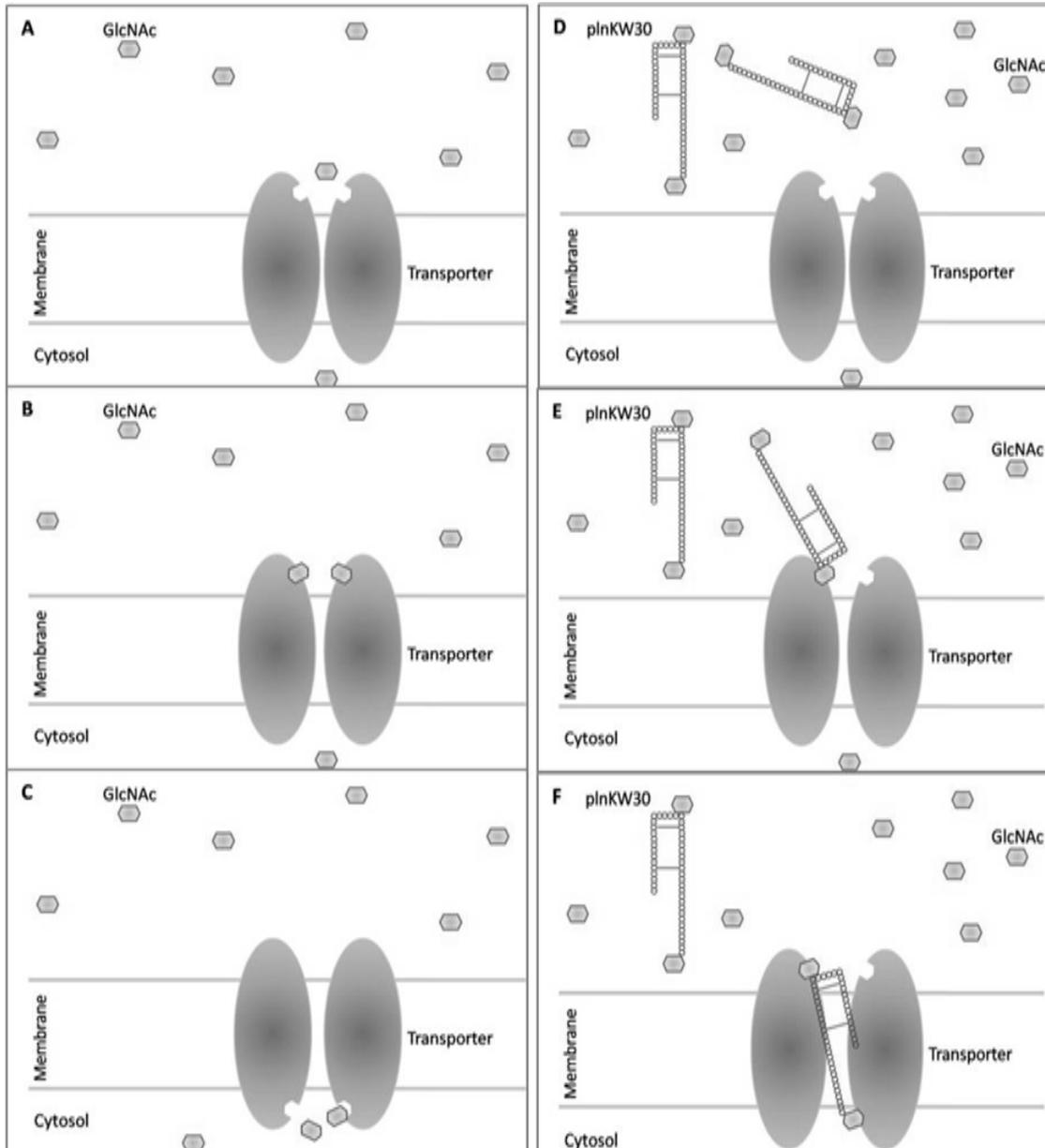
The effect of different sugars on the activity of PlnKW30 was tested, because work by Diep *et al.* (2007) showed that some class II bacteriocins use components of the mannose phosphotransferase system (man-PTS) of susceptible cells as receptors and because the unusual post-translational modifications 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 EIICBA 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  $\mu$ M 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  $\mu$ M) 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.

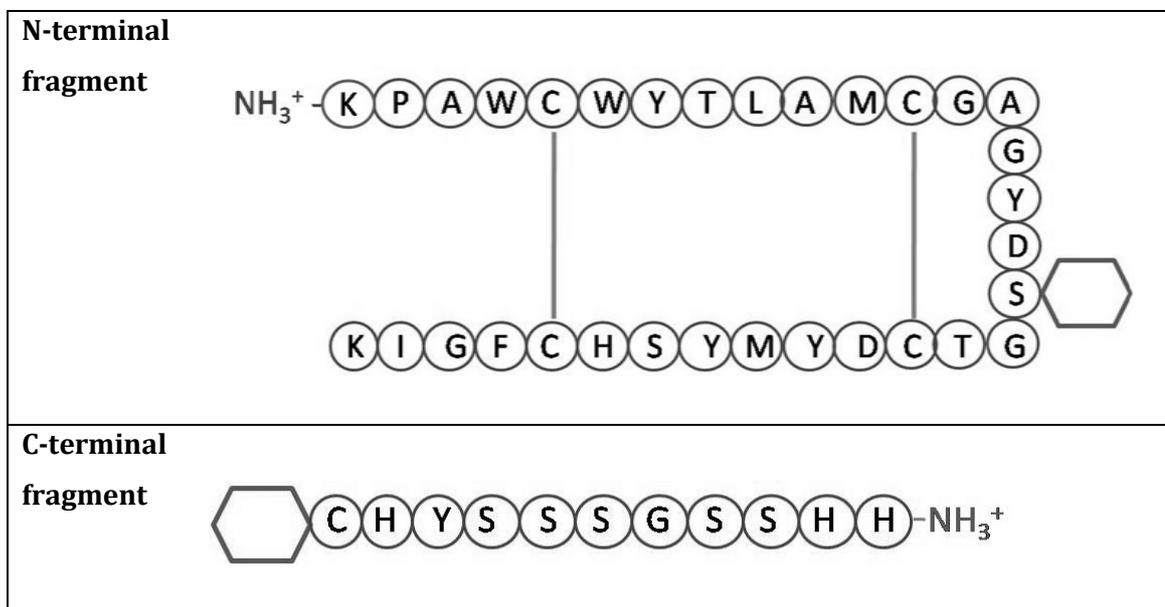
It is unlikely that the model for PlnKW30 proposed in figure 4.8 can fully explain the bacteriostatic activity of PlnKW30. Analysis of the interaction of PlnKW30 with other components of the cell membrane, and of the orientation, organisation and position of PlnKW30 in the membrane, might also give insight 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 GlcNAcs 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.

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 secrete *O*-glycosylhydrolases as a defence against *O*-glycosylated bacteriocins, such as the *N*-acetyl- $\beta$ -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

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  $\beta$ -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- $\beta$ -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  $\beta$ -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  $\beta$ -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  $\beta$ -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

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 C $\beta$ -C $\gamma$  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.

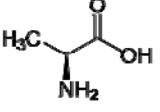
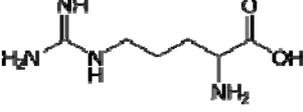
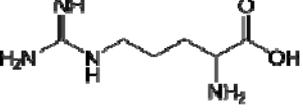
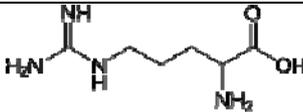
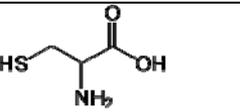
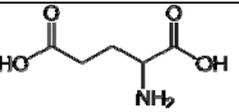
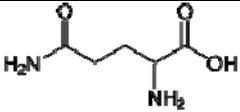
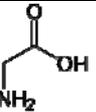
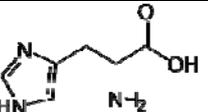
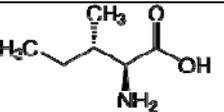
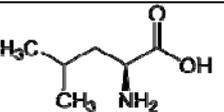
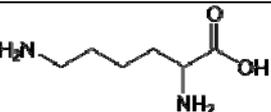
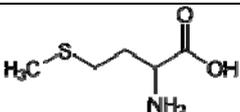
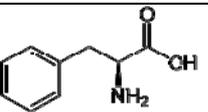
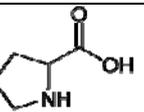
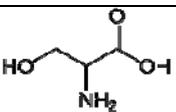
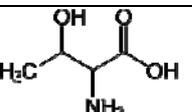
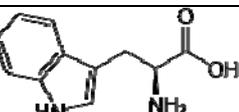
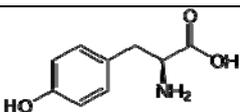
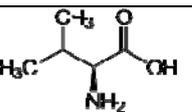
## 5 APPENDIX

### APPENDIX 1: THE GENETIC CODE

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

		Second letter				
		U	C	A	G	
First letter	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA Stop UAG Stop	UGU } Cys UGC } UGA Stop UGG Trp	U C A G
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA } CGG }	U C A G
	A	AUU } AUC } Ile AUA } AUG Met	ACU } ACC } Thr ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }	U C A G
	G	GUU } GUC } Val GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } GGC } Gly GGA } GGG }	U C A G
						Third letter

## APPENDIX 2: STRUCTURES AND SYMBOLS OF THE STANDARD AMINO ACIDS

		
Alanine (Ala / A)	Arginine (Arg / R)	Asparagine (Asn / N)
		
Aspartic acid (Asp / D)	Cysteine (Cys / C)	Glutamic acid (Glu / E)
		
Glutamine (Gln / Q)	Glycine (Gly / G)	Histidine (His / H)
		
Isoleucine (Ile / I)	Leucine (Leu / L)	Lysine (Lys / K)
		
Methionine (Met / M)	Phenylalanine (Phe / F)	Proline (Pro / P)
		
Serine (Ser / S)	Threonine (Thr / T)	Tryptophan (Trp / W)
		
Tyrosine (Tyr / Y)	Valine (Val / V)	

## APPENDIX 3: LISTS OF STRAINS AND VECTORS USED IN THIS THESIS

Table 5.1: List of strains used in this thesis

Strain	Relevant Genotype <sup>a</sup>	Reference <sup>b</sup>
<i>E. coli</i> TOP10	F <sup>-</sup> , <i>araD139</i> $\Delta$ ( <i>ara</i> , <i>leu</i> )7697, $\Delta$ <i>lacX74</i> , <i>galU</i> , <i>galK</i> , <i>rpsL</i> , <i>deoR</i> , $\Phi$ 80 <i>dlacZ</i> $\Delta$ M15, <i>endA1</i> , <i>nupG</i> , <i>recA1</i> , <i>mcrA</i> , $\Delta$ ( <i>mrr</i> <i>hsdRMS</i> <i>mcrBC</i> )	Invitrogen
<i>E. coli</i> XL1-Blue	<i>recA1</i> , <i>endA1</i> , <i>gyrA96</i> , <i>thi-1</i> , <i>hsdR17</i> ( <i>r<sub>k</sub></i> <sup>-</sup> , <i>m<sub>k</sub></i> <sup>+</sup> ), <i>supE44</i> , <i>relA1</i> , -, <i>lac</i> [F <sup>-</sup> , <i>proAB</i> , <i>lacI<sup>q</sup></i> , <i>lacZ</i> $\Delta$ M15, Tn10(Tc <sup>r</sup> )]	(Bullock <i>et al.</i> 1987)
<i>E. coli</i> DH5 $\alpha$	<i>supE44</i> <i>hsdR17</i> <i>recA1</i> <i>gyrA96</i> <i>thi-1</i> <i>relA1</i>	(Hanahan 1983)
<i>E. coli</i> EC1000	RepA1 MC1000, Km <sup>r</sup> , carrying a single copy of the pWV01 <i>repA</i> gene in the <i>glgB</i> gene; host for pORI28-based plasmids	(Law <i>et al.</i> 1995)
<i>E. coli</i> BL21 (DE3)	F <sup>-</sup> , <i>ompT</i> <i>hsdS<sub>B</sub></i> ( <i>r<sub>B</sub></i> <sup>-</sup> <i>m<sub>B</sub></i> <sup>-</sup> ) <i>gal</i> <i>dcm</i> (DE3)	Novagen
<i>E. coli</i> Origami (DE3)	$\Delta$ <i>ara-leu7697</i> $\Delta$ <i>lacX74</i> $\Delta$ <i>phoA</i> <i>PvuII</i> <i>phoR</i> <i>araD139</i> <i>ahpC</i> <i>galE</i> <i>galK</i> <i>rpsL</i> F '[ <i>lac</i> <sup>+</sup> <i>lacI<sup>q</sup></i> <i>pro</i> ] <i>gor522</i> ::Tn10 (Tc <sup>r</sup> ) <i>trxB</i> ::kan (DE3)	Novagen
<i>E. coli</i> Rosetta (DE3)	F <sup>-</sup> <i>ompT</i> <i>hsdS<sub>B</sub></i> ( <i>r<sub>B</sub></i> <sup>-</sup> <i>m<sub>B</sub></i> <sup>-</sup> ) <i>gal</i> <i>dcm</i> <i>lacY1</i> (DE3) pRARE (Cm <sup>R</sup> )	Novagen
<i>E. coli</i> Rosetta (DE3) pLysS	F <sup>-</sup> <i>ompT</i> <i>hsdS<sub>B</sub></i> ( <i>r<sub>B</sub></i> <sup>-</sup> <i>m<sub>B</sub></i> <sup>-</sup> ) <i>gal</i> <i>dcm</i> <i>lacY1</i> (DE3) pLysSRARE (Cm <sup>r</sup> )	Novagen
<i>E. coli</i> Rosetta Blue (DE3)	<i>endA1</i> <i>hsdR17</i> ( <i>r<sub>K12</sub></i> <sup>-</sup> <i>m<sub>K12</sub></i> <sup>+</sup> ) <i>supE44</i> <i>thi-1</i> <i>recA1</i> <i>gyr96</i> <i>relA1</i> <i>lac</i> F'[ <i>proA</i> <sup>+</sup> <i>B</i> <sup>+</sup> <i>lacI<sup>q</sup></i> <i>Z</i> $\Delta$ M15 ::Tn10 (Tc <sup>R</sup> )] (DE3) pRARE (Cm <sup>r</sup> )	Novagen
<i>E. coli</i> Rosetta-gami (DE3)	$\Delta$ <i>ara-leu7697</i> $\Delta$ <i>lacX74</i> $\Delta$ <i>phoA</i> <i>PvuII</i> <i>phoR</i> <i>araD139</i> <i>ahpC</i> <i>galE</i> <i>galK</i> <i>rpsL</i> F '[ <i>lac</i> <sup>+</sup> ( <i>lacI<sup>q</sup></i> ) <i>pro</i> ] <i>gor522</i> ::Tn10 (Tc <sup>r</sup> ) <i>trxB</i> ::kan (DE3) pRARE (Cm <sup>r</sup> )	Novagen
<i>E. coli</i> TB1	F <sup>-</sup> <i>ara</i> $\Delta$ ( <i>lac-proAB</i> ) [ $\Phi$ 80 <i>dlac</i> $\Delta$ ( <i>lacZ</i> )M15] <i>rpsL</i> (Str <sup>r</sup> ) <i>thi</i> <i>hsdR</i>	New England Biolabs
<i>L. plantarum</i> KW30	PlnKW30 producer strain	(Kelly <i>et al.</i> 1996)
<i>L. plantarum</i> ATCC 8014	PlnKW30 indicator strain	ATCC
<i>L. plantarum</i> NC8	Host strain, silage isolate	(Aukrust & Blom 1992)
<i>L. sakei</i> Lb790	Host strain, meat isolate	(Schillinger & Lucke 1989)
<i>L. lactis</i> MG1363	Derivative of <i>L. lactis</i> NCD0712	(Gasson 1983)
<i>L. lactis</i> NZ9000	Model strain MG1363 with <i>nisRK</i> integrated in chromosome; host strain for nisin inducible vectors	(Kuipers <i>et al.</i> 1998)

<i>L. lactis</i> NZ9800	NZ9700 DnisA, not able to produce nisin, but still immune; host strain for nisin inducible vectors	(Kuipers <i>et al.</i> 1998)
<i>Streptococcus pyogenes</i>		IMBS CC
<i>Oenococcus oeni</i>		IMBS CC
<i>Saccharomyces cerevisiae</i>		IMBS CC

<sup>a</sup> Cm<sup>r</sup>, Chloramphenicol resistance; Km<sup>r</sup>, Kanamycin resistance; Str<sup>r</sup>, Streptomycin resistance; Tc<sup>r</sup>, Tetracycline resistance.

<sup>b</sup> 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**

Vector	Size [bp]	Selection and other features <sup>a</sup>	Reference <sup>b</sup>
pT7-7	2469	Amp <sup>r</sup>	(Tabor & Richardson 1985)
pPROExHTb	4778	Amp <sup>r</sup>	Invitrogen
pMAL_p2G	6721	Amp <sup>r</sup> ; <i>lacI</i> <sup>F</sup> ; P <sub>tac</sub> ; <i>malE</i> ; <i>lacZα</i> ; M13 origin	Novagen
pMAL_c2G		Amp <sup>r</sup> ; <i>lacI</i> <sup>F</sup> ; P <sub>tac</sub> ; <i>malE</i> (signal sequence deleted); <i>lacZα</i> ; M13 origin	Novagen
pET32a(+)	5900	Amp <sup>r</sup> ; T7 promoter; optional C-terminal or internal His-tag; internal S-tag; N-terminal Trx-tag	Novagen
pETDuet_TrxA_HST_Nco	5704	Amp <sup>r</sup> ; two MCS preceded by P <sub>T7</sub> & RBS; pBR322 origin; <i>lacI</i> ; His-tag & rTEV site into site 1; TrxA inserted into site 2 ( <i>NdeI/XhoI</i> )	Novagen
pETDuet_DsbC_HST_Nco	6025	Amp <sup>r</sup> ; two MCS preceded by P <sub>T7</sub> & RBS; pBR322 origin; <i>lacI</i> ; His-tag & rTEV site into site 1; DsbC inserted into site 2 ( <i>NdeI/XhoI</i> )	Novagen
pETDuet_MalE_HST_Nco	6478	Amp <sup>r</sup> ; two MCS preceded by P <sub>T7</sub> & RBS; pBR322 origin; <i>lacI</i> ; His-tag & rTEV site into site 1; MalE inserted into site 2 ( <i>NdeI/XhoI</i> )	Novagen
pETDuet_SUMO_HST_Nco	5705	Amp <sup>r</sup> ; two MCS preceded by P <sub>T7</sub> & RBS; pBR322 origin; <i>lacI</i> ; His-tag & rTEV site into site 1; SUMO inserted into site 2 ( <i>NdeI/XhoI</i> )	Novagen

pORI28	1597	Em <sup>r</sup> , ori (pWV01), replicates only with <i>repA</i> provided <i>in trans</i>	(Law <i>et al.</i> 1995)
pTRK669	1462	ori (pWV01), Cm <sup>r</sup> , RepA1	(Russell & Klaenhammer 2001)
pSIP409	5500	spp-based expression vector, Em <sup>r</sup> ; SppKR expression driven by ermL read-through; with P <sub>orfX</sub> :: <i>gusA</i>	(Sorvig <i>et al.</i> 2005)
pSIP412	5500	spp-based expression vector, Em <sup>r</sup> ; SppKR expression driven by ermL read-through; with SH71rep and P <sub>orfX</sub> :: <i>pepN</i>	(Sorvig <i>et al.</i> 2005)
pNZ8112	3478	Broad host range vector with <i>nisA</i> -promoter and mcs; terminator sequence behind the mcs; <i>usp45</i> signal sequence translationally fused to the <i>nisA</i> -promoter and <i>NaeI</i> site for gene fusion; possibility for C-terminal His-tag; Cm <sup>r</sup>	NIZO
pNZ8113	3368	Broad host range vector with <i>nisA</i> -promoter and mcs; terminator sequence behind the mcs; possibility for C-terminal His-tag; Cm <sup>r</sup>	NIZO
pNZ8148	3165	Broad host range vector with <i>nisA</i> -promoter and mcs; terminator sequence behind the mcs; with a deletion of 184 bp to remove a small <i>B. subtilis</i> DNA fragment; Cm <sup>r</sup>	NIZO
pNZ5319	7028	Low copy vector pLL252 with <i>nisR</i> and <i>nisK</i> genes; compatible with pNZ8*** vectors in other hosts than <i>Lactococcus lactis</i> ; Em <sup>r</sup>	NIZO

<sup>a</sup> Amp<sup>r</sup>, Ampicillin resistance; Cm<sup>r</sup>, Chloramphenicol resistance; Em<sup>r</sup>, Erythromycin resistance; Gm<sup>r</sup>, Gentamycin resistance; Km<sup>r</sup>, Kanamycin resistance; MCS, multiple cloning site; RBS, ribosome binding site; rTEV.

<sup>b</sup> 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

No.	Primer name	Length (nt)	Sequence (5'-3')
<b>Classification of <i>L. plantarum</i> KW30</b>			
1	Lab16	19	CCTTTCCTCACGGTACTG
2	Lacto1	21	GAGTTTGATCCTGGCTCAGGA
3	Lacto2	21	AAGGCGGCTGTCTGGTCTGTA
<b><i>plnKW30</i> cloning</b>			
4	PreBacExF ( <i>NdeI</i> )	33	GGTATTACATATGAGTAAATTGTTAAGACACTT
5	MatBacEx/F ( <i>NdeI</i> )	34	AAACAACCATATGAAACCTGCATGGTGTGGTAT
6	BacExR ( <i>BamHI</i> )	30	TCGGATCCTTAACAATGATAATGCTACTAC
<b>NICE system</b>			
7	Pln_NaeI_F	39	GGTGGTGCCGGCATGGGTAAACCTGCATGGTGTGGTAT
8	Pln_HindIII_R	33	GGTAAGCTTCTAACAAATGATAACTGCTACTACC
9	Pln_XbaI_R	31	GTCTAGACTAACAAATGATAACTGCTACTACC
10	MatBacDPF	22	CCAAAACCTGCATGGTGTGGT
11	PreBacDPF	24	CCAATGAGTAAATTGGTTAAGACA
<b>Modification of bacteriocin</b>			
12	plnMatModNcoI	35	GGTGGTCCATGGGTAAACCTGCATGGTGTGGTAT
13	PlnMatModHindIII (Met)	31	GGTCAAGCTTCTACATATGATAACTGCTACT
14	plnModSer_HindIII (Ser)	31	GGTCAAGCTTCTACGAATGATAACTGCTACT
<b>Cloning of RR (former putative FTase)</b>			
15	FTase.RRHindIIIR	36	GGTAAGCTTTTAAAATGTCTTCATCGGGATTAAATC
16	FTase.RRNcoIF	32	GGTGGTCCATGGCTTATCAAGCAGAAGGAGAA
17	FTaseRRBglR	37	GGTAGATCTATTAATAATGTCTTCATCGGGATTAAATC
18	FTaseRRNdeF	33	GGTGGTCATATGGCTTATCAAGCAGAAGGAGAA
19	RR_R	27	CTAAAATGTCTTCATCGGGATTAAATC
20	RR_int_R	22	CGACATTGATGATGAACCTTGTT
21	RR_int_F	20	TAGTTCCACAGACTAGAATA
<b>NICE system</b>			
22	RR_NaeI_F	36	GGTGGTGCCGGCATGGCTTATCAAGCAGAAGGAGAA
23	RR_XbaI_R	35	GTCTAGAATTAATAATGTCTTCATCGGGATTAAATC
<b>Yeast system</b>			
24	RRNdeF_Yeast		GGTGGTCATATGGGTGGCGGTGGCGGTATGGCTTATCAAGC AGAAGGAGAA
25	RRBamR_Yeast		ATATCGGATCCCTAAAATGTCTTCATCGGGATTAAATC
<b>Cloning of new putative FTase target (lp_1715; PTase; <i>L. plantarum</i> WCFS1)</b>			
26	Lp_1715NcoI	37	GGTGGTCCATGGGTTTTAAAAAATGGCTAACTGGCC
27	Lp_1715HindIII	24	GGTAAGCTTTTACAACCAAGTCCC
<b>Cloning of GTase</b>			
28	GlycoT_KW30_BamHI_F	39	ATATCGGATCCATGAAAAATAGACAAAATGAAATTGACA
29	GlycoT_KW30_HindIII_R	57	TGCAATATACTATATTGAATCATTTAAAATTAATGAAATAT AGTAAAAGCTTGTCAT
30	GT_R	27	TTAATTATTCTCATTTTTGCTGTCACC
31	GTase_int_F	27	GTCGTCCTTTCATTATTATGCAAAAAT
32	GTase_int_R	26	TTTGAGTTTGTCTGAGGCAGTACAT

	<b>NICE system</b>		
33	GTase_NcoI_F	35	GGCCATGGGTAAAAATAGACAAAATGAAATTGACA
34	GTase_HindIII_R	32	GGAAGCTTTTAATTATTCTCATTTTTGCTGTC
35	GTase_NaeI_F	40	GGTGGTGCCGGCATGAAAAATAGACAAAATGAAATTGACA
36	GTase_XbaI_R	31	GTCTAGATTAATTATTCTCATTTTTGCTGTC
	<b>Yeast system</b>		
37	GTNdeF_Yeast		GGTGGTCATATGGGTGGCGGTGGCGGTATGAAAAATAGACA AAATGAAATTGACA
38	GTBamR_Yeast		ATATCGGATCCTTAATTATTCTCATTTTTGCTGTC
	<b>Other target genes of PlnKW30 gene cluster</b>		
39	ABC1KW30_Bam_F	39	GCGGGATCCAGAATAATTAACAATGACCAAAATGAT
40	ABC1KW30_Not_R	38	TAGCGGCCGCTCAAATGAATCTACTACTAACCGCTTCT
41	ABC2KW30_NcoI_F	27	TATCACCATGGGCTTGTTTAAATCACA
42	ABC2KW30_XhoI_R	26	ACCGCTCGAGTCATAACAATGCCTCC
43	Trx1KW30_NcoI_F	26	TATCACCATGGGCAAAAAAACCATTA
44	Trx1KW30_XhoI_R	26	ACCGCTCGAGCTAAAGATCTCCCGTG
45	Trx2KW30_Bam_F	43	GCGGGATCCAATCTATTAAAAAAAGAAAAGTAGTATTAAT TT
46	Trx2KW30_HindIII_R	34	CGCCAAGCTTCTAAATTTTCTCCTTCTGCTTGAT
	<b>Modification of pET vector</b>		
47	TEVNcoIEcoRIHindIII	47	ACCGAAAACCTGTATTTTCAGGGCGCCATGGGAATTCAAGC TTGGTC
	<b>Gene disruption using Russell &amp; Klaenhammer 2001</b>		
48	KW30FTaseKOBamF	29	CCAAAACCTTTATATGTTCACTCTCTATCG
49	KW30FTaseKOEcoR	29	CGGACTTGATATTCATCATATTTCAAATG
	<b>Gene disruption using cre-lox system (Lambert <i>et al.</i> 2007)</b>		
50	Pln_KO_1F	24	AAGGGATGGACTTCATGGGTACAG (GC 50%, Tm 57°C)
51	Pln_KO_1R	26	AACCAATTTACTCATTCTAATACCC (GC 35%, Tm 53°C)
52	Pln_KO_2F	22	AGTTATCATTGTTAGTTTGTG (GC 27%, Tm 41°C)
53	Pln_KO_2R	21	TCCAATAAGGGACTTTCATTG (GC 38%, Tm 49°C)
54	Pln_KO_1F_XhoI	33	GGTCTCGAGAAGGGATGGACTTCATGGGTACAG
55	Pln_KO_2R_BglII	30	ACCAGATCTTCCAATAAGGGACTTTCATTG
56	GTase_KO_1F	22	TACCGAAAAGACGAAGGGCACC
57	GTase_KO_1R	25	GTCTATTTTTTTCATTCTCATTCTGTC
58	GTase_KO_2F	25	GGTGACAGCAAAAATGAGAATAATT
59	GTase_KO_2R	25	ATATCTTGGCTAAATGTTTCCTGTG
60	GTase_KO_1F_XhoI	31	GGTCTCGAGTACCGAAAAGACGAAGGGCACC
61	GTase_KO_2R_BglII	34	ACCAGATCTATATCTTGGCTAAATGTTTCCTGTG
	<b>Cre-lox verifying primer</b>		
62	85	24	GTTTTTTTCTAGTCCAAGCTCACA
63	87	22	GCCGACTGTACTTTCGGATCCT
	<b>Housekeeping genes</b>		
64	gap_F	25	AAATTGGTATTAATGGTTTCGGACG
65	gap_R	20	AATGCTAATGGTGCAAGGCA
	<b>Sequencing of <i>plnKW30</i> gene cluster</b>		
66	BacLin/F (degenerate primer for first sequencing)	26	ATGAARCCNGCNTGGTGTGGTATAC

67	BacCon/R (degenerate primer for first sequencing)	24	ATGATGTTTRATNCCAAARCARTG
68	F6ii	26	AGATGATTGAACGAAAGACGGGGATT
69	F6(CT)	22	GAACGAAAGACGGGGATTAGTA
70	LplgDNAF5ii	28	CAAATTGAGTTACTAAAATTGGCAGGAG
71	LplgDNAF5	28	CGAGGCCTTAGCCCATTTATCCAAGAGT
72	LplgDNA/F4	28	CTATGGCTTTCTGATACCAATTCATTCG
73	LplgDNA/F3	29	AATGCACCAATTGTACGTGACTAGAACAA
74	LplgDNA/F2	30	CAATGGACGTTCTAAATTTGGATTTGATAC
75	LplgDNA/F	28	ATGTGTGGTGCTGGTTATGATTTTCGGAA
76	LplgDNA/R	28	TTCCGAAATCATAACCAGCACCACACAT
77	LplgDNARgap1	28	GCTGAATAGCACTGGCTGTTTCGGAGTT
78	LplgDNARgap2	28	GTTGTGACCAAATCTTGCTAAAAAGCTA
79	LplgRgap2ii	33	ACCAAATCTTGCTAAAAAGCTATTGAGAATGTC
80	LplgDNA/R2	30	AGCGCTAACTGCTAATTATTCACCTAGAGA
81	LplgDNAR3	30	ATTAGAATGTTACAAATATCAGCAGCCAAT
82	LplgDNAR3ii	29	CGGACTTGATATTCATCATATTTCAAATG
83	R4	25	CATCTAAAGATCTCCCGTGTACCCT
84	R4i(CT)	22	CGAAGTACGTTCTCATGTTTCAT
85	R4ii(CT)	22	CCATAGTGTTTGTCTTTGATTC
86	R5(CT)	25	CCAAGAGTTATATTATTCATAACTG
87	R5ii(CT)	25	TCAAAACCAATACTACTAGCAATCC
88	R6	25	TGAATCTACTACTAACCCTTCTTT
89	R7	26	ATCACTGTCTTCCCCACCTGTTATCT
90	R8	23	CATTCTCATTCCTGCTTTCGTGG
91	R8a	21	GCAAGGAGTTCGGAACATCTC
92	R9	23	GTCATCAAACGTTTGATTAACCG
93	R10	26	CACGAAATTAGCTTGCATACTGTTTG
94	R11	26	CTTCTTGCTGAATCGCATCATCTTTG
95	R12		CAAGAGATATGGATTGAAGTC
96	R13		ACCGAAGCTGCAGTGGTTGA
97	R14	22	CCAAGAGCAACTTATATTTATG
98	FR9	27	GAAAATATCTTACTGTTAGGAACTTC
99	FR10	22	GTATAGCAATTCCTTATAAGC
100	FR11	21	CGTGAGATTCATCCTGGTGAA
101	FR12	24	TCAAATAAGTTATTAATCGCGATT
102	FR15	24	CTTGATTTAGAACAATTGCAACAA
103	20103_R	21	GACAGCGTCAATCCATAAATG (1173-1193 bp; 8 mismatches)
104	Rsub13	23	CAGTGGTCGAGAAATCCTGCTTA
105	Ex18R1	23	GGGAATGGATTATCGTTGTAGTC

## 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 PlnKW30 was actually an *N*-acetylglucosamine rather than a farnesyl group.

### 5.1 INTRODUCTION INTO FARNESYLATION

Prenylation is a lipid modification where the C<sub>1</sub> 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 carboxymethylated 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).

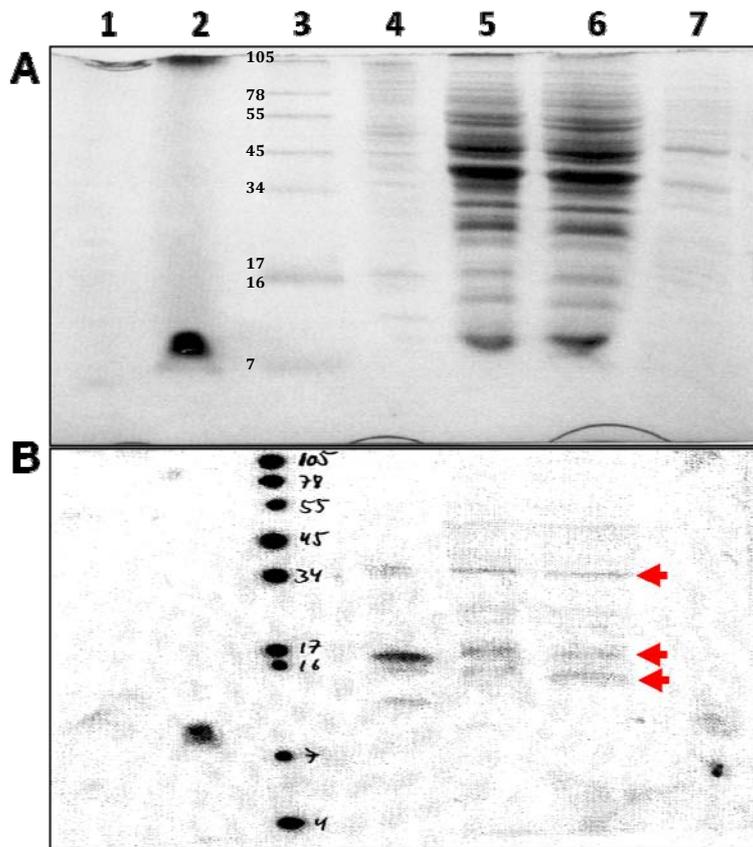
To date, no bacterial FTases have been described, although a recent publication shows that the structure of the lanthionine cyclase NisC produced by *L. lactis* is similar to the  $\beta$  subunit of mammalian FTase. NisC is involved in the biosynthesis of the five cyclic thioethers of the lantibiotic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



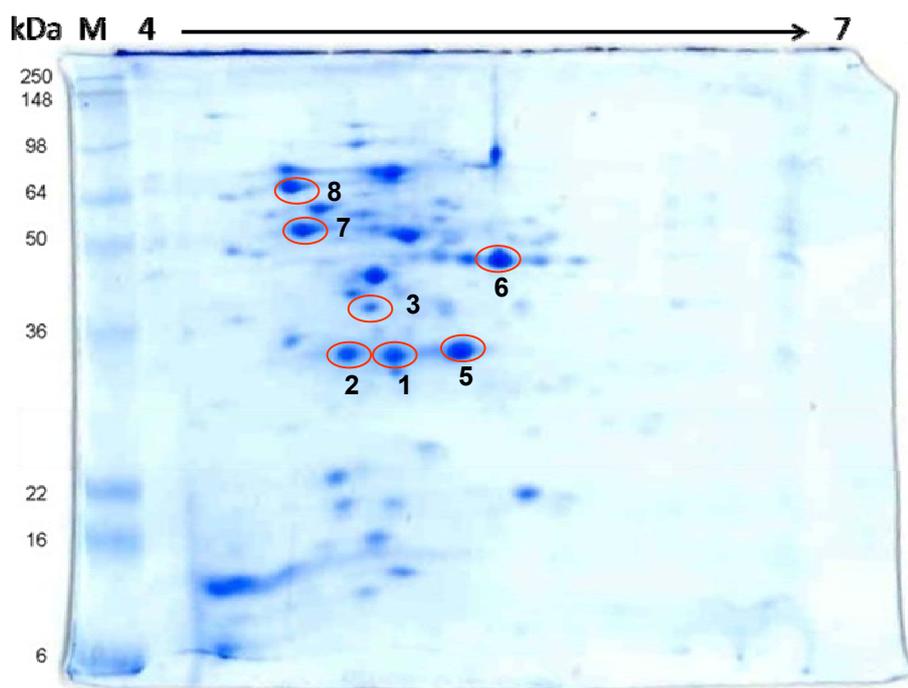
**Figure 5.1: SDS-PAGE (A) and Western blot (B) using anti-farnesyl antibodies.**

**1:** recombinant PlnKW30; **2:** native PlnKW30; **3:** prestained protein standard; **4:** whole cell extract of MDA human cell line; **5:** whole cell extract of log phase *L. plantarum* KW30; **6:** whole cell extract of stationary-phase *L. plantarum* KW30; **7:** whole cell extract of *Streptococcus pyogenes*.

The native PlnKW30 (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).



**Figure 5.2: 2D-electrophoresis gel of whole cell extract of *L. plantarum* KW30.**

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.

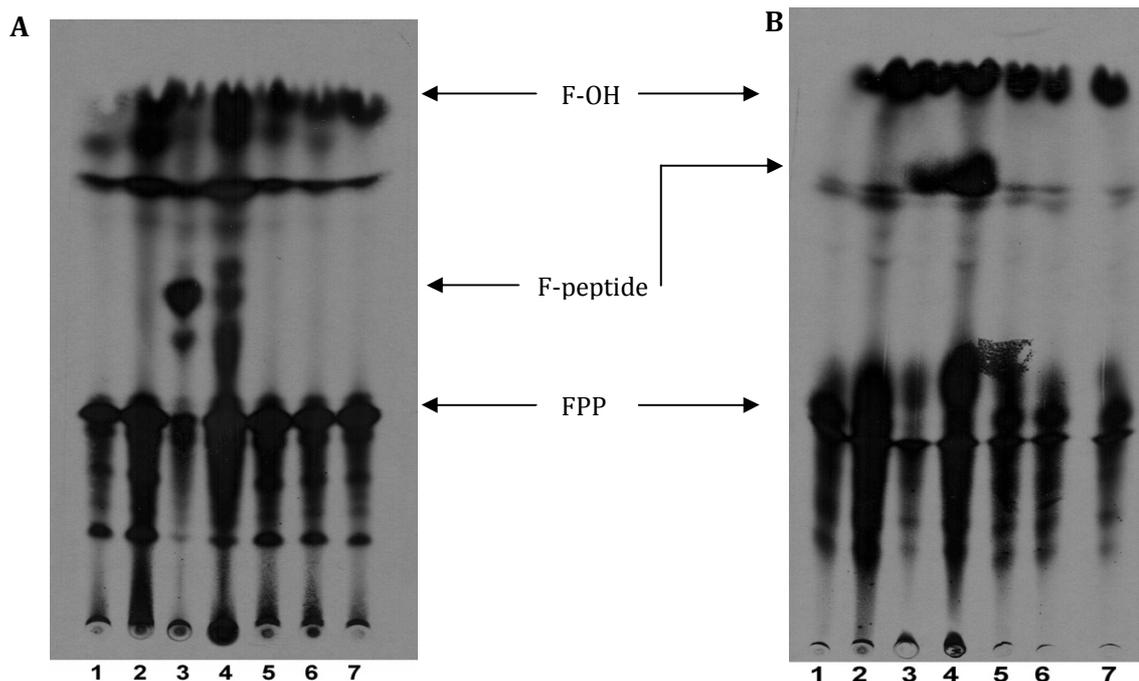
### 5.2.3 FARNESYLTRANSFERASE ACTIVITY ASSAY

In order to detect the putative farnesyltransferase (FTase) activity in different *L. plantarum* KW30 cell extracts, an *in vitro* FTase activity assay was developed. Firstly, the assay was developed using crude *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 *et al.* 2001). Later, *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 ( $^3\text{H}$ -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  $^3\text{H}$ -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 *et al.* (1995), it was possible to assign identities to spots in figures 5.3A and B.

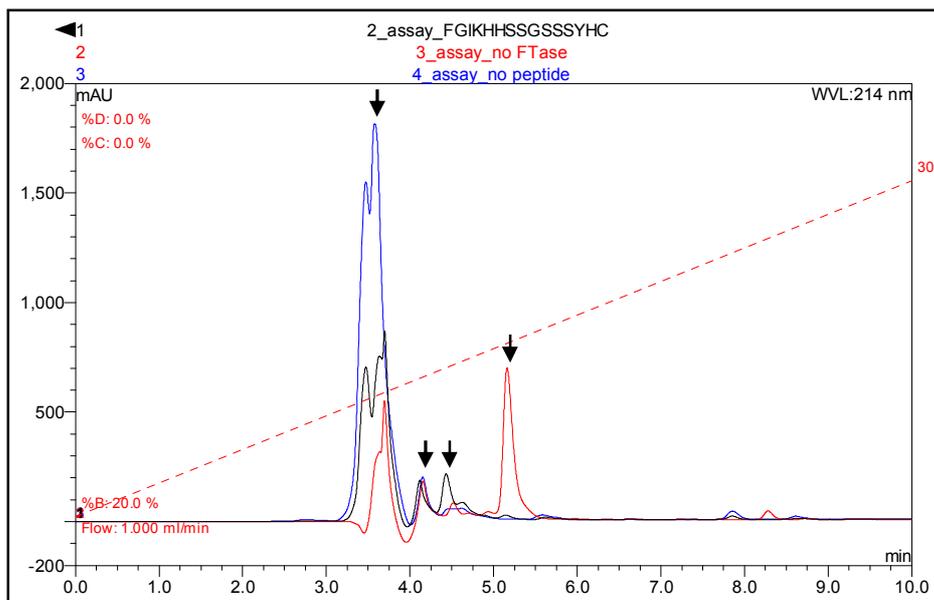
The TLC plates shown in figure 5.3 were sprayed with ninhydrin (0.2 % in ethanol) and heated at 110 °C for 5 – 10 minutes. All spots seen on the autoradiography films were also visible after ninhydrin treatment.



**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; 1  $\mu$ L  $^3$ H-FPP; **2:** Pure FTase; 3  $\mu$ L  $^3$ H-FPP; **3:** Yeast extract; 1  $\mu$ L  $^3$ H-FPP; **4:** Yeast extract; 3  $\mu$ L  $^3$ H-FPP; **5:** No FTase; **6:** No peptide; **7:**  $^3$ H-FPP only, 1  $\mu$ L.

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.

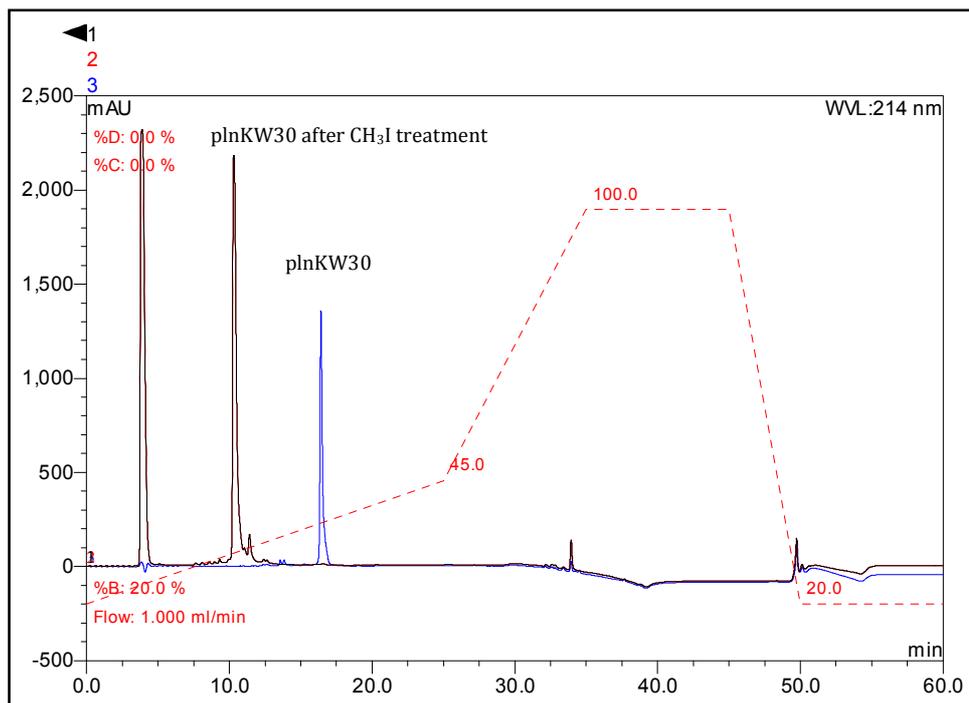


**Figure 5.4: FTase activity assay separated by RP-HPLC.**

A Jupiter C18 column was used with solutions A: H<sub>2</sub>O, 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 METHYLIODIDE TREATMENT

PlnKW30 was treated with methyl iodide (CH<sub>3</sub>I) 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 CH<sub>3</sub>I-treated sample had a lower retention time than PlnKW30, eluting at 28 - 29 % B compared to 36 - 37 % B, respectively.



**Figure 5.5: RP-HPLC chromatogram of native PlnKW30 and CH<sub>3</sub>I-treated PlnKW30.**

A Jupiter C4 column was used with the mobile phase a step gradient made up from A: H<sub>2</sub>O, 0.1 % TFA and B: Acetonitrile, 0.08 % TFA. **1 (black)**: PlnKW30 (eluted at 36 – 37 % B); **2 (blue)**: CH<sub>3</sub>I-treated PlnKW30 (eluted at 28 – 29 % B).

Purified samples of native PlnKW30 and CH<sub>3</sub>I-treated PlnKW30 were subjected to mass spectrometry (MALDI-TOF, Micromass). The mass spectrometry results showed that the disulfide bonds of PlnKW30 were broken during the methyl iodide treatment, which would result in deactivation of the bacteriocin (Section 3.6.1).

### 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-*S*-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 farnesyltransferase 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 farnesyltransferase, although this gene had no sequence homology to other farnesyltransferases 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 pyrophosphate. 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-*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<sub>3</sub>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.**

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

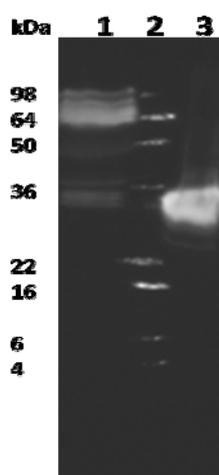
Different growth temperatures and IPTG concentrations (or the respective induction peptide) were examined for each expression strain. One construct (pMAL\_C2G\_HST::RR) produced a small amount of soluble protein in *E. coli* TB1. However, cleavage of the maltose

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.



**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 *plnKW30*, RR and GT in this thesis**

Construct	Restriction sites used	Modifications in vector	Expression host tried	Expression <sup>a</sup>	Primer No. <sup>b</sup>
pT7-7 ::PlnKW30	NdeI – BamHI	-	<i>E. coli</i> BL21 (DE3)	-	5, 6
pT7-7::RR	NdeI - BglII	-	<i>E. coli</i> BL21 (DE3)	-	17, 18
pT7-7 ::PlnKW30	NdeI – BamHI	-	<i>E. coli</i> BL21 (DE3) Rosetta	-	5, 6
pT7-7::RR	NdeI - BglII	-	<i>E. coli</i> BL21 (DE3) Rosetta	-	17, 18
pT7-7::RR	NdeI - BglII	-	<i>E. coli</i> BL21 (DE3) Origami pLysS	-	17, 18
pET32a(+)_HST ::RR	NcoI - HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> BL21 (DE3)	-	15, 16
pET32a(+)_HST ::RR	NcoI - HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> BL21 (DE3) Origami	-	15, 16
pET32a(+)_HST ::PlnKW30	NcoI - HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> BL21 (DE3) Origami	-	8, 12
pMAL_C2G_HST ::RR	NcoI – HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> BL21 (DE3)	-	15, 16
pMAL_C2G_HST ::RR	NcoI – HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> TB1	+	15, 16
pMAL_C2G_HST	NcoI – HindIII	Addition of His-tag, spacer, rTEV	<i>E. coli</i> BL21	-	15, 16

::RR		cleavage site	(DE3) Rosetta		
pMAL_C2G_HST ::RR	NcoI – HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> BL21 (DE3) Rosetta- gami	-	15, 16
pMAL_P2G_HST ::RR	NcoI – HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> BL21 (DE3)	-	15, 16
pMAL_P2G_HST ::RR	NcoI – HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> TB1	-	15, 16
pETDuet_HST_Trx A::PlnKW30	NcoI – HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> BL21 (DE3)	-	8, 12
pETDuet_HST_ TrxA::RR	NcoI – HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> BL21 (DE3)	++, IB	15, 16
pETDuet_HST_ MalE::PlnKW30	NcoI – HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> BL21 (DE3)	-	8, 12
pETDuet_HST_ MalE::RR	NcoI – HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> BL21 (DE3)	+, IB	15, 16
pETDuet_HST_ MalE::RR	NcoI – HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> BL21 (DE3) Rosetta	+, IB	15, 16
pETDuet_HST_ DsbC::PlnKW30	NcoI – HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> BL21 (DE3)	+	8, 12
pETDuet_HST_ DsbC::RR	NcoI – HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> BL21 (DE3)	-	15, 16
pETDuet_HST_ SUMO::PlnKW30	NcoI – HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> BL21 (DE3)	-	8, 12
pETDuet_HST_ SUMO::RR	NcoI – HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> BL21 (DE3)	+, IB	V
pETDuet_HST_ SUMO::RR	NcoI – HindIII	Addition of His-tag, spacer, rTEV cleavage site	<i>E. coli</i> BL21 (DE3) Rosetta	++, IB	15, 16

pSIP409 ::PlnKW30	NcoI – HindIII	<i>L. plantarum</i> NC8	-	8, 12
pSIP409 ::PlnKW30	NcoI – HindIII	<i>L. sakei</i> Lb709	-	8, 12
pSIP409::RR	NcoI – HindIII	<i>L. plantarum</i> NC8	-	15, 16
pSIP409::GT	NcoI – HindIII	<i>L. plantarum</i> NC8	-	33, 34
pSIP409::GT	NcoI – HindIII	<i>L. sakei</i> Lb709	-	33, 34
pNZ8148 ::PlnKW30	NcoI – HindIII	<i>L. lactis</i> NZ9000	-	8, 12
pNZ8148::GT	NcoI – HindIII	<i>L. lactis</i> NZ9000	-	33, 34
pNZ8112::RR	NaeI - XbaI	<i>L. lactis</i> NZ9000	-	22, 23
pNZ8112::GT	NaeI - XbaI	<i>L. lactis</i> NZ9000	-	35, 36

<sup>a</sup> '-': no expression; '+': expression; '++': good expression; IB: inclusion bodies

<sup>b</sup> numbers refer to primers listed in table 5.3, appendix 4

## 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 PlnC 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.

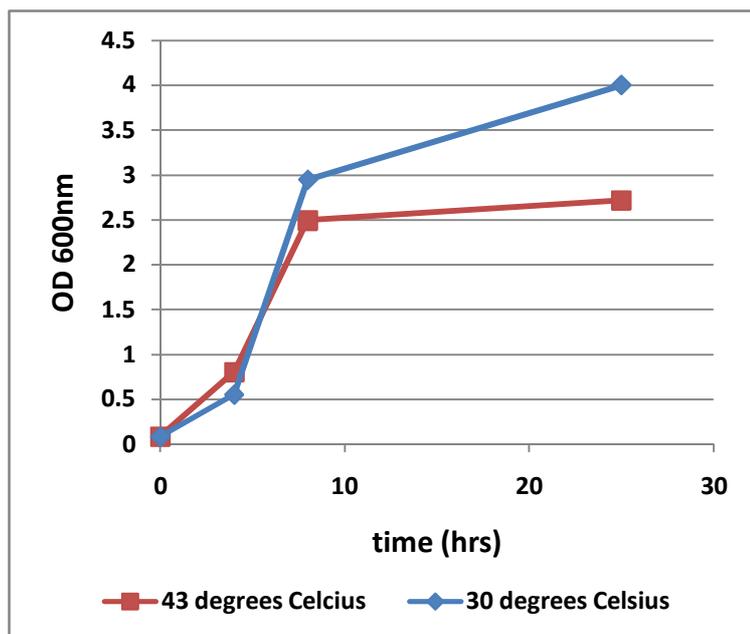
### 5.6 TARGETED GENE DISRUPTION USING A TEMPERATURE SENSITIVE PLASMID

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<sub>600</sub> reaches only ~2.6 compared to an OD<sub>600</sub> 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.



**Figure 5.7: Growth curves of *L. plantarum* KW30 at 30 °C and 43 °C.**  
The results shown are the average of three independent experiments.

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.**

<b>1. test</b>			
Temperature tested	30 °C	43 °C	43 °C
Number of transfers	3	None	None
OD <sub>600</sub>	3.857	Only minimal growth after two days	Only minimal growth after two days
<b>2. test</b>			
Temperature tested	30 °C	40 °C	40 °C
Number of transfers	3	6	6
OD <sub>600</sub>	3.791	3.370	3.378
<b>3. test</b>			
Temperature tested	30 °C	41 °C	41 °C
Number of transfers	11	3	3
OD <sub>600</sub>	4.0	2.919	2.928
Number of transfers	15	7	7
OD <sub>600</sub>	3.842	2.912	2.930

## 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 *PmeI* restriction site and a ~1 kbp fragment downstream of the *gccA* (*plnKW30*) gene into the *Ecl136II* restriction site. The *PmeI*-

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 *L. plantarum* KW30 could be solved, as this system allows multiple gene disruptions with the same antibiotic resistance-cassette.

### 5.8.1 CONCLUSIONS

Despite several attempts to disrupt the *plnKW30* (*gccA*) or putative FTase (now putative response regulator; *gccF*) genes using the two different systems, no gene disruption was observed.

APPENDIX 8: 16-23SRDNA SEQUENCE OF *L. PLANTARUM* KW30

ATTGGTGCTTGCATCATGATTTAAGCTACTAAGGTGGCGAACTGGTGAGTAACACGTGGGAAACCTCCTTAAAAGCG  
 GGGGATAACACCTGGAAACAGATGCTAATACCGCATAACAACCTTGGACCGCATGGTCCGAGTTTGAAAGATGGCTTC  
 GGCTATCACTTTTGGATGGTCCC GCGGCGTATTAGCTAGATGGTGGGGTAACGGCTCACCATGGCAATGATACGTAGC  
 CGACCTGAGAGGGTAATCGGCCACATTGGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTAGGGAATCTT  
 CCACAATGGACGAAAGTCTGATGGAGCAACGCCGCGTGAGTGAAGAAGGGTTTCGGCTCGTAAAACCTCTGTTGTAA  
 AGAAGAACATATCTGAGAGTAACTGTTTCAGGTATTGACGGTATTTAACCAGAAAGCCACGGCTAACTACGTGCCAGC  
 AGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTTTTAAGTC  
 TGATGTGAAAGCCTTCGGCTCAACCGAAGAAGTGCATCGGAAACTGGGAAACTTGAGTGCAGAAGAGGACAGTGGAA  
 CTCCATGTGTAGCGGTGAAATGCGTAGATATATGGAAGAACCAGTGGCGAAGGCGGCTGTCTGGTCTGTAAGTGA  
 CGCTGAGGCTCGAAAAGTATGGGTAGCAAACAGGATTAGATACCCTGGTAGTCCATACCGTAAACGATGAATGCTAAG  
 TGTTGGAGGGTTTCCGCCCTTCAGTGTGCGACTAACGCATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGCTGA  
 AACTCAAAGGAATTGACGGGGGCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCTACGCGAAGAACCTTAC  
 CAGGTCTTGACATACTATGCAAATCTAAGAGATTAGACGTTCCCTTCGGGGACATGGATACAGGTGGTGCATGGTTG  
 TCGTCAGCTCGTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTATTATCAGTTGCCAGCATTAAAGT  
 TGGGCACTCTGGTGAGACTGCCGGTGACAAAACGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGACC  
 TGGGCTACACACGTGCTACAATGGATGGTACAACGAGTTGCGAACTCGCGAGAGTAAGCTAATCTCTTAAAGCCATTC  
 TCAGTTCCGATTGTAGGCTGCAACTCGCTACATGAAGTCGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGA  
 ATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGAGAGTTTGTAAACCCAAAGTCGGTGGGGTAACCTTT  
 TAGGAACCAGCCGCTAAGGTGGGACAGATGATTAGGGTGAAGTCGTAACAAGGTAGCCGTAGGAGAACCTGCGGCT  
 GGATCACCTCCTTTCTAAGGAATATTACGGAAACCTACACACTCGTCGAAACTTTGTTTAGTTTTGAGAGATTTAACT  
 CTCAAAACTTGTCTTTGAAAACCTAGATAATATCAAATATATTTTTTTCATAATGAAACCGAGAACACCGCGTTTTTT  
 GAGTTTTTTTATTGAAGTTTAATTATCGCTAAACTCATTAAATCGCATTACCGTTAGGTAAATGAGGTAAAGTTAACA  
 AGGGCGCATGGTGAATGCCTTGGCACTAGGAGCCGATGAAGGACGGGACTAACACCGATATGCTTCGGGGAGCTGTA  
 CGTAAGCTATGATCCGGAGATTTCCGAATGGGGCAACCCAGCAGTTTTAATCAACTGTTACCACTAGATGAATTCATA  
 GTCTAGTTGGAGGTAAACGCTGTGAACTGAAACATCTCATTAGCAGCAGGAATATAAAGAAATTTTCGATTCCTAAG  
 TAGCGGCGAGCGAACGGGGAACAGCCCAAACCAAAGTGCTTGCACCTTTGGGGTTGTAGGACTGAACATTTGAGTTAC  
 CAAAGAACTTGATAGTCGAAGGATTTGGGAAAATCCGCCATAGATGGTGATAGCCAGTAGATTAATCAAATTCCTC  
 TCAGTTCAGGATCCAGTCTAATTGCCGGAACACGTGAAATTCGCTCGGAATCCGGGAGGACCATCT

## 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):

atgaaaaatagacaaaatgaaattgacagttattgaaccttcatttaaggcctgtccataaatcatttgattttggaaatttaacaaacattgatcaatttagacat  
 catatttatgtatcataattgtaatatgtaaaaaattcctaagcaactattgaaagatgtgtgaaattcaattgctcaaaaatggaaaatggagatgaaattgattgtt  
 ttgacactggttcaactgatgaaactgtgcatctgtttaaagaatagccacaagcaaaaatcagtaacaaaatggaaaaatgatttctcagaggttaga  
 aataaggcactcaaattagcatctaaagactgggtgttctatgttgatagtgatgagtggttgatgtgatggagctcagctaaagaaaattttattaaagt  
 tcaagcaaaaatttaagttcgtaatcaaccaacattttctgacattcaggacagatataccaacagtagggagaatcttccctaaaaagtcgctcttcatt  
 attatgcaaaaattcacgaagaagtaagaaaagaagtaaaaaactaggttacgatgtaggacacttcgcatgtgatgataattctttatcatgatggataga  
 taaagaagtattacagataaagataaaaataagagaaaacttcgtttattacaagagatgactgtgaggagcctcaaaatgctagatggccgttttactagc  
 acgtgatggatttgatgtagctcctcaggacaaactcaaacagttagtaaaaaggaccttagacttagttgctccgacagcttcaagaaaaatattcgccattt  
 gcaaaaaattgctggccgtattttattaagagaaggtaaaaccaactcaagcagattattcttttaagatgttttacagataacaggtggggaagacagtgat  
 gcaatatactatattgaaatctttaaaataatgaaattatagcgaagtaagtctatagaaagtaaaaatgttgagatattaaataaacacaaaggaatgattg  
 atgttaatagcgtatatactggaaactactatcatatagctcaggttattttagaatgtgacataaataagtgccaactacagccatctttcccttaatatcagaaa  
 ttccataaaaactttcaggcgatataaaaagtagtgcaaatctgcagtgaaactgtattcaaaattacaaggtgacagcaaaaatgagaataattaa

ABC-transporter nucleotide sequence, 2121 bp (5361 – 7481 bp in appendix 10):

gtgacagcaaaaaatgagaataattaacaaattgacaaaatgattgtggaccagcagcagtagcaacaatcattggatgattagtaaaatataaattgatga  
 ctggcagataaggcaaatatcacaacaacaaatattgggactaattttatcagataataaaacggataaagtcagttcaggtgagctgttgtagaaaagtgt  
 gtaaagaatccagatgtgtttgatgaaattgaattccagtgctgacacaaatgatgcaaaaatggctatttactttgtagttctcagaaaggttctggttcaaaa  
 ttatattgggcagaccaggtagtggaataatagaaagattgacaaggctaaattatgatgattggacgcccttatcttaacgatacctagatcacaacaaat  
 ttcaaatctaaaaacaaaagttgaaagcttaaaagtcctattttgtggttagtttaggaaacactgggcattgctgttaataatattggatgctttcaataat  
 atttgctacaacactattcttagcaactatgtattctgcctatttaaccgattgttccaaataagttatttcagcgattcctactataactctcgtatattgactgtt  
 caagtagcatcaatttttggaagttattaatgcaatcctattgcctatactggtaattcaatcagtaagaattatgaagacatcttaaggctttcccaaaa  
 agaagcggtagtagtagattcattgaagatggtgaaattattacaagattatcgattatcaaatgcaggtagacaactgtgcttttaggtgacttacc  
 attgacattactatggggataacccttttcttctgtagaatcaatccattttgacgattttattgtttatccaatgatattatggcattattttatctaa  
 ttacagcgaattaaaaaagaagtacaaattgtttaaatcacaggaacatttagccaagatattattgacaacataaaaaattttcaactataaaaaactttt  
 catgctgtagctacatcaattccaaaataacaaaagtagtggaaaagtatactgcagcaaacctgattttggcttatgatagattacagcagatagcaaga  
 aatggtgttcacaactgtttggaattttattctcgggtggggctttctagtaataaagggcaatattgtgcttggaaactgcttagttttatgctttaacggg  
 aatgttctaaatccgtttgtgaaaataataatgcaagtaaatctgctcaaggaaagtagcaacagaacgctacttagatctattactttctcgtgaagac  
 aataacaataaagcaaaactaacactttcaaatctattgaaataagaaggttagcattttcatatgatggggttacaccaataataaaaaacatgaacggtggtt  
 tctcaagtttcaactgcaataattggcaatagtgaggagtggaagtagcaactattgctaaactaatggcaggttttatctcagagggctctatcattaggaaa  
 cataccttattgcaataggggcagaaatcttaggaatacaattacatgttgacaacatcaccacaaatatttccgatacagttatgaaataataactcttgg  
 taggaaaaatcacagttaaatcaatattaggattgctagtagtatggttttgacgaatgtttgaatctttacaaaacggaaatatacacaaataggatcct  
 ctggatattcattatctggaggtcaaaaaatattaaatattataaggtctatggaattccaagtaggtttatcttttgatgaaataacgaatgggctggaca  
 tcaatacaaaaattgaaagtagaaaattttattatcacaacaaactaaatattttattactcagatttaacattggctataagatgtgatgatattttca  
 ctgttaaaaatggtgagatgcacaaaatagaaactcaaaaaactctactgaatcattgaggaggcattgttatga

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

atgaaaaaaaccattatcactattattgccttagtgttatctcaataataggagatattcttggcaaaggctcactttgtacaagagccattatcactcaata  
 aaggcgacacaattgcaagcagacataaaaaataagaaggatgatgttagttacttctacaaaaaggtgtgaaggatgtgaacatagctacctccatattgaa  
 ccgcttattcgaaataaacgtgtcactgtaaagcaatcgatataaatgcagatccaagcaaagattttatttgggaacattgggaatatcatcaaccaacc  
 atattgtttataccatggaaaaattgttcaaaggataaatcaagctttggatagaggacttaacaactaaataaaggggtacacgggagatctttag

TRX 2 nucleotide sequence, 534 bp (7898 – 8431 bp in appendix 10):

atgaatctataaaaaaagaaaactagattattttgctaacttctcattatttttagtaattggatttttcgggtattattacgaacggaattcttatga  
 acatgagaacgtacttcgggctgaaattattcagaatcaaagacaacactatggttacaaggaattaaattcaataacaattccaacttttggaaaaggaaatt  
 tcggataaaaagagattttctagatataattggacggcctactgttcagattgtaattttatgatcctattttgtaataatgaattgaaaaagaagatacgaccagc  
 aatgtgtcttttcaacgtagcatgggaacccaaaaggatggacttcaggtgacagtttaagaaaaaatggattcaagcagacacctgcaataattcac  
 atcacaatggcaaaagtattatacagtgggggaataataaagggataagtaaggagatttacactatggcttatcaagcagaaggagaaaaattta  
 g

LytTR nucleotide sequence, 777 bp (8401 – 9177 bp in appendix 10):

atggcttatcaagcagaaggagaaaatttagttccacagactagaatattgttttaacaaataaacgggaatgccaaaatttaataacatataattacaata  
 atgctcaattatcgaatcatatcaagtttgcacccaataaaaactatcatcaagctctgaaaatttccaaaactttatattgctcctctatcgtattttaattt  
 agatgaagaaatagacttaataaaaaaattagattgctgtcaagaggcttctaattcacggaaatacttattacattggatcaagtagacatgttttgaa  
 aacttaatttcaatcattatagctttatgatgactaatgacaaaactttcaaaaagatgatggaaatattggctgctgatattgtaacattctaatacaa  
 attctattcccttaaatagttcaaacactatacaaatctgtcaaaaagacagaagtttttctaagattggatgatataattttgtaataactgacaatcaaccgca  
 taagtaataatttcaattgaaatgatgaatatcaagtcctgatgactttaaacaagttcatcatcaatgtcgtaatcttactgcttgccatgaaggagccttagt  
 taacctaaaaaacactaaatttaacaagagcaataggatcctataatttgaaaatggtgatacttgcgaggtttcaagaagatacaaaaaagattagtgat  
 ttaatacccgatgaagacattttag

PlnKW30 nucleotide sequence, 195 bp (9255 – 9449 bp in appendix 10):

atgagtaattggttaagacacttactataagtgaaatttctaaggctcaaaacaacggtgaaaaactgcatggtgttggtatacttttagcaatgtgtggtgctg  
 gttatgattcgggaacctgtgattatagttatcgattgttttggtataaagcatcatagtagtggtagtagcagttatcattgttag

GTase amino acid sequence, 422aa

MKNRQNEIDSYLNLHLRPVHKSFDGFLTNIDQFRHHIYVSYIVICKNSQATIERCVNSIAQNMENGDELIVLDTGSTDE  
 TVHLVKKNMPQAKISVTNWKNDFSEVRNKALKLASKDWVYVDSDEWLDVDDGAQLKKILFKVQAKNFKFVINPTF  
 SDHSGQIYQTVGRIFPKKSSFHYAKIHHEVRKEDQKLGVDVRHFACDDIILYHDGYDKEVLRDKDKIKRNIRLLQEMTC  
 EEPQARWPFLLARDGFDVLPQDKLKQLVKRTLDELVASDSLQEKYSPFAKKLLGRILLREGKTTQAVLSFKDVLQITGG  
 ESDAIYYIESFKINEIIAEAKSIEVKMLRYLNKHKGMIDVNSDISGNYHIAQVILECDIISANYSHLFLPLISEIPKNFSGDI  
 KSSVKS AVKLYSKLQGDSKNENN

ABC-transporter amino acid sequence, 706aa:

MTAKMRIKQIDQNDCGPAAVATIIGMISKIKIDDWQIRQIITTTNYGTNFISIINGLKSQVESVVEKCVKNPDPVDEIEF  
 PVL TQIMQNGYLHFVVLTKCSGSKLYWADPGSGKIESIDKAKFMMHWTPILITIPRSQNFQILKTKVESLKVPIWFSF  
 RKHWA CLLIYWMLSIHIFATTLFLATMYSAYFNRIVPNKFISAIPTITLVYLVQVASFLEVINAILIAYTGNSISKELYEDI  
 FKAFPKKKRLVVDSEFEDGEIITRFNAIYQIAGRQLVLLGDLPIDIITMGITLFLARINPFLTILLFIPMILMALLFYLSSDAI  
 KKRSTNLFKSQETFSQDIIEQIKNFSTIKTFHAGSYINSKITKSMKEYTAAANRSFGLYDRLQIARNGVSQLFGIILFSVGA  
 FLVIKGNIVLGTLLSFYALTGNVLPFVKIITMQVNLSQGVATERYLDLLSPEDNNKSKLTLNSIEIRKVAFSYDGV  
 TPIIKNMNGVFSKFPTAIGNSGSGKSTIAKLMAGFYPAEGLSLGNIPYYAIGAESLGNTITYVEQSPQIFSDTVMNNITLG  
 RKNITVKSINRIASSIGFDECLNHLPNGIYTQLGSSGYSLSGGQKQLLNIRSMVIPSRIIFDEITNGLDINTKLKVENYLLS  
 QNKPKLIFITHDLTLAIRCDDIFTVKNEMHQIETNKNSTESLREALL

TRX 1 amino acid sequence, 139aa:

MKKTIIITHALSVISIIGVYSWQRLTLVQEPLYHSIKATQLQADIKNKKDDVVYFYQKGCEGCEHSTPILNRFIRNKRVTVK  
 AIDINADPSKDFILGTLGISSTPTILFIHHGKIVQRINQVFGYEDLTTKYKGYTGDL

TRX 2 amino acid sequence, 177 aa:

MNLLKKRKLVLILLILLIILLVIGLFSVIITKRNSYEHENVLRAEIISESKTNTMVYKELNSITPTFEKEISDKRDFLVYIGR  
 PTCSDCNLFDPILVNELKKEDMTSNVVFLNVAWERQKGWTSWVQFKKKYGFKQTPAIIHYHNGKVLSHIQWGNKKGIS  
 KGDHLHLWLIKQKEKI

LytTR amino acid sequence, 258 aa:

MAYQAEGENLVPQTRIFVLTNKPECQNLNNHILQIMLNLSNHIKFCHPIKLSSSENFQNFICLSIVILNLDEEIDLIKKIS  
 IAVKRLPNSRKYFIYIGSSRHVFENLIFNHMYLYDALMTKTFQKDDGNILAADICNILIHNSIPLNSSNTIQISVKRQKFFL  
 RLDDILFVITDNQPHKVIFHLKYDEYQVRMTLQVHHQCRNLTAHEGALVNLKNITKFNKSNRILYFENGDTCEVSR  
 RYTKKISDLIPMKTF

PlnKW30 precursor protein, 64 aa:

MSKLVKTLTISEISKAQNNGGKPAWCWYTLAMCGAGYDSGTCDYMYSHCFGIKHHSSGSSSYHC

APPENDIX 10: COMPLETE SEQUENCE OF *PLNKW30* GENE CLUSTER (14149 BP)

```

1 ATGTGCCGAA TCCTCGGTGT TTCCAGAGCT CAGTATTATC GTTATCGATC
51 CCCCAAACCT TCAAAACGCC GGGCCGAAGA TGC GGACTTG AAACAACGAA
101 TTCTGCGGAT CTTTGCGGAA TTTAAGCAGC GATACGGTGT TATGAAGATC
151 CACCATGAAT TGAATCTGGA ACTTCAACCA CTGCAGCTTC GGTGCAGTCC
201 ACGACGGATT TCCC GGCTCA TGAAGAACT GGATATCCAC TCCGTTACCG
251 TCAATAAAGT GAAAGCGGCT TCGGCTTCCA AAACCAAGGT TGAACAGCGT
301 CCAAACCTGC TTAAGCAGGA TTTCTCGACC ACTGGTTTAA ATCAAAAATG
351 GACCGCTGAT ATGACCTATA TTCAAAACGAA GCGTAATGCT GGTGTTACTT
401 ATCAACCATC ATGGACCTGC ACTCACGACG GATTATCGGC TATTCGTTCT
451 CAAAAAAGAT GGCTACTGAT TTAGTCTTAA AGACCCTTGA AAGCGCGGTT
501 AAAAATTGAA CCATTACTGG GGACCTGATT ATCCATACGG ATTTAGGATC
551 ACAGTATACC AGCGATGATT ACAATCAACG TTTAACTGAG CTACATATCC
601 GCCACTCATA CAGCTGTAAG GGTGTCCGT ATGATAATGC GCCAATGGAA
651 TCCTTTTACG CTTCCCTCAA AAAGGAATGT GTTTATCCAG TGCCGGTCTT
701 TGAAGATTAT GAAACTGCCG CTGCCGTCC TTTTGAATAT GTGCATGCTT
751 TCTACAATAG GAAGAGAATT CATAGTTCAC TGGGCTACCA GACCCCTTA
801 CAAGTTGAAA TTGCAACACT TACGAGCCAA ATGGCCGCCCT GATTTAATGC
851 TTTCCAGGGT TCAATAAAGT TATTAATCGC GATTAATGAT TTATTTGAGC
901 TGCGGAAGGT AAACGCGGTT CTGAATGTCT CTTAAAAGCT GTCTCAAATA
951 TTGACTTCAA TCCATATCTC TTGAAAATCA CTACAATATC ATTGGGC TTT
1001 TTTTAAACAT TTTTATTTT AGGT TAGTCA AAACATATTT TATGGTTT CAT
1051 CTCAAAACAT TGTTTATACT TAATTTGTAT AGAAAGGAAG CAGTAAAATG
1101 ACTGAAAAGC ACTTAGAACT TAGAGATTAT GGCCACAAC CATTGCTCC
1151 AGATTTAAAG GCCGCTGCCT TAGCTAATGA CAATTATCGA ACTACTTTGT
1201 GGACGGGTGA TCATTTTCAA GTCAC TTTGA TGGCAATTCC AGCTGGTGGT
1251 GGAGATATCG GAATGGAGAT TCACCATGGT AATGACCAGT TCATTTATTT
1301 GGTAGACGGT GTTGGTCACG TTCAAATGGG CAAAGATAAA AATAAATTGA
1351 CGATTGACCG TGAGATTCAT CCTGGTGAAG CAGTAATTAT TCCAGATAAT
1401 ACTTGGCACA ATGTAATTAA CGCTGGTAGT AAAACGATGA AGGTCTTCTC
1451 AATTTATTCT CCAGTAAAC ACGCCAAGGG AACTGATGAA AAGACCAAAG
1501 ATGATGCGAT TCAGCAAGAA GGCCCACTAG AGGGCACTGG TGAATAAAGT
1551 CATGAGATTT TTTT TGACGA TGCTAATTTA GCTTCTCTTA ACCTGATTTA
1601 TTCACAAAAG TATGCAAAAT ACCCATCAA CACATGCATC ACGCGTTTGA
1651 TGGGTATTTT GCATTTTTCAC CTTTTAAGTG CACAAAAGA GTCTCAGCTC
1701 TGGTATAGTT AAATTGGCAA ACAAACCAC ATGAGAGAAG AGGACTCTAC
1751 ATTGAAGACT TTACAGGAAA TGGCGTTCAA TTTCAACAAA AACATTTT TAG
1801 TATCGCATA GGGTGGTCAA TTATCTTCCG ATGGCGGGCT GACATTGTGT
1851 GTAGACTGA TGGAAAAGTT TCAGTTCACC AACTTGGCCG ACAAGCTATT
1901 GAGGCTCAAT GACCAGCGGA GATATTGTCA ACTTTT TAAA GCAGCTGGCA
1951 CTGACAAAAC ATGATTCCGG CTGTGCGTT AGCACATTGC GGATCCGTTT
2001 ATTCAAAGTT GCGGCACGTG TTGTGCATAC GGGCCGACGC ATTCAGTTGC
2051 GGCTGAGCTC GCATCATGTT TACCATCGGC TGTCTACCA GGCTCTACAG
2101 CGTATCCAGG CTATTGAATA ACCGGCAAAA TTA AAAAAAAT CTGTACAACC
2151 AAGGGACAAG TGTATCCAAA AGTTCAAAAGC AATTCGAGT TAGTTGAAAG
2201 ATCTCACTAG CTAGA ACTGA TAAACATCGT ATTT CAGCAA AAGATAAAGA
2251 CATACCCGTC ATTTTTCGTG TTTTTCGAA AACTGACGGG ATTTT TATGA
2301 ACTATGAATA ATTCAGGTTA AAGTACACAA ATCTCAAGAT ATTAATAATT
2351 TAAGATTGGG AGATGTATAA GCTCTTGGGG TATATTTT TA CTAATTTTTC
2401 ATGAAAAGAAA ATTGAACAAA TTTGTATAGC AATTCCTTA TAAGCAAGGC
2451 CAATAGGGCA TACCCTATTG GCTTTACTTG GCTGTATGAA CTGATGCCAC
2501 CTGCAATAGT TCAGTAGGAT TGTC AAAATT AGGCGATATT TATCAAACAG
2551 TATGCAAGCT AATTTCTGTG TTGAAGATGG TAAAGGTTT ATATACATGC
2601 TGCTGTACC ACACAGAGGG CACCAGCCCC GAATAGCCAG ATTCAGTTCC
2651 AGGTCTGCCG GATCCAAGAC GACGAGCAGG GTTGTGCAA TAATTACCAC
2701 GAAAGAAGGT TGCCAAGTTT TTGACAAGCA TCAACGTGCA ATAAGGACGG
2751 TAAGCTAACC GTAACGCCAA AAGCTGATAA AACCGCGAG TATGCTGGTA

```

```

2801 ACCCAGCAAA ACCAACGACT CCAGTACCTA AGTCAACGGT TACTAAAGTA
2851 GTTAAGATGG CCCAAACTGG TCATCACCAA ACTTGGTTTG ATCGATTGGT
2901 GGCTTGGTTC AAATAATTGC TAGTTTGAGG GCTCAGCAGC TATACTAATG
2951 GTGTTAAGAA ATGACTACTG GCCAATAATG GGAGGAATCA ACATGGCTGA
3001 TAATGTGTTA ATGGCTATC ATATTGTGCA TGATCCTGAT GAGCGGGCAA
3051 AACATGTTTT GAATACGAAA AAGTTGTATA AGTGGCGGAT TACCGAAAAG
3101 ACGAAGGGCA CCCCAGTAGT TGGTAATGTG GCATTAGTCC AAACCCAGTT
3151 TGCGAAGCGC ACCCCAGTTA TGATTTATGC GACGAAGGAA GTTGCCAATG
3201 ATTTGAGTGA CTTACAACCA GTTAAAGTCT TTACCAACAA TCGTGATCAA
3251 GAGACGGTTA ATCAAACGTT TGATGACCTG ATGAGATAAT CTAACGGCTT
3301 GTACCAACAT GATTTTTTTC AAGTTGGTAC AAGCCTTTTT TTATAATTTA
3351 ATACTTAGTA AAATAGTGCA TTTTAAATATA TCTGGGTGGT CATTAACAAG
3401 TGCATTAGTGT AAATAATTTG CTTAATGATT TTAAC TACAA AAAGCCAAAG
3451 AGATTTGGCT CTCTTTGGCT TTTTGTAGTT AATGCAGACT TGAATCCCAC
3501 TTAGCATTTTT AAGAGAAGCA CGTGGTAATT TAAAACCGAT GATTTTTGTTT
3551 AATCAAATCA TTAAAAATAT TTGAATATTG ATTGTTTTTCA ATCCAGAAAA
3601 CAAGGAACAA AAGAACTGTT ACAGAAATTA TGGATCTTAA TATCTAGGAT
3651 TGAATTGAAG CATGCTGTGC ATGTAATAAC TAAAATAATA TCTTCTGATG
3701 TACGATGTAA CTAAGCATCA TATTGAAATA ATTTTATTTT AATTGCATGA
3751 GATAATTCCT TGAGATCTGA CGAAAATATC TTACTGTTAG GAAACTTCAC
3801 GTACTTAGCT TTGTTTAAAA AAGAGGCGGG TAAAACATGA TACACGTTAT
3851 CTGGGAAAAT ATATATAGTA TTTGGATTTG GCTCCCAATT TGACTCGAAT
3901 TCTGATAGAG TAAAAATTTT TGAGATGTTT GGAAC TCC TT GCAACGACTT
3951 TTTTAGCCGG TGAGTAAAAAT GTGCTGAGAT TGCTGAGTCT ATTATCATA
4001 TGCTGATAT ACTACTTTCT TGAATACCAA CGACAGCAGA AGGAATAATA
4051 CAAATATTAG TAACCATTTT GTTACCTCTT TCTTAAATA CATGATACCA
4101 CGAAAGCAGG AATGAGAATG AAAAAATAGAC AAAATGAAAT TGACAGTTAT
4151 TTGAACCTTC ATTTAAGGCC TGTCCATAAA TCATTTGATT TTGGAATTTT
4201 AACAAACATG GATCAATTTA GACATCATAT TTATGTATCA TATATTTTAA
4251 TATGTAAAAA TTCTCAAGCA ACTATTGAAA GATGTGTGAA TTCAATTGCT
4301 CAAAATATGG AAAATGGAGA TGAATTGATT GTTTTAGACA CTGGTTCAAC
4351 TGATGAAACT GTGCATCTGG TAAAAAGAA TATGCCACAA GCAAAAATAT
4401 CAGTAAACAAA TTGGAAAAAT GATTTCTCAG AGGTTAGAAA TAAGGCACTC
4451 AAATTAGCAT CTAAAGACTG GGTGTTCTAT GTTGATAGTG ATGAGTGGTT
4501 GGATGTTGAT GATGGAGCTC AGCTAAAGAA AATTTTATTT AAAGTTCAAG
4551 CAAAAAATTT TAAGTTCGTA ATTAACCCAA CATTTTCTGA TCATTCAGGA
4601 CAGATATACC AAACAGTAGG GAGAACTTTC CCTAAAAGT CGTCTTTTCA
4651 TTATTATGCA AAAATTCACG AAGAAGTAAG AAAAGAAGAT CAAAACTAG
4701 GTTACGATGT TAGGCACTTC GCATGTGATG ATATTATTTCT TTATCATGAT
4751 GGATATGATA AAGAAGTATT ACGAGATAAA GATAAAAATA AGAGAAACAT
4801 TCGTTTATTA CAAGAGATGA CTTGTGAGGA GCCTCAAAAT GCTAGATGGC
4851 CGTTTTTACT AGCACGTGAT GGATTTGATG TACTGCC TCA GGACAACTC
4901 AAACAGTTAG TAAAAAGGAC CTTAGACTTA GTTGCTTCCG ACAGTCTTCA
4951 AGAAAAATAT TCGCCATTTG CAAAAAAATT GCTTGGCCGT ATTTTATTAA
5001 GAGAAGGTAA AACCCTCAA GCAGTATTAT CTTTTAAAGA TGTTTTACAG
5051 ATAACAGGTG GGGAAGACAG TGATGCAATA TACTATATTG AATCATTTAA
5101 AATTAATGAA ATTATAGCGG AAGCTAAGTC TATAGAAGTA AAAATGTTGA
5151 GATATTTAAA TAAACACAAA GGAATGATTG ATGTTAATAG CGATATATCT
5201 GGAACTACT ATCATATAGC TCAGGTTATT TTAGAATGTG ACATAATAAG
5251 TGCCAAC TAC AGCCATCTTT TCCCTTTAAT ATCAGAAATT CCTAAAACT
5301 TTTTAGGCGA TATAAAAAGT AGTGTCAAAT CTGCAGTGAA ACTGTATTCA
5351 AAATTACAAG GTGACAGCAA AAATGAGAAT AATTAAACAA ATTGACCAA
5401 ATGATTGTGG ACCAGCAGCA GTAGCAACAA TCATTTGGTAT GATTAGTAAA
5451 ATTAAAATTT ATGACTGGCA GATAAGGCAA ATTATCACAA CAACAAATTA
5501 TGGGACTAAT TTTATCAGTA TAATAAACGG ATTAAAGTCA GTTCAGGTTG
5551 AGTCTGTTGT AGAAAAGTGT GTAAAGAATC CAGATGTGTT TGATGAAATT
5601 GAATTTCCAG TGCTGACACA AATTATGCAA AATGGCTATT TACACTTTGT
5651 AGTTCTCACG AAGTGTCTTG GTTCAAAATT ATATTGGGCA GACCCAGGTA
5701 GTGGAAAAAT AGAAAATATT GACAAGGCTA AATTTATGAT GCATTGGACG

```

```

5751 CCGCTTATCT TAACGATACC TAGATCACAA AATTTTCAAA TTCTAAAAAC
5801 AAAAGTTGAA AGCTTAAAAAG TCCCTATTTT GTGGTTTAGT TTTAGGAAAC
5851 ACTGGGCATG CCTGTTAATA TATTGGATGC TTTCAATAAT TATATTTGCT
5901 ACAACACTAT TCTTAGCAAC TATGTATTCT GCCTATTTTA ACCGCATTGT
5951 TCCCAATAAG TTTATTTFCAG CGATTCCCTAC TATAACTCTC GTATATTTGA
6001 CTGTTCAAGT AGCATCAATT TTTTTGGAAG TTATTAATGC AATCCTTATT
6051 GCCTATACTG GTAATTCAAT CAGTAAAGAA TTATATGAAG ACATCTTTAA
6101 GGCTTTTCCC AAAAAAGAAGC GGTAGTAGT AGATTCATTT GAAGATGGTG
6151 AAATTATTAC AAGATTTAAT GCGATTTATC AAATTCAGG TAGACAACCT
6201 GTGCTTTTAG GTGACTTACC AATFGACAT ATTACTATGG GGATAACCCT
6251 TTTTCTTCTT GCTAGAATCA ATCCATTTT GACGATTTTA TTGTTTATTC
6301 CAATGATATT AATGGCATT TATTTTTATC TAAGTTCAGA CGCAATTAAA
6351 AAAAGAAGTA CAAATTTGTT TAAATCACAG GAAACATTTA GCCAAGATAT
6401 TATTGAACAA ATAAAAAAT TTTCAACTAT AAAAACTTTT CATGCTGGTA
6451 GCTACATCAA TTCCAAAAATA AAAAAAGTA TGGAAAAGTA TACTGCAGCA
6501 AACCGTAGTT TTGGCTTATA TGATAGATTA CAGCAGATAG CAAGAAATGG
6551 TGTTTCACAA CTGTTTGGAA TTATTTTATF CTCGGTGGGG GCTTTTCTAG
6601 TAATAAAGGG CAATATTGTG CTTGGAACAT TGCTTAGTTT TTATGCTTTA
6651 ACGGGAAATG TTCTAAATCC GTTTGTGAAA ATAATAACTA TGCAAGTAAA
6701 TCTGTCTCAA GGAAAAGTAG CAACAGAACG CTACTIONGAT CTATTACTTT
6751 CTCCTGAAGA CAATAACAAT AAAAGCAAAC TAACACTTTC AAATTCCTATT
6801 GAAATAAGAA AGGTAGCATT TTCATATGAT GGGGTTACAC CAATAATTAA
6851 AAACATGAAC GGTGTTTTCT CAAAGTTTCC AACTGCAATA ATTGGCAATA
6901 GTGGGAGTGG CAAGTCAACT ATTGCTAAAC TAATGGCAGG TTTTATCCCT
6951 GCAGAGGGTC TATCATTAGG AAACATACCT TATTATGCAA TAGGGGCAGA
7001 ATCTTTAGGT AATACAATTA CCTATGTTGA ACAATCACCA CAAATATTTT
7051 CGGATACAGT TATGAATAAT ATAACCTCTG GTAGGAAAAA TATCACAGTT
7101 AAATCAATTA ATAGGATTGC TAGTAGTATT GGTTTTGACG AATGTTTGAA
7151 TCATTTACCA AACGGAATAT ATACACAATT AGGATCCCTT GGATATTCAT
7201 TATCTGGAGG TCAAAAAACA TTATTAAATA TTATAAGGTC TATGGTAATT
7251 CCAAGTAGGT TTATCATTTT TGATGAAATA ACGAATGGGC TGGACATCAA
7301 TACAAAATTG AAAGTAGAAA ATTTATTTAT ATCACAAAAC AAACCTAAAT
7351 TAATATTTAT TACTCACGAT TTAACATTGG CTATAAGATG TGATGATATT
7401 TTCACTGTTA AAAATGGTGA GATGCACCAA ATAGAAACTA AAAAAAATC
7451 TACTGAATCA TTGAGGGAGG CATTGTTATG AAAAAAACCA TTATCACTAT
7501 TATTGCGCTT AGTGTTATCT CAATAATAGG AGTATATTC TGGCAAAGGC
7551 TCACTTTGGT ACAAGAGCCA TTATATCACT CAATAAAGGC GACACAATTG
7601 CAAGCAGACA TAAAAAATAA GAAGGATGAT GTAGTTTACT TCTACCAAAA
7651 GGGTTGTGAA GGATGTGAAC ATAGTACTCC CATATTGAAC CGCTTTATTC
7701 GAAATAAACG TGTCACTGTA AAAGCAATCG ATATAAATGC AGATCCAAGC
7751 AAAGATTTTA TTTTGGGAAC ATTGGGAATA TCATCAACAC CAACCATATT
7801 GTTTATACAC CATGGAAAAA TTGTTCAAAG GATAAATCAA GTCCTTGGAT
7851 ATGAGGACTT AACAACATAA TATAAGGGGT ACACGGGAGA TCTTTAGATG
7901 AATCTATTAA AAAAAAGAAA ACTAGTATTA ATTTTGCTAA TACTTCTCAT
7951 TATTTTATTA GTAATTGGAT TATTTTCGGT GATTATTACG AAACGGAATT
8001 CTTATGAACA TGAGAACGTA CTTGCGGCTG AAATTATTTT AGAATCAAAG
8051 ACAAACTACTA TGGTTTACAA GGAATTAAT TCAATAACAA TTCCAACTTT
8101 TGAAAAGGAA ATTTCCGGATA AAAGAGATTT TCTAGTATAT ATTGGACGGC
8151 CTACTIONGTT AGATTGTAAT TTATTTGATC CTATTTTAGT TAATGAATTG
8201 AAAAAAGAAG ATATGACCAG CAATGTTGTC TTTCTCAACG TAGCATGGGA
8251 ACGCCAAAAG GGATGGACTT CATGGGTACA GTTTAAGAAA AAATATGGAT
8301 TCAAGCAGAC ACCTGCAATA ATTCATATC ACAATGGCAA AGTATTAAGT
8351 ATTATACAGT GGGGAATAA TAAAGGGATA AGTAAGGGAG ATTTACACTT
8401 ATGGCTTATC AAGCAGAAGG AGAAAATTTA GTTCCACAGA CTAGAATATT
8451 TGTTTTAACA AATAAACCGG AATGCCAAA TTTAAATAAC CATATATTAC
8501 AAATAATGCT CAATTTATCG AATCATATCA AGTTTTGTCA TCCAATAAAA
8551 CTATCATCAA GCTCTGAAAA TTTCCAAAAC TTTATATGTT CACTCTCTAT
8601 CGTAATTTTA AATTTAGATG AAGAAAATAGA CTTAATAAAA AAAATTAGTA
8651 TTGCTGTCAA GAGGCTTCTT AATTCACGGA AATACTTCAT TTACATTGGA

```

```

8701 TCAAGTAGAC ATGTTTTTTGA AACTTAAATT TTCAATCATT ATATGCTTTA
8751 TGATGCACTA ATGACCAAAA CTTTTCAAAA AGATGATGGA AATATATTGG
8801 CTGCTGATAT TTGTAACATT CTAATACACA ATTCATATCC CTTAAATAGT
8851 TCAAACACTA TACAAATTTT TGTCAAAAGA CAGAAGTTTT TTCCTAAGATT
8901 GGATGATATA TTATTTGTAA TAACTGACAA TCAACCGCAT AAGGTAATAT
8951 TTCATTTGAA ATATGATGAA TATCAAGTCC GTATGACTTT AAAACAAGTT
9001 CATCATCAAT GTCGTAATCT TACTGCTTGC CATGAAGGAG CCTTAGTTAA
9051 CCTAAAAAAC ATCACTAAAT TTAACAAGAG CAATAGGATC CTATATTTTG
9101 AAAATGGTGA TACTTGCGAG GTTCAAGAA GATATACAAA AAAGATTAGT
9151 GATTTAATCC CGATGAAGAC ATTTTAGTTA CCGAATAAGA TAGTTGCTCT
9201 CCAATATGTT TGTTTTTGGT TTATAATTTT CTTATTAATA GGAGGGTATT
9251 AGGAATGAGT AAATTGGTTA AGACACTTAC TATAAGTGAA ATTTCTAAGG
9301 CTCAAAACAA CCGTGGAAAA CCTGCATGGT GTTGGTATAC TTTAGCAATG
9351 TGTGGTGCTG GTTATGATTC GGGAACCTGT GATTATATGT ATTCGCATTG
9401 TTTTGGTATA AAGCATCATA GTAGTGGTAG TAGCAGTTAT CATTGTTAGT
9451 TTTGTGAATG TTTTAGACTT ATTAAGTTAC GTTGATAGAG TATTTATCT
9501 TTGATTTTGT AATTTCCACA ATAAAGAGTT ATTTGGAAAA TTTATGACGC
9551 AAATAAGGAA GACAATAACA TAATAATTAT TGTCTTCCTT TAAATTTCTC
9601 TAAGTGAATA ATTAGCAGTT AGCGCTATAA AAAGAGTGT TCTGATCTTG
9651 ACTATTTAGT ATTAAATCA AGATAGAAAG GTACAGGATA CATTTATGGA
9701 CTAACATTAA TGTGAACAAT GGCTTATCTT GATGCTGTTT TTAATAGAAC
9751 CAAAATCAC AAGGAGATGT TATTTATGGG AATATTCGAA GATCTAAATA
9801 ATAATGGTGA CGACTTTGTA AAGCCAATAG ATTTGGGGCA AATATTTTCC
9851 AATTTGAATC AAGTTCAAGG TTATGATTAT CCTCGTTCCTG TTCAATCAGA
9901 CTTTTTTAAA GAATGGTATT CAAGAAGAAG TGACAAAGAT TTAATAGGAA
9951 TTCTAGGTAC TGGAACAGGG AAAACATTTG TAGGATTAAT AGTTTTACAA
10001 TCTTATCTTA ATGATAATTC AGGGCCATGT CTGTATTTAT GCCCAACAAA
10051 ACAGTTAGTT AGGCAAACAA TTATTTGATGC TGGTAATTTT GGTGTTCAAA
10101 CAGTTGAATT TAGTAGCTCG GAGTTTCCCT CGGAGTTTAT CAACTCAGAG
10151 GCCATATTAG TTACAACTTT CGATAAAGTT TTCAATGGAC GTTCTAAATT
10201 TGGATTTGAT ACGCCGGCAT ATCAAAAAAT AGGAGGTTTA GTAGTTGATG
10251 ATGCACATGC CGCTATTAAT CTTGTTAAAC AAGCTTGCAC GCTAAAAATC
10301 CCCAGAGAAG ATAAGCGATA CCAACAAATT TTATCTTTAT ATGACACCTC
10351 ACTAAATACT CAGAGACCTT CAAAACAAA AAATATAAAG TACGTTTCCA
10401 ATTCAAATTC TTCAATGAAA GTCCCTTATT GGACAGTGGG TGATAACTTG
10451 GATGAATTGA TAAATATAAT AAGCAACTCC GAAACAGCCA GTGCTATTCA
10501 GCTACCGTTT GTTCTAGAAA ATGAAGAATT GCTGGATGTT TTCATATCTT
10551 ATAAGGAAAT ATCCGTCAGG CCTCAGCATG CTCCAATCAA TTTAATATCT
10601 TCGTTCCTCAG AAGCGAAGCA TCGTCTGTTT TTATCGGCTA CATTAGCTGA
10651 TGCAGGGGAC TTTATTTCCG AACTAGGGGT AGAAAAGGGT GCGGTACTTA
10701 ACCCCATAAA AATCAATTCA TTAACGACA CAGGGCAAAA ATGGTACTTT
10751 GATTTTCTA GAGTAATTC GACAGTAAAT ACTGATAAGA TCCGGGAATA
10801 TGTGCGAACA ATATCAACTA CTTATAAAAA GAATGTTCTG GTACTTACCA
10851 CGTCAAATTA TTACGCTGGT TTTTGGGAGG AAAAAGGCGC TAAACTATAT
10901 TCGGCCGGAG ATCTTCCTGA GCTTAAAAGA GATTTAAAAA ATAAAAATCC
10951 AATGATTGCA GTTTTGCAA ATAAATACGA CGGGATAGAT TTACCAGGTG
11001 ATTTATGTCA TATAACAGTT ATTGACAAA CTCCACGTCA ATCATCTACT
11051 GCTGACAAAAG CAAATATTCA CCGTCTTCC T AATTCCGATT CATTAATATGC
11101 ACCAATTGTA CGTGACATAG AACAGGCAT GGGGAGAGCT GTACGTTCAA
11151 AAGCAGACTA CTCTTTGGTA TTTGTATTGG GGCAAAATTT AAGCAACTTT
11201 ATATACTCAC ATGATGCTCT TCTAAGTACA GAAACCTCAG CACAATGGGA
11251 GTTTTCTAAA AACATCACAG ATAAAAATCAA ATTCATTTAT TCTAACAAA
11301 AATCCCGGCT AGAACAATTT CAGAGTATCA TAACTGGAGT ACTGAACCAA
11351 GAAGAAAAAT GGACTAGTTT ATATAAAACC GAAATATCAA AAATATACAA
11401 CAAGCTTGCT TCGGATAGCT CAACACAACA CTATACTTTT AATGTTGCTT
11451 TAAATAACGC ATGGCTTGCA GCACAACAAG GCCGTTATAG CTTTCGCTGTA
11501 AATGAGCTTC AAGAAAAAAT ATCAAGCTTT AATAAATCAG ACCAAGCAAT
11551 CTTATATGAA TGTATGGCAC GCTATGAATA CAAAGTTAAC AAACAACATT
11601 CATCTGATTT ACAAATTAAT GCACATGAAC TAGAAACACA TCTTTTAA

```

```

11651 CCGCTCAATG TTTTGTATGA AAAACACACC TTGCCAAAAA AAGAACAAGC
11701 TACGGCATCC CTTAATTATC TAAAATCAAG GAATTTAAAA GACAGCAACG
11751 AATTAGCCAG TTATATTGAG TCAATTAATA GTAATCTAAA ATATTCCTCA
11801 TTTTCAAGCG AAGAAGATTT CAGGGAAGCA ATACAAGATC TTGGTAGCCT
11851 TTTAGGCTGT AATTCTAGTC AACCTGAAGC AGAAGCAATC ATAAAGGATG
11901 GTAGCCCGA CAATCTATGG CTTTCTGATA CCAATTCATT CGTTATAGAA
11951 GATAAAAAATC AAGTTACTAC TTCCAAGATT AGCAGAGGCG ATATTCAGCA
12001 AATAACTGGA TCTTCTTCAT GGTTTGAAAA AACATATAGT ATTAACATATG
12051 TTCCCGTACT TTTTCATCGT TCAAATATTA TGAATACAGA TGCCAATACT
12101 AACTTAGAAA TCTACGTGGT AAACGAAAAA AAATTAATAA CTTTAAACTA
12151 CTCTCTTCAT AAACCTGCTC AAATTATTGG TCAAAAACCT CTTAATGAAT
12201 GGAATTCATC CACACTGATG CAAGCATTC AAGATAGTGA CTTATTAATA
12251 AACGACTTTC GACAAAACTA TTGTGTCAA GCAACTAAGG CTTAAGTTAT
12301 CTGTGTGTAT TTTGAAGAAA ATTGACCATA ATATTACAAT ACAGACGGTT
12351 GAAATTAATA AAATTGATTA TCCAAACTTA GATTTTTTAT CTCGAAAGAG
12401 TTTTACGGAC AAGCTATTAA TATACAAAAA GTATTAAAAA ACTAATGACA
12451 TTCTCAATAG CTTTTTAGCA AGATTTGGTC ACAACTTAAC AGTTGTGGCC
12501 TCTTTTTTGA TACTGTATAA TTAAAGGGTT AGATCCGTGT CATTACTTAA
12551 TGACATAAAT AAGGAGATAA AAAGATGAAA TATGGCTATG CGCGGGTCAG
12601 TACCACTGAT CAAAAATTAG CAAATCAAAT TGAGTTACTA AAATTGGCAG
12651 GAGCAGAAAA AATCTTTCAA GAAAAGTTTA CCGGCACAAC TACCGAACGA
12701 CCGGAGTTTC AAAAAGTGT GCGCGTCTA AAAACAGGCG ATACTTTGAT
12751 TGTCACTAAG TTGGATCGGT TTGCGCGGAA CACACGCGAG GCCTTAGCCA
12801 TTATCCAAGA GTTGTTTAAA GAAAAATGTC AAGTTAACAT TTTGAATATG
12851 GGCTTAATTG ACAATACGCC GACTGGACAA CTAGTCTTTA CAATATTTAG
12901 CGCCTTTGCG CAGTTTGAAC GCGATATGAT TGTFCACGCG ACACAAGAAG
12951 GTAAAGTGTA TGCCAAGCAA CATGATCCGT TGTTCGGGA AGGGCGACCG
13001 AAAACGTATT CTGAGGAACA AATCAGATTT GCCTACGAGT TACGAAAACA
13051 AGGCATGACC TATAAATGA TTGAACGAAA GACGGGGATT AGTAAACGCA
13101 CGCAACAGCG ACGATTTAAA TTAGTTAAAA ACAATATATT TTAAAACGTG
13151 TTACACCATA CCAAAAGGGG CTTCGAATAC GAGGCCCTT TTAGATTTAG
13201 CTAGTAATCC GTATTAAAGT TAATGGCTAG CTCCCAGTTC AACTTTTTTAA
13251 TGGCTTACGA TCTGTCAACA AAGCACTTGG TGACAGAATC AAAACAATCA
13301 CCGCTCCTAC TAAGTAGGTA TGACTGAAGG CCTGGCTGAT CTGTCGATTC
13351 TTAGTATGCT TAATTTTTCT CAAAGCTGTC GTTAAAGCTT CTTTTTGTGC
13401 CAATTGTTTG AGACCTACTA ACAATTTTTT CTGACCCTGG CTTATTTTCT
13451 GATTACCCAA ATTAATTTTT TGAGAAGCTA CTGACAGTTT AGATATAGCG
13501 GTATTTTGAT TTTCTAACTT GGTAAAAGTA AGTGCTAGTT GCTGGTTACC
13551 CGCTTTTAAG TTGTCAGTAG CTAGTGCTAA CTTCTTCTGG GCTTGGTAGA
13601 CCTGGTGAAG ATCGGTTCCC TTTTGGGGCT ATGGTAAATC AACTTGAGAT
13651 TTGGCCGCTT GTTTAAGGGC ACGTCTTAGG TTCGTCTGTT TGTTTTTAGT
13701 TGTTAATTTG GCTGATCCAG TAGTGGAAAA TAGGCTGTTC AATTGTTGAT
13751 GAGCCACTTT CTTTACTTTA GGTGATAAAT GCTGTCGATC AACAAATTTGA
13801 ATGGTATCTT GATGCATATG CTGCTTAGCT GTATTGACAT TATTATCAAA
13851 GACCATTACT AGAATTGCAA TTCTTAAACA GGTGCCAATT TGACGCGCTG
13901 CATTAAACAAT TCCTGAAACA ATCCCCTTT TATTTTTGGG CAAATGCTTA
13951 ACCAAGACCA CTAACGAAAC CGAAGAGAAC CAAAACCAA CACCGTTGAT
14001 CACTAGGAAA GTAATGATAA CAGCTTTAGG CGTTGTGGTG GTGATAAACG
14051 ACAAGAGTAG TAAACTAGCT GCCATGACTA AGAGTCCAAT CAAAGTAACC
14101 GAAACTGCGC CAACTTGATC AAATAACTTA GTTCCCAAAG GCATCGCTA

```

## APPENDIX 11: SEQUENCE ALIGNMENTS

```

          10          20          30          40          50
3bcv ( 3 )      lip-----kVSVIVpIynve
GT            MKNRQNEIDSYLNHLRPVHKSFDGNTNIDQFRHHIYVSYIVICKNSQ
                bbbbbb

          60          70          80          90          100
3bcv ( 18 )    kyLdqCVqaLlaQtLsdIeIILIDDeSpdnCpkiCddyaaqypnIkVihk
GT            ATIERCVNSIAQNMENGDELIVLDTGSTDET'VHLVKKNMPQ-AKISVTNW
                aaaaaaaaaa bbbbbb aaaaaaaaaa bbbb

          110         120         130         140         150
3bcv ( 68 )    knaglGAcnsGLdvAtGeYVafCdSddyvdsdY-----tynvAqkytCD
GT            KNDFSEVRNKALKLASKDWVfYVDSDEWLDVDDGAQLKKILFKVQAKNFK
                aaaaaa bbbb aaaaaa

          160         170         180         190         200
3bcv ( 116 )   AVFtfkly-----knkneIhtlLkdlIasdpyaReerAi
GT            FVINPTFSDHSGQIYQTVGRIFPKKSSFHYAKIHEEVRKEDQKLGVDVR
                bbb bbb aaaaaaa333 aaaaaaa

          210         220         230         240         250
3bcv ( 170 )   qVsakvVLYrrnLiekKhLrFvseriLpSedLIFNVDVlanSniVCvlp
GT            HFACDDIILYHDGYDKEVLR-----DKDKIKRNIRLLQEMT--CEEPQ
                aa bbbbbaaaaaa aaaaaaa bbb

          260         270         280         290         300
3bcv ( 4 )     ipkVSVIVpIynvekyLdqCVqaLlaQtLsdIeIILIDDeSpdnCpkiCd
GT            NARWPFLLARDGFDVLPQDKLKQLVKRTLD-----LVA---SDSLQEKYS
                bbbbbb aaaaaaaaaa bbbbbb aaaaaa

          310         320         330         340         350
3bcv ( 54 )    dyaaqypnIkVih--kknaGLGAcnsGLdvAtGeYVafCdsdd-yvdsdY
GT            PFAKLLGRILLREGKTTQAVLSFKDVLQITGGE-----DSDAIYYIESF
                aaaaa bbbbb aaaaaa bbbb

          360         370         380         390         400
3bcv ( 104 )   tynvAqkytCDAVf'fklyknkneIhtlLkdlIasdpyaReerAiqVsak
GT            KINEIIAEAKSIEVKMLRYLNKHKGMIDVNSDISGNYHIAQVILECDII
                aaaaaa bbb bbb aaaaaaa333 aaaaaaaaaa

          410         420         430         440
3bcv ( 175 )   vVLYr--rnLiekKhLrFvseriLpSedLIFNVDVlanSniVCvlp
GT            SANYSHLFPLISEIPKNFSGDIKSSVKS AVKLYSKLQGD SKNENN-
                bbbb aaaaaa aaaaaaaaaa 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<sub>10</sub> helix; lower case: solvent accessible; UPPER CASE: solvent inaccessible).

```

          10      20      30      40      50
hs1qg8a  (  2 )                pkVSVIMtSynks
QUERY GT      MKNRQNEIDSYLNLHLRPVHKSFDFGNLTNIDQFRHHIYVSYIVICKNSQ
                bbbbbbb

          60      70      80      90      100
hs1qg8a  ( 15 ) dyVakSIsIlsQtfsdFeLFIMDDnSneeTlnvIrpflndnrVrfyqsd
QUERY GT      ATIERCVNSIAQNMENGDELIVLDTGSTDETVHLVKKNMPQAKISVTNWK
                aaaaaaaaaa      bbbbbbb      aaaaaaa33333      bbbbb

          110     120     130     140     150
hs1qg8a  ( 65 ) isgvkerTekTryAALInqAIemAeGeYITYAtDdNiYmpdRLlkMVreL
QUERY GT      N-----DFSEVRNKALKLASKDWVFYVDSEWLDVDDGAQLKKIL
                aaaaa      aaaaaaaaaaa      bbbbbbb      bbb      aaaaaaaaa

          160     170     180     190     200
hs1qg8a  ( 115 ) dthpekaVIYSASktyhlndivketvrpAaqvtwnAPçaIdhCSVMHrys
QUERY GT      FKVQAKNFKFVINPTFSDHSGQIYQTVG-----RIFPKKSSFHYYA
                aa      bbbbbbbbbbb      bbbbbbb      333bbbb      a

          210     220     230     240     250
hs1qg8a  ( 168 ) VLekVkekfgsyWdespaFyriGDArFFwrVNhfypFypldeeLDlnyit
QUERY GT      KIHEEVRKEDQKLGYDVRHFACDDIILYHD-----GYDKEVLRDKDKI
                aaaaaaaaa      333      aaaaaaaaa      bbbbbbbbbbbbbb

          260     270     280     290     300
hs1qg8a  ( 232 ) efvrnLppqrnçreLresLklgmg
QUERY GT      KRNIRLLQEMTCEEPQNARWPFLLARDGFDVLPQDKLKQLVKRTLDLVAS
                333      aaaaaaaaaaa

          310     320     330     340     350
hs1qg8a
QUERY GT      DSLQEKYSPFAKKLLGRILLREGKTTQAVLSFKDVLQITGGEDSDAIYYI

          360     370     380     390     400
hs1qg8a
QUERY GT      ESFKINEIIAEAKSIEVKMLRYLNKHKGMIDVNSDISGNYYHIAQVILEC

          410     420     430     440
hs1qg8a
QUERY GT      DIISANYSHLFPLISEIPKNFSGDIKSSVKSAVKLYSKLQGDSKNENN

```

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; 3:3<sub>10</sub> 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

a	b	c	#	Res:	#	x	y"	z"
101.108	129.103	146.129	1	Lys	43	-	-	-
198.161	226.156	243.182	2	Pro	42	5099.955	5075.991	5058.965
269.198	297.193	314.219	3	Ala	41	5002.902	4978.939	4961.912
455.277	483.272	500.299	4	Trp	40	4931.865	4907.901	4890.875
558.286	586.281	603.308	5	Cys	39	4745.786	4721.822	4704.796
744.366	772.36	789.387	6	Trp	38	4642.776	4618.813	4601.786
907.429	935.424	952.45	7	Tyr	37	4456.697	4432.734	4415.707
1008.477	1036.471	1053.498	8	Thr	36	4293.634	4269.67	4252.644
1121.561	1149.556	1166.582	9	Leu	35	4192.586	4168.623	4151.596
1192.598	1220.593	1237.619	10	Ala	34	4079.502	4055.539	4038.512
1323.638	1351.633	1368.66	11	Met	33	4008.465	3984.501	3967.475
1426.647	1454.642	1471.669	12	Cys	32	3877.424	3853.461	3836.434
1483.669	1511.664	1528.69	13	Gly	31	3774.415	3750.452	3733.425
1554.706	1582.701	1599.727	14	Ala	30	3717.394	3693.43	3676.404
1611.727	1639.722	1656.749	15	Gly	29	3646.357	3622.393	3605.367
1774.791	1802.786	1819.812	16	Tyr	28	3589.335	3565.372	3548.345
1889.818	1917.813	1934.839	17	Asp	27	3426.272	3402.308	3385.282
2179.929	2207.924	2224.951	18	Ser	26	3311.245	3287.281	3270.255
2236.951	2264.946	2281.972	19	Gly	25	3021.134	2997.17	2980.144
2337.998	2365.993	2383.02	20	Thr	24	2964.112	2940.149	2923.122
2441.007	2469.002	2486.029	21	Cys	23	2863.064	2839.101	2822.074
2556.034	2584.029	2601.056	22	Asp	22	2760.055	2736.092	2719.065
2719.098	2747.093	2764.119	23	Tyr	21	2645.028	2621.065	2604.038
2850.138	2878.133	2895.16	24	Met	20	2481.965	2458.001	2440.975
3013.202	3041.196	3058.223	25	Tyr	19	2350.924	2326.961	2309.934
3100.234	3128.228	3145.255	26	Ser	18	2187.861	2163.898	2146.871
3237.292	3265.287	3282.314	27	His	17	2100.829	2076.866	2059.839
3340.302	3368.297	3385.323	28	Cys	16	1963.77	1939.807	1922.78
3487.37	3515.365	3532.392	29	Phe	15	1860.761	1836.798	1819.771
3544.392	3572.386	3589.413	30	Gly	14	1713.693	1689.729	1672.703
3657.476	3685.471	3702.497	31	Ile	13	1656.671	1632.708	1615.681
3785.571	3813.565	3830.592	32	Lys	12	1543.587	1519.624	1502.597
3922.629	3950.624	3967.651	33	His	11	1415.492	1391.529	1374.502
4059.688	4087.683	4104.71	34	His	10	1278.433	1254.47	1237.443
4146.72	4174.715	4191.742	35	Ser	9	1141.374	1117.411	1100.384
4233.752	4261.747	4278.774	36	Ser	8	1054.342	1030.379	1013.352
4290.774	4318.769	4335.795	37	Gly	7	967.31	943.347	926.32
4377.806	4405.801	4422.827	38	Ser	6	910.289	886.325	869.299
4464.838	4492.833	4509.859	39	Ser	5	823.257	799.293	782.267
4551.87	4579.865	4596.891	40	Ser	4	736.225	712.261	695.235
4714.933	4742.928	4759.955	41	Tyr	3	649.193	625.229	608.203
4851.992	4879.987	4897.014	42	His	2	486.129	462.166	445.139
-	-	-	43	Cys	1	349.07	325.107	308.08

## APPENDIX 13: LISTS OF ENZYMES AND OTHER MATERIALS

**Table 5.8: List of polymerases and restriction enzymes**

<b>Enzyme</b>	<b>Manufacturer</b>
KOD Hot Start DNA Polymerase	Novagen, Merck KGaA, Darmstadt, GER
Platinum <i>Taq</i> DNA Polymerase	Invitrogen, CA, USA
<i>Taq</i> DNA Polymerase	Roche, Mannheim, GER
T4 DNA Ligase	Invitrogen, CA, USA
Tth DNA Polymerase	Roche, Mannheim, GER
<i>Bam</i> HI	Invitrogen, CA, USA
<i>Bgl</i> II	Roche, Mannheim, GER
<i>Ecl</i> 136II	Fermentas
<i>Eco</i> RI	Roche, Mannheim, GER
<i>Eco</i> RV	Invitrogen, CA, USA
<i>Hind</i> II	Roche, Mannheim, GER
<i>Kpn</i> I	Roche, Mannheim, GER
<i>Msc</i> I	New England Biolabs, MA, USA
<i>Nae</i> I	Roche, Mannheim, GER
<i>Nco</i> I	Invitrogen, CA, USA
<i>Nde</i> I	Roche, Mannheim, GER
<i>Pme</i> I	New England Biolabs, MA, USA
<i>Sac</i> I	Roche, Mannheim, GER
<i>Sal</i> I	Invitrogen, CA, USA
<i>Sma</i> I	Invitrogen, CA, USA
<i>Xba</i> I	Roche, Mannheim, GER
<i>Xho</i> I	Roche, Mannheim, GER

**Table 5.9: List of commercial kits**

<b>Kit</b>	<b>Manufacturer</b>
BM Chemiluminescence Blotting Substrate (POD)	Roche, Mannheim, GER
RTS 100 <i>E. coli</i> HY cell free expression kit	Roche, Mannheim, GER
Complete, Mini, EDTA-free	Roche, Mannheim, GER
High pure PCR product purification kit	Roche, Mannheim, GER
High pure plasmid isolation kit	Roche, Mannheim, GER
Illustra RNAspin mini RNA isolation kit	GE Healthcare, UK
LIVE/DEAD BacLight Bacterial Viability Kit	Invitrogen, CA, USA
Perfectprep gel cleanup	Eppendorf, Hamburg, GER
ProbeQuant™ G-50 micro columns	Pharmacia Biotech, Uppsala, Sweden
Random primers DNA labelling system	Invitrogen, CA, USA
TRIzol® Reagent	Invitrogen, CA, USA
TURBO DNA-free™ Kit	Ambion, Applied Biosystems, CA, USA
Wizard Genomic DNA purification kit	Promega, Wisconsin, USA

**Table 5.10: List of consumables**

<b>Consumable</b>	<b>Manufacturer</b>
Filter membranes	Millipore, MA, USA
Chelating Sepharose Fast Flow	GE Healthcare, SE
Gloves, latex	Ansell, Muenchen, GER
Gloves, nitrile	LabServ, Biolab Ltd., Auckland, NZ
Immobiline Dry Strips	GE Healthcare, UK
Immobiline Dry Strip Cover Fluid	GE Healthcare, UK
Immobilon-P PVDF membrane, 0.45 µm	Millipore, MA, USA
Maximum Recovery tubes	Axygen, CA, USA
Nitrocellulose membrane	Sigma-Aldrich, MO, USA
Parafilm	Pechiney Plastic Packaging, USA
Petri dishes	LabServ, Biolab Ltd., Auckland, NZ
PP-Test tubes, 15 mL	Greiner bio-one, Frickenhausen, GER
PP-Test tubes, 50 mL	Greiner bio-one, Frickenhausen, GER
Quartz SUPRASIL® precision cell, 0.1 mm path length	Hellma GmbH & Co. KG, Müllheim, GER
Quartz SUPRASIL® precision cell holder	Hellma GmbH & Co. KG, Müllheim, GER
Screw cap tubes, sterile	NUNC™ <i>CyroTubes</i>
Syringes, sterile	Terumo Corporation, Tokyo, Japan
Syringe filters	Millipore, MA, USA
Tips, 5 mL	Eppendorf, Hamburg, GER
Tips, 1 mL	Axygen, CA, USA
Tips, 200 µL	Axygen, CA, USA
Tips, 10 µL	Axygen, CA, USA
Tubes, 1.5 mL	Axygen, CA, USA
Tubes, 0.7 mL	Axygen, CA, USA
Whatman 3 mm paper	Whatman International Ltd., Kent, UK
ZipTip <sub>C18</sub> pipette tips	Millipore, MA, USA

**Table 5.11: List of equipment used**

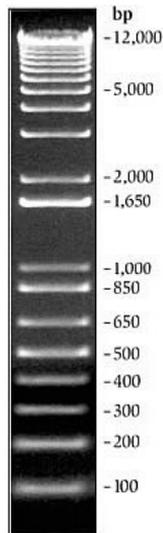
<b>Equipment</b>	<b>Manufacturer</b>
ABI Prism 377-64 sequencer	Perkin Elmer, MA, USA
ABI Prism 3730 capillary sequencer	Perkin Elmer, MA, USA
Bead mill	Fast prep cell disruptor, Thermo Savant, Qbiogene, Carlsbad, USA
Biofuge fresco centrifuge	Heraeus, Thermo Scientific, MA, USA
Chirascan Circular Dichroism spectrometer	Applied Photophysics, UK
CoverWell, 20-mm diameter, 0.5 mm deep	Molecular Probes, Invitrogen, CA, USA
ESI-MS, micromass ZMD	Micromass, Manchester, UK
Ettan IPGphor II Isoelectric Focusing Unit	GE Healthcare, UK
GASTIGHT syringe	Hammilton CO., NV, USA
Hamamatsu ORCA-ER C4742-80 digital charge-coupled device camera	Hamamatsu Corporation, Japan
Intelligent Dark Box II, LAS-1000	Fujifilm Holdings Corporation, Tokyo, Japan
Jupiter 5u C <sub>4</sub> 300A columns, 250 x 10 mm	Phenomenex, CA, USA
Jupiter 5u C <sub>4</sub> 300A columns, 250 x 4.6 mm	Phenomenex, CA, USA
Liquid scintillation & luminescence counter 1450 Micro Beta, Trilux, Wallac	Perkin Elmer, MA, USA
Macrotrap, Peptide 6PK	Michrom Bioresources, Inc., CA, USA
MALDI-MS, micromass M@LDI™	Micromass, Manchester, UK
Micro Pulser (electroporator)	BioRad Laboratories, Milan, I
Mini-Protean II system	BioRad Laboratories, Milan, I
MiniSpin plus centrifuge	Eppendorf, Hamburg, GER
Mini Trans-blot electrophoresis transfer cell	BioRad Laboratories, Milan, I
Nano drop ND-1000 Spectrophotometer	Thermo Scientific, MA, USA
NANOpure II filtration system	Sybron/Barnstead, Maryland, USA
Olympus IX 71 microscope	Olympus Europa Holding GmbH, Hamburg, GER
PCR cycler T gradient	Biometra, Goettingen, GER
Rotation shaker Model G25	New Brunswick Scientific, New Jersey, USA
RP-HPLC Waters 2790	Waters Corporation, MA, USA
RTS ProteoMaster	Roche, Mannheim, GER

Smart Spec™ Plus Spectrophotometer	BioRad Laboratories, Milan, I
Sonicator	Soniclean Pty. Ltd., Adelaide, AUS
Sorvall Evolution RC	Thermo Fisher Scientific, MA, USA
Sorvall RT7	Thermo Fisher Scientific, MA, USA
SpeedVac Concentrator	Savant, Global Medical Instrumentation, Inc., Minnesota, USA
Summit HPLC	Dionex, Corporation, CA, USA
UltiMate 3000 HPLC	Dionex, Corporation, CA, USA
UV-Trans-Illuminator, Bio Rad Gel Doc	BioRad Laboratories, Milan, I

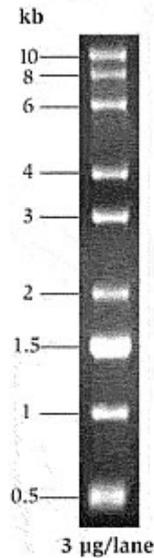
---

APPENDIX 14: DNA, RNA AND PROTEIN MARKERS

1kb Plus DNA Ladder (Invitrogen) on 0.9 % agarose gel stained with ethidium bromide.



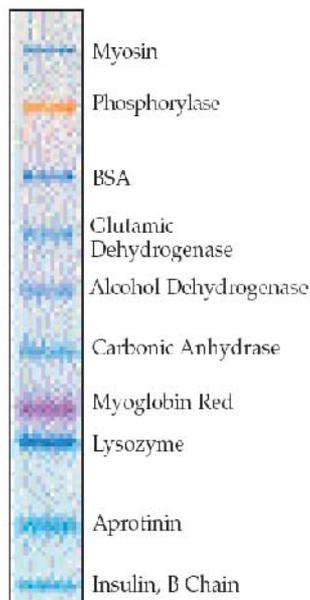
0.5-10 kb RNA Ladder (Invitrogen) on 1.2 % formaldehyde agarose gel stained with ethidium bromide.



Pre-stained SeeBlue Plus2 protein marker (Invitrogen)

Protein

Approximate Molecular Weights (kDa)



	Tris-Glycine	Tricine	NuPAGE® MES	NuPAGE® MOPS
Myosin	250	210	188	191
Phosphorylase	148	105	98	97
BSA	98	78	62	64
Glutamic Dehydrogenase	64	55	49	51
Alcohol Dehydrogenase	50	45	38	39
Carbonic Anhydrase	36	34	28	28
Myoglobin Red	22	17	17	19
Lysozyme	16	16	14	14
Aprotinin	6	7	6	n/a
Insulin, B Chain	4	4	3	n/a

NuPAGE® Novex® Bis-Tris 4-12% Gel

## 6 REFERENCES

Abee, T., T. R. Klaenhammer & L. Letellier (1994). "Kinetic-studies of the action of lactacin F, a bacteriocin produced by *Lactobacillus johnsonii* that forms poration complexes in the cytoplasmic membrane." *Applied and Environmental Microbiology* **60**(3): 1006-1013.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers & D. J. Lipman (1990). "Basic Local Alignment Search Tool." *Journal of Molecular Biology* **215**(3): 403-410.

Amsterdam, D. (1996). Susceptibility testing of antimicrobials in liquid media. *Antibiotics in laboratory medicine*. V. Loman. Baltimore, Williams and Wilkins: 52-111.

Anderssen, E. L., D. B. Diep, I. F. Nes, V. G. H. Eijsink & J. Nissen-Meyer (1998). "Antagonistic activity of *Lactobacillus plantarum* C11: Two new two-peptide bacteriocins, plantaricins EF and JK, and the induction factor plantaricin A." *Applied and Environmental Microbiology* **64**(6): 2269-2272.

Apweiler, R., T. K. Attwood, A. Bairoch, A. Bateman, E. Birney, M. Biswas, P. Bucher, T. Cerutti, F. Corpet, M. D. R. Croning, R. Durbin, L. Falquet, W. Fleischmann, J. Gouzy, H. Hermjakob, N. Hulo, I. Jonassen, D. Kahn, A. Kanapin, Y. Karavidopoulou, R. Lopez, B. Marx, N. J. Mulder, T. M. Oinn, M. Pagni, F. Servant, C. J. A. Sigrist & E. M. Zdobnov (2001). "The InterPro database, an integrated documentation resource for protein families, domains and functional sites." *Nucleic Acids Research* **29**(1): 37-40.

Archibald, F. S. & I. Fridovich (1981). "Manganese and defenses against oxygen-toxicity in *Lactobacillus plantarum*." *Journal of Bacteriology* **145**(1): 442-451.

Arner, E. S. J. & A. Holmgren (2000). "Physiological functions of thioredoxin and thioredoxin reductase." *European Journal of Biochemistry* **267**(20): 6102-6109.

Ashby, M. N. (1998). "CaaX converting enzymes." *Current Opinion in Lipidology* **9**(2): 99-102.

Asmundson, R. V., D. G. Moore, W. Winiata, W. J. Kelly, C. M. Huang & A. Hardacre (1994). Kaanga wai: a New Zealand maori corn fermentation - general characteristics. *Advances in Microbial Processes for the Utilization of Tropical Raw Materials in the Production of Food Products*. P. C. Sanchez, UNESCO Regional Network for Microbiology in Southeast Asia: 123-134.

Aukrust, T. & H. Blom (1992). "Transformation of *Lactobacillus* strains used in meat and vegetable fermentations." *Food Research International* **25**(4): 253-261.

Bartolucci, S., G. De Simone, S. Galdiero, R. Improta, V. Menchise, C. Pedone, E. Pedone & M. Saviano (2003). "An integrated structural and computational study of the thermostability of two thioredoxin mutants from *Alicyclobacillus acidocaldarius*." *Journal of Bacteriology* **185**(14): 4285-4289.

Bauer, R. & L. M. T. Dicks (2005). "Mode of action of lipid II-targeting lantibiotics." *International Journal of Food Microbiology* **101**(2): 201-216.

Bendtsen, J. D., H. Nielsen, G. von Heijne & S. Brunak (2004). "Improved prediction of signal peptides: SignalP 3.0." *Journal of Molecular Biology* **340**(4): 783-795.

## References

- Bernardeau, M., J. Vernoux, S. Henri-Dubernet & M. Guéguen (2008). "The *Lactobacillus* genus." *International Journal of Food Microbiology* **126**(3): 278-285.
- Bernsel, A., H. Viklund, J. Falk, E. Lindahl, G. von Heijne & A. Elofsson (2008). "Prediction of membrane protein topology from first principles." *Proceedings of the National Academy of Sciences of the United States of America* **105**(20): 7177-7181.
- Bernsel, A., H. Viklund, A. Hennerdal & A. Elofsson (2009). "Consensus prediction of membrane protein topology." *Nucleic Acids Research* **in press**.
- Bhunia, A. K., M. C. Johnson, B. Ray & N. Kalchayanand (1991). "Mode of action of pediocin Ach from *Pediococcus acidilactici* H on sensitive bacterial strains." *Journal of Applied Bacteriology* **70**(1): 25-33.
- Bierbaum, G. & H.-G. Sahl (1985). "Induction of autolysis of staphylococci by the basic peptide antibiotics Pep5 and nisin and their influence on the activity of autolytic enzymes." *Archives of Microbiology* **142**: 249-254.
- Boix, E., G. J. Swaminathan, Y. N. Zhang, R. Natesh, K. Brew & K. R. Acharya (2001). "Structure of UDP complex of UDP-galactose :beta-galactoside-alpha-1,3-galactosyltransferase at 1.53-angstrom resolution reveals a conformational change in the catalytically important C terminus." *Journal of Biological Chemistry* **276**(51): 48608-48614.
- Boix, E., Y. N. Zhang, G. J. Swaminathan, K. Brew & K. R. Acharya (2002). "Structural basis of ordered binding of donor and acceptor substrates to the retaining glycosyltransferase, alpha-1,3-galactosyltransferase." *Journal of Biological Chemistry* **277**(31): 28310-28318.
- Bourne, Y. & B. Henrissat (2001). "Glycoside hydrolases and glycosyltransferases: families and functional modules." *Current Opinion in Structural Biology* **11**(5): 593-600.
- Bradford, M. M. (1976). "Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding." *Analytical Biochemistry* **72**(1-2): 248-254.
- Breton, C. & A. Imberty (1999). "Structure/function studies of glycosyltransferases." *Current Opinion in Structural Biology* **9**(5): 563-571.
- Breton, C., J. Mucha & C. Jeanneau (2001). "Structural and functional features of glycosyltransferases." *Biochimie* **83**(8): 713-718.
- Breukink, E. & B. de Kruijff (2006). "Lipid II as a target for antibiotics." *Nature Reviews Drug Discovery* **5**(4): 321-332.
- Breukink, E., I. Wiedemann, C. van Kraaij, O. P. Kuipers, H. G. Sahl & B. de Kruijff (1999). "Use of the cell wall precursor lipid II by a pore-forming peptide antibiotic." *Science* **286**(5448): 2361-2364.
- Brock, T. D. (1999). *Brock Biology of Microorganisms*, Prentice Hall.
- Brotz, H., G. Bierbaum, K. Leopold, P. E. Reynolds & H. G. Sahl (1998). "The lantibiotic mersacidin inhibits peptidoglycan synthesis by targeting lipid II." *Antimicrobial Agents and Chemotherapy* **42**(1): 154-160.

## References

Brown, T., K. Machery & T. Du (2004). Analysis of RNA by Northern and slot blot hybridisation. *Current Protocols in Molecular Biology*. Massachusetts, Wiley Interscience.

Bryson, K., L. J. McGuffin, R. L. Marsden, J. J. Ward, J. S. Sodhi & D. T. Jones (2005). "Protein structure prediction servers at University College London." *Nucleic Acids Research* **33**: W36-W38.

Bullock, W. O., J. M. Fernandez & J. M. Short (1987). "Xl1-Blue - a high-efficiency plasmid transforming *Escherichia coli* strain with beta-galactosidase selection." *Biotechniques* **5**(4): 376-379.

Caldéron-Santoyo, M., P. G. Mendonza-García, M. A. García-Alvarado & B. I. Escudero-Abarca (2001). "Effect of physical factors on the production of bacteriocin from *Pediococcus acidilactici* ITV26. ." *J. Ind. Microbiol. Biotechnol.* **26**: 191-195.

Caldinia, G., F. Trottaa, M. Villarinib, M. Morettib, R. Pasquinib, G. Scassellati-Sforzolinib & G. Cenci (2005). "Screening of potential lactobacilli antigenotoxicity by microbial and mammalian cell-based tests." *International Journal of Food Microbiology* **102**(1): 37-47.

Campbell, J. A., G. J. Davies, V. Bulone & B. Henrissat (1997). "A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities." *Biochemical Journal* **326**: 929-939.

Casey, P. J., P. A. Solski, C. J. Der & J. E. Buss (1989). "P21ras is modified by a farnesyl isoprenoid." *Proceedings of the National Academy of Sciences of the United States of America* **86**(21): 8323-8327.

Casey, P. J., J. A. Thissen & J. F. Moomaw (1991). "Enzymatic modification of proteins with a geranylgeranyl isoprenoid." *Proceedings of the National Academy of Sciences of the United States of America* **88**(19): 8631-8635.

Castric, P., F. J. Cassels & R. W. Carlson (2001). "Structural characterization of the *Pseudomonas aeruginosa* 1244 pilin glycan." *Journal of Biological Chemistry* **276**(28): 26479-26485.

Chagnaud, P., K. Machinis, L. A. Coutte, A. Marecat & A. Mercenier (2001). "Rapid PCR-based procedure to identify lactic acid bacteria: application to six common *Lactobacillus* species." *Journal of Microbiological Methods* **44**(2): 139-148.

Chaillou, S., M. C. Champomier-Verges, M. Cornet, A. M. Crutz-Le Coq, A. M. Dudez, V. Martin, S. Beaufiles, E. Darbon-Rongere, R. Bossy, V. Loux & M. Zagorec (2005). "The complete genome sequence of the meat-borne lactic acid bacterium *Lactobacillus sakei* 23K." *Nature Biotechnology* **23**(12): 1527-1533.

Charnock, S. J. & G. J. Davies (1999a). "Cloning, crystallization and preliminary X-ray analysis of a nucleotide-diphospho-sugar transferase *spsA* from *Bacillus subtilis*." *Acta Crystallographica Section D-Biological Crystallography* **55**: 677-678.

Charnock, S. J. & G. J. Davies (1999b). "Structure of the nucleotide-diphospho-sugar transferase, *SpsA* from *Bacillus subtilis*, in native and nucleotide-complexed forms." *Biochemistry* **38**(20): 6380-6385.

Chen, Y., R. D. Ludescher & T. J. Montville (1997a). "Electrostatic interactions, but not the YGNGV consensus motif, govern the binding of pediocin PA-1 and its fragments to phospholipid vesicles." *Applied and Environmental Microbiology* **63**: 4770-4777.

## References

- Chen, Y., R. D. Ludescher & T. J. Montville (1998). "Influence of lipid composition on pediocin PA-1 binding to phospholipid vesicles." *Applied and Environmental Microbiology* **64**: 3530-3532.
- Chen, Y., R. Shapira, M. Eisenstein & T. J. Montville (1997b). "Functional characterization of pediocin PA-1 binding to liposomes in the absence of a protein receptor and its relationship to a predicted tertiary structure." *Applied and Environmental Microbiology* **63**: 524-531.
- Cheng, X. G. & G. W. Hart (2001). "Alternative O-glycosylation/O-phosphorylation of serine-16 in murine estrogen receptor beta - Post-translational regulation of turnover and transactivation activity." *Journal of Biological Chemistry* **276**(13): 10570-10575.
- Chesneau, V., K. Vekrellis, M. R. Rosner & D. J. Selkoe (2000). "Purified recombinant insulin-degrading enzyme degrades amyloid beta-protein but does not promote its oligomerization." *Biochemical Journal* **351**: 509-516.
- Chikindas, M. L., M. J. Garciagarcerá, A. J. M. Driessen, A. M. Ledebóer, J. Nissenmeyer, I. F. Nes, T. Abee, W. N. Konings & G. Venema (1993). "Pediocin Pa-1, a bacteriocin from *Pediococcus acidilactici* Pac1.0, forms hydrophilic pores in the cytoplasmic membrane of target-cells." *Applied and Environmental Microbiology* **59**(11): 3577-3584.
- Chiu, C. P. C., A. G. Watts, L. L. Lairson, M. Gilbert, D. Lim, W. W. Wakarchuk, S. G. Withers & N. C. J. Strynadka (2004). "Structural analysis of the sialyltransferase CstII from *Campylobacter jejuni* in complex with a substrate analog." *Nature Structural & Molecular Biology* **11**(2): 163-170.
- Christensen, D. P. & R. W. Hutkins (1992). "Collapse of the proton motive force in *Listeria monocytogenes* caused by a bacteriocin produced by *Pediococcus acidilactici*." *Applied and Environmental Microbiology* **58**(10): 3312-3315.
- Christensson, C., H. Bratt, L. J. Collins, T. Coolbear, R. Holland, M. W. Lubbers, P. W. O'Toole & J. R. Reid (2002). "Cloning and expression of an oligopeptidase, PepO, with novel specificity from *Lactobacillus rhamnosus* HN001 (DR20)." *Applied and Environmental Microbiology* **68**(1): 254-262.
- Claesson, M. J., D. van Sinderen & P. W. O'Toole (2007). "The genus *Lactobacillus* - a genomic basis for understanding its diversity." *Fems Microbiology Letters* **269**(1): 22-28.
- Claros, M. G. & G. von Heijne (1994). "Toppred-Ii - an improved software for membrane protein structure predictions." *Computer Applications in the Biosciences* **10**(6): 685-686.
- Cochu, A., C. Vadeboncoeur, S. Moineau & M. Frenette (2003). "Genetic and biochemical characterization of the phosphoenolpyruvate : glucose/mannose phosphotransferase system of *Streptococcus thermophilus*." *Applied and Environmental Microbiology* **69**(9): 5423-5432.
- Cotter, P. D., C. Hill & R. P. Ross (2005). "Bacteriocins: developing innate immunity for food." *Nature Reviews Microbiology* **3**(10): 777-788.
- Coutinho, P. M., E. Deleury, G. J. Davies & B. Henrissat (2003). "An evolving hierarchical family classification for glycosyltransferases." *Journal of Molecular Biology* **328**(2): 307-317.
- Craven, R. A., D. J. F. Griffiths, K. S. Sheldrick, R. E. Randall, I. M. Hagan & A. M. Carr (1998). "Vectors for the expression of tagged proteins in *Schizosaccharomyces pombe*." *Gene* **221**(1): 59-68.

## References

- D'Andrea, L. D. & L. Regan (2003). "TPR proteins: the versatile helix." *Trends in Biochemical Sciences* **28**(12): 655-662.
- Dalet, K., C. Briand, Y. Cenatiempo & Y. Hechard (2000). "The rpoN gene of *Enterococcus faecalis* directs sensitivity to subclass IIa bacteriocins." *Current Microbiology* **41**(6): 441-443.
- Dalet, K., Y. Cenatiempo, P. Cossart, Y. Hechard & C. European Listeria Genome (2001). "A sigma(54)-dependent PTS permease of the mannose family is responsible for sensitivity of *Listeria monocytogenes* to mesentericin Y105." *Microbiology-Sgm* **147**: 3263-3269.
- Das, A. K., P. T. W. Cohen & D. Barford (1998). "The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions." *Embo Journal* **17**(5): 1192-1199.
- Davidson, A. L., E. Dassa, C. Orelle & J. Chen (2008). "Structure, function, and evolution of bacterial ATP-binding cassette systems." *Microbiology and Molecular Biology Reviews* **72**(2): 317-364.
- Davies, G. J., T. M. Gloster & B. Henrissat (2005). "Recent structural insights into the expanding world of carbohydrate-active enzymes." *Current Opinion in Structural Biology* **15**(6): 637-645.
- Davis, R. W., D. Botstein & J. R. Roth (1980). *Advanced bacterial genetics, a manual for genetic engineering*. New York, USA, Cold Spring Harbour Laboratory, Cold Spring Harbour.
- Dawson, R. J. P. & K. P. Locher (2006). "Structure of a bacterial multidrug ABC transporter." *Nature* **443**(7108): 180-185.
- de Jong, A., S. van Hijum, J. J. E. Bijlsma, J. Kok & O. P. Kuipers (2006). "BAGEL: a web-based bacteriocin genome mining tool." *Nucleic Acids Research* **34**: W273-W279.
- DeLano, W. L. (2002). The PyMOL Molecular Graphics System. Palo Alto, CA, USA, DeLano Scientific.
- Delfederico, L., A. Hollmann, M. Martinez, N. G. Iglesias, G. De Antoni & L. Semorile (2006). "Molecular identification and typing of lactobacilli isolated from kefir grains." *Journal of Dairy Research* **73**(1): 20-27.
- Delgado, A., F. N. Arroyo Lopez, D. Brito, C. Peres, P. Fevereiro & A. Garrido-Fernandez (2007). "Optimum bacteriocin production by *Lactobacillus plantarum* 17.2b requires absence of NaCl and apparently follows a mixed metabolite kinetics." *Journal of Biotechnology* **130**(2): 193-201.
- Delgado, A., D. Brito, C. Peres, F. Noe-Arroyo & A. Garrido-Fernandez (2005). "Bacteriocin production by *Lactobacillus pentosus* B96 can be expressed as a function of temperature and NaCl concentration." *Food Microbiology* **22**(6): 521-528.
- Demel, R. A., T. Peelen, R. J. Siezen, B. deKruiff & O. P. Kuipers (1996). "Nisin Z, mutant nisin Z and lactacin 481 interactions with anionic lipids correlate with antimicrobial activity - A monolayer study." *European Journal of Biochemistry* **235**(1-2): 267-274.
- Desper, R. & O. Gascuel (2004). "Theoretical foundation of the balanced minimum evolution method of phylogenetic inference and its relationship to weighted least-squares tree fitting." *Molecular Biology and Evolution* **21**(3): 587-598.

## References

- Diep, D. B., L. Axelsson, C. Grefslı & I. F. Nes (2000). "The synthesis of the bacteriocin sakacin A is a temperature-sensitive process regulated by a pheromone peptide through a three-component regulatory system." *Microbiology-Uk* **146**: 2155-2160.
- Diep, D. B., L. S. Havarstein & I. F. Nes (1995). "A bacteriocin-like peptide induces bacteriocin synthesis in *Lactobacillus plantarum* C11." *Molecular Microbiology* **18**(4): 631-639.
- Diep, D. B., L. S. Havarstein & I. F. Nes (1996). "Characterization of the locus responsible for the bacteriocin production in *Lactobacillus plantarum* C11." *Journal of Bacteriology* **178**(15): 4472-4483.
- Diep, D. B., L. S. Havarstein, J. Nissenmeyer & I. F. Nes (1994). "The gene encoding plantaricin-a, a bacteriocin from *Lactobacillus plantarum* C11, is located on the same transcription unit as an Agr-like regulatory system." *Applied and Environmental Microbiology* **60**(1): 160-166.
- Diep, D. B., O. Johnsborg, P. A. Risoen & I. F. Nes (2001). "Evidence for dual functionality of the operon plnABCD in the regulation of bacteriocin production in *Lactobacillus plantarum*." *Molecular Microbiology* **41**(3): 633-644.
- Diep, D. B., R. Myhre, O. Johnsborg, A. Aakra & I. F. Nes (2003). "Inducible bacteriocin production in *Lactobacillus* is regulated by differential expression of the pln operons and by two antagonizing response regulators, the activity of which is enhanced upon phosphorylation." *Molecular Microbiology* **47**(2): 483-494.
- Diep, D. B., M. Skaugen, Z. Salehian, H. Holo & I. F. Nes (2007). "Common mechanisms of target cell recognition and immunity for class II bacteriocins." *Proceedings of the National Academy of Sciences of the United States of America* **104**(7): 2384-2389.
- Diep, D. B., D. Straume, M. Kjosaa, C. Torresb & I. F. Nes (2009). "An overview of the mosaic bacteriocin pln loci from *Lactobacillus plantarum*." **In Press**.
- Dobson, A. E., R. B. Sanozky-Dawes & T. R. Klaenhammer (2007). "Identification of an operon and inducing peptide involved in the production of lactacin B by *Lactobacillus acidophilus*." *Journal of Applied Microbiology* **103**: 1766-1778.
- Dodd, H. M., N. Horn & M. J. Gasson (1990). "Analysis of the genetic determinant for production of the peptide antibiotic nisin." *Journal of General Microbiology* **136**: 555-566.
- Doherty, H., C. Condon & P. Owen (1982). "Resolution and *in vitro* glycosylation of membrane-glycoproteins in *Micrococcus luteus* (Lysodeikticus)." *Fems Microbiology Letters* **15**(4): 331-336.
- Dong, D. L. Y. & G. W. Hart (1994). "Purification and characterization of an O-GlcNAc selective N-acetyl-beta-D-glucosaminidase from rat spleen cytosol." *Journal of Biological Chemistry* **269**(30): 19321-19330.
- Dorenbos, R., T. Stein, J. Kabel, C. Bruand, A. Bolhuis, S. Bron, W. J. Quax & J. M. van Dıjl (2002). "Thiol-disulfide oxidoreductases are essential for the production of the lantibiotic sublancin 168." *Journal of Biological Chemistry* **277**(19): 16682-16688.
- Drıder, D., G. Fimland, Y. Hechard, L. M. McMullen & H. Prevost (2006). "The continuing story of class IIa bacteriocins." *Microbiology and Molecular Biology Reviews* **70**(2): 564-582.

## References

Dubnau, D. & C. M. Lovett, Jr. (2002). *Bacillus subtilis* and its closest relatives: From genes to cells. *Bacillus subtilis and Its Closest Relatives: From Genes to Cells*. A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC, Am. Soc. Microbiol.: 453-471.

Dunny, G. M. & B. A. B. Leonard (1997). "Cell-cell communication in gram-positive bacteria." *Annual Review of Microbiology* **51**: 527-564.

Eijsink, V. G. H., L. Axelsson, D. B. Diep, L. S. Havarstein, H. Holo & I. F. Nes (2002). "Production of class II bacteriocins by lactic acid bacteria; an example of biological warfare and communication." *Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology* **81**(1-4): 639-654.

Eijsink, V. G. H., M. B. Brurberg, P. H. Middelhoven & O. F. Nes (1996). "Induction of bacteriocin production in *Lactobacillus sake* by a secreted peptide." *Journal of Bacteriology* **178**(8): 2232-2237.

Eijsink, V. G. H., M. Skeie, P. H. Middelhoven, M. B. Brurberg & I. F. Nes (1998). "Comparative studies of class IIa bacteriocins of lactic acid bacteria." *Applied and Environmental Microbiology* **64**(9): 3275-3281.

Ennahar, S., T. Sashihara, K. Sonomoto & A. Ishizaki (2000). "Class IIa bacteriocins: biosynthesis, structure and activity." *Fems Microbiology Reviews* **24**(1): 85-106.

Fath, M. J. & R. Kolter (1993). "ABC-transporters - bacterial exporters." *Microbiological Reviews* **57**(4): 995-1017.

Fimland, G., R. Jack, G. Jung, I. F. Nes & J. Nissen-Meyer (1998). "The bactericidal activity of pediocin PA-1 is specifically inhibited by a 15-mer fragment that spans the bacteriocin from the center toward the C terminus." *Applied and Environmental Microbiology* **64**(12): 5057-5060.

Fimland, G., L. Johnsen, L. Axelsson, M. B. Brurberg, I. F. Nes, V. G. H. Eijsink & J. Nissen-Meyer (2000). "A C-terminal disulfide bridge in pediocin-like bacteriocins renders bacteriocin activity less temperature dependent and is a major determinant of the antimicrobial spectrum." *Journal of Bacteriology* **182**(9): 2643-2648.

Fimland, G., L. Johnsen, B. Dalhus & J. Nissen-Meyer (2005). "Pediocin-like antimicrobial peptides (class IIa bacteriocins) and their immunity proteins: biosynthesis, structure, and mode of action." *Journal of Peptide Science* **11**(11): 688-696.

Finn, R. D., J. Tate, J. Mistry, P. C. Coghill, S. J. Sammut, H. R. Hotz, G. Ceric, K. Forslund, S. R. Eddy, E. L. L. Sonnhammer & A. Bateman (2008). "The Pfam protein families database." *Nucleic Acids Research* **36**: D281-D288.

Fleury, Y., M. A. Dayem, J. J. Montagne, E. Chaboisseau, J. P. LeCaer, P. Nicolas & A. Delfour (1996). "Covalent structure, synthesis, and structure-function studies of mesentericin Y 105(37), a defensive peptide from gram-positive bacteria *Leuconostoc mesenteroides*." *Journal of Biological Chemistry* **271**(24): 14421-14429.

Franke, C. M., K. J. Leenhouts, A. J. Haandrikman, J. Kok, G. Venema & K. Venema (1996). "Topology of LcnD, a protein implicated in the transport of bacteriocins from *Lactococcus lactis*." *Journal of Bacteriology* **178**(6): 1766-1769.

## References

- Franke, C. M., J. Tiemersma, G. Venema & J. Kok (1999). "Membrane topology of the lactococcal bacteriocin ATP-binding cassette transporter protein LcnC - Involvement of LcnC in lactococcal A maturation." *Journal of Biological Chemistry* **274**(13): 8484-8490.
- Franz, C., M. E. Stiles & M. J. van Belkum (2000b). "Simple method to identify bacteriocin induction peptides and to auto-induce bacteriocin production at low cell density." *Fems Microbiology Letters* **186**(2): 181-185.
- Franz, C., M. J. van Belkum, R. W. Worobo, J. C. Vederas & M. E. Stiles (2000a). "Characterization of the genetic locus responsible for production and immunity of carnobacteriocin A: the immunity gene confers cross-protection to enterocin B." *Microbiology-Uk* **146**: 621-631.
- Fremaux, C., Y. Hechard & Y. Cenatiempo (1995). "Mesentericin Y105 gene clusters in *Leuconostoc mesenteroides* Y105." *Microbiology-Uk* **141**: 1637-1645.
- Fukuchi, S. & K. Nishikawa (2001). "Protein surface amino acid compositions distinctively differ between thermophilic and mesophilic bacteria." *Journal of Molecular Biology* **309**(4): 835-843.
- Galperin, M. Y. (2006). "Structural classification of bacterial response regulators: Diversity of output domains and domain combinations." *Journal of Bacteriology* **188**(12): 4169-4182.
- Galperin, M. Y. (2008). "Telling bacteria: Do not LytTR." *Structure* **16**(5): 657-659.
- Gambello, M. J. & B. H. Iglewski (1991). "Cloning and characterization of the *Pseudomonas aeruginosa* Lasr gene, a transcriptional activator of elastase expression." *Journal of Bacteriology* **173**(9): 3000-3009.
- Gao, R., T. R. Mack & A. M. Stock (2007). "Bacterial response regulators: versatile regulatory strategies from common domains." *Trends in Biochemical Sciences* **32**: 225-234.
- Gasson, M. J. (1983). "Plasmid complements of *Streptococcus lactis* Ncd0-712 and other lactic streptococci after protoplast-induced curing." *Journal of Bacteriology* **154**(1): 1-9.
- Gilbreth, S. E. & G. A. Somkuti (2005). "Thermophilin 110: A bacteriocin of *Streptococcus thermophilus* ST110." *Current Microbiology* **51**(3): 175-182.
- Gooley, A. A., B. J. Classon, R. Marschalek & K. L. Williams (1991). "Glycosylation sites identified by detection of glycosylated amino acids released from Edman degradation - the identification of Xaa-Pro-Xaa-Xaa as a motif for Thr-O-glycosylation." *Biochemical and Biophysical Research Communications* **178**(3): 1194-1201.
- Gostick, D. O., J. Green, A. S. Irvine, M. J. Gasson & J. R. Guest (1998). "A novel regulatory switch mediated by the FNR-like protein of *Lactobacillus casei*." *Microbiology-Sgm* **144**: 705-717.
- Gostick, D. O., H. G. Griffin, C. A. Shearman, C. Scott, J. Green, M. J. Gasson & J. R. Guest (1999). "Two operons that encode FNR-like proteins in *Lactococcus lactis*." *Molecular Microbiology* **31**(5): 1523-1535.
- Grebe, T. W. & J. B. Stock (1999). "The histidine protein kinase superfamily." *Advances in Microbial Physiology, Vol 41* **41**: 139-227.

## References

Greis, K. D., B. K. Hayes, F. I. Comer, M. Kirk, S. Barnes, T. L. Lowary & G. W. Hart (1996). "Selective detection and site-analysis of *O*-GlcNAc-modified glycopeptides by beta-elimination and tandem electrospray mass spectrometry." *Analytical Biochemistry* **234**(1): 38-49.

Gross, E. & J. L. Morell (1971). "Structure of nisin." *Journal of the American Chemical Society* **93**(18): 4634-4635.

Guest, J. R., J. Green, A. S. Irvine & S. Spiro (1995). The FNR modulon and FNR-regulated gene expression. *Regulation of gene expression in Escherichia coli*.

Guyonnet, D., C. Fremaux, Y. Cenatiempo & J. M. Berjeaud (2000). "Method for rapid purification of class IIa bacteriocins and comparison of their activities." *Applied and Environmental Microbiology* **66**(4): 1744-1748.

Hall, G., M. Shah, P. A. McEwan, C. Laughton, M. Stevens, A. Westwell & J. Emsley (2006). "Structure of *Mycobacterium tuberculosis* thioredoxin C." *Acta Crystallographica Section D-Biological Crystallography* **62**: 1453-1457.

Haltiwanger, R. S., G. D. Holt & G. W. Hart (1990). "Enzymatic addition of *O*-GlcNAc to nuclear and cytoplasmic proteins - identification of a uridine diphospho-*N*-acetylglucosamine-peptide beta-*N*-acetylglucosaminyltransferase." *Journal of Biological Chemistry* **265**(5): 2563-2568.

Han, C. S., G. Xie, J. F. Challacombe, M. R. Altherr, S. S. Bhotika, D. Bruce, C. S. Campbell, M. L. Campbell, J. Chen, O. Chertkov, C. Cleland, M. Dimitrijevic, N. A. Doggett, J. J. Fawcett, T. Glavina, L. A. Goodwin, K. K. Hill, P. Hitchcock, P. J. Jackson, P. Keim, A. R. Kewalramani, J. Longmire, S. Lucas, S. Malfatti, K. McMurry, L. J. Meincke, M. Misra, B. L. Moseman, M. Mundt, A. C. Munk, R. T. Okinaka, B. Parson-Quintana, L. P. Reilly, P. Richardson, D. L. Robinson, E. Rubin, E. Saunders, R. Tapia, J. G. Tesmer, N. Thayer, L. S. Thompson, H. Tice, L. O. Ticknor, P. L. Wills, T. S. Brettin & P. Gilna (2006). "Pathogenomic sequence analysis of *Bacillus cereus* and *Bacillus thuringiensis* isolates closely related to *Bacillus anthracis*." *Journal of Bacteriology* **188**(9): 3382-3390.

Hanahan, D. (1983). "Studies on transformation of *Escherichia coli* with plasmids." *Journal of Molecular Biology* **166**(4): 557-580.

Hartmann, E. & H. König (1989). "Uridine and dolichyl diphosphate activated oligosaccharides are intermediates in the biosynthesis of the S-layer glycoprotein of *Methanothermobacter feravidus*." *Archives of Microbiology* **151**(3): 274-281.

Hartmann, E., P. Messner, G. Allmeier & H. König (1993). "Proposed pathway for biosynthesis of the S-layer glycoprotein of *Bacillus alvei*." *Journal of Bacteriology* **175**(14): 4515-4519.

Harty, D. W. S., Y. J. Chen, C. L. Simpson, T. Berg, S. L. Cook, J. A. Mayo, N. Hunter & N. A. Jacques (2004). "Characterisation of a novel homodimeric *N*-acetyl-beta-D-glucosaminidase from *Streptococcus gordonii*." *Biochemical and Biophysical Research Communications* **319**(2): 439-447.

Hasper, H. E., B. de Kruijff & E. Breukink (2004). "Assembly and stability of nisin-Lipid II pores." *Biochemistry* **43**(36): 11567-11575.

Hasper, H. E., N. E. Kramer, J. L. Smith, J. D. Hillman, C. Zachariah, O. P. Kuipers, B. de Kruijff & E. Breukink (2006). "An alternative bactericidal mechanism of action for lantibiotic peptides that target lipid II." *Science* **313**(5793): 1636-1637.

## References

- Havarstein, L. S., D. B. Diep & I. F. Nes (1995). "A family of bacteriocin ABC-transporters carry out proteolytic processing of their substrates concomitant with export." *Molecular Microbiology* **16**(2): 229-240.
- Havarstein, L. S. & D. A. Morrison (1999). "Quorum sensing and peptide pheromones in streptococcal competence for genetic transformation." *Cell-Cell Signaling in Bacteria*: 9-26.
- Hechard, Y., C. Jayat, F. Letellier, R. Julien, Y. Cenatiempo & M. H. Ratinaud (1992). "Online visualization of the competitive behavior of antagonistic bacteria." *Applied and Environmental Microbiology* **58**(11): 3784-3786.
- Hechard, Y., C. Pelletier, Y. Cenatiempo & J. Frere (2001). "Analysis of sigma(54)-dependent genes in *Enterococcus faecalis*: a mannose PTS permease (EIIMan) is involved in sensitivity to a bacteriocin, mesentericin Y105." *Microbiology-Sgm* **147**: 1575-1580.
- Hechard, Y. & H. G. Sahl (2002). "Mode of action of modified and unmodified bacteriocins from Gram-positive bacteria." *Biochimie* **84**(5-6): 545-557.
- Hecker, M., J. Pane-Farre & U. Volker (2007). "SigB-dependent general stress response in *Bacillus subtilis* and related gram-positive bacteria." *Annual Review of Microbiology* **61**: 215-236.
- Hecker, M. & U. Volker (2001). "General stress response of *Bacillus subtilis* and other bacteria." *Advances in Microbial Physiology, Vol 44* **44**: 35-91.
- Hegge, F. T., P. G. Hitchen, F. E. Aas, H. Kristiansen, C. Lovold, W. Egge-Jacobsen, M. Panico, W. Y. Leong, V. Bull, M. Virji, H. R. Morris, A. Dell & M. Koomey (2004). "Unique modifications with phosphocholine and phosphoethanolamine define alternate antigenic forms of *Neisseria gonorrhoeae* type IV pili." *Proceedings of the National Academy of Sciences of the United States of America* **101**(29): 10798-10803.
- Henderson, J. T., A. L. Chopko & P. D. Vanwassenaar (1992). "Purification and primary structure of pediocin Pa-1 produced by *Pediococcus acidilactici* Pac-10." *Archives of Biochemistry and Biophysics* **295**(1): 5-12.
- Heng, N. C. K. & J. R. Tagg (2006). "What's in a name? Class distinction for bacteriocins." *Nature Reviews Microbiology* **4**(2).
- Heng, N. C. K., P. A. Wescombe, J. P. Burton, R. W. Jack & J. R. Tagg (2007). The diversity of bacteriocins produced by Gram-positive bacteria. *Bacteriocins - Ecology and Evolution*. R. M. C. MA. Heidelberg, Germany, Springer: 45-92.
- Henrissat, B., G. Sulzenbacher & Y. Bourne (2008). "Glycosyltransferases, glycoside hydrolases: surprise, surprise!" *Current Opinion in Structural Biology* **18**(5): 527-533.
- Herranz, C. & A. J. M. Driessen (2005). "Sec-mediated secretion of bacteriocin enterocin P by *Lactococcus lactis*." *Applied and Environmental Microbiology* **71**(4): 1959-1963.
- Hightower, K. E., P. J. Casey & C. A. Fierke (2001). "Farnesylation of nonpeptidic thiol compounds by protein farnesyltransferase." *Biochemistry* **40**(4): 1002-1010.

## References

- Holo, H. & I. F. Nes (1989). "High-frequency transformation, by electroporation, of *Lactococcus lactis* subsp. *cremoris* grown with glycine in osmotically stabilized media." *Applied and Environmental Microbiology* **55**(12): 3119-3123.
- Hsu, S. T. D., E. Breukink, E. Tischenko, M. A. G. Lutters, B. de Kruijff, R. Kaptein, A. Bonvin & N. A. J. van Nuland (2004). "The nisin-lipid II complex reveals a pyrophosphate cage that provides a blueprint for novel antibiotics." *Nature Structural & Molecular Biology* **11**(10): 963-967.
- Hu, Y. N. & S. Walker (2002). "Remarkable structural similarities between diverse glycosyltransferases." *Chemistry & Biology* **9**(12): 1287-1296.
- Hurtado-Guerrero, R., H. C. Dorfmüller & D. M. Van Aalten (2008). "Molecular mechanisms of O-GlcNAcylation." *Current Opinion in Structural Biology* **18**(5): 551-557.
- Igura, M., N. Maita, T. Obita, J. Kamishikiryo, K. Maenaka & D. Kohda (2007). "Purification, crystallization and preliminary X-ray diffraction studies of the soluble domain of the oligosaccharyltransferase STT3 subunit from the thermophilic archaeon *Pyrococcus furiosus*." *Acta Crystallographica Section F-Structural Biology and Crystallization Communications* **63**: 798-801.
- Ingham, A. B., K. W. Sproat, M. L. V. Tizard & R. J. Moore (2005). "A versatile system for the expression of nonmodified bacteriocins in *Escherichia coli*." *Journal of Applied Microbiology* **98**(3): 676-683.
- Ingram, L. C. (1969). "Synthesis of antibiotic nisin - formation of lanthionine and beta-methylanthionine." *Biochimica Et Biophysica Acta* **184**(1): 216-219.
- Innis, M. A., D. H. Gelfand, J. J. Sninsky & T. J. White (1990). "PCR protocols: a guide to methods and applications." *PCR protocols: a guide to methods and applications*: xviii + 482 pp.
- Iuchi, S., T. Fujiwara & E. C. C. Lin (1990). "The *Arca* gene of *Escherichia coli* encodes a sensor-regulator protein for anaerobic repression of the *Arc* modulon." *Molecular Microbiology* **4**(5): 715-727.
- Iuchi, S. & E. C. C. Lin (1988). "Arca (Dye), a global regulatory gene in *Escherichia coli* mediating repression of enzymes in aerobic pathways." *Proceedings of the National Academy of Sciences of the United States of America* **85**(6): 1888-1892.
- Jacob, F., A. Lwoff, A. Siminovitch & E. Wollman (1953). "Definition de quelques termes relatifs a la lysogenie." *Annales De L Institut Pasteur* **84**(1): 222-224.
- Jaffe, H., Veeranna & H. C. Pant (1998). "Characterization of serine and threonine phosphorylation sites in beta-elimination ethanethiol addition-modified proteins by electrospray tandem mass spectrometry and database searching." *Biochemistry* **37**(46): 16211-16224.
- Jahreis, K., E. F. Pimentel-Schmitt, R. Bruckner & F. Titgemeyer (2008). "Ins and outs of glucose transport systems in eubacteria." *Fems Microbiology Reviews* **32**(6): 891-907.
- Jetten, A. M., F. Dewindt & G. D. Vogels (1972). "Production and purification of a *Staphylococcus epidermidis* bacteriocin." *Journal of Bacteriology* **112**(1): 235-&.
- Jetten, A. M. & G. D. Vogels (1972a). "Mode of action of a *Staphylococcus epidermidis* bacteriocin." *Antimicrobial Agents and Chemotherapy* **2**(6): 456-&.

## References

- Jetten, A. M. & G. D. Vogels (1972b). "Nature and properties of a *Staphylococcus epidermidis* bacteriocin." *Journal of Bacteriology* **112**(1): 243-&.
- Jimenezdiaz, R., R. M. Riossanchez, M. Desmazeaud, J. L. Ruizbarba & J. C. Piard (1993). "Plantaricins S and T, 2 new bacteriocins produced by *Lactobacillus plantarum* Lpc010 isolated from a green olive fermentation." *Applied and Environmental Microbiology* **59**(5): 1416-1424.
- Johnsen, L., G. Fimland, D. Mantzilas & J. Nissen-Meyer (2004). "Structure-function analysis of immunity proteins of pediocin-like bacteriocins: C-terminal parts of immunity proteins are involved in specific recognition of cognate bacteriocins." *Applied and Environmental Microbiology* **70**(5): 2647-2652.
- Johnsen, L., G. Fimland & J. N. Meyer (2005). "The C-terminal domain of pediocin-like antimicrobial peptides (class IIa bacteriocins) is involved in specific recognition of the C-terminal part of cognate immunity proteins and in determining the antimicrobial spectrum." *Journal of Biological Chemistry* **280**(10): 9243-9250.
- Jones, D. T. (1999). "GenTHREADER: An efficient and reliable protein fold recognition method for genomic sequences." *Journal of Molecular Biology* **287**(4): 797-815.
- Kaneko, Y., H. Kobayashi, P. Kiatpapan, T. Nishimoto, R. Napitupulu, H. Ono & Y. Murooka (2000). "Development of a host-vector system for *Lactobacillus plantarum* L137 isolated from a traditional fermented food produced in the Philippines." *Journal of Bioscience and Bioengineering* **89**(1): 62-67.
- Kawai, Y., R. Kemperman, J. Kok & T. Saito (2004). "The circular bacteriocins gassericin A and circularin A." *Current Protein & Peptide Science* **5**(5): 393-398.
- Kawai, Y., T. Saito, M. Suzuki & T. Itoh (1998). "Sequence analysis by cloning of the structural gene of gassericin A, a hydrophobic bacteriocin produced by *Lactobacillus gasseri* LA39." *Bioscience Biotechnology and Biochemistry* **62**(5): 887-892.
- Kelly, S. M., T. J. Jess & N. C. Price (2005). "How to study proteins by circular dichroism." *Biochimica Et Biophysica Acta-Proteins and Proteomics* **1751**(2): 119-139.
- Kelly, W. J., R. V. Asmundson & C. M. Huang (1996). "Characterization of plantaricin KW30, a bacteriocin produced by *Lactobacillus plantarum*." *Journal of Applied Bacteriology* **81**(6): 657-662.
- Kelly, W. J., R. V. Asmundson, D. G. Moore & C. M. Huang (1994). Microbiological investigation of Kaanga wai. *Advances in Microbial Processes for the Utilization of Tropical Raw Materials in the Production of Food Products*. P. C. Sanchez, UNESCO Regional Network for Microbiology in Southeast Asia: 135-146.
- Kemperman, R., M. Jonker, A. Nauta, O. P. Kuipers & J. Kok (2003). "Functional analysis of the gene cluster involved in production of the bacteriocin circularin A by *Clostridium beijerinckii* ATCC 25752." *Applied and Environmental Microbiology* **69**(10): 5839-5848.
- Kjos, M., O. F. Nes & W. B. Diep (2009). "Class II one-peptide bacteriocins target a phylogenetically defined subgroup of mannose phosphotransferase systems an sensitive cells." *Microbiology and Molecular Biology Reviews* **epub ahead of print**.

## References

Klaenhammer, T. R. (1993). "Genetics of bacteriocins produced by lactic-acid bacteria." *Fems Microbiology Reviews* **12**(1-3): 39-86.

Kleerebezem, M., J. Boekhorst, R. van Kranenburg, D. Molenaar, O. P. Kuipers, R. Leer, R. Tarchini, S. A. Peters, H. M. Sandbrink, M. Fiers, W. Stiekema, R. M. K. Lankhorst, P. A. Bron, S. M. Hoffer, M. N. N. Groot, R. Kerkhoven, M. de Vries, B. Ursing, W. M. de Vos & R. J. Siezen (2003). "Complete genome sequence of *Lactobacillus plantarum* WCFS1." *Proceedings of the National Academy of Sciences of the United States of America* **100**(4): 1990-1995.

Kleerebezem, M., L. E. N. Quadri, O. P. Kuipers & W. M. deVos (1997). "Quorum sensing by peptide pheromones and two-component signal-transduction systems in Gram-positive bacteria." *Molecular Microbiology* **24**(5): 895-904.

Koponen, O., M. Tolonen, M. Q. Qiao, G. Wahlstrom, J. Helin & P. E. J. Saris (2002). "NisB is required for the dehydration and NisC for the lanthionine formation in the post-translational modification of nisin." *Microbiology-Sgm* **148**: 3561-3568.

Koretke, K. K., A. N. Lupas, P. V. Warren, M. Rosenberg & J. R. Brown (2000). "Evolution of two-component signal transduction." *Molecular Biology and Evolution* **17**(12): 1956-1970.

Kotelnikova, E. A. & M. S. Gelfand (2002). "Bacteriocin production by Gram-positive bacteria and the mechanisms of transcriptional regulation." *Russian Journal of Genetics* **38**(6): 628-641.

Krier, F., A. M. Revol-Junelles & P. Germain (1998). "Influence of temperature and pH on production of two bacteriocins by *Leuconostoc mesenteroides* subsp. *mesenteroides* FR52 during batch fermentation." *Applied Microbiology and Biotechnology* **50**(3): 359-363.

Kristiansen, P. E., G. Fimland, D. Mantzilas & J. Nissen-Meyer (2005). "Structure and mode of action of the membrane-permeabilizing antimicrobial peptide pheromone plantaricin A." *Journal of Biological Chemistry* **280**(24): 22945-22950.

Kuipers, A., E. de Boef, R. Rink, S. Fekken, L. D. Kluskens, A. J. M. Driessen, K. Leenhouts, O. P. Kuipers & G. N. Moll (2004). "NisT, the transporter of the lantibiotic nisin, can transport fully modified, dehydrated, and unmodified prenisin and fusions of the leader peptide with non-lantibiotic peptides." *Journal of Biological Chemistry* **279**(21): 22176-22182.

Kuipers, A., J. Wierenga, R. Rink, L. D. Kluskens, A. J. M. Driessen, O. P. Kuipers & G. N. Moll (2006). "Sec-mediated transport of posttranslationally dehydrated peptides in *Lactococcus lactis*." *Applied and Environmental Microbiology* **72**(12): 7626-7633.

Kuipers, O. P., M. M. Beerthuyzen, P. Deruyter, E. J. Luesink & W. M. Devos (1995). "Autoregulation of nisin biosynthesis in *Lactococcus lactis* by signal-transduction." *Journal of Biological Chemistry* **270**(45): 27299-27304.

Kuipers, O. P., P. de Ruyter, M. Kleerebezem & W. M. de Vos (1998). "Quorum sensing-controlled gene expression in lactic acid bacteria." *Journal of Biotechnology* **64**(1): 15-21.

Kunst, F., N. Ogasawara, I. Moszer, A. M. Albertini, G. Alloni, V. Azevedo, M. G. Bertero, P. Bessieres, A. Bolotin, S. Borchert, R. Borriss, L. Boursier, A. Brans, M. Braun, S. C. Brignell, S. Bron, S. Brouillet, C. V. Bruschi, B. Caldwell, V. Capuano, N. M. Carter, S. K. Choi, J. J. Codani, I. F. Connerton, N. J. Cummings, R. A. Daniel, F. Denizot, K. M. Devine, A. Dusterhoft, S. D. Ehrlich, P. T. Emmerson, K. D. Entian, J. Errington, C. Fabret, E. Ferrari, D. Foulger, C. Fritz, M. Fujita, Y. Fujita, S. Fuma, A. Galizzi, N.

## References

- Galleron, S. Y. Ghim, P. Glaser, A. Goffeau, E. J. Golightly, G. Grandi, G. Guiseppi, B. J. Guy, K. Haga, J. Haiech, C. R. Harwood, A. Henaut, H. Hilbert, S. Holsappel, S. Hosono, M. F. Hullo, M. Itaya, L. Jones, B. Joris, D. Karamata, Y. Kasahara, M. KlaerrBlanchard, C. Klein, Y. Kobayashi, P. Koetter, G. Koningstein, S. Krogh, M. Kumano, K. Kurita, A. Lapidus, S. Lardinois, J. Lauber, V. Lazarevic, S. M. Lee, A. Levine, H. Liu, S. Masuda, C. Mauel, C. Medigue, N. Medina, R. P. Mellado, M. Mizuno, D. Moestl, S. Nakai, M. Noback, D. Noone, M. O'Reilly, K. Ogawa, A. Ogiwara, B. Oudega, S. H. Park, V. Parro, T. M. Pohl, D. Portetelle, S. Porwollik, A. M. Prescott, E. Presecan, P. Pujic, B. Purnelle, G. Rapoport, M. Rey, S. Reynolds, M. Rieger, C. Rivolta, E. Rocha, B. Roche, M. Rose, Y. Sadaie, T. Sato, E. Scanlan, S. Schleich, R. Schroeter, F. Scoffone, J. Sekiguchi, A. Sekowska, S. J. Seror, P. Serror, B. S. Shin, B. Soldo, A. Sorokin, E. Tacconi, T. Takagi, H. Takahashi, K. Takemaru, M. Takeuchi, A. Tamakoshi, T. Tanaka, P. Terpstra, A. Tognoni, V. Tosato, S. Uchiyama, M. Vandenbol, F. Vannier, A. Vassarotti, A. Viari, R. Wambutt, E. Wedler, H. Wedler, T. Weitzenegger, P. Winters, A. Wipat, H. Yamamoto, K. Yamane, K. Yasumoto, K. Yata, K. Yoshida, H. F. Yoshikawa, E. Zumstein, H. Yoshikawa & A. Danchin (1997). "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*." *Nature* **390**(6657): 249-256.
- Kuroda, Y., N. Suzuki & T. Kataoka (1993). "The effect of posttranslational modifications on the interaction of Ras2 with adenylyl cyclase." *Science* **259**(5095): 683-686.
- Lagos, R., M. Baeza, G. Corsini, C. Hetz, E. Strahsburger, J. A. Castillo, C. Vergara & O. Monsterio (2001). "Structure, organization and characterization of the gene cluster involved in the production of microcin E492, a channel-forming bacteriocin." *Molecular Microbiology* **42**(1): 229-243.
- Lairson, L. L., C. P. C. Chiu, H. D. Ly, S. M. He, W. W. Wakarchuk, N. C. J. Strynadka & S. G. Withers (2004). "Intermediate trapping on a mutant retaining alpha-galactosyltransferase identifies an unexpected aspartate residue." *Journal of Biological Chemistry* **279**(27): 28339-28344.
- Lairson, L. L., B. Henrissat, G. J. Davies & S. G. Withers (2008). "Glycosyltransferases: Structures, functions, and mechanisms." *Annual Review of Biochemistry* **77**: 521-555.
- Lambert, J. M., R. S. Bongers & M. Kjeerebezem (2007). "Cre-lox-based system for multiple gene deletions and selectable-marker removal in *Lactobacillus plantarum*." *Applied and Environmental Microbiology* **73**(4): 1126-1135.
- Lämmli, U. K. (1970). "Cleavage of structural proteins during assembly of the head of bacteriophage T4." *Nature* **227**: 680-685.
- Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson & D. G. Higgins (2007). "Clustal W and clustal X version 2.0." *Bioinformatics* **23**: 2947-2948.
- Law, J., G. Buist, A. Haandrikman, J. Kok, G. Venema & K. Leenhouts (1995). "A system to generate chromosomal mutations in *Lactococcus lactis* which allows fast analysis of targeted genes." *Journal of Bacteriology* **177**(24): 7011-7018.
- Lazazzera, B. A., I. G. Kurtser, R. S. McQuade & A. D. Grossman (1999). "An autoregulatory circuit affecting peptide signaling in *Bacillus subtilis*." *Journal of Bacteriology* **181**(17): 5193-5200.
- Lechner, J. & F. Wieland (1989). "Structure and biosynthesis of prokaryotic glycoproteins." *Annual Review of Biochemistry* **58**: 173-194.

## References

- Letunic, I., R. R. Copley, B. Pils, S. Pinkert, J. Schultz & P. Bork (2006). "SMART 5: domains in the context of genomes and networks." *Nucleic Acids Research* **34**: D257-D260.
- Lewus, C. B., S. Sun & T. J. Montville (1992). "Production of an amylase-sensitive bacteriocin by an atypical *Leuconostoc paramesenteroides* strain." *Applied and Environmental Microbiology* **58**(1): 143-149.
- Li, B., J. P. J. Yu, J. S. Brunzelle, G. N. Moll, W. A. van der Donk & S. K. Nair (2006). "Structure and mechanism of the lantibiotic cyclase involved in nisin biosynthesis." *Science* **311**(5766): 1464-1467.
- Liang, C. F., M. C. Yan, T. C. Chang & C. C. Lin (2009). "Synthesis of S-linked alpha(2 -> 9) octasialic acid via exclusive alpha S-glycosidic bond formation." *Journal of the American Chemical Society* **131**(9): 3138-3139.
- Liu, J. & A. Mushegian (2003). "Three monophyletic superfamilies account for the majority of the known glycosyltransferases." *Protein Science* **12**(7): 1418-1431.
- Liu, W. & J. N. Hansen (1993). "The antimicrobial effect of a structural variant of subtilin against outgrowing *Bacillus cereus* T-spores and vegetative cells occurs by different mechanisms." *Applied and Environmental Microbiology* **59**(2): 648-651.
- Lorca, G. L., R. D. Barabote, V. Zlotopolski, C. Tran, B. Winnen, R. N. Hvorup, A. J. Stonestrom, E. Nguyen, L. W. Huang, D. S. Kim & M. H. Saier (2007). "Transport capabilities of eleven Gram-positive bacteria: Comparative genomic analyses." *Biochimica Et Biophysica Acta-Biomembranes* **1768**(6): 1342-1366.
- Lote, C. J. & J. B. Weiss (1971a). "Identification of digalactosylcysteine in a glycopeptide isolated from urine by a new preparative technique." *Febs Letters* **16**(2): 81-85.
- Lote, C. J. & J. B. Weiss (1971b). "Identification in urine of a low-molecular-weight highly polar glycopeptide containing cysteinyl-galactose." *Biochemical Journal* **123**(4): 25.
- Lovering, A. L., L. H. de Castro, D. Lim & N. C. J. Strynadka (2007). "Structural insight into the transglycosylation step of bacterial cell-wall biosynthesis." *Science* **315**(5817): 1402-1405.
- Ly, H. D., B. Lougheed, W. W. Wakarchuk & S. G. Withers (2002). "Mechanistic studies of a retaining alpha-galactosyltransferase from *Neisseria meningitidis*." *Biochemistry* **41**(16): 5075-5085.
- Maftah, A., D. Renault, C. Vignoles, Y. Hechard, P. Bressollier, M. H. Ratinaud, Y. Cenatiempo & R. Julien (1993). "Membrane permeabilization of *Listeria monocytogenes* and mitochondria by the bacteriocin mesentericin-Y105." *Journal of Bacteriology* **175**(10): 3232-3235.
- Makarova, K., A. Slesarev, Y. Wolf, A. Sorokin, B. Mirkin, E. Koonin, A. Pavlov, N. Pavlova, V. Karamychev, N. Polouchine, V. Shakhova, I. Grigoriev, Y. Lou, D. Rohksar, S. Lucas, K. Huang, D. M. Goodstein, T. Hawkins, V. Plengvidhya, D. Welker, J. Hughes, Y. Goh, A. Benson, K. Baldwin, J. H. Lee, I. Diaz-Muniz, B. Dosti, V. Smeianov, W. Wechter, R. Barabote, G. Lorca, E. Altermann, R. Barrangou, B. Ganesan, Y. Xie, H. Rawsthorne, D. Tamir, C. Parker, F. Breidt, J. Broadbent, R. Hutkins, D. O'Sullivan, J. Steele, G. Unlu, M. Saier, T. Klaenhammer, P. Richardson, S. Kozyavkin, B. Weimer & D. Mills (2006). "Comparative genomics of the lactic acid bacteria." *Proceedings of the National Academy of Sciences of the United States of America* **103**(42): 15611-15616.

## References

- Maldonado, A., R. Jimenez-Diaz & J. L. Ruiz-Barba (2004). "Induction of plantaricin production in *Lactobacillus plantarum* NC8 after coculture with specific gram-positive bacteria is mediated by an autoinduction mechanism." *Journal of Bacteriology* **186**(5): 1556-1564.
- Maqueda, M., A. Galvez, M. M. Bueno, M. J. Sanchez-Barrena, C. Gonzalez, A. Albert, M. Rico & E. Valdivia (2004). "Peptide AS-48: Prototype of a new class of cyclic bacteriocins." *Current Protein & Peptide Science* **5**(5): 399-416.
- Maqueda, M., M. Sánchez-Hidalgo, M. Fernández, M. Montalbán-López, E. Valdivia & M. Martínez-Bueno (2008). "Genetic features of circular bacteriocins produced by Gram-positive bacteria." *FEMS Microbiology Reviews* **32**(1): 2-22.
- Marco, M. L. & M. Kleerebezem (2008). "Assessment of real-time RT-PCR for quantification of *Lactobacillus plantarum* gene expression during stationary phase and nutrient starvation." *Journal of Applied Microbiology* **104**(2): 587-594.
- Martin-Visscher, L. A., T. Sprules, L. J. Gursky & J. C. Vederas (2008). "Nuclear magnetic resonance solution structure of Pisl, a group B immunity protein that provides protection against the type IIa bacteriocin piscicolin 126, Pisa." *Biochemistry* **47**(24): 6427-6436.
- Martinez-Fleites, C., M. S. Macauley, Y. He, D. L. Shen, D. J. Vocadlo & G. J. Davies (2008). "Structure of an O-GlcNAc transferase homolog provides insight into intracellular glycosylation." *Nature Structural & Molecular Biology* **15**(7): 764-765.
- Marugg, J. D., C. F. Gonzalez, B. S. Kunka, A. M. Ledebøer, M. J. Pucci, M. Y. Toonen, S. A. Walker, L. C. M. Zoetmulder & P. A. Vandenberg (1992). "Cloning, expression, and nucleotide-sequence of genes involved in production of pediocin-Pa-1, a bacteriocin from *Pediococcus acidilactici* Pac1.0." *Applied and Environmental Microbiology* **58**(8): 2360-2367.
- McGuffin, L. J. & D. T. Jones (2003). "Improvement of the GenTHREADER method for genomic fold recognition." *Bioinformatics* **19**(7): 874-881.
- Mescher, M. F. & J. L. Strominger (1976). "Purification and characterization of a prokaryotic glycoprotein from cell-envelope of *Halobacterium salinarium*." *Journal of Biological Chemistry* **251**(7): 2005-2014.
- Messner, P. & U. B. Sleytr (1988). "Asparaginyl-rhamnose - a novel type of protein-carbohydrate linkage in a eubacterial surface-layer glycoprotein." *FEBS Letters* **228**(2): 317-320.
- Mierau, I. & M. Kleerebezem (2005). "10 years of the nisin-controlled gene expression system (NICE) in *Lactococcus lactis*." *Applied Microbiology and Biotechnology* **68**(6): 705-717.
- Miller, K. W., P. Ray, T. Steinmetz, T. Hanekamp & B. Ray (2005). "Gene organization and sequences of pediocin Ach/PA-1 production operons in *Pediococcus* and *Lactobacillus* plasmids." *Letters in Applied Microbiology* **40**(1): 56-62.
- Miller, K. W., R. Schamber, Y. L. Chen & B. Ray (1998a). "Production of active chimeric pediocin Ach in *Escherichia coli* in the absence of processing and secretion genes from the *Pediococcus* pap operon." *Applied and Environmental Microbiology* **64**(1): 14-20.

## References

- Miller, K. W., R. Schamber, O. Osmanagaoglu & B. Ray (1998b). "Isolation and characterization of pediocin Ach chimeric protein mutants with altered bactericidal activity." *Applied and Environmental Microbiology* **64**(6): 1997-2005.
- Moens, S. & J. Vanderleyden (1997). "Glycoproteins in prokaryotes." *Archives of Microbiology* **168**(3): 169-175.
- Moll, G., H. Hildeng-Hauge, J. Nissen-Meyer, I. F. Nes, W. N. Konings & A. J. M. Driessen (1998). "Mechanistic properties of the two-component bacteriocin lactococcin G." *Journal of Bacteriology* **180**(1): 96-99.
- Moll, G., T. UbbinkKok, H. HildengHauge, J. NissenMeyer, I. F. Nes, W. N. Konings & A. J. M. Driessen (1996). "Lactococcin G is a potassium ion-conducting, two-component bacteriocin." *Journal of Bacteriology* **178**(3): 600-605.
- Moll, G. N., W. N. Konings & A. J. M. Driessen (1999). "Bacteriocins: mechanism of membrane insertion and pore formation." *Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology* **76**(1-4): 185-198.
- Morelli, L., P. S. Cocconcelli, V. Bottazzi, G. Damiani, L. Ferretti & V. Sgaramella (1987). "Lactobacillus protoplast transformation." *Plasmid* **17**(1): 73-75.
- Morera, S., L. Lariviere, J. Kurzeck, U. Aschke-Sonnenborn, P. S. Freemont, J. Janin & W. Ruger (2001). "High resolution crystal structures of T4 phage beta-glucosyltransferase: Induced fit and effect of substrate and metal binding." *Journal of Molecular Biology* **311**(3): 569-577.
- Moretro, T., I. M. Aasen, I. Storro & L. Axelsson (2000). "Production of sakacin P by *Lactobacillus sakei* in a completely defined medium." *Journal of Applied Microbiology* **88**(3): 536-545.
- Motlagh, A. M., A. K. Bhunia, F. Szostek, T. R. Hansen, M. C. Johnson & B. Ray (1992). "Nucleotide and amino-acid-sequence of Pap-gene (pediocin Ach production) in *Pediococcus acidilactici* H." *Letters in Applied Microbiology* **15**(2): 45-48.
- Muller, S., S. Goletz, N. Packer, A. Gooley, A. M. Lawson & F. G. Hanisch (1997). "Localization of O-glycosylation sites on glycopeptide fragments from lactation-associated MUC1 - All putative sites within the tandem repeat are glycosylation targets in vivo." *Journal of Biological Chemistry* **272**(40): 24780-24793.
- Murray, B. W., S. Takayama, J. Schultz & C. H. Wong (1996). "Mechanism and specificity of human alpha-1,3-fucosyltransferase V." *Biochemistry* **35**(34): 11183-11195.
- Naterstad, K., I. Rud, I. Kvam & L. Axelsson (2007). "Characterisation of the gap operon from *Lactobacillus plantarum* and *Lactobacillus sakei*." *Current Microbiology* **54**(3): 180-185.
- Navarro, L., B. Rojo-Bezares, Y. Sáenz, L. Díez, M. Zarazaga, F. Ruiz-Larrea & C. Torres (2008). "Comparative study of the pln locus of the quorum-sensing regulated bacteriocin-producing *L. plantarum* J51 strain." *International Journal of Food Microbiology* **128**(2): 390-394.
- Nealson, K. H. & J. W. Hastings (1979). "Bacterial bioluminescence - Its control and ecological significance." *Microbiological Reviews* **43**(4): 496-518.

## References

- Nealson, K. H., T. Platt & J. W. Hastings (1970). "Cellular control of synthesis and activity of bacterial luminescent system." *Journal of Bacteriology* **104**(1): 313-322.
- Nes, I. F., D. B. Diep, L. S. Havarstein, M. B. Brurberg, V. Eijsink & H. Holo (1996). "Biosynthesis of bacteriocins in lactic acid bacteria." *Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology* **70**(2-4): 113-128.
- Nes, I. F., D. B. Diep & H. Holo (2007). "Bacteriocin diversity in *Streptococcus* and *Enterococcus*." *Journal of Bacteriology* **189**(4): 1189-1198.
- Nes, I. F. & V. G. H. Eijsink (1999). "Regulation of group II peptide bacteriocin synthesis by quorum-sensing mechanisms." *Cell-Cell Signaling in Bacteria*: 175-192.
- Neuhoff, V., N. Arold, D. Taube & W. Ehrhardt (1988). "Improved staining of proteins in polyacrylamide gels including isoelectric-focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant blue G-250 and R-250." *Electrophoresis* **9**(6): 255-262.
- Niederhoffer, E. C. (2007). Biochemistry and molecular biology resource. G. linkages.
- Nielsen, H., J. Engelbrecht, S. Brunak & G. vonHeijne (1997). "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites." *Protein Engineering* **10**(1): 1-6.
- Nieto Lozano, J. C., J. Nissen-Meyer, K. Sletten, C. Pelaz & I. F. Nes (1992). "Purification and amino acid sequence of a bacteriocin produced by *Pediococcus acidilactici*." *Journal of Genetic Microbiology* **138**: 1985-1990.
- Nikolskaya, A. N. & M. Y. Galperin (2002). "A novel type of conserved DNA-binding domain in the transcriptional regulators of the AlgR/AgrA/LytR family." *Nucleic Acids Research* **30**(11): 2453-2459.
- Nissen-Meyer, J., H. Holo, L. S. Havarstein, K. Sletten & I. F. Nes (1992). "A novel lactococcal bacteriocin whose activity depends on the complementary action of 2 peptides." *Journal of Bacteriology* **174**(17): 5686-5692.
- Nissen-Meyer, J., P. Rogne, C. Oppegard, H. S. Haugen & P. E. Kristiansen (2009). "Structure-function relationships of the non-lanthionine-containing peptide (class II) bacteriocins produced by Gram-positive bacteria." *Current Pharmaceutical Biotechnology* **10**(1): 19-37.
- Nissen, H., H. Holo, L. Axelsson & H. Blom (2001). "Characterization and growth of *Bacillus* spp. in heat-treated cream with and without nisin." *Journal of Applied Microbiology* **90**(4): 530-534.
- Nolan, E. M. & C. T. Walsh (2009). "How Nature Morphs Peptide Scaffolds into Antibiotics." *ChemBiochem* **10**(1): 34-53.
- Novick, P. & P. Brennwald (1993). "Friends and family - the role of the Rab GTPases in vesicular traffic." *Cell* **75**(4): 597-601.
- Novick, R. P. (2003). "Autoinduction and signal transduction in the regulation of staphylococcal virulence." *Molecular Microbiology* **48**(6): 1429-1449.

## References

- O'Sullivan, D. J. & T. R. Klaenhammer (1993). "Rapid mini-prep isolation of high-quality plasmid DNA from *Lactococcus* and *Lactobacillus* Spp." *Applied and Environmental Microbiology* **59**(8): 2730-2733.
- Ochsner, U. A., A. Koch, A. Fiechter & J. Reiser (1994). "Isolation and characterization of a regulatory gene affecting rhamnolipid biosurfactant synthesis in *Pseudomonas aeruginosa*." *Journal of Bacteriology* **176**(7): 2044-2054.
- Oppegard, C., P. Rogne, L. Emanuelsen, P. E. Kristiansen, G. Fimland & J. Nissen-Meyer (2007). "The two-peptide class II bacteriocins: Structure, production, and mode of action." *Journal of Molecular Microbiology and Biotechnology* **13**(4): 210-219.
- Paik, S. H., A. Chakicherla & J. N. Hansen (1998). "Identification and characterization of the structural and transporter genes for, and the chemical and biological properties of, sublancin 168, a novel lantibiotic produced by *Bacillus subtilis* 168." *Journal of Biological Chemistry* **273**(36): 23134-23142.
- Pak, J. E., P. Arnoux, S. H. Zhou, P. Sivarajah, M. Satkunarajah, X. K. Xing & J. M. Rini (2006). "X-ray crystal structure of leukocyte type core 2 beta 1,6-N-acetylglucosaminyltransferase - Evidence for a convergence of metal ion-independent glycosyltransferase mechanism." *Journal of Biological Chemistry* **281**(36): 26693-26701.
- Papagianni, M. (2003). "Ribosomally synthesized peptides with antimicrobial properties: biosynthesis, structure, function, and applications." *Biotechnology Advances* **21**(6): 465-499.
- Parkinson, J. S. & E. C. Kofoid (1992). "Communication modules in bacterial signaling proteins." *Annual Review of Genetics* **26**: 71-112.
- Paulsen, I. T., L. Banerjee, G. S. A. Myers, K. E. Nelson, R. Seshadri, T. D. Read, D. E. Fouts, J. A. Eisen, S. R. Gill, J. F. Heidelberg, H. Tettelin, R. J. Dodson, L. Umayam, L. Brinkac, M. Beanan, S. Daugherty, R. T. DeBoy, S. Durkin, J. Kolonay, R. Madupu, W. Nelson, J. Vamathevan, B. Tran, J. Upton, T. Hansen, J. Shetty, H. Khouri, T. Utterback, D. Radune, K. A. Ketchum, B. A. Dougherty & C. M. Fraser (2003). "Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*." *Science* **299**(5615): 2071-2074.
- Pedersen, L. C., T. A. Darden & M. Negishi (2002). "Crystal structure of beta 1,3-glucuronyltransferase I in complex with active donor substrate UDP-GlcUA." *Journal of Biological Chemistry* **277**(24): 21869-21873.
- Pedersen, L. C., J. Dong, F. Taniguchi, H. Kitagawa, J. M. Krahn, L. G. Pedersen, K. Sugahara & M. Negishi (2003). "Crystal structure of an alpha 1,4-N-acetylhexosaminyltransferase (EXTL2), a member of the exostosin gene family involved in heparan sulfate biosynthesis." *Journal of Biological Chemistry* **278**(16): 14420-14428.
- Pena, J. A., S. Y. Li, P. H. Wilson, S. A. Thibodeau, A. J. Szary & J. Versalovic (2004). "Genotypic and phenotypic studies of murine intestinal lactobacilli: Species differences in mice with and without colitis." *Applied and Environmental Microbiology* **70**(1): 558-568.
- Persson, K., H. D. Ly, M. Dieckelmann, W. W. Wakarchuk, S. G. Withers & N. C. J. Strynadka (2001). "Crystal structure of the retaining galactosyltransferase LgtC from *Neisseria meningitidis* in complex with donor and acceptor sugar analogs." *Nature Structural Biology* **8**(2): 166-175.

## References

- Porcelli, F., B. A. Buck-Koehntop, S. Thennarasu, A. Ramamoorthy & G. Veglia (2006). "Structures of the dimeric and monomeric variants of magainin antimicrobial peptides (MSI-78 and MSI-594) in micelles and bilayers, determined by NMR spectroscopy" *Biochemistry* **45**(18): 5793-5799.
- Qiao, L., B. W. Murray, M. Shimazaki, J. Schultz & C. H. Wong (1996). "Synergistic inhibition of human alpha-1,3-fucosyltransferase V." *Journal of the American Chemical Society* **118**(33): 7653-7662.
- Quadri, L. E. N. (2002). "Regulation of antimicrobial peptide production by autoinducer-mediated quorum sensing in lactic acid bacteria." *Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology* **82**(1-4): 133-145.
- Ramnath, M., S. Arous, A. Gravesen, J. W. Hastings & Y. Hechard (2004). "Expression of mptC of *Listeria monocytogenes* induces sensitivity to class IIa bacteriocins in *Lactococcus lactis*." *Microbiology-Sgm* **150**: 2663-2668.
- Rauch, P. J. G. & W. M. Devos (1992). "Characterization of the novel nisin-sucrose conjugative transposon Tn5276 and its insertion in *Lactococcus lactis*." *Journal of Bacteriology* **174**(4): 1280-1287.
- Reading, N. C. & V. Sperandio (2006). "Quorum sensing: the many languages of bacteria." *Fems Microbiology Letters* **254**(1): 1-11.
- Rees, D. C., E. Johnson & O. Lewinson (2009). "ABC transporters: the power to change." *Nature Reviews Molecular Cell Biology* **10**(3): 218-227.
- Richard, C., R. Canon, K. Naghmouchi, D. Bertrand, H. Prevost & D. Drider (2006). "Evidence on correlation between number of disulfide bridge and toxicity of class IIa bacteriocins." *Food Microbiology* **23**(2): 175-183.
- Riley, M. A. & J. E. Wertz (2002a). "Bacteriocin diversity: ecological and evolutionary perspectives." *Biochimie* **84**(5-6): 357-364.
- Riley, M. A. & J. E. Wertz (2002b). "Bacteriocins: Evolution, ecology, and application." *Annual Review of Microbiology* **56**: 117-137.
- Risoen, P. A., M. B. Brurberg, V. G. H. Eijsink & I. F. Nes (2000). "Functional analysis of promoters involved in quorum sensing-based regulation of bacteriocin production in *Lactobacillus*." *Molecular Microbiology* **37**(3): 619-628.
- Risoen, P. A., L. S. Havarstein, D. B. Diep & I. F. Nes (1998). "Identification of the DNA-binding sites for two response regulators involved in control of bacteriocin synthesis in *Lactobacillus plantarum* C11." *Molecular and General Genetics* **259**(2): 224-232.
- Risoen, P. A., O. Johnsborg, D. B. Diep, L. Hamoen, G. Venema & I. F. Nes (2001). "Regulation of bacteriocin production in *Lactobacillus plantarum* depends on a conserved promoter arrangement with consensus binding sequence." *Molecular Genetics and Genomics* **265**(1): 198-206.
- Robichon, D., E. Gouin, M. Debarbouille, P. Cossart, Y. Cenatiempo & Y. Hechard (1997). "The rpoN (sigma(54)) gene from *Listeria monocytogenes* is involved in resistance to mesentericin Y105, an antibacterial peptide from *Leuconostoc mesenteroides*." *Journal of Bacteriology* **179**(23): 7591-7594.

## References

- Rogne, P., G. Fimland, J. Nissen-Meyer & P. E. Kristiansen (2008). "Three-dimensional structure of the two peptides that constitute the two-peptide bacteriocin lactococcin G." *Biochimica Et Biophysica Acta-Proteins and Proteomics* **1784**(3): 543-554.
- Rojo-Bezares, B., Y. Saenz, L. Navarro, R. Jimenez-Diaz, M. Zarazaga, F. Ruiz-Larrea & C. Torres (2008). "Characterization of a new organization of the plantaricin locus in the inducible bacteriocin-producing *Lactobacillus plantarum* J23 of grape must origin." *Archives of Microbiology* **189**(5): 491-499.
- Rost, B., P. Fariselli & R. Casadio (1996). "Topology prediction for helical transmembrane proteins at 86% accuracy." *Protein Science* **5**(8): 1704-1718.
- Rost, B., G. Yachdav & J. F. Liu (2004). "The PredictProtein server." *Nucleic Acids Research* **32**: W321-W326.
- Ruhr, E. & H. G. Sahl (1985). "Mode of action of the peptide antibiotic nisin and influence on the membrane-potential of whole cells and on cytoplasmic and artificial membrane-vesicles." *Antimicrobial Agents and Chemotherapy* **27**(5): 841-845.
- Russell, W. M. & T. R. Klaenhammer (2001). "Efficient system for directed integration into the *Lactobacillus acidophilus* and *Lactobacillus gasseri* chromosomes via homologous recombination." *Applied and Environmental Microbiology* **67**(9): 4361-4364.
- Ruyter, P. G. G. A. d., O. P. Kuipers, M. M. Beerthuyzen, I. v. Alen-Boerrigter & W. M. d. Vos (1996). "Functional analysis of promoters in the nisin gene cluster of *Lactococcus lactis*." *Journal of Bacteriology* **178**(12): 3434-3439.
- Saguir, F. M. & M. C. M. de Nadra (2007). "Improvement of a chemically defined medium for the sustained growth of *Lactobacillus plantarum*: Nutritional requirements." *Current Microbiology* **54**(6): 414-418.
- Sahl, H. G. (1994). "Staphylococcin-1580 is identical to the lantibiotic epidermin - implications for the nature of bacteriocins from Gram-positive bacteria." *Applied and Environmental Microbiology* **60**(2): 752-755.
- Sambrook, J., E. F. Fritsch & T. Maniatis (1989). *Molecular cloning: a laboratory manual*. New York, USA, Cold Spring Harbour Laboratory Press, Cold Spring Harbour.
- Schaefer, B. C. (1995). "Revolutions in rapid amplification of cDNA ends: New strategies for polymerase chain reaction cloning of full-length cDNA ends." *Analytical Biochemistry* **227**(2): 255-273.
- Schaffer, C., M. Graninger & P. Messner (2001). "Prokaryotic glycosylation." *Proteomics* **1**(2): 248-261.
- Schägger, H. (2006). "Tricine-SDS-PAGE." *Nature Protocols* **1**: 16-22.
- Schillinger, U. & F. K. Lucke (1989). "Antibacterial activity of *Lactobacillus sake* isolated from meat." *Applied and Environmental Microbiology* **55**(8): 1901-1906.

## References

- Schmidt, M. A., L. W. Riley & I. Benz (2003). "Sweet new world: glycoproteins in bacterial pathogens." *Trends in Microbiology* **11**(12): 554-561.
- Schultz, J., F. Milpetz, P. Bork & C. P. Ponting (1998). "SMART, a simple modular architecture research tool: Identification of signaling domains." *Proceedings of the National Academy of Sciences of the United States of America* **95**(11): 5857-5864.
- Schved, F., A. Lalazar, Y. Henis & B. J. Juven (1993). "Purification, partial characterization and plasmid-linkage of pediocin Sj-1, a bacteriocin produced by *Pediococcus acidilactici*." *Journal of Applied Bacteriology* **74**(1): 67-77.
- Scopes, R. K. (1974). "Measurement of protein by spectrophotometry at 205 nm." *Analytical Biochemistry* **59**(1): 277-282.
- Serizawa, M., K. Kodama, H. Yamamoto, K. Kobayashi, N. Ogasawara & J. Sekiguchi (2005). "Functional analysis of the YvrGHb two-component system of *Bacillus subtilis*: Identification of the regulated genes by DNA microarray and northern blot analyses." *Bioscience Biotechnology and Biochemistry* **69**(11): 2155-2169.
- Shi, J. Y., T. L. Blundell & K. Mizuguchi (2001). "FUGUE: Sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties." *Journal of Molecular Biology* **310**(1): 243-257.
- Sidote, D. J., C. M. Barbieri, T. Wu & A. M. Stock (2008). "Structure of the *Staphylococcus aureus* AgrA LytTR domain bound to DNA reveals a beta fold with an unusual mode of binding." *Structure* **16**(5): 727-735.
- Siegers, K. & K. D. Entian (1995). "Genes involved in immunity to the lantibiotic nisin produced by *Lactococcus lactis* 6f3." *Applied and Environmental Microbiology* **61**(3): 1082-1089.
- Sinensky, M. (2000). "Functional aspects of polyisoprenoid protein substituents: roles in protein-protein interaction and trafficking." *Biochimica Et Biophysica Acta-Molecular and Cell Biology of Lipids* **1529**(1-3): 203-209.
- Solomon, J. M., R. Magnuson, A. Srivastava & A. D. Grossman (1995). "Convergent sensing pathways mediate response to 2 extracellular competence factors in *Bacillus subtilis*." *Genes & Development* **9**(5): 547-558.
- Song, Y. L., N. Kato, C. X. Liu, Y. Matsumiya, H. Kato & K. Watanabe (2000). "Rapid identification of 11 human intestinal *Lactobacillus* species by multiplex PCR assays using group- and species-specific primers derived from the 16S-23S rRNA intergenic spacer region and its flanking 23S rRNA." *Fems Microbiology Letters* **187**(2): 167-173.
- Sorvig, E., G. Mathiesen, K. Naterstad, V. G. H. Eijsink & L. Axelsson (2005). "High-level, inducible gene expression in *Lactobacillus sakei* and *Lactobacillus plantarum* using versatile expression vectors." *Microbiology-Sgm* **151**: 2439-2449.
- Spiro, R. G. (2002). "Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds." *Glycobiology* **12**(4): 43R-56R.
- Spiro, S. & J. R. Guest (1990). "Fnr and its role in oxygen-regulated gene-expression in *Escherichia coli*." *Fems Microbiology Reviews* **75**(4): 399-428.

## References

- Stimmel, J. B., R. J. Deschenes, C. Volker, J. Stock & S. Clarke (1990). "Evidence for an S-farnesylcysteine methyl-ester at the carboxyl terminus of the *Saccharomyces cerevisiae* Ras2 protein." *Biochemistry* **29**(41): 9651-9659.
- Stimson, E., M. Virji, K. Makepeace, A. Dell, H. R. Morris, G. Payne, J. R. Saunders, M. P. Jennings, S. Barker, M. Panico, I. Blench & E. R. Moxon (1995). "Meningococcal pilin - a glycoprotein substituted with digalactosyl 2,4-diacetamido-2,4,6-trideoxyhexose." *Molecular Microbiology* **17**(6): 1201-1214.
- Straume, D., L. Axelsson, I. F. Nes & D. B. Diep (2006). "Improved expression and purification of the correctly folded response regulator PlnC from lactobacilli." *Journal of Microbiological Methods* **67**(2): 193-201.
- Sturme, M. H. J., C. Francke, R. J. Siezen, W. M. de Vos & M. Kleerebezem (2007). "Making sense of quorum sensing in lactobacilli: a special focus on *Lactobacillus plantarum* WCFS1." *Microbiology-Sgm* **153**: 3939-3947.
- Sumper, M. (1987). "Halobacterial glycoprotein-biosynthesis." *Biochimica Et Biophysica Acta* **906**(1): 69-79.
- Szymanski, C. M. & B. W. Wren (2005). "Protein glycosylation in bacterial mucosal pathogens." *Nature Reviews Microbiology* **3**(3): 225-237.
- Tabor, S. & C. C. Richardson (1985). "A bacteriophage-T7 RNA-polymerase promoter system for controlled exclusive expression of specific genes." *Proceedings of the National Academy of Sciences of the United States of America* **82**(4): 1074-1078.
- Tannock, G. W., A. Tilsala-Timisjarvi, S. Rodtong, J. Ng, K. Munro & T. Alatosava (1999). "Identification of *Lactobacillus* isolates from the gastrointestinal tract, silage, and yoghurt by 16S-23S rRNA gene intergenic spacer region sequence comparisons." *Applied and Environmental Microbiology* **65**(9): 4264-4267.
- Tarbouriech, N., S. J. Charnock & G. J. Davies (2001). "Three-dimensional structures of the Mn and Mg dTDP complexes of the family GT-2 glycosyltransferase SpsA: A comparison with related NDP-sugar glycosyltransferases." *Journal of Molecular Biology* **314**(4): 655-661.
- Taylor, C. M. (1998). "Glycopeptides and glycoproteins: Focus on the glycosidic linkage." *Tetrahedron* **54**(38): 11317-11362.
- Thibodeaux, C. J., C. E. Melancon & H. W. Liu (2007). "Unusual sugar biosynthesis and natural product glycodiversification." *Nature* **446**(7139): 1008-1016.
- Thomas, X., D. Destoumieux-Garzon, J. Peduzzi, C. Afonso, A. Blond, N. Birlirakis, C. Goulard, L. Dubost, R. Thai, J. C. Tabet & S. Rebuffat (2004). "Siderophore peptide, a new type of post-translationally modified antibacterial peptide with potent activity." *Journal of Biological Chemistry* **279**(27): 28233-28242.
- Ulrich, L. E., E. V. Koonin & I. B. Zhulin (2005). "One-component systems dominate signal transduction in prokaryotes." *Trends in Microbiology* **13**(2): 52-56.

## References

- Uden, G., S. Becker, J. Bongaerts, J. Schirawski & S. Six (1994). "Oxygen-regulated gene-expression in facultatively anaerobic-bacteria." *Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology* **66**(1-3): 3-22.
- Upreti, G. C. & R. D. Hinsdill (1973). "Isolation and characterization of a bacteriocin from a homofermentative *Lactobacillus*." *Antimicrobial Agents and Chemotherapy* **4**(4): 487-494.
- Upreti, G. C. & R. D. Hinsdill (1975). "Production and mode of action of lactocin 27 - bacteriocin from a homofermentative *Lactobacillus*." *Antimicrobial Agents and Chemotherapy* **7**(2): 139-145.
- Upreti, R. K., M. Kumar & V. Shankar (2003). "Bacterial glycoproteins: Functions, biosynthesis and applications." *Proteomics* **3**(4): 363-379.
- Uteng, M., H. H. Hauge, P. R. L. Markwick, G. Fimland, D. Mantzilas, J. Nissen-Meyer & C. Muhle-Goll (2003). "Three-dimensional structure in lipid micelles of the pediocin-like antimicrobial peptide sakacin P and a sakacin P variant that is structurally stabilized by an inserted C-terminal disulfide bridges." *Biochemistry* **42**(39): 11417-11426.
- van Belkum, M. J., R. W. Worobo & M. E. Stiles (1997). "Double-glycine-type leader peptides direct secretion of bacteriocins by ABC transporters: Colicin V secretion in *Lactococcus lactis*." *Molecular Microbiology* **23**(6): 1293-1301.
- Viklund, H. & A. Elofsson (2008). "OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar." *Bioinformatics* **24**(15): 1662-1668.
- Vogelstein, B. & D. Gillespie (1979). "Preparative and analytical purification of DNA from agarose." *Proceedings of the National Academy of Sciences of the United States of America* **76**(2): 615-619.
- Vogt, A., Y. M. Qian, M. A. Blaskovich, R. D. Fossum, A. D. Hamilton & S. M. Sehti (1995). "A nonpeptide mimetic of Ras-Caax - selective-inhibition of farnesyltransferase and Ras processing." *Journal of Biological Chemistry* **270**(2): 660-664.
- von Heijne, G. (1992). "Membrane protein structure prediction - hydrophobicity analysis and the positive-inside rule." *Journal of Molecular Biology* **225**(2): 487-494.
- Vrieling, A., W. Ruger, H. P. C. Driessen & P. S. Freemont (1994). "Crystal-structure of the DNA modifying enzyme beta-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose." *Embo Journal* **13**(15): 3413-3422.
- Waite, B. L., G. R. Siragusa & R. W. Hutkins (1998). "Bacteriocin inhibition of two glucose transport systems in *Listeria monocytogenes*." *Journal of Applied Microbiology* **84**(5): 715-721.
- Wang, Y. J., M. E. Henz, N. L. F. Gallagher, S. Y. Chai, A. C. Gibbs, L. Z. Yan, M. E. Stiles, D. S. Wishart & J. C. Vederas (1999). "Solution structure of carnobacteriocin B2 and implications for structure-activity relationships among type IIa bacteriocins from lactic acid bacteria." *Biochemistry* **38**(47): 15438-15447.
- Weiss, J. B., C. J. Lote & H. Bobinski (1971). "New low molecular weight glycopeptide containing triglycosylcysteine in human erythrocyte membrane." *Nature-New Biology* **234**(44): 25-26.
- West, A. H. & A. M. Stock (2001). "Histidine kinases and response regulator proteins in two-component signaling systems." *Trends in Biochemical Sciences* **26**(6): 369-376.

## References

Whitmore, L. & B. A. Wallace (2008). "Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases." *Biopolymers* **89**(5): 392-400.

Wiedemann, I., E. Breukink, C. van Kraaij, O. P. Kuipers, G. Bierbaum, B. de Kruijff & H. A. Sahl (2001). "Specific binding of nisin to the peptidoglycan precursor lipid II combines pore formation and inhibition of cell wall biosynthesis for potent antibiotic activity." *Journal of Biological Chemistry* **276**(3): 1772-1779.

Wolanin, P. M., P. A. Thomason & J. B. Stock (2002). "Histidine protein kinases: key signal transducers outside the animal kingdom." *Genome Biology* **3**(10): 3013.1-3013.8.

Woody, R. W. (1994). "Contributions of tryptophan side-chains to the far-ultraviolet circular-dichroism of proteins." *European Biophysics Journal with Biophysics Letters* **23**(4): 253-262.

Worobo, R. W., T. Henkel, M. Sailer, K. L. Roy, J. C. Vederas & M. E. Stiles (1994). "Characteristics and genetic determinant of a hydrophobic peptide bacteriocin, carnobacteriocin A, produced by *Carnobacterium piscicola* Lv17a." *Microbiology-Uk* **140**: 517-526.

Yang, W. H., J. E. Kim, H. W. Nam, J. W. Ju, H. S. Kim, Y. S. Kim & J. W. Cho (2006). "Modification of p53 with O-linked N-acetylglucosamine regulates p53 activity and stability." *Nat Cell Biol* **8**(10): 1074-1083.

Yokoyama, K., G. W. Goodwin, F. Ghomashchi, J. A. Glomset & M. H. Gelb (1991). "A protein geranylgeranyltransferase from bovine brain - Implications for protein prenylation specificity." *Proceedings of the National Academy of Sciences of the United States of America* **88**(12): 5302-5306.

Yoshida, A., M. Suzuki, H. Ikenaga & M. Takeuchi (1997). "Discovery of the shortest sequence motif for high level mucin-type O-glycosylation." *Journal of Biological Chemistry* **272**(27): 16884-16888.

Young, J. D., D. Tsuchiya, D. E. Sandlin & M. J. Holroyde (1979). "Enzymatic O-glycosylation of synthetic peptides from sequences in basic myelin protein." *Biochemistry* **18**: 4444-4448.

Young, N. M., J. R. Brisson, J. Kelly, D. C. Watson, L. Tessier, P. H. Lanthier, H. C. Jarrell, N. Cadotte, F. S. Michael, E. Aberg & C. M. Szymanski (2002). "Structure of the N-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, *Campylobacter jejuni*." *Journal of Biological Chemistry* **277**(45): 42530-42539.

Zachara, N. E. & G. W. Hart (2002). "The emerging significance of O-GlcNAc in cellular regulation." *Chemical Reviews* **102**(2): 431-438.

Zechel, D. L. & S. G. Withers (2000). "Glycosidase mechanisms: Anatomy of a finely tuned catalyst." *Accounts of Chemical Research* **33**(1): 11-18.

Zhang, F. L. & P. J. Casey (1996). "Protein prenylation: Molecular mechanisms and functional consequences." *Annual Review of Biochemistry* **65**: 241-269.

Zhu, B. C. R., R. R. Drake, H. Schweingruber & R. A. Laine (1995). "Inhibition of glycosylation by amphomycin and sugar nucleotide analogs Pp36 and Pp55 indicates that *Haloflex volcanii* beta-glucosylates both glycoproteins and glycolipids through lipid-linked sugar intermediates - evidence for 3 novel glycoproteins and a novel sulfated dihexosyl-archaeol glycolipid." *Archives of Biochemistry and Biophysics* **319**(2): 355-364.

## References

Zhu, X. M., T. Haag & R. R. Schmidt (2004). "Synthesis of an S-linked glycopeptide analog derived from human Tamm-Horsfall glycoprotein." *Organic & Biomolecular Chemistry* 2(1): 31-33.