INVESTIGATION INTO THE STRUCTURE AND FUNCTION OF THE GLYCOSYLATED BACTERIOCIN GccF AND THE GLYCOSYLTRANSFERASE GccA FROM *LACTOBACILLUS PLANTARUM* KW30

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Science

in

Biological Sciences

at Massey University, Palmerston North

New Zealand

Evelyn Marianne Poulson

2011

ABSTRACT

Bacteriocins are typically antimicrobial proteins or peptides produced by Gram-positive and G–negative bacteria, that are capable of inhibiting the growth of other bacteria. Glycocin F (GccF) is a 43 amino acid bacteriocin produced by *Lactobacillus plantarum* KW30 which is post-translationally modified by two N-acetylglucosamine residues (GlcNAc). One of these residues is linked to a serine side-chain (*O*-linked), while the other is linked through the thiol sulphur at the C-terminal cysteine (*S*-linked). Within the gene cluster encoding GccF are a set of genes thought to be required for the maturation and secretion of GccF. The GccF gene cluster consists of six genes encoding a family 2 glycosyltransferase (GTase) thought to be responsible for the addition of either one or both of these GlcNAc groups, an ABC transporter involved in the secretion of the bacteriocin across the cellular membrane, two thioredoxin-like genes which may be responsible for the disulfide bonding pattern of GccF, a gene of unknown function, and GccF itself.

Within *L. plantarum* KW30 no other proteins modified by a GlcNAc residue were identified in the present study, making GccF the only known GlcNAcylated protein produced by this organism. Methods were developed to pull-down the proteins involved in the maturation and secretion of GccF, and to find its binding target(s) in strains suceptible to its activity. Although proteins were found to bind tightly to GccF during pull-down experiments, those that bound were mostly involved in glycolysis/gluconeogenesis which does not fit into the hypothesised mechanism of action for GccF. Fluorescent microscopy experiments on wild-type GccF and GccF that contained only the *O*-linked or *S*-linked GlcNAc residue found that localisation of the modified GlcNAcylated GccF on susceptible strains was different to what is seen for wild-type in that they appeared randomly along the cells, whereas wild-type GccF appeared to localise at the point of cell division and at the tips of the cells. These microscopy results show that the post-translational modifications appear to play a role in targeting of GccF to susceptible cells. Assays to detect and test the activity of the GTase found that it may be located within the cytosol of *L. plantarum* KW30 instead of the membrane which is where it was proposed to be due to the presence of a predicted transmembrane spanning region identified during bioinformatic analysis.

ACKNOWLEDGEMENTS

I would like to thank my supervisors Dr Gill Norris and Dr Mark Patchett for all their support, guidance and encouragement throughout the duration of my thesis. Without their enthusiasm, positive attutidue, and willingness to answer my queries I would have been very lost.

I would like to thank everyone in the X-lab especially Trevor Loo and Natalie Burr, for their help, advice and good humour which made working so much fun.

I would also like to thank all staff in IMBS for their help and support during my time here.

Finally I would like to thank all of my family and friends for their continued support thoughout the duration of my studies. Special thanks to my mum and Rob for putting up with me when my emotions took the better of me, when I was stressing over finishing and when things just weren't working.

TABLE OF CONTENTS

LIST OF FIGURES

PART 1: INTRODUCTION

PART 2: GENERAL METHODS

PART 3: FUNCTIONAL CHARACTERIZATION OF THE *L. PLANTARUM* KW30 BACTERIOCIN, GCCF.

PART 4: STRUCTURAL AND FUNCTIONAL CHARACTERISATION OF THE *L. PLANTARUM* KW30 GLYCOSYLTRANSFERASE, GCCA.

LIST OF TABLES

ABBREVIATIONS

Å	Ångström (10^{-10} m)
aa	Amino acid
ABC transporter	ATP-binding cassette transporter
ACN	Acetonitrile
Amp	Ampicillin
Amu	Atomic mass unit
APS	Ammonium persulfate
ATCC	American type Culture Collection
ATP	Adenosine-5'-triphosphate
Bis-Tris	Bis-(2-hydroxy-ethyl)-amino-tris(hydroxymethyl)-methane
BLAST	Basic Local Alignment Search Tool
BSA	Bovine serum albumin
°C	Degrees Celcius
C11	The 11 C-terminal amino acids of GccF, unmodified.
C15	The 15 C-terminal amino acids of GccF, unmodified.
CAZy	Carbohydrate-Active enzyme database
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
Cm	Chloramphenicol
CV	Column Volume
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxyribose nucleotide triphosphate
DTT	Dithiothreitol
EDTA	Ethylene diamine tetraacetic acid
Efc	*Enterococcus faecalis*
et al.	*et alteri* (and others)
EtBr	Ethidium bromide
EtOH	Ethanol
FITC	Fluorescein isothiocyante
g	Gram; standard gravity ($9.81 m/s^2$)
GccF	Glycocin F
$GccF^{deFFL}$	Glycocin F without the *O*-linked *N*-acetylglucosamine residue
$GccF^{-HC}$	Glycocin F without the *S*-linked *N*-acetylglucosamine residue and the last two C-terminal amino acids
GlcNAc	*N*-acetylglucosamine
GlcNAcylated	Modified by an *N*-acetylglucosamine
GT2	Family 2 glycosyltransferase
GTase	Glycosyltransferase
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HFBA	Heptafluorobutyric Acid
HPLC	High Pressure Liquid Chromatography

hrs	hours
IEF	Isoelectric Focusing
IMAC	Immobilised Metal Affinity Chromatography
IPTG	Isopropyl-β-D-Thiogalactopyranoside
Kan	Kanamycin
kb	kilobases
kDa	kilo daltons
L	Litre
LAB	Lactic Acid Bacteria
LB	Luria Bertani broth
M	Molar (mol/L)
m	Metre
min	Minute
mol	6.023×10^{23} molecules (Avogandro's constant)
MRS	de Man, Rogosa and Sharpe broth
MW	Molecular mass
m/z	Mass-to-charge ratio
NCBI	National Centre for Biotechnology Information
NH_2	Amine
Sulfo-NHS biotin	*N*-hydroxysulfosuccinimide biotin
NMR	Nuclear magnetic resonance
OD	Optical Density
o/n	Overnight
Pa	Pascal (= bar 10^{-5} = 145.04×10^{-6} psi)
PCR	Polymerase Chain Reaction
pH	Negative decadal logarithm of the proton concentration
PMSF	Phenylmethylsulfphonyl fluoride
ppm	parts per million
PTM	Post-translational modification
RNA	ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
RT	Room temperature
RP-HPLC	Reverse-Phase High Pressure Liquid Chromatography
RR	Response regulator
sec	Secretory system
s/sec	Second
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis
SOB	Super optimal broth
SOC	SOB with catabolite repression, indicative of the presence of glucose
sp.	Species

subsp.	Subspecies
T	Time
TAE	Tris-Acetate-EDTA
TCEP	Tris(2-carboxyethyl)phosphine
TE	Tris-EDTA
TET	Tetracycline
TFA	Trifluoroacetic Acid
TMD	Transmembrane domain
TOF/TOF	Tandem time-of-flight
Tris	Tris(hydroxymethyl)aminomethane
trx	Thioredoxin
UDP	Uridine diphosphate
UDP-GlcNAc	Uridine diphosphate *N*-acetylglucosamine
USA	United States of America
UV	Ultraviolet light
V	Volts
v/v	Volume per volume
w/v	Weight per volume
w/w	Weight per weight
WGA	Wheatgerm agglutinin
WT	Wild-type
Yfr	*Yersinia frederiksenii*