



<http://researchspace.auckland.ac.nz>

ResearchSpace@Auckland

Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage.

<http://researchspace.auckland.ac.nz/feedback>

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form.

pu/A, a thermostable pullulanase
from an extreme thermophile
Caldocellum saccharolyticum

Gregory D. Albertson

Department of Cellular and Molecular Biology,

University of Auckland

Thesis submitted in fulfilment of the
requirements for the degree of
Doctor of Philosophy

November 1992

UNIVERSITY OF AUCKLAND LIBRARY
BIOLOGY
THESIS

93-038

Abstract

The pullulanase gene from *Caldocellum saccharolyticum*, an obligate thermophilic anaerobe, was sequenced and expressed in *E. coli*. Expression and substrate induction studies in *E. coli* showed that while gene expression was substrate inducible and the enzyme was exported into the growth medium in *C. saccharolyticum*, expression was non-inducible in *E. coli* and the enzyme remained in the cytoplasm. The nucleotide sequence of the *pulA* gene was shown to be 2478 basepairs (bp) in length, coding for a protein of 96 kDa. The proposed promoter sequences showed homology to both the standard *E. coli* sequences and the consensus sequences obtained from other *C. saccharolyticum* genes.

The enzyme from the native organism was purified from the growth medium and shown to have a molecular mass of approximately 120 kDa. Periodic acid-Schiffs staining showed that this enzyme was glycosylated and substrate characterisation revealed that the enzyme debranched pullulan to produce only maltotriose, but hydrolysed amylopectin, amylose and β -limit dextran to produce a number of smaller oligosaccharides.

The enzyme was expressed in *E. coli* from its own promoters and was purified from the cytoplasmic fraction. Substrate characterisation revealed that the enzyme debranched pullulan to produce only maltotriose, but had only limited activity on β -limit dextran and amylopectin, and no activity on amylose.

The pullulanase gene was also expressed under the control of a heat-inducible overexpression system in *E. coli* and a copper-inducible expression system in yeast.

Amino acid homology comparisons of the pullulanase to other pullulanase sequences and related enzymes revealed a high degree of homology, particularly around three highly conserved regions. In α -amylases amino acids in these regions are involved in catalytic activity, substrate binding and metal ion binding.

Overview

Abstract	i	
Table of Contents	iv	
Acknowledgements	xiii	
Abbreviations	xv	
Chapter One	Introduction	1
Chapter Two	Materials and Methods	47
Chapter Three	The Pullulanase expressed from <i>Caldocellum saccharolyticum</i>	87
Chapter Four	DNA Sequence of the Pullulanase Gene from <i>C. saccharolyticum</i>	113
Chapter Five	Expression of the Pullulanase in <i>Escherichia coli</i>	141
Chapter Six	Inducible Enzyme Expression in <i>E. coli</i>	173
Chapter Seven	Pullulanase Expression Studies using Yeast	187
Chapter Eight	Analysis of the Deduced Peptide Sequence of the pullulanase	191
Chapter Nine	Discussion	211
Appendices		231
References		289

Table of Contents

Chapter One	Introduction.....	1
1.1	Thermophiles.....	1
1.1.1	Thermophilic organisms.....	1
1.1.1.1	History and introduction.....	1
1.1.1.2	Temperature and environmental range.....	1
1.1.1.3	Taxonomy.....	2
1.1.1.4	The molecular basis of thermophily.....	3
1.1.2	Uses of thermophilic/thermostable enzymes.....	5
1.1.2.1	Biotechnological advantages of using thermophilic enzymes.....	5
1.1.2.2	Increased enzymatic reaction rates.....	6
1.1.2.3	Unique properties of thermophilic enzymes.....	6
1.1.2.4	Purification of thermophilic enzymes expressed in mesophilic hosts. ...	7
1.1.3	Applications of thermophilic enzymes.....	7
1.1.3.1	Enzymatic bleaching of wood pulp.....	7
1.1.3.2	Biomass fermentation.....	7
1.1.3.3	Thermophilic enzymes used in laboratory research.....	8
1.1.4	Protein thermostability.....	8
1.1.4.1	History and introduction.....	8
1.1.4.2	Random and site-directed mutagenesis.....	9
1.1.4.3	Statistical analysis.....	9
1.1.4.4	3-dimensional model systems.....	10
1.1.4.5	Chimeric genes.....	11
1.2	Starch.....	12
1.2.1	Chemical structure of starch and related carbohydrates.....	12
1.2.1.1	Amylose.....	12
1.2.1.2	Amylopectin.....	13
1.2.1.3	Pullulan.....	14
1.2.1.4	Cyclodextrin.....	14
1.2.2	The starch industry.....	15
1.2.2.1	Economic uses and value of starch products.....	15
1.2.2.2	Economic uses and value of starch hydrolysis products.....	15
1.2.2.3	Current hydrolysis methods.....	16
1.2.2.4	Usefulness of thermophilic enzymes.....	18
1.2.3	Usefulness of a thermostable debranching enzyme.....	19
1.2.3.1	Theory and practice in amylopectin hydrolysis.....	19
1.2.3.2	Useful properties of debranching enzymes.....	21
1.3	Genetic Analysis of Pullulanases and Related Genes.....	23
1.3.1	<i>E. coli</i> maltose regulon.....	23
1.3.2	<i>Klebsiella</i> maltose regulon.....	24

1.3.3 Thermophilic enzymes	26
1.4 Pullulanase and Related Enzymes	28
1.4.1 Non-branching enzymes	29
1.4.1.1 α -amylase (E.C. 3.2.1.1) 1,4- α -D-Glucan glucanohydrolase	29
1.4.1.2 β -amylase (E.C. 3.2.1.2) 1,4- α -D-Glucan maltohydrolase	30
1.4.1.3 Glucoamylase (E.C. 3.2.1.3) 1,4- α -D-Glucan glucohydrolase	30
1.4.1.4 Alpha glucosidase.....	30
1.4.1.5 "Non-classical" α -amylase.....	31
1.4.2 Debranching enzymes	32
1.4.2.1 Type I or true pullulanase.....	33
1.4.2.2 Type II pullulanase	34
1.4.2.3 α -amylase-pullulanase.....	34
1.4.2.4 Neopullulanase and isopullulanase	35
1.4.2.5 Isoamylase (E.C. 3.2.1.68) Glycogen 6-glucanohydrolase.....	36
1.5 <i>Caldocellum saccharolyticum</i>	37
1.5.1 <i>Caldocellum saccharolyticum</i> bacteria.....	37
1.5.1.1 Growth conditions.....	37
1.5.1.2 Taxonomic characteristics.....	37
1.5.2 <i>C. saccharolyticum</i> enzymes transferred and expressed in <i>E. coli</i>	38
1.5.2.1 β -glucosidase	38
1.5.2.2 Cellulase	38
1.5.2.3 Hemicellulase.....	39
1.6 Previous Work Done and Aims of Thesis.....	41
1.6.1 Gene isolation and expression in <i>E. coli</i>	41
1.6.2 Bal 31 deletions	41
1.6.3 The aims of research for this thesis.....	45
Chapter Two Materials and Methods	47
2.1 Materials.....	47
2.1.1 Chemicals.....	47
2.1.2 Radioactivity.....	47
2.1.3 Enzymes	47
2.1.3.1 Restriction enzymes	47
2.1.4 Common buffers and general solutions	48
2.2 Organisms.....	49
2.2.1 <i>E. coli</i> strains.....	49

2.2.2 Growth media and conditions.....	49
2.2.2.1 Rich media.....	51
2.2.2.2 Minimal media	51
2.2.2.3 Antibiotics	52
2.2.3 Measurement of cell growth.....	52
2.2.3.1 Optical density	52
2.2.3.2 Viable cell counts	52
2.2.3.3 Plasmid copy number determination	53
2.2.4 <i>Caldocellum saccharolyticum</i>	53
2.2.5 Yeast strains.....	54
2.3 DNA.....	55
2.3.1 Plasmids	55
2.3.2 Small scale plasmid isolation.....	55
2.3.2.1 Standard alkaline lysis method.....	55
2.3.2.2 Phenol/chloroform method	56
2.3.3 Large scale plasmid isolation.....	57
2.3.3.1 CsCl-EtBr density gradient purification.....	57
2.3.3.2 Large-Scale LiCl method	57
2.3.4 Standard agarose gel electrophoresis.....	58
2.3.5 Preparative agarose gel electrophoresis	58
2.3.5.1 Separation	58
2.3.5.2 Recovery of DNA fragments from agarose gels	59
2.3.6 Ligation and transformation	59
2.3.6.1 Ligation	59
2.3.6.2 Heat shock transformation.....	59
2.3.6.3 DMSO transformation method	60
2.3.7 Radioactively labelled nucleic acid probes.....	60
2.3.7.1 Nick translation	60
2.3.7.2 Random priming.....	61
2.3.7.3 Gel filtration.....	61
2.3.7.4 Scintillation counting.....	61
2.3.8 Transfer of DNA to charged membranes	61
2.3.8.1 Southern blotting - alkaline Transfer	61
2.3.8.2 Dot blots	62
2.3.8.3 Colony hybridisation.....	62
2.3.9 Hybridisation, washing and autoradiography.....	62
2.3.9.1 Prehybridisation and hybridisation	62
2.3.9.2 Stringency calculations.....	63
2.3.9.3 Autoradiography.....	63
2.3.9.4 Removal of probes	64
2.3.10 Genomic DNA preparations	64
2.3.11 Polymerase chain reaction.....	64

2.3.12	Yeast DNA methods	65
2.3.12.1	Yeast transformations.....	65
2.3.12.2	Plasmid isolation from yeast.....	65
2.4	Sequencing.....	66
2.4.1	Single stranded DNA isolation	66
2.4.1.1	M13 phage.....	66
2.4.1.2	Phagemids	66
2.4.2	Oligonucleotide primers.....	66
2.4.3	Sequencing methods.....	68
2.4.3.1	Single strand sequencing.....	68
2.4.3.2	Plasmid sequencing.....	68
2.4.4	Acrylamide sequencing gels	68
2.4.5	Sequencing library.....	69
2.4.5.1	Pseudorandom library.....	69
2.4.5.2	Random library.....	69
2.5	Enzyme Purification.....	70
2.5.1	Isolation from <i>E. coli</i>	70
2.5.1.1	Cell harvesting	70
2.5.1.2	Heat purification	70
2.5.1.3	Ammonium Sulphate Precipitation.....	70
2.5.1.4	Enzyme location within the <i>E. coli</i> cell.....	71
2.5.2	Isolation from <i>Caldocellum saccharolyticum</i>	73
2.5.3	Isolation from yeast.....	73
2.5.4	Chromatography	74
2.5.4.1	Dialysis and desalting	74
2.5.4.2	Ion exchange	74
2.5.4.3	Gel filtration	74
2.5.4.4	Affinity chromatography.....	75
2.5.5	Estimation of total protein concentrations.....	75
2.6	Enzyme Characteristics	77
2.6.1	Enzyme assays	77
2.6.1.1	Reaction Buffers	77
2.6.1.2	PAHBAH reducing sugar assay	79
2.6.1.3	Starch-iodine complex.....	79
2.6.1.4	Dye-linked substrates.....	80
2.6.1.5	Substrate preparation	80
2.6.2	Protein gel electrophoresis.....	80
2.6.2.1	SDS-PAGE.....	80
2.6.2.2	Non-denaturing PAGE.....	81
2.6.2.3	Coomassie Blue Stain	81

2.6.2.4 Phastsystem	81
2.6.2.5 Periodic acid-Schiff's reagent.....	81
2.6.2.6 Activity gels	82
2.6.2.7 Agarose G.E.	82
2.6.3 Paper chromatography.....	82
2.7 Computer Analysis.....	84
2.7.1 Standard Programs	84
2.7.2 Sequence analysis	84
2.7.3 Database searches.....	84
2.7.4 Homology comparisons	85
2.7.5 Restriction fragment analysis.....	85
Chapter Three The Pullulanase expressed from <i>Caldocellum saccharolyticum</i>	87
3.1 Introduction.....	87
3.2 Substrate Induction of Pullulanase Expression.....	87
3.2.1 The effect of carbohydrate source on pullulanase synthesis.....	88
3.2.2 Cellular location of the <i>C. saccharolyticum</i> pullulanase.....	89
3.2.3 Validation of strain identity.....	90
3.3 Purification of the Pullulanase From <i>Caldocellum saccharolyticum</i>	92
3.3.1 Ammonium sulphate precipitation	94
3.3.2 Ion exchange	94
3.3.3 Gel filtration.....	96
3.3.4 Affinity binding.....	99
3.4 Enzymatic Characterisation	102
3.4.1 Glycosylation and activity studies	102
3.4.2 The effect of metal ions and inhibitors on pullulanase activity	104
3.4.2.1 Initial studies	104
3.4.2.2 Calcium ion studies.....	105
3.4.2.3 Monovalent cations.....	106
3.4.3 Thermostability and temperature and pH optima of the pullulanase	107
3.4.4 Substrate characterisation.....	108
3.5 Summary.....	111

Chapter Four	DNA Sequence of the Pullulanase Gene from <i>C. saccharolyticum</i>	113
4.1	Introduction	113
4.2	Restriction endonuclease map of pNZ1452 and pNZ1038.....	113
4.3	Sequencing of pNZ1452	117
4.3.1	Sequencing of previously constructed recombinants	117
4.3.2	Pseudo random library	117
4.3.3	Random shear library.....	119
4.3.4	Recombinant plasmids	121
4.4	Analysis of the Sequence of pNZ1452	122
4.4.1	Determination of open reading frames (ORF)	122
4.4.2	Incomplete gene on pNZ1452.....	124
4.4.3	Sequencing the remainder of pNZ1038	126
4.5	Genetic Structural elements of pNZ1038.....	129
4.5.1	ORF 1, 2, 3 and 4 gene assignment.....	129
4.5.2	Repeat sequences	132
4.5.3	Terminator sequences	133
4.5.4	Leader sequence analysis	135
4.5.5	Codonpreference.....	136
4.6	Summary	139
Chapter Five	Expression of the Pullulanase in <i>Escherichia coli</i>	141
5.1	Summary	141
5.2	Enzyme Expression and Cell Growth.....	141
5.2.1	Expression levels	141
5.2.2	Substrate induction of enzyme synthesis.....	143
5.2.3	Dependence of cellular growth phase on enzyme expression.....	144
5.2.4	Consequences of pullulanase expression on the <i>E. coli</i> cells.....	145
5.2.4.1	Growth rate studies.....	145
5.2.4.2	Microscopic examination.....	146
5.2.4.3	Plasmid stability tests.....	149
5.2.5	Cellular location of the enzyme.....	150
5.2.6	Proof that the pullulanase gene in all constructs was expressed from a <i>C. saccharolyticum</i> promoter.....	151
5.3	Purification of the <i>C. saccharolyticum</i> Pullulanase Expressed in <i>E. coli</i>	153
5.3.1	Initial purification studies	155

5.3.2 Heat purification and ammonium sulphate precipitation	155
5.3.3 Ion Exchange.....	156
5.3.4 Gel filtration.....	158
5.3.5 Purification using ultrafiltration.....	160
5.4 Enzymatic Characterisation	161
5.4.1 Activity gels and N-terminal sequencing.....	161
5.4.2 The effect of metal ions and inhibitors on pullulanase activity	162
5.4.2.1 Initial studies	162
5.4.2.2 Calcium ion studies.....	163
5.4.2.3 Monovalent cations.....	164
5.4.3 Thermostability and temperature and pH optima of the pullulanase	165
5.4.4 Substrate characterisation.....	168
5.5 Summary.....	171
 Chapter Six Inducible Enzyme Expression in <i>E. coli</i>	173
6.1 Introduction.....	173
6.2 Optimisation of Expression Levels.....	173
6.2.1 Expression plasmid construction	173
6.2.1.1 Expression plasmid choice	173
6.2.1.2 Structure of pJLA602	175
6.2.2 Optimisation of induction.....	176
6.2.2.1 28 - 42 heat shift induction	176
6.2.2.2 28 - 37 heat shift induction	178
6.2.2.3 28 - 42 - 37 heat shift induction.....	178
6.2.2.4 Alkaline pH shift induction.....	179
6.2.2.5 Optimal induction procedure	179
6.2.3 Enzyme location within the cell	180
6.2.4 Microscopic examination of heat induced cultures.....	181
6.3 Attempted Purification of Overexpressed Enzyme	183
6.4 Enzymatic Characterisation	184
6.5 Summary.....	185

Chapter Seven	Pullulanase Expression Studies using Yeast.....	187
7.1	Plasmid Construction.....	187
7.2	Enzyme Induction and Expression Studies.....	189
7.3	Plasmid Stability Tests.....	190
7.4	Summary.....	190

Chapter Eight	Analysis of the Deduced Peptide Sequence of the <i>C. saccharolyticum</i> Pullulanase.....	191
---------------	--	-----

8.1	Introduction.....	191
8.2	Comparison of the Putative Peptide Sequence of the <i>C. saccharolyticum</i> Pullulanase to Other Related Enzymes.....	192
8.3	The Presence of Three Conserved Regions in Pullulanases.....	195
8.4	Comparison of the Peptide Sequence of the <i>C. saccharolyticum</i> Pullulanase to other Pullulanases.....	200
8.5	Pullulanase Secondary Structure Predictions.....	204
8.6	Summary.....	208

Chapter Nine	Discussion.....	211
--------------	-----------------	-----

9.1	Comparison of the DNA sequence of the pNZ1038 <i>Caldocellum saccharolyticum</i> pullulanase to other known <i>Caldocellum saccharolyticum</i> genes and the DNA sequences of other related enzymes.....	211
-----	--	-----

9.2	Comparison of the peptide sequence of the <i>C. saccharolyticum</i> pullulanase to the peptide sequences of other related enzymes.....	214
9.2.1	Comparison of the amino acid sequences.....	214
9.2.2	Differences related to thermostability.....	214
9.2.3	Differences related to substrate characterisation.....	215

9.3	Enzymatic properties of the <i>Caldocellum saccharolyticum</i> pullulanase compared to the enzyme expressed in <i>E. coli</i>	217
-----	---	-----

9.4 Enzymatic properties of the <i>Caldocellum saccharolyticum</i> pullulanase compared to the enzymatic properties of other related enzymes	221
9.5 The potential biotechnological applications of the <i>Caldocellum saccharolyticum</i> pullulanase.....	223
9.6 Future work solving expression problems.....	224
9.6.1 Possible methods to improve pullulanase expression in <i>E. coli</i> -based expression systems.....	224
9.6.2 Possible methods to improve pullulanase expression in expression systems other than <i>E. coli</i>	225
9.7 Future work on the study and modification of enzyme activity.....	227
Appendices	231
One Results of Sequencing pCGN Recombinants.....	231
Two Analysis of pNZ1038.....	234
Three Peptide Comparisons	241
Four Genetic Analysis of <i>Caldocellum saccharolyticum</i> Genes.....	254
Five Three-dimensional Structure of α -amylase.....	260
Six Possible Errors in Debranching Enzyme Sequences.....	268
Seven Frame-shift Detection Algorithm.....	283
References.....	289

Acknowledgements

I would like to thank my supervisors, Professor Peter Bergquist, for his excellent advice and support during the course of this thesis, and Dr. Richard Gardiner, my associate supervisor. Thank you also to Dr. Ronnie McHale for his useful advice and discussion.

I am grateful to the other members of the Thermophile Genetics group, and other members of the Cell Biology Department as a whole, for creating a friendly atmosphere in which to work. Special thanks go to my fellow PhD student, Gordon Munro, for encouragement and the loan of a hard drive in my hour of need.

The following people from other departments have given time and technical assistance towards the completion of this work. Dr. Dave Christie, Biochemistry Department; Prof. Ted Baker, Massey University and the Thermophile group, Waikato University.

I would like to thank my family and friends for their support and encouragement, especially my wife, Carol, and my daughter, Rebekah, for typing my thesis.

Finally I wish to thank Pacific Enzyme Ltd for a PhD scholarship and funding of consumables for this research.

Abbreviations

A	adenine
BCIG	5-bromo-4-chloro-3-indolyl- β -galactopyranoside
bp	base pair(s)
C	cytosine
cpm	counts per minute
DEAE	O-(Diethylaminoethyl)-cellulose
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease I
dpm	disintegrations per minute
EDTA	ethylenediaminetetraacetate
EtBr	ethidium bromide
EtOH	ethanol
G	guanine
g	acceleration of gravity
kb	1000 base pairs
kDa	kilo Daltons
L-broth	Luria broth
MES	4-morpholineethanesulphonic acid
min	minutes
OD ₆₀₀	optical density (absorbance) at 600 nm
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
RNA	ribonucleic acid
RNase	ribonuclease A
SDS	sodium dodecyl sulphate
T	thymine
T _m	melting temperature
Tris	2-amino-2-hydroxymethylpropane-1,3-diol
Tris-HCl	Tris solution, pH adjusted with HCl
tRNA	transfer RNA
U	uracil
UV	ultraviolet light
v/v	volume per volume
vol.	volume(s)
w/v	weight per volume
°C	degrees Celsius

S.I. (Système Internationale (d'Unités)) abbreviations for units and standard notations for chemical elements, formulae, and chemical abbreviations (e.g. M \equiv moles per litre) are used throughout this work.