Isolation and Characterization of *Pseudobutyrivibrio ruminis* Gene Promoters

Tobias Schoep

B. Biol. Sc., Hons. Biotech., MU.

This thesis is presented for the degree of

Doctor of Philosophy

at Murdoch University

July, 2004

I declare that this thesis is my own account of my research and contains as its main content work which has not previously been submitted for a degree at any tertiary education institution.

____

Tobias Delavilla Schoep

ABSTRACT

A family of *E. coli* - *P. ruminis* shuttle-plasmids was constructed to allow the isolation and characterization of gene promoters from the rumen bacterium *P. ruminis*. The promoter rescue plasmid pBK was used to isolate a total of 4 genomic DNA fragments that promoted transcription in *P. ruminis* strains 0/10. These promoters, and an additional promoter, previously isolated from *P. ruminis* strain OR38 (Schoep, 1999), were identified by their ability to initiate expression of a promoterless *erm*AM gene in *P. ruminis*. Within 4 of the fragments, a total of 5 transcription start sites were identified in *P. ruminis* using a novel, fluorescent-primer extension analysis protocol. Comparison of promoters isolated in this and previous studies revealed a strong consensus RNA polymerase DNA-binding motif, including the well characterized –35 and –10 elements. Consensus sequences established for these elements were: TTgacA and **A**tAAT**A**ta respectively, where bold upper-case font, regular upper-case, and lower-case fonts represent conservation in 100%, 80%, and 70% of promoters respectively. The −10 and −35 motifs were interspaced by 16 – 18 nt. Among the newly identified promoters, the consensus for the –10 element was extended one nucleotide upstream and downstream of the standard hexamer (boxed). These motifs were similar to those recognized by eubacterial RNA polymerase containing the σ^{70}-like factor. Promoters also contained possible UP elements, and were significantly more curved than protein-coding regions. Additional plasmid vectors were constructed, to allow the use of both the quantitative SYBR green real time PCR and ß-glucuronidase assays, to examine 4 promoters in depth. This showed a wide range of promoter strengths within the group. However, no correlation was found between the composition and context of elements within *P. ruminis* promoters, and promoter strength. A mutation within the –35 element of one promoter revealed that promoter strength, and the choice of transcription start site were both sensitive to single nucleotide changes.

TABLE OF CONTENTS

General Acknowledgments……………....………………………………………..……..……i
List of Figures……………..……………………….…………………..………….….……ii
List of Tables……………………………………………………………………..…….....iv
Commonly Used Abbreviations……………………..…………………………………….vi
Chapter 1 : General Introduction 1
1.1 Biology and Molecular Biology of *P. ruminis* and Closely Related Species 1
1.1.1 Molecular Phylogeny 2
1.1.2 Molecular Tools for the Study of *P. ruminis* and Closely Related Species.....3
1.2 Bacterial Gene Promoters 4
1.2.1 Mechanisms of Transcription...5
1.2.2 Factors Affecting Transcription 7
1.2.2.1 Intrinsic Factors Affecting Transcription..9
1.2.2.2 Extrinsic Factors Affecting Transcription..16
1.2.3 Diversity of Promoter Recognition Regions 18
1.2.3.1 Intracellular Variation 18
1.2.3.2 Inter-Species Variation 22
1.2.4 Relationship Between Promoter Recognition Region Structure and Strength
....26
1.3 Gene Promoters Isolated from *P. ruminis* and Closely Related Species 27
1.3.1.1 Phylogenetic Reclassification of *B. fibrisolvens*: Implications for This Study 27
1.3.1.2 Reported Gene Promoters 28
1.4 Project Objectives 36
Chapter 2 : Materials And Methods...38
2.1 Selected Chemicals and Enzymes..38
2.2 Commonly Used Solutions 40
2.3 Bacterial Growth Media..42
2.4 Bacterial Strains, Plasmids and Primers Used 44
2.5 General Methods 45
2.5.1 Handling of *P. ruminis*...45
2.5.2 Agarose Gel Electrophoresis...45
2.5.3 DNA Mini-Preparations...49

2.5.3.1 Alkaline Lysis 49
2.5.4 Medium Scale Preparation of Plasmid DNA 50
2.5.5 Ethanol Precipitation 51
2.5.6 Butanol Precipitation 51
2.5.7 Bacterial Genomic DNA Preparations 51
2.5.8 DNA Extraction From Gels and Solutions 52
2.5.9 Dephosphorylation of DNA 5' Termini 53
2.5.10 DNA "Blunt Ending" or "End Polishing" Reaction 53
2.5.11 Restriction Endonuclease Reaction 54
2.5.12 DNA Ligations 54
2.5.13 Bacterial Transformations 55
2.5.13.1 Transformation of *Pseudobutyrivibrio ruminis* 55
2.5.13.2 Transformation of *Escherichia coli* 56
2.5.14 Polymerase Chain Reaction (PCR) 58
2.5.15 Screening *E. coli* and *P. ruminis* Colonies for Plasmid 59
2.5.16 DNA Sequencing Reaction 60
2.5.17 Protein Quantification 60
2.5.18 Total RNA Extraction 62
2.5.19 Measurement of DNA concentration using Spectrophotometry 63
2.5.20 Statistical Tests 64
2.5.21 General Computer Analyses 64
2.5.21.1 Searching for Conserved Motifs 64
2.5.21.2 Searching for Palindromes 65
Chapter 3 : Promoter Rescue Plasmid Construction 66
3.1 Introduction 66
3.1.1 The Promoter Rescue Plasmid pBHE 66
3.2 Methods 68
3.2.1 Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) 68
3.2.2 Phylogenetic Classifications 70
3.3 Results 71
3.3.1 Phylogenetic Classification of *P. ruminis* 0/10 71
3.3.2 Construction of Promoter Rescue Plasmids 72
3.3.3 Plasmid Characterization and Analysis 79
3.3.3.1 pBK3 79
3.3.3.2 pBK5 82

3.3.3.3 Transformation Studies 87
3.4 Discussion 87
3.4.1 Read-through transcription of the *erm*AM gene 87
3.4.1.1 pBK3 87
3.4.1.2 pBK5 88
3.4.1.3 Possible Sources of Read-through transcription 90
3.4.2 Comments 91
Chapter 4 : Isolation of Novel Gene Promoters 92
4.1 Introduction 92
4.2 Methods 93
4.2.1 Plasmid Based Promoter Rescue 93
4.2.2 Gene Library Construction and Auto Cre-Mediated Subcloning 94
4.2.3 Identification of Potential Gene Promoters 96
4.2.4 Transfer of Known and Potential Promoters to pBK6 96
4.2.5 Identification of Transformants 97
4.3 Results 98
4.3.1 Isolation of Promoter Rescue Plasmid Transformants 98
4.3.1.1 Analysis of "Pseudo-transformants" 99
4.3.2 Characterization of Promoter Rescue Plasmid Transformants 101
4.3.3 Identification of Promoters From a *P. ruminis* Gene Library 103
4.3.3.1 Identification of Promoters From Other *P. ruminis* Strains 107
4.3.4 Activity of Potential Promoters in *P. ruminis* 107
4.3.4.2 Activity of Potential *xyn*A Promoter in *P. ruminis* 109
4.3.5 Transformation Efficiencies 109
4.4 Discussion 109
4.4.1 Identification of Genes Under the Control of Gene Promoters 109
4.4.2 Identification of Potential Gene Promoters 111
4.4.3 Inactivity of Potential Gene Promoters 114
4.4.4 Possible Mechanisms of Non-Plasmid-Borne Erythromycin Resistance 115
4.4.5 Factors Affecting Transformation Efficiencies 116
4.4.6 Comments 118
Chapter 5 : Analysis of Gene Promoter Structure 119
5.1 Introduction 119
5.2 Methods 120
5.2.1 Growth Curves of *P. ruminis* 120

5.2.2 Primer Extension Analysis 121
5.2.2.1 Primer Extension Protocol 121
5.2.2.2 Migration Analysis of HEX Labeled DNA Molecules 123
5.2.2.3 Validation of Primer Extension Method 124
5.2.3 Searching for DNA motifs 125
5.2.4 Searching for Palindromic Sequences 126
5.2.5 Analysis of DNA curvature 126
5.2.6 Derivation of Consensus Sequences 127
5.2.7 Identification of Potential Promoters in other *P. ruminis* and Closely Related Species 127
5.3 Results 129
5.3.1 Development and Validation of Primer Extension Analysis Protocol 129
5.3.1.1 Migration Analysis of HEX and Big Dye Labeled DNA Molecules 129
5.3.1.2 Control Primer Extension Analysis 134
5.3.2 Primer Extension Analysis of *P. ruminis* Promoters 137
5.3.3 Analysis of Gene Promoter Structure 137
5.3.3.1 Position of Transcription Start Site Relative to Putative Coding Regions 137
5.3.3.2 Identification of Conserved Nucleotide Sequences 140
5.3.3.3 General Structural Characteristics of Gene Promoters 147
5.3.4 Analysis of Putative Promoters 149
5.3.4.1 Promoter Identification using RNAP sigma factor, DNA-binding consensus motif 149
5.4 Discussion 154
5.4.1 Fluorescent-Primer Extension Method 154
5.4.2 Characteristics of *P. ruminis* and Closely Related Species Gene Promoters 156
5.4.2.1 Promoters 18 and 54 156
5.4.2.2 Initial Transcribed Sequence 157
5.4.2.3 -10 Element 157
5.4.2.4 EX –10 Element 158
5.4.2.5 Interspacer 159
5.4.2.6 -35 Element 159
5.4.2.7 UP Element 160
5.4.2.8 DNA Curvature 160

5.4.3 Search for Conserved Sequences ... 162
5.4.3.1 Search for Common Motifs ... 162
5.4.3.2 Search for Common Structures ... 163
5.4.4 Application of Derived Consensus as a Searching Tool ... 164
5.4.4.1 Consensus Derivation and Searching ... 164
5.4.4.2 Search for Promoters ... 165
Chapter 6 : Analysis of Gene Promoter Activity ... 168
6.1 Introduction ... 168
6.2 Methods ... 170
6.2.1 Quantitative PCR (qPCR) ... 170
6.2.1.1 RT-qPCR to Determine Promoter Activity ... 172
6.2.1.2 qPCR to Determine Plasmid Copy Number ... 173
6.2.2 Glucuronidase Assay (GUS Assay) ... 174
6.2.3 Promoter Induction Experiments ... 175
6.2.4 Search for Regulatory Protein DNA-Binding Motifs ... 176
6.2.4.1 Database Searches ... 176
6.2.4.2 Algorithm Based Searches ... 176
6.2.5 Search for Palindromes ... 177
6.3 Results ... 177
6.3.1 Plasmid Construction ... 177
6.3.2 Promoter Activity: Measured Using Quantitative PCR (qPCR) ... 179
6.3.2.1 Standard Curve ... 179
6.3.2.2 Copy Number of pBGT Variants ... 181
6.3.2.3 Activity of Experimentally Identified Promoters ... 182
6.3.3 Promoter Activity: Measured Using the GUS Assay ... 185
6.3.3.1 Standard Curve ... 185
6.3.3.2 Activity of Experimentally Identified Promoters ... 186
6.3.3.3 Activity of Putative Promoters ... 188
6.3.4 Motifs Possibly Mediating Transcription Initiation ... 190
6.3.4.1 Protein Binding Sites ... 190
6.3.4.2 Palindromic Sequences ... 190
6.4 Discussion ... 191
6.4.1 Construction of vector pBG ... 191
6.4.2 Plasmid Copy Number Experiments ... 195
6.4.3 Inducible Promoters ... 195

6.4.4 Relationship Between Promoter Strength and Structure............................. 196
6.4.4.1 Intrinsic Factors... 196
6.4.4.2 Extrinsic Factors 199
6.4.4.3 Comments On Other Isolated Promoters 204
6.4.5 Promoter Activity: Measured Using GUS and RT-qPCR Assays 205
6.4.6 Disparity Between GUS Assay and RT-qPCR Analyses............................ 206
6.4.6.1 Possible Causes of Disparity Between Assays.................................... 210
6.4.7 Concluding Comments.. 212
Chapter 7 General Conclusions and Remarks.. 213
7.1 General Characteristics of *P. ruminis* and Closely Related Species Gene Promoters 213
7.1.1 Results of Recent Studies... 216
7.2 Promoter Strengths.. 216
7.3 Experimental Considerations 217
7.3.1 Tools Developed 217
7.3.2 Difficulties Cloning Rumen Bacterial DNA 218
7.4 Future Studies.. 219
7.5 Conclusion... 223
References.. 225
Appendix 1.. 241

ACKNOWLEDGEMENTS

I would first and foremost like to thank Prof. Keith Gregg, who guided me but always encouraged my own ideas. I feel that it allowed me to complete a study of philosophy, rather than just an experimental project. Keith was also diligent, beyond that of my co-workers supervisors, in proofing work quickly and finding additional funding for my project. I would also like to thank my partner Juliana Hamzah, who studying a PhD herself was very understanding of my work hours. Thanks to my family for always providing pleasant company and escape, and finally I thank the one remaining member of Rumen Biotech, Shawn Seet for his comradery in the laboratory.

I would like to thank members of the State Agricultural Biotechnology Centre for supplying me with feedback on ideas, and the occasional reagent. Especially, the Centre for Rhizobium Studies (Beau Fenner and Wayne Greeve) and Australian Centre for Necrotrophic Fungi Pathogens. Socially, people in SABC always provided friendly faces and happy conversation, which made coming to work a pleasure.

Finally thanks to all those international academics who fielded my questions by email and supplied me with laboratory materials.

LIST OF FIGURES

Figure 1.0 Interaction between activated RNA polymerase and the promoter recognition region at initial binding. 5
Figure 1.1 The pathway from transcriptional initiation to elongation as proposed by Murakami and Darst (2003). 8
Figure 1.2 A simplistic model of initial RNAP-promoter interactions (Bashyam and Tyagi, 1998; De Haseth et al., 1998; Vo et al., 2003). 10
Figure 1.3 Interactions between transcriptional activators, promoter recognition regions and RNAP. 20
Figure 1.4 Interactions between transcriptional repressors, promoter recognition regions (PRRs) and RNAP. 21
Figure 2.0 Typical standard curve used in protein assays, showing the relationship between concentration of bovine serum albumin and absorbance at 595 nm. 61
Figure 3.0 Phylogenetic tree showing the relationships between the 16S rDNA sequence of *Pseudobutyrivibrio ruminis* strains OR38 and 0/10. 71
Figure 3.1 The characteristics of the promoter rescue plasmid pBHE. 72
Figure 3.2 Diagrammatic representation of the construction of the plasmid pBK1. 73
Figure 3.3 A diagrammatic representation of the construction of the plasmid pBK2. 74
Figure 3.4 A diagrammatic representation of the construction of the plasmid pBK3. 75
Figure 3.5 A diagrammatic representation of the construction of the plasmid pBK4. 77
Figure 3.6 A diagrammatic representation of the construction of the plasmid pBK5. 78
Figure 3.7 A diagrammatic representation of the construction of the plasmid pBK6. 79
Figure 3.8 A diagrammatic representation of plasmids, showing the primer binding sites used in RT-PCR studies. 81
Figure 3.9 RT-PCR analysis to determine the origin of transcripts in pBK3. 83
Figure 3.10 A diagrammatic representation of pBK5(10), showing the primer binding sites used in RT-PCR studies. 84
Figure 3.11 RT-PCR analysis to determine whether transcription could originate upstream of the promoter insert in pBK5 and pBK6. 86
Figure 4.0 The positions of primer combinations used to amplify regions of plasmid pBK6 and its variants. 100
Figure 4.1 PCR analysis of pseudo-transformant pBK6(4). 101
Figure 4.2 Organization of sequenced regions of *P. ruminis* 0/10 genomic DNA cloned into vector pSC. 108
Figure 5.0 Position and sizes of primers used for transcriptional analysis of the pUC18 *bla*A gene. 125
Figure 5.1 The relationship between the mobility of HEX labeled DNA molecules, relative to: **(A)** Big Dye and **(B)** rhodamine labeled molecules. 131
Figure 5.2 The affect of HEX labeled DNA molecule size on electrophoretic mobility, relative to molecules labeled with rhodamine dye terminators. 132
Figure 5.3 The affect of HEX labeled DNA molecule size on electrophoretic mobility relative to DNA molecules labeled with Big Dye terminators. 133
Figure 5.4 Identification of the pUC18 *bla*A transcription start site using fluorescently labeled primer extension products. 135
Figure 5.5 Identification of the pUC18 *bla*A transcription start site (TSS) using fluorescently labeled primer extension products. 136
Figure 5.6 Optical densities of *P. ruminis* 0/10 cultures, in the presence (n = 4) and absence of plasmid (n = 3), during growth at 39°C. 138

Figure 5.7 Primer extension analysis of promoters, cloned in pBK6, from *P. ruminis* strains 0/10 and OR38 139
Figure 5.8 Manual alignment of 7 promoters identified in *P. ruminis* 142
Figure 5.9 The distribution of W_4-tracts in *P. ruminis* promoter regions and initial transcribed sequence (ITS) 143
Figure 5.10 A graphical representation of nucleotide sequence conservation within *P. ruminis* promoters, aligned at –10 elements. 144
Figure 5.11 Distribution of potential stem loop structure in *P. ruminis* gene promoters. 146
Figure 5.12 Distribution of palindromic sequences in native and randomized promoters from *P. ruminis*. 147
Figure 5.13 Local DNA curvature and G/C content of promoter regions, analyzed between –200 and +100 using the Bend.it program 152
Figure 6.0 Position of RT-qPCR amplicon relative to *gus*A translational start site and promoter insert region. 172
Figure 6.1 Diagrammatic representation of vector pBG2 construction. Tc^R: tetracycline resistance factor encoded within pFUS1 180
Figure 6.2 The use of SYBR green based qPCR to determine the relationship between Ct values and copy number of vector pBGT(10). 181
Figure 6.3 Copy numbers of plasmid pBGT variants in *P. ruminis* 0/10. 182
Figure 6.4 The use of SYBR green based qPCR to determine the relationship between Ct values and single stranded *gus*A gene copy number. 183
Figure 6.5 *gus*A transcript copy numbers in *P. ruminis* 0/10. 184
Figure 6.6 Relationship between absorbance (405nm) and concentration of 4-nitrophenol. 186
Figure 6.7 The activity of different promoters in *P. ruminis* 0/10 measured using the GUS assay. 187
Figure 6.8 Induced activity from putative gene promoters isolated from *P. ruminis* 0/10, measured by the GUS assay. 189
Figure 6.9 Analysis of potential transcription factor binding sites within promoter regions. 194

LIST OF TABLES

Table 1.0 The diversity of factors in *E. coli* affecting transcription initiation through interactions with RNA polymerase. 18
Table 1.1 Consensus DNA-binding motifs of various sigma factors in bacteria. IUB codes have been used where consensus motifs have a variable nucleotide composition (Appendix I). 23
Table 1.2 Approximate number of genes regulated by each type of sigma factor in *E. coli*. 24
Table 1.3 Comparison of bacterial promoter sequences likely to bind core RNAP/σ^{70}-like complexes. 25
Table 1.4 Methods previously used for defining *P. ruminis* and closely related species promoters. 30
Table 1.5 Various promoter elements suggested, or shown, to be constituents of active gene promoter recognition regions in *P. ruminis* and closely related species. 32
Table 1.6 Expression of the *xyn*A under the control of various versions of the C.. *proteoclasticum* Bu49 *xyn*A gene promoter in *P. species* strain OB156 (Xue et al., 1997). 35
Table 2.0 Sources of important chemicals. 38
Table 2.1 The geno/pheno types of bacterial strains used in this study 46
Table 2.2 The characteristics of the plasmids used in this study. 47
Table 2.3 Oligonucleotide primers used in this study. 48
Table 2.4 Electroporation settings for transformation of *Escherichia coli* and *Pseudobutyrivibrio ruminis*. 55
Table 2.5 The thermal cycle characteristics for PCR 58
Table 2.6 Thermal cycle conditions for dye terminator cycle sequencing reactions 60
Table 2.7 MEME and AlignAce settings used to analysis *P. ruminis* gene promoter regions for potentially common motifs. 65
Table 3.0 Temperatures at which PCR amplification of reverse transcription products were performed for different primer sets. 69
Table 4.0 Primers used for sequencing and cloning of both potential and experimentally proven gene promoters into plasmid pBK6. 97
Table 4.1 Showing genes downstream of and experimentally determined gene promoters in *P fibrisolvens* 0/10. 104
Table 4.2 Substrates or conditions considered likely to induce the promoter activity of potential gene promoters from the *P. ruminis* 0/10 gene library. 117
Table 5.0 Size standard PCR products used for calibration of primer extension analysis. 124
Table 5.1 Regions upstream of genes analyzed for the presence of gene promoters, and those used as background for calculation of DNA curvature. 128
Table 5.2 Comparison between theoretical, observed and corrected lengths (nt) of HEX labeled primer extension products co-electrophoresed with rhodamine and Big Dye based sequence. 134
Table 5.3 Length of HEX labeled primer extension products co-electrophoresed with rhodamine based sequence. 140
Table 5.4 Common motifs detected in promoter sequences using the programs MEME and AlignAce. 145
Table 5.5 Potential stem-loop structures in *P. ruminis* promoter regions. 146
Table 5.6 The curvature of *P. ruminis* coding and promoter regions. 148

Table 5.7 The program FindPatterns was used, as described in section 5.2.7, to identify potential gene promoters, performed using variations of consensus sequences defined in Figure 5.8. 153
Table 6.0 Composition of a typical qPCR reaction. 170
Table 6.1 Thermocycle conditions and data acquisition points used in qPCR. 171
Table 6.2 Conditions tested for induction of potential gene promoters in *P. ruminis* . 176
Table 6.3 Statistical comparison of mean promoter activities in *P. ruminis* after 6 h growth, measured using the RT-qPCR assays. 185
Table 6.4 Statistical comparison of mean promoter activities in *P. ruminis*, measured using GUS assays, after 6 h and 10 h growth. 188
Table 6.5 Analysis of potential transcription factor binding sites within promoters tested for activity using the GUS or RT-qPCR assay. 192
Table 6.6 Potentially significant palindromes in promoters of which the activities were examined.. 194
Table 6.7 Characteristics of various promoter elements thought to mediate promoter strength. 198
Table 6.8 Comparison of GUS assay and RT-qPCR methods for measuring promoter strength in *P. ruminis*. 208
Table 6.9 Comparison promoter strengths in *P. ruminis* at the mid-exponential growth stage, measured using GUS assay and RT-qPCR methods. 209
Table A1 The International Union of Biochemists (IUB) codes for nucleotides......... 241

COMMONLY USED ABBREVIATIONS

Abbreviation	Expansion
ABI	Applied Biosystems Inc.
AMV	Avian myeloblastosis virus
ATP	adenosine-5'-triphosphate
AR	analytical reagent
b	bases
bp	base pair
DEPC	diethyl pyrocarbonate
DNA	deoxyribonucleic Acid
DNAse	deoxyribonuclease
dNTP	deoxynucleotides
ds	double stranded
EDTA	ethylenediaminetetra-acetic acid
IPTG	isopropyl-β-D- thiogalactopyranoside
MCS	multiple cloning site
MMLV	Moloney murine leukemia virus
OD	optical density
PCR	polymerase chain reaction
pers. comm.	personal communication
pNPG	*p*-Nitrophenyl ß-D-Glucuronide
RNAse	ribonuclease
SDS	sodium dodecyl sulphate
ss	single stranded
TAE	Tris-acetate EDTA buffer
TBE	Tris-borate EDTA buffer
X-Gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside