

Characterisation of class II and III fructose
biphosphatases in *Clostridium*
acetobutylicum ATCC 824

Ali Abdullah Almohaish

A thesis submitted in partial fulfilment of
the requirements of Edinburgh Napier
University, for the award of Doctor of
Philosophy

March 2012

Abstract

In low GC Gram-positive bacteria, the phenomenon of carbon catabolite repression is dependent on a metabolite-activated bifunctional protein kinase/phosphorylase. In *Clostridium acetobutylicum* ATCC 824, the *hprK* gene encodes the bifunctional protein kinase/phosphorylase that is dependent on fructose 1,6-bisphosphate for activity. However, a putative *glpX* class II gene that might encode for fructose 1,6-bisphosphatase is located upstream of *hprK* and to date this is a unique gene arrangement. Therefore, the product of the putative *glpX* gene might have a vital role in regulating the activity of bifunctional protein kinase/phosphorylase by hydrolysing fructose 1,6-bisphosphate which is needed for the kinase activity. In the present work, experimental evidence is presented that the putative *glpX* gene encodes a fructose 1,6-bisphosphatase which can hydrolyse fructose 1,6-bisphosphate to fructose 6-phosphate and phosphate. FBPase activity was first demonstrated by cloning and transforming the *glpX* gene into an *E. coli* *fbp* mutant which cannot grow on gluconeogenic substrates such as glycerol. The transformed *glpX* gene was able to complement the *fbp* mutation of *E. coli* for growth on glycerol. GlpX was overexpressed as a GST-fusion protein and purified, and activity was demonstrated using fructose 1,6-bisphosphate as substrate. Activity was assayed at pH 8.0, and in the presence of Mn^{2+} , but the enzyme was inhibited completely by 1 mM phosphate. A putative *fbp* class III gene was also overexpressed and the encoded protein was purified as a GST-fusion product in order to compare Fbp with GlpX. To our knowledge, this is the first study that was able to purify a Fbp class III enzyme. The Fbp protein showed almost the same behaviour towards inhibitors compared to GlpX, but had a considerably higher specific activity than GlpX under the conditions of the experiments. The *glpX* gene was shown by RT-PCR to be transcribed together with *hprK* on the same mRNA during growth on glucose, and this indicates that *glpX* is unlikely to be a gluconeogenic gene. The results suggest that GlpX may play a novel and specific role in regulating the kinase activity activity of bifunctional protein kinase/phosphorylase and carbon catabolite repression in *C. acetobutylicum*.

Acknowledgement

This research project would not have been possible without the support of many people. It is a pleasure to express my gratitude to them all in my humble acknowledgment. First and foremost, I would like to express my sincere gratitude to my academic supervisor Prof. Martin Tangney for the continuous support of my PhD study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better supervisor and mentor for my PhD research.

Also, I would like to record my gratitude to my second academic supervisor Dr. Wilf Mitchell for his supervision, advice, and guidance from the very early stage of this PhD research as well as giving me extraordinary experiences throughout the work in the lab. Most of all, he gave me constant encouragement and support in various ways which contributed to create an excellent scientist that can handle novel researches. He gave me the chance to help, guide and supervise many undergraduate and postgraduate students and to present my work in prestigious conferences. I am really indebted to him more than he actually knows.

I would also acknowledge my Saudi Aramco advisers Christine Evans for helping me to get my extension and Saad Nassar, for his advice and support throughout my study. I am very grateful to Dr. Eve Bird for being the first person who taught me basic skills in microbiology. My special thanks go to my parents for giving birth to me in the first place, supporting and praying for me to succeed in this life. Finally, words fail me to express my gratitude to my wife for her patience, support and persistent confidence in me, which has reduced the load of the PhD on my shoulders.

List of abbreviations

ABE	Acetone Butanol Ethanol
ADP	Adenosine diphosphate
Amp	Ampicillin
ATP	Adenosine 5'-triphosphate
bp	Base pairs
cac	(referring to gene on chromosome) <i>C. acetobutylicum</i>
CcpA	Catabolite control protein A
CCR	Carbon catabolite repression
<i>cre</i>	Catabolite repression HPr-like protein
DNA	Deoxyribonucleic acid
EI	Enzyme One (phosphotransferase system)
EII	Enzyme Two (phosphotransferase system)
FBP	Fructose 1,6-bisphosphate (sugar).
FBPase	Any class of fructose 1,6-bisphosphatase enzyme
<i>fbp</i>	Referring to gene that encodes fructose 1,6-bisphosphatase
GlpX	Referring to class II fructose 1,6-bisphosphatase
<i>glpX</i>	Referring to gene that encodes class II fructose 1,6-bisphosphatase
g	Grams
GC	Guanine Cytosine content
HPr	Heat stable histidine-phosphorylatable protein
HPr K/P	Bifunctional HPr protein kinase/ phosphorylase
<i>hprK</i>	Referring to gene that encodes HPr K/P enzyme
Lac	(referring to operon) Lactose
LB	Luria-Bertoni medium
L	Litres
min	Minutes
nm	Nanometres
OD	Optical Density
ORF	Open Reading Frame
PEP	PhosphoEnolPyruvate
PTS	PhosphoTransferase System
P-Ser-HPr	HPr phosphorylated at a serine site

P-His-HPr	HPr phosphorylated at a histidine site
S	Second
U	units
μg	Micrograms
V	Volts
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Table of Contents

Abstract.....	i
Acknowledgements.....	ii
List of abbreviations.....	iii
Table of contents.....	v
List of figures.....	x
List of Tables.....	xiii
1 Chapter 1 General introduction.....	1
1.1 <i>Clostridium acetobutylicum</i>	2
1.2 The history of acetone-butanol-ethanol (ABE) fermentation.....	2
1.3 Revival of ABE fermentation.....	4
1.4 Biofuels.....	6
1.5 The physiology and biochemistry of the ABE process.....	7
1.6 The carbohydrate phosphotransferase system (PTS).....	11
1.7 Elements of Carbon Catabolite Repression (CCR).....	14
1.8 CCR in Gram negative bacteria.....	15
1.9 CCR in Gram positive bacteria.....	16
1.9.1 HPr and Crh.....	18
1.9.2 HPrK/P.....	19
1.9.3 CcpA.....	21
1.10 Elements of carbon catabolite repression in clostridia.....	23
1.11 Fructose 1,6-bisphosphatase.....	26
1.11.1 Fbpase class I.....	28
1.11.2 Fbpase class II (GlpX).....	28
1.11.2.1 In <i>E. coli</i>	28
1.11.2.2 In <i>Corynebacterium glutamicum</i>	29
1.11.2.3 In <i>Mycobacterium tuberculosis</i>	29
1.11.2.4 In <i>Bacillus</i> species.....	30
1.11.3 Fbpase class III.....	30
1.11.4 Fbpase class IV and V.....	30
1.12 Project aim.....	32
2 Chapter 2 General materials and methods.....	33
2.1 Bacterial strains, growth conditions, primer and vector details.....	34
2.1.1 <i>E. coli</i> strain growth conditions.....	36

2.1.2	Primer sequences	38
2.1.3	Growth conditions for <i>C. acetobutylicum</i> ATCC 824	39
2.2	Chromosomal DNA extraction from <i>C. acetobutylicum</i> ATCC 824.....	39
2.3	Polymerase Chain Reaction (PCR) for TOPO™ TA cloning	40
2.4	Gel electrophoresis.....	40
2.5	TOPO™ cloning procedure	40
2.6	Transformation of chemically competent <i>E. coli</i>	41
2.7	Screening of clones	41
2.7.1	PCR analysis	41
2.7.2	Plasmid purification.....	42
2.7.3	DNA Sequencing	42
2.7.4	Restriction analysis	43
2.7.5	Transformation of <i>E. coli</i> mutant strains	43
2.7.6	Fructose 1,6-bisphosphatase assay	45
2.7.6.1	Microburet protien assay	46
2.8	Cloning of the <i>glpX</i> and <i>fbp</i> genes in the Gateway expression system	48
2.8.1	BP cloning reaction.....	48
2.8.2	Transformation of chemically competent cells.....	49
2.8.3	LR reaction	49
2.8.4	Transformation into expression host.....	49
2.9	Cloning of <i>glpX</i> and <i>fbp</i> by Ligation-Independent Cloning (LIC)	50
2.9.1	LIC cloning reaction	50
2.9.2	Purification and precipitation of PCR product	51
2.9.3	T4 DNA Polymerase treatment of the purified PCR product.....	52
2.9.4	Transformation into the cloning strain.....	52
2.9.5	Transformation into the expression strain	53
2.9.6	PCR Screening.....	53
2.9.7	Expression of the Fbp and GlpX proteins.....	53
2.9.8	Purification of protien under native conditions using immobilizing metal-affinity chromatography (IMAC)	54
2.9.9	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE).....	55
2.9.10	Dialyzing of the recombinant proteins.....	56
2.9.11	Cleaving the fusion protein with recombinant enterokinase	57
2.10	Growth of <i>C. acetobutylicum</i> ATCC 842 for RNA extraction	57

2.10.1	RNA extraction for <i>C. acetobutylicum</i> ATCC 824.....	58
2.10.2	Analysis of purified RNA.....	58
2.10.3	Reverse transcriptase PCR.....	59
2.11	Proteomic analysis.....	61
2.12	Bioinformatics analysis.....	61
3	Chapter 3 Characterisation of <i>glpX</i> gene	62
3.1	Introduction.....	63
3.2	Results.....	63
3.2.1	Bioinformatic analysis.....	63
3.2.2	The <i>cac1088</i> gene.....	65
3.2.3	The <i>cac1572</i> gene.....	69
3.3	Cloning <i>glpX</i> and <i>fbp</i> genes using TOPO cloning system.....	74
3.3.1	Cloning of the <i>C. acetobutylicum</i> class II <i>glpX</i> gene.....	74
3.3.2	Determination of the orientation of the insertion using PCR analysis.....	77
3.3.3	Growth of <i>E. coli</i> strains on different selective media for phenotype studies.....	83
3.3.4	Chemical transformation and complementation of different <i>E. coli</i> mutants.....	85
3.3.5	Fructose bisphosphatase assay.....	90
3.3.6	Analysis of sequenced recombinant plasmid.....	92
3.4	Cloning <i>fbp</i> gene using TOPO cloning system.....	93
3.5	Discussion.....	97
3.5.1	Cloning of <i>glpX</i> gene using TOPO cloning system.....	97
3.5.2	Cloning of <i>fbp</i> gene using TOPO cloning system.....	103
3.6	Conclusion.....	104
4	Chapter 4 <i>GlpX</i> and <i>Fbp</i> overexpression and purification	105
4.1	Gateway cloning system.....	106
4.1.1	Introduction.....	106
4.1.3	Results.....	108
4.1.3.1	Cloning of the <i>glpX</i> gene into the donor vector pDONR TM 221.....	108
4.1.3.2	Transferring the <i>glpX</i> gene into the expression vector pET-60-DEST	111
4.1.3.3	Cloning of the <i>fbp</i> gene into the donor vector pDONR TM 221.....	113

4.2	Cloning and expression of the <i>glpX</i> and <i>fbp</i> genes using Ligation	
	Independent Cloning system (LIC).....	116
4.2.1	Introduction.....	116
4.2.2	Results.....	117
4.2.2.1	Amplification and purification of the <i>glpX</i> and <i>fbp</i> PCR products..	117
4.2.2.2	Cloning the <i>glpX</i> into the pET-41 Ek/LIC vector.....	122
4.2.2.3	Expression of the <i>glpX</i> gene in the recombinant pET-41 Ek/LIC vector.....	124
4.2.2.4	Cleavage of the recombinant fusion GlpX protein.....	127
4.2.2.5	Assay of the fructose 1,6 biphosphatase activity of the purified recombinant GlpX protein.....	128
4.2.2.6	Overexpression of the <i>glpX</i> gene in the mutant JB108.....	128
4.2.2.7	Complementation of JB108 mutant.....	130
4.2.2.8	Assay the fructose 1,6-biphosphatase activity of the recombinant GlpX from JB108 ₁ /pNova- <i>glpX</i> -3 crude extract.....	131
4.2.2.9	Cloning and expression of the <i>fbp</i> gene.....	133
4.2.2.10	Expression of the <i>fbp</i> gene in the recombinant pET-41 Ek/LIC vector.....	136
4.2.2.11	Expression of the <i>fbp</i> gene in the mutant JB108.....	138
4.2.2.12	Complementation of JB108 mutant via <i>fbp</i> gene.....	140
4.2.2.13	Assay of the fructose 1,6 biphosphatase activity of the recombinant Fbp protein from a crude extract of the JB108 ₃ / pNova- <i>fbp</i> -3.....	141
4.2.2.14	Comparision of the FBPase activities of recombinant GlpX (native and tagged) and recombinant Fbp proteins.....	143
4.3	Discussion.....	146
4.3.1	Overexpression of GlpX and Fbp proteins by Gateway cloning system.....	146
4.3.2	Overexpression of GlpX and Fbp proteins by LIC system.....	147
4.3.3	Assay of the purified Fbp and GlpX recombinant fusion proteins.....	150
4.4	Conclusion.....	151
5	Chapter 5 Transcriptional analysis.....	152
5.1	Reverse transcriptase- Polymerase chain reaction (RT-PCR).....	153

5.1.1	Introduction.....	153
5.1.2	Results.....	153
5.1.2.1	Determination the expression of <i>glpX</i> and <i>hprK</i>	153
5.1.2.2	QIAGEN [®] One-Step RT-PCR Kit.....	158
5.1.2.3	One Step RT-PCR Master Mix Kit from Novagen [®]	160
5.2	Discussion.....	163
5.3	Conclusion	165
6	Chapter 6 General discussion	166
6.1	The importance of intracellular FBP levels in CCR and strain improvements.....	167
6.2	Conclusions and future work	171
7	Appendix	172
7.1	Sequence of the recombinant fusion tags.....	173
7.2	Bioline Hyperladders	174
7.3	HyperPAGE Prestained Protein Marker	175
7.4	pCR 2.1 TOPO Cloning Vector	176
7.5	The donor vector pDONR TM 221.....	177
7.6	The Gateway [®] Nova pET-60-DEST TM vector	178
7.7	The pET-41 Ek/LIC vector	179
7.8	Amino acids abbreviation	180
7.9	Calibration curve of protein	181
7.10	FBPase enzyme assay curve	182
7.11	Proteomic analysis of FBPase class II (GlpX) protein	183
8	References	184

List of Figures

Figure 1.1: Biochemical Pathways in <i>C. acetobutylicum</i>	8
Figure 1.2: The Phosphoenolpyruvate dependent phosphotransferase system.....	12
Figure 1.3: Model of catabolite activation or repression in <i>B. subtilis</i>	17
Figure 1.4: The model of the phosphotransferase system (PTS) mediated by the catabolite repression in <i>C. acetobutylicum</i> based on <i>B. subtilis</i> model.....	24
Figure 1.5: Pathway of gluconeogenesis. Adapted from Berg et al., (2002).	27
Figure 2.1: Schematic illustration of a protein with a His•Tag attached at the N-terminal binding to a metal-chelated affinity support (NTA-Ni).	54
Figure 3.1: Unrooted phylogenetic tree of characterised FBPsases and the two putative clostridial FBPsases	64
Figure 3.2: Alignment of the putative <i>C. acetobutylicum</i> fructose 1,6-bisphosphatase II (cac1088) sequence with the closest homologues.	68
Figure 3.3: Multiple alignment of the putative <i>C. acetobutylicum</i> fructose 1,6- bisphosphatase (cac1572) sequence with the top 10 results from the BLAST search.....	73
Figure 3.4: Two agarose gels of PCR amplification and screening of <i>glpX</i> gene.	76
Figure 3.5: Illustration of the strategy to determine the orientation of the insert in the pCR2.1-TOPO vector.....	78
Figure 3.6: Agarose gel of PCR analysis to determine the orientation of the <i>glpX</i> gene in the vector.	79
Figure 3.7: Illustration of the strategy to determine the orientation of the insert in the vector by restriction enzymes.....	81
Figure 3.8: Agarose gel showing <i>Hind</i> III (A) and <i>Eco</i> RI (B) digests of purified plasmid.	82
Figure 3.9: Growth of <i>E. coli</i> strains on M9 minimal medium containing 0.4% glycerol.	84
Figure 3.10: Growth of JB108/pM ₁ W strains on two M9 glycerol plates.	86
Figure 3.11: Growth of JB108/pM ₁ W strains on two LB plates.....	87
Figure 3.13: Growth of JB103/pM ₁ W and JLD2402/pM ₁ B on M9 glycerol plate.....	89
Figure 3.14: Growth of JB103/ M ₁ B on M9 glycerol plate (+Tet).	89
Figure 3.14: Two agarose gels of PCR amplification and purification of <i>fbp</i> gene.....	96
Figure 4.1: Illustration of the two BP and LR reactions of Gateway® cloning Technology.....	107

Figure 4.2: Agarose gel of PCR screening of the <i>glpX</i> gene in MAX- <i>glpX</i> isolates.	110
Figure 4.3: Agarose gel of the purified plasmids (pMAX- <i>glpX</i>).	110
Figure 4.4: Agarose gel of PCR screening of the <i>glpX</i> gene in Rosetta- <i>glpX</i> isolates.	111
Figure 4.5: Agarose gel of PCR screening of the <i>fbp</i> gene in MAX- <i>fbp</i> isolates.	115
Figure 4.6: Agarose gel of the purified plasmids (pMAX- <i>fbp</i>).	115
Figure 4.7: Diagram of the Ek/LIC strategy.	116
Figure 4.8: Agarose gel of PCR amplification of the <i>glpX</i> and <i>fbp</i> genes.	117
Figure 4.9: Agarose gel of purified PCR products of the <i>glpX</i> and <i>fbp</i> genes using agarose gel purification (Ultrafree-DA).	120
Figure 4.10: Agarose gel of purified PCR products of the <i>glpX</i> and <i>fbp</i> genes following chloroform extraction and ethanol precipitation.	121
Figure 4.11: Agarose gel of PCR screening of the <i>glpX</i> gene in Nova- <i>glpX</i> isolates.	122
Figure 4.12: Agarose gel of the purified pNova- <i>glpX</i> plasmids.	123
Figure 4.13: Agarose gel of PCR screening of the <i>glpX</i> after purification from Nova- <i>glpX</i> -3 and Nova- <i>glpX</i> -4 isolates.	124
Figure 4.14: Agarose gel of PCR screening of the <i>glpX</i> gene in BL21/ pNova- <i>glpX</i> -3 isolates.	125
Figure 4.15: SDS-PAGE of purified GlpX proteins in the eluted fractions after His-tag purification under native conditions.	126
Figure 4.16: Illustration of GlpX protein attached to tagged proteins at the N- and C-terminus. Enterokinase cleavage site is shown by the horizontal arrow.	127
Figure 4.17: Agarose gel of PCR screening of the <i>glpX</i> gene in the JB108/ pNova- <i>glpX</i> -3 isolates.	128
Figure 4.18: SDS-PAGE of purified GlpX proteins in the eluted fractions after His-tag purification under native conditions.	129
Figure 4.19: Growth of JB108 ₁ /pNova- <i>glpX</i> -3 and JB108 ₂ /pNova- <i>glpX</i> -3 strain on two M9 glycerol plates.	130
Figure 4.20: Agarose gel of PCR screening of the <i>fbp</i> gene in Nova- <i>fbp</i> isolates.	134
Figure 4.21: Agarose gel of the purified pNova- <i>fbp</i> plasmids.	134
Figure 4.22: Agarose gel of PCR screening of the <i>fbp</i> gene after purification.	135
Figure 4.23: Agarose gel of PCR screening of the <i>fbp</i> gene in BL21/pNova- <i>fbp</i> -3 isolates.	136
Figure 4.24: SDS-PAGE of the purified Fbp proteins in the eluate fractions after His-tag purification under native conditions.	137

Figure 4.25: Agarose gel of PCR screening of the <i>fbp</i> gene in JB108 ₃ /pNova- <i>fbp</i> -3 and JB108 ₄ /pNova- <i>fbp</i> -3 isolates.....	138
Figure 4.26: SDS-PAGE of the purified Fbp proteins in the eluate fractions after His-tag purification under native conditions.....	139
Figure 4.27: Growth of JB108 ₃ /pNova- <i>fbp</i> -3 and JB108 ₄ /pNova- <i>fbp</i> -3 strains on two M9 glycerol plates.....	140
Figure 5.1: Illustration of <i>glpX</i> , <i>hprK</i> and 5-formyltetrahydrofolate cyclo-ligase gene arrangement.....	156
Figure 5.2: Agarose gel electrophoresis of PCR amplification of fragments between <i>hprK</i> , <i>glpX</i> and a gene that encodes for 5-formyltetrahydrofolate cyclo-ligase.....	157
Figure 5.3: Agarose gel of purified total RNA from a culture of <i>C. acetobutylicum</i> grown in glucose minimal medium.....	158
Figure 5.4: Agarose gel electrophoresis of RT-PCR attempt to amplify fragments between <i>hprK</i> , <i>glpX</i> and a gene that encodes for 5-formyltetrahydrofolate cyclo-ligase.....	159
Figure 5.5: Agarose gel electrophoresis of a control RT-PCR reaction from the Novagen [®] One Step RT-PCR Master Mix Kit.....	161
Figure 5.6: Agarose gel electrophoresis of RT-PCR reaction to amplify fragments between <i>hprK</i> , <i>glpX</i> and a gene that encodes for 5-formyltetrahydrofolate cyclo-ligase.....	162

List of Tables

Table 2.1: Details of <i>E. coli</i> strains used in this study and their genotypes.....	35
Table 2.2: Plasmids used in this study	36
Table 2.3: Growth media composition.....	36
Table 2.4: List of antibiotics	37
Table 2.5: Primer details	38
Table 2.6: Composition of the restriction digest mixture	43
Table 2.7: The composition of the fructose bisphosphatase assay reaction mixture	46
Table 2.8: Different amounts of BSA and dH ₂ O that were needed in the 1 ml assay mixture.....	47
Table 2.9: Hot start DNA polymerase reaction setup.	50
Table 2.10: Temperatures and times that were used for amplification by KOD Hot Start DNA Polymerase.....	51
Table 2.11: T4 DNA polymerase reaction setup.....	52
Table 2.12: Reaction components for small scale optimization incubated in room or 37°C temperature.....	57
Table 2.13: Reaction mix for one-step RT-PCR.....	59
Table 2.14: RT-PCR Cycling steps (QIAGEN [®] One-Step RT-PCR Kit).....	60
Table 2.15: Reaction mix for one-step RT-PCR.....	60
Table 2.16: RT-PCR Cycling steps (One Step RT-PCR Master Mix Kit Novagen [®])....	61
Table 3.1: BLAST homology results for the deduced amino acid sequence of the putative fructose 1,6-bisphosphatase II (cac1088).	66
Table 3.2 BLAST homology results for the deduced amino acid sequence of the putative fructose 1,6-bisphosphatase III (cac1572).	70
Table 3.3: Locations of TOPOglpX primers on the sequence of cac1088 (<i>glpX</i> gene)..	75
Table 3.4: Growth of the <i>E.coli</i> mutants on LB, glycerol and their phynotype.....	84
Table 3.5: Growth of different <i>E. coli</i> mutant strains in LB and glycerol medium supplemented with or without IPTG.	87
Table 3.6: The effect of different metabolites (effectors) on the reaction rate of the crude GlpX enzyme compare to the control (no effector).....	91
Table 3.7: Locations of TOPOfbp primers on the sequence of cac1572 (<i>fbp</i> gene).....	94
Table 3.8: Comparision of the effect of various metabolites on the activity of different GlpX enzymes.	102

Table 4.1: Locations of GATEglpX primers on the sequence of cac1088 (<i>glpX</i> gene).	109
Table 4.2: Locations of GATEfbp primers on the sequence of cac1572 (<i>fbp</i> gene).....	114
Table 4.3: Locations of LICglpX primers on the sequence of cac1088 (<i>glpX</i> gene)....	118
Table 4.4: Locations of LICfbp primers on the sequence of cac1572 (<i>fbp</i> gene).....	119
Table 4.5: The effect of different metabolites (effectors) on the reaction rate of the crude recombinant GlpX enzyme compare to the control (no effector).....	132
Table 4.6: The effect of different metabolites on the reaction rate of the recombinant Fbp enzyme compare to the control (no effector).	142
Table 4.7: Comparisons of different effectors on the FB Pase activity of GlpX (native and tagged) and Fbp tagged of <i>C. acetobutylicum</i>	143
Table 4.8: Comparisons of different divalent cations require for FB Pase activity of GlpX (native and tagged) and Fbp tagged of <i>C. acetobutylicum</i>	144
Table 5.1: Locations of RNAglpX primers on the sequences of cac1088-cac1089 (<i>glpX-hprK</i> genes).	154
Table 5.2: Locations of RNAhprK primers on the sequences of cac1089-cac1090 (<i>hprK-5-formyltetrahydrofolate cyclo-ligase</i> genes).....	155

Chapter 1:

1 General introduction

1.1 *Clostridium acetobutylicum*

The taxonomy of the *Clostridium acetobutylicum* is Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium. It is a rod-shaped, Gram positive, obligate anaerobe that cannot grow in the presence of oxygen, and is capable of producing endospores under environmental stress conditions (Vos *et al.*, 2009). *C. acetobutylicum* belongs to a large genus that includes several important pathogens such as *C. botulinum* which produces the most neurotoxin proteins that cause botulism illness, *C. perfringens*, *C. difficile*, and *C. tetani* (Montecucco and Molgó, 2005). Also within the same genus is *C. thermocellum* which is capable of converting cellulosic substrate into ethanol (Ram and Seenayya, 2005). *C. acetobutylicum* was once known as the Weizmann organism, named after chemist Chaim Weizmann who originally utilised the bacteria to produce acetone, butanol and ethanol (ABE) in a large scale fermentation (Jones and Woods, 1986). The strain *C. acetobutylicum* ATCC 824 was isolated from garden soil in Connecticut in 1924, and has the ability to ferment a wide range of biomass. This strain has now been shown to be related to the Weizmann strain (Nolling *et al.*, 2001). The genome of the solvent-producing bacterium *C. acetobutylicum* ATCC 824 consists of a 4-Mbp chromosome and a megaplasmid (around 200 kbp in size) that carries the genes involved in the production of solvent, hence, this plasmid was named pSOL1 (Nolling *et al.*, 2001).

1.2 The history of acetone-butanol-ethanol (ABE) fermentation

The first butanol production from a microbial fermentation was reported in 1861 by Louis Pasteur (Pasteur, 1862). The shortage of natural rubber in the early 20th century stimulated an interest in producing butanol due to synthetic rubber was needed in manufacturing of automobile tyres (Gabriel, 1928). This problem stimulated the research of many scientists - one of them the chemist Chaim Weizmann, who isolated and studied many microorganisms between 1912 and 1914. One of the isolates was *C. acetobutylicum* (Jones and Woods, 1986). The First World War changed the future of the ABE fermentation. The smokeless gun powder (cordite) was required in large quantities for the production of munitions, and vast quantities of acetone were required for the manufacture process (Killeffer, 1927; Gabriel, 1928; Jones and Woods, 1986). Also, the automobile industry needed a lacquer, and butanol was the best solvent for the production of lacquer. The acetone production during World War II was needed again for military use (Jones and Woods, 1986).

In 1945 one-tenth of the acetone and two-thirds of the butanol in the USA were still produced in USA by fermentation (Rose, 1961). However due to many political, economical and scientific reasons around the 1950's the large scale fermentation for solvent production began to decline.

Several of the major drawbacks that played an important role in the decline of the ABE fermentation industry were firstly, the destructive problem of bacteriophage infection of the bioreactors, which could cause the production of ABE to be cut by half for the year (Jones and Woods, 1986). Secondly, solvent toxicity, especially butanol which is very toxic for bacteria even at low concentrations (1-2% v/v). This resulted in low final product concentrations, and subsequently high solvent recovery costs. Thirdly, the growing cost of biomass substrates and finally, the ever expanding petrochemical industry began to produce solvents at competitive prices compared to fermentatively produced ABE (Hasting, 1971).

1.3 Revival of ABE fermentation

As a result of the dramatic price increase of petrochemicals, interest in the ABE fermentation process has been revived, which has encouraged the development of new technologies to produce fuels from cheap, secure and renewable resources. The USA aims to increase the production of biofuels by three fold by 2025 (Hansen *et al.*, 2005). There has been much research conducted on many fronts (such as in clostridial genetics, bioreactor technology and identification of cheap, renewable feedstocks) to enable the ABE fermentation to once again become economically competitive against fossil fuel (Jones and Woods, 1986). On the molecular front, many of the genomes of solventogenic clostridia have been sequenced such as *C. beijerinckii* NCIMB 8052, *C. beijerinckii* BA101, and *C. acetobutylicum* ATCC 824 (Nolling *et al.*, 2001), *C. beijerinckii* BA101, a mutant strain which is derived from *C. beijerinckii* NCIMB 8052, is a hyper-butanol producing strain, that can tolerate butanol better than the parental strain, and also butanol production is increased by 100% (Qureshi and Blaschek, 2001).

Ingenious bioreactor design, which can remove ABE during the fermentation, has mitigated the problems caused by solvent toxicity (Nakas *et al.*, 1983). Many substrates have been identified which can replace molasses and grains, including, microalgae *Dunaliella* which can be utilized, with the addition of 4% glycerol, by *C. pasteurianum* with the production of 16 g/L of solvents (Nakas *et al.*, 1983). Bioinformatics analysis of *C. acetobutylicum* ATCC 824 identified 11 genes encoding for proteins that are involved in cellulose utilization (Nolling *et al.*, 2001). Whey, which is a by-product of the dairy industry, can be used as an economical fermentation feedstock for solvent-producing clostridia such as *C. acetobutylicum* (Qureshi and Maddox, 2005). Also, uptake of lactose which is the major sugar in whey has been confirmed to be catalyzed by a phosphotransferase system (PTS) (see section 1.6) (Yu *et al.*, 2007). Crude glycerol which is the by-product of the biodiesel industry can be used as an alternative substrate. This low quality crude glycerol contains a mixture of contaminated glycerol, water, methanol, free fatty acids, and both organic and inorganic salts (Thomson and He 2006).

Glycerol is produced in huge quantities during the production of biodiesel. 10 lbs of crude glycerol is made from every 100 lbs of biodiesel generated from vegetable oils and animal fats (Thompson and He, 2006). Due to the low price, crude glycerol is very competitive with other sugars that are used in biofuel production by microbial fermentation (Yazdani and Gonzalez, 2007). Up to date, there is no economical viable use for the crude glycerol and it cannot be used in cosmetics and pharmaceutical industries due to the purification process not being economical (Dobson *et al.*, 2011). However, in the USA the price of crude glycerol is constantly decreasing from US \$0.20 / lb in 2001 to US \$0.01 / lb in 2006 which makes it in the future a most promising, economically viable biological fermentation substrate (Dasari, 2007). Around 50% of the production cost of biofuels is due to the cost of the fermentation substrate, hence, if crude glycerol can be used the total production cost of the biofuel would be lower (Willke and Vorlop, 2008; Dobson *et al.*, 2011).

Several organisms are able to grow in glycerol, however, the best microorganisms that able to grow in glycerol and produce a valuable product belong to the genus *Clostridium*. Recent study indicates that a newly isolated *Clostridium butyricum* AKR102a from soil was able to grow in low-quality crude glycerol and produce 1,3-propanediol (1,3-PD) which can be used in cars as a coolant (Ringel *et al.*, 2011). When the strain was grown in crude glycerol, 76.2 g/L of 1,3-PD was produced making it the best non-engineered strain able to produce a valuable product in such a large quantity (Wilkens *et al.*, 2011). It has also been reported that *C. acetobutylicum* was able to co-utilize glucose and glycerol (Vasconcelos *et al.*, 1993). Carlos *et al.*, 2003 reported that *C. acetobutylicum* can grow on crude glycerol in the presence of glucose and produce the same amount of solvents as when grown on pure glycerol in the presence of glucose. A continuous culture of *C. acetobutylicum* over 70 days was able to produce 0.34 mol/mol of butanol with overall butanol productivity of 0.42 g/L/h with no sign of degeneration (loss of pSOL1 plasmid leads the cells to not produce solvents) (Carlos *et al.*, 2003). Biodiesel is produced in many EU countries, with Germany the largest producer and consumer of this biofuel with the annual production exceeding 2.5 billion liters. However, these countries have severe problems in disposal of the excessive crude glycerol generated from this industry and also the disposal process is quite expensive (Silva *et al.*, 2009). Therefore, establishing a biobutanol industry together with biodiesel leads not only to increase the profit, but also solves the disposal problem.

1.4 Biofuels

Biofuels are fuels produced from biomass (Demirbas, 2007). Biomass is a term which can refer to any organic biological material, however in this context, biomass usually refers to matter such as crops, wood and waste residues produced by farming, the food and drink processing industry and so forth. This biomass can be used for the production of renewable energy. Lignocellulose, which is the most abundant biomass on the Earth, is mainly composed of three sugars, glucose, arabinose, and xylose (Xiao *et al.*, 2011). Crop waste and forestry residues are examples of lignocellulosic biomass which does not compete with animal feedstock and food industry (Weber *et al.*, 2010). Biomass usually needs to be degraded into manageable mono or disaccharides before microbial conversion into solvents or biofuels can occur, and thermochemical, enzymatic (or a combination of the two) pretreatment is usually employed (Brodeur *et al.*, 2011). Biofuel is the most promising alternative energy that will achieve the aims of reducing CO₂ emissions and the dependence on fossil fuel, also production and the trade of biofuel can create new jobs and support the economy of the developing countries (Groom *et al.*, 2008).

There are two main types of liquid biofuel used for transport fuel:

- Biodiesel is derived from biological sources such as animal fats and vegetable oil. It has low emissions and is readily biodegradable, compared with petrodiesel. The term biodiesel was introduced by the National Soy Diesel Development Board in USA during 1992. Biodiesel can be used in the current diesel engine with moderate or no modification to it. However, there are some problems associated with the use of biodiesel in the current diesel engine, for example, high viscosity and gum formation which at the end damages the diesel engine (Ramadhas *et al.*, 2003).

- Bioalcohols, for example bioethanol produced by yeasts, are already being used in Brazil, the USA and some European countries on a large scale, and have potential to be one of the leading renewable biofuels in the coming 20 years due to the increase in crude oil price. Bioethanol is produced from sugar cane in Brazil and from corn in USA, however this biomass stream can also be utilised for human and animal nutrition, which causes these feedstocks to be subject to erratic price fluctuations and become a target for the growing food versus fuel movement (Hägerdal, 2006).

Biobutanol, which is butanol that is derived from biomass, is produced by the solventogenic clostridia such as *C. acetobutylicum*. The World market for butanol is currently 350 million gallons per year. The USA alone consumes 220 million gallons per year (Shapovalov and Ashkinazi, 2008). It is mainly used as solvent and an intermediate in chemical production (Lee *et al.*, 2008). However, Butanol has several advantages as a transport fuel over ethanol. Butanol has a higher energy content compared to ethanol, it blends better with gasoline at any concentration, and it can be used in current cars without any modification to the engine. It also evaporates six times less than ethanol, is not corrosive compared to ethanol, hence, can be distributed by pipelines. (Dürre, 2007; Shapovalov and Ashkinazi, 2008).

1.5 The physiology and biochemistry of the ABE process

The solvent-producing clostridia have two main metabolic phases. The initial growth phase - the acidogenic phase - produces hydrogen, carbon dioxide, acetate, and butyrate (Figure 1.1). Production of the acids results in a decrease of the pH of the medium.

The second phase of the fermentation is the solventogenic phase, in which the acetate and butyrate are taken up from the medium then converted into solvents, and this subsequently results in an increase of the pH of the medium (Davies and Stephenson, 1941; Johnson *et al.*, 1931; Reilly *et al.*, 1920; Rose, 1961).

The build-up of large amounts of acetic and butyric acids puts the cell in danger by dissipating the proton gradient across the cytoplasmic membrane. As a result, the cells stop the formation of acids and shift towards the production of neutral solvents, the acids being mostly reassimilated and converted to acetone, butanol and ethanol (Dürre, 2002).

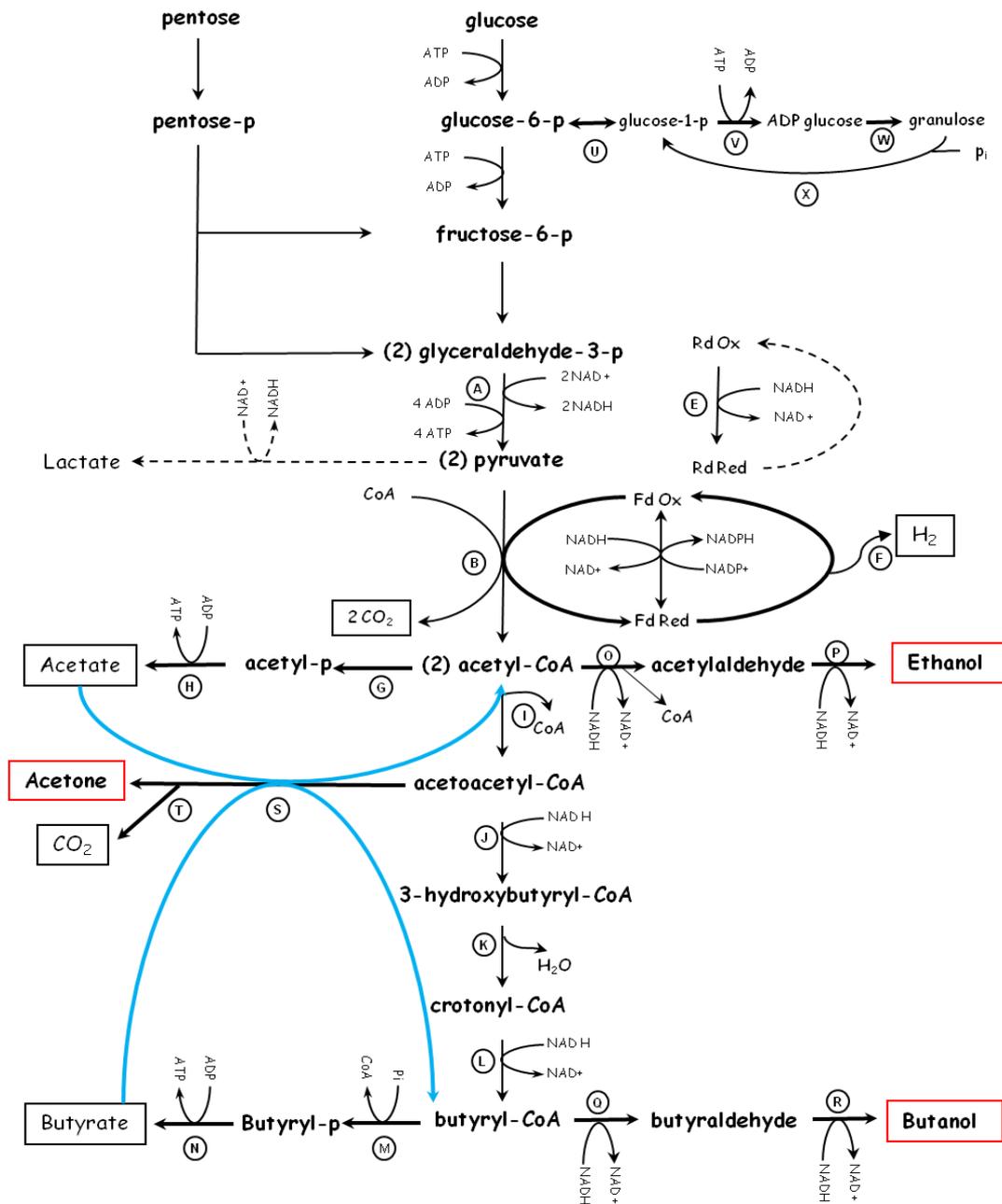


Figure 1.1: Biochemical Pathways in *C. acetobutylicum*

Figure adapted from Jones & Woods, 1986. Reactions which predominate during the acidogenic phase and the solventogenic phase of the fermentation are shown by thin and thick arrows, respectively. Enzymes are indicated by letters as follows: (A) glyceraldehyde 3-phosphate dehydrogenase; (B) pyruvate-ferredoxin oxidoreductase; (C) NADH-ferredoxin oxidoreductase; (D) NADPH-ferredoxin oxidoreductase; (E) NADH-rubredoxin oxidoreductase; (F) hydrogenase; (G) phosphate acetyltransferase (phosphotransacetylase);

(H) acetatekinase; (I) thiolase (acetyl-CoA acetyltransferase); (J) 3-hydroxybutyryl-CoA dehydrogenase; (K) crotonase; (L) butyryl-CoA dehydrogenase; (M) phosphatebutyryltransferase (phosphotransbutyrylase); (N) butyrate kinase; (O) acetaldehyde dehydrogenase; (P) ethanol dehydrogenase; (Q) butyraldehyde dehydrogenase; (R) butanol dehydrogenase; (S) acetoacetyl-CoA:acetate/butyrate:CoA transferase; (T) acetoacetate decarboxylase; (U) phosphoglucomutase; (V) ADP-glucose pyrophosphorylase; (W) granulose (glycogen) synthase; (X) granulose phosphorylase.

Under certain conditions *C. acetobutylicum* can convert pyruvate into lactate. Lactic acid is not produced under normal conditions, but this occurs when the hydrogenase activity is inhibited by carbon monoxide or the microbe runs out of iron (Jones and Woods, 1986). This pathway only operates as an alternative pathway in order to allow the continuation of NADH oxidation and energy generation. In the acidogenic pathway, butyrate and acetate are produced from butyryl-CoA and acetyl-CoA which results in the production of ATP at end of the pathway.

There are four main enzymes responsible for the formation of acids; phosphate acetyltransferase, and acetate kinase, encoded by *pta* and *ack* genes, for the production of acetic acid, and phosphate butyryl-transferase and butyrate kinase encoded by *ptb* and *buk* genes for the production of butyrate (Jones and Woods, 1986). In the acidogenic phase the highest amount of hydrogen production occurs, concurrent with a high rate of glucose consumption. The first decrease of hydrogen production is associated with a reduction in the metabolic activity and cell growth in the culture. Thus in this stage, the decrease in the hydrogen production is due to the reduction in the metabolic rate rather than inhibition of hydrogenase activity (Kim and Zeikus, 1985). The start of the solventogenic phase coincides with a fall in the pH of the medium which is linked to accumulation of acids. The addition of butyrate and acetate to a culture of *C. acetobutylicum* at pH 5.0 was reported to accelerate the induction of solventogenesis (Bernhauer and Kurschner, 1935; Reilly *et al.*, 1920), which was followed by a decrease in the growth rate and hydrogen production. The acetate and butyrate uptake only occurs when the sugars in the medium have been depleted.

The genes that are required for solvent production in *C. acetobutylicum* are *aad*, *ctfA*, *ctfB*, and *adc* which are located on a 192 kbp megaplasmid pSOL1. *C. acetobutylicum* may lose pSOL1 resulting in the loss of the ability to produce solvents. Strains that have lost the plasmid are said to have degenerated (Scotcher and Bennett, 2005). Solvent production serves as a detoxification process in response to the accumulation of acids which result in conditions that reduce the cells growth rate (Hartmanis *et al.*, 1984). The shift from acidogenic phase to solventogenic phase is accompanied by a change in the production ratios of hydrogen and CO₂ (a decrease in hydrogen production and an increase in CO₂ production. In the solventogenic phase, the cell metabolism continues until the concentration of solvents reaches an inhibitory point which is usually around 20 g/L. The most toxic solvent is butanol, and the solvent production is inhibited when the concentration of butanol reaches 13 g/L. Recent studies have confirmed that growth is completely inhibited by the addition of 12 to 16 g/L of butanol. However, the addition of ethanol and acetone to the medium reduces the growth by 50% at a concentration of 40 g/L and growth was totally inhibited at a concentration of 70 g of acetone and 50 g to 60 g of ethanol per litre (Leung and Wang, 1981; Jeanine and Antonio, 1983).

Butanol toxicity is related to the hydrophobic property of this molecule. The main effect of this compound appears to be the disruption of the phospholipid bilayer of the cell membrane which results in an increase in membrane fluidity. Furthermore, a recent study has observed an increase in the ratio of saturated to unsaturated fatty acids in the membranes both from solventogenic phase cells and vegetative cells grown in the presence of butanol (Schneck *et al.*, 1984). Hence, this increase in the saturated fatty acids in the membrane is due to the physiological response of the cell to tolerate the effect of increased membrane fluidity, which is similar to the response of cells grown at high temperature. Butanol tolerance could be enhanced by the addition of saturated fatty acids to the medium which results in an increase in the content of saturated fatty acids in the membrane. Also, a decrease in fermentation temperature may increase the butanol tolerance during the solventogenic phase (Rogers, 1984).

1.6 The carbohydrate phosphotransferase system (PTS)

Transport from the environment to inside the cell is the first step in any sugar metabolism pathway and the PTS system is the main sugar transport system in many Gram positive and Gram negative microorganisms (Postma *et al.*, 1993; Deutscher *et al.*, 2006). In 1964, this system was firstly reported in *E. coli* by Kundig *et al.* (1964). The PTS is a predominant transport system in many facultative and obligate anaerobic bacteria and involves the uptake and phosphorylation of a large number of carbon sources. Although its main function is as a transport system, it can also be involved in movement towards carbon sources (chemotaxis), and the regulation of gene expression (Deutscher *et al.*, 2006; Postma *et al.*, 1993). All the PTS systems that have been reported catalyze the following process:

$$\text{P-enolpyruvate (inside cell) + carbohydrate (outside cell)} \xrightarrow{\text{PTS}} \text{pyruvate (inside cell) + carbohydrate-P (inside cell)}.$$

Sugar uptake in clostridia relies heavily but not completely on the phosphoenolpyruvate-dependent PTS (Figure 1.2). The main mechanism of carbohydrate transport in *C. acetobutylicum* and *C. beijerinckii* is the PTS (Mitchell and Tangney, 2005). Many sugars such as sucrose, maltose, glucose, lactose and fructose have been reported to be PTS substrates in *C. acetobutylicum* and the genes of these systems have been characterised (Tangney and Mitchell, 2000; Tangney *et al.*, 2001; Tangney and Mitchell, 2005; Yu *et al.*, 2007). As in many bacteria, the expression of PTS genes is induced by their substrates and they are subject to carbon catabolite repression (CCR) by glucose. Neither sucrose nor maltose can be utilized by the growing bacteria when cultured in the presence of glucose (Tangney and Mitchell, 2007).

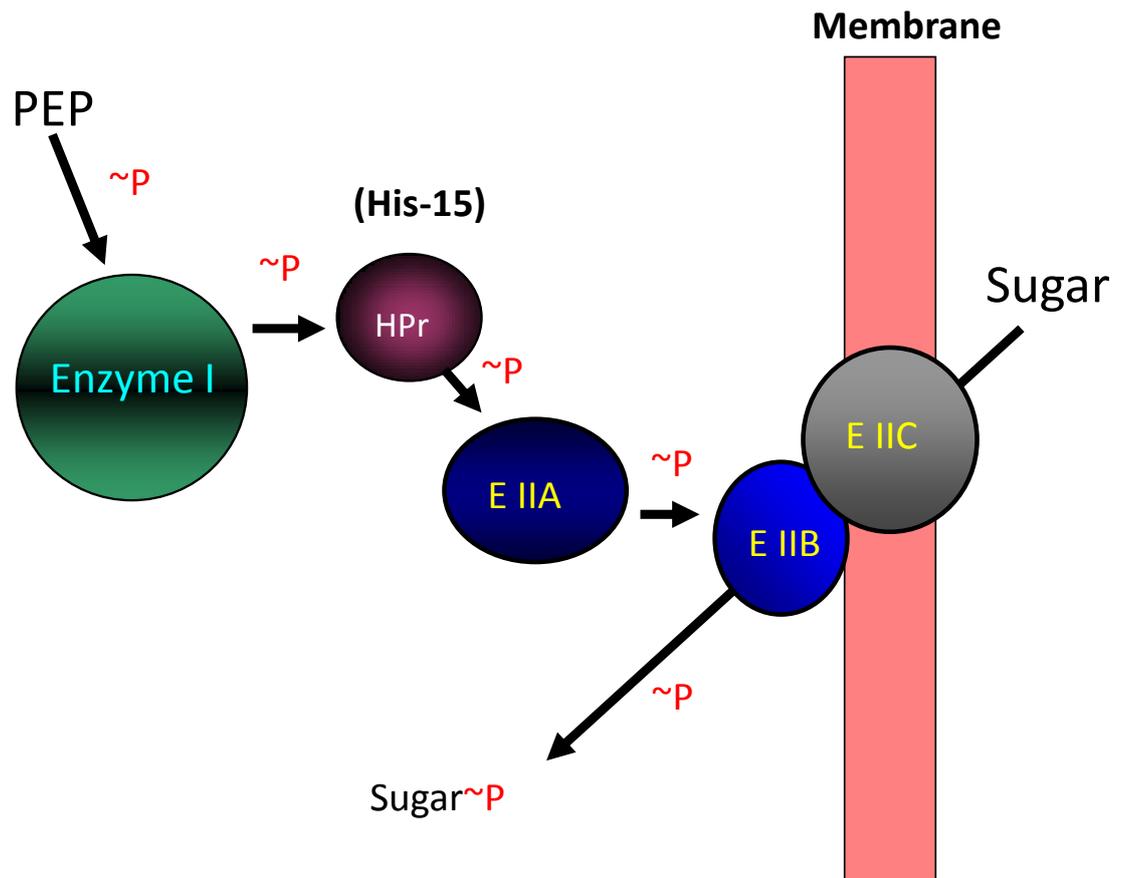


Figure 1.2: The Phosphoenolpyruvate dependent phosphotransferase system

Organisation of the phosphoenolpyruvate dependent phosphotransferase system (PEP: PTS). Figure adapted from Mitchell and Tangney, 2005. Enzyme I and heat stable histidine-phosphorylatable protein (HPr) are the general PTS proteins for all PTS's. The substrate specific enzyme II can have highly variable domain structure, present as a single polypeptide or individual proteins. A phosphate (~p) is donated from PEP to enzyme I, HPr, enzyme II, and then the sugar.

The PTS is composed of a multiprotein phosphoryl transfer chain. The first two proteins, enzyme I (EI) and HPr are soluble proteins that are involved in the phosphorylation of all of the PTS carbohydrates in any microorganism; hence, they have been called the general PTS proteins (Deutscher *et al.*, 2006; Tangney and Mitchell, 2005; Postma *et al.*, 1993). Bioinformatics analysis has revealed that most of the EIs from Gram-positive and Gram-negative bacteria show strong identity and consist of around 570 amino acids with size about 63 kDa (Kuang-Yu and Saier Jr, 2002). EI is phosphorylated in the presences of Mg^{2+} at a conserved histidine site which is located in the N-terminal region (His-189) in *E. coli* (Weigel *et al.*, 1982). Moreover, several studies reported Mg^{2+} is needed for EI phosphorylation, but is not needed for transfer of the phosphate from EI to the conserved His-15 in HPr (Weigel *et al.*, 1982; Chauvin *et al.*, 1996). The C terminal of the enzyme which contains the PEP binding site is needed for dimerization of EI. Only dimer can be phosphorylated by PEP, which may imply that monomer to dimer transition is involved in regulation of the enzyme activity (Chauvin *et al.*, 1996).

HPr, which is the heat stable histidine-phosphorylatable protein and EI are encoded by the genes *ptsH* and *ptsI*, respectively. The other PTS proteins are the substrate-specific enzyme II (EII) domains, which might exist separately or connected within a single polypeptide chain (Tangney and Mitchell, 2007). These domains are referred to as the EIIA, EIIB, and EIIC, whereas a small number of systems were reported to have an extra EIID domain such as the fructose PTS system in *B. subtilis* (Martin-Verstraete *et al.*, 1990; and Reizer, 1992). The first two domains IIA and IIB are hydrophilic and both contain sites for phosphorylation, whereas IIC and IID are membrane bound proteins and their function is to translocate the substrate. The EIIC domain is more hydrophobic than the IID (Saier Jr and Reizer, 1992; Postma *et al.*, 1993).

The clostridial system consists of functional proteins equivalent to those in other bacteria. Hence, the soluble part of the *C. beijerinckii* NCIMB 8052 PTS complemented the membranes prepared from *Bacillus subtilis*, *C. pasteurianum* and *E. coli* for glucose phosphorylation (Tangney and Mitchell, 2005).

Also, clostridial extracts restored the PTS activity of the extracts prepared from *ptsI* and *ptsH* mutants of *B. subtilis*, *Staphylococcus aureus* and *E. coli* (Saier and Reizer, 1992; Tangney and Mitchell, 2005). The phosphotransferase system (PTS) has been shown to play a main role in bacterial CCR through HPr (Zeng and Burne, 2010; Deutscher *et al.*, 2006).

1.7 Elements of Carbon Catabolite Repression (CCR)

More than a century ago, it was reported that growth on glucose can decrease the activity of specific enzymes in yeast and bacteria. This event became known as the glucose effect. The first description of this phenomenon was by Dienert in 1900 in France. He reported that *Saccharomyces cerevisiae* cells that had been adapted to galactose, lost this adaptation directly after they were exposed to glucose or fructose (Dienert, 1900). Another study was done in Wageningen, Netherlands by Sohngen and Coolhaas, who measured the influence of temperature on the glucose effect in *S. cerevisiae* (Sohngen and Coolhaas, 1924). At the beginning of the 1940s, Jacques Monod studied *B. subtilis* and *E. coli* and concluded that “in each organism, a specific hierarchy existed for the utilization of carbon sources, with glucose usually being in the top of it” (Monod, 1942). Later, it was found that the preferred carbon sources such as glucose, fructose and sucrose could repress the synthesis of the enzymes that are needed for the transport and metabolism of less favourable carbon sources. This observable fact became known as carbon catabolite repression (CCR) (Deutscher *et al.*, 2006).

When the preferred sugar in a culture is depleted, cells begin to synthesize the necessary enzymes for the transport and metabolism of the less favourable sugar, and usually show a lag phase before they can resume growth again. Therefore, the glucose effect phenomenon reported in the 1900s can be considered as carbon catabolite repression. Carbon catabolite repression can be defined as the inhibitory effect of a particular sugar in the growth medium on gene expression or the activity of enzymes involved in the catabolism of other sugars (Deutscher *et al.*, 2006). CCR can be mediated by various mechanisms, which can repress the transcription of catabolic genes via specific or global regulators (Deutscher, 2008). The main CCR mechanism in firmicutes is mediated by the global regulator catabolite control protein A (CcpA) as will be explained later in greater detail.

1.8 CCR in Gram negative bacteria

A major global regulator in *E. coli* is the cyclic AMP receptor protein (CRP), also called the catabolite activator protein (CAP) which is a homodimeric transcription protein with each monomer composed of two functional domains. The first domain is located in the N-terminal region which contains a cAMP binding site and the second domain is located in the C-terminal which contains a specific DNA binding site (Cheng and Lee, 1998; Busby and Ebright, 1999; Li and Lee, 2011). CRP activates transcription at many promoters in the presence of cAMP by binding to specific palindromic sequence of 22 bp located in or upstream of the promoter (Busby and Ebright, 1999; Lawson *et al.*, 2004). In absence of glucose, cAMP is present at a high concentration which stimulates CRP binding to the specific CRP-site to activate transcription. However, under glucose growth conditions, cAMP is present at a low level, hence, no transcriptional activation occurs (Lawson *et al.*, 2004; Crasnier-Mednansky *et al.*, 1997).

The catabolite repressor /activator (Cra) protein is another global regulator that mediates the central metabolic pathways in Gram negative bacteria (Chavarría *et al.*, 2011). Cra was initially known as the fructose repressor (FruR) for repressing of the genes that are involved in fructose metabolism, however, more studies reported that FruR is not specific for repressing fructose utilizing genes but has a global regulator function which can repress or activate many catabolic genes. Therefore, it was termed Cra (Ramseier *et al.*, 1993; Shimada *et al.*, 2010; Chavarría *et al.*, 2011). Cra belongs to the GalR-LacI family of proteins which contain two domains, an N-terminal helix-turn-helix domain which recognises a specific imperfect palindromic DNA sequence of around 12 bp and binds to it, and a second C-terminal domain which is the inducer binding domain by which fructose 1,6-bisphosphate (FBP) and fructose 1-phosphate bind to the protein (Saier Jr and Ramseier, 1996; Shimada *et al.*, 2010).

Repression and activation of the catabolic genes are dependent on the location of the Cra-binding sites. Therefore, activation occurs when the Cra-binding site is located within or downstream of the promoter, but, gene repression occurs when the Cra-binding site is located upstream of the promoter (Saier Jr and Ramseier, 1996).

Cra is regulated by FBP and/ or fructose 1-phosphate (effectors) via stimulating dissociation of Cra from the Cra-binding sites. As a result, the genes that are associated with a Cra-binding site located upstream of the promoter are activated by Cra and since the effector binds to Cra, this leads to displacement of the protein from the Cra-binding site, hence, these genes are subject to catabolite repression. Whereas, the genes are associated with a Cra-binding site that is located within or downstream of the promoter are repressed by Cra and when the effector binds to Cra, this leads to displacement of the protein from the Cra-binding site, and hence these genes are subject to catabolite activation (Saier Jr and Ramseier, 1996).

To date, it is unclear why the cell has two different global regulators, Cra and CRP, for controlling carbon metabolism (Shimada *et al.*, 2010). However, recent studies indicate the *crp* gene encoding for CRP is under the control of the Cra regulator which makes it the true global regulator in Gram negative bacteria (Shimada *et al.*, 2010; Chavarría *et al.*, 2011).

1.9 CCR in Gram positive bacteria

Catabolite control protein A (CcpA) is the the global regulator that controls the central metabolic pathways in Gram positive bacteria (Singh *et al.*, 2008). It also can regulate genes that are involved in sporulation, toxin production and biofilm formation (Varga *et al.*, 2004; Varga *et al.*, 2008; Antunes *et al.*, 2010). Like Cra in Gram negative bacteria, CcpA belongs to the GalR-LacI family (Shimada *et al.*, 2010; Deutscher *et al.*, 2008). CCR occurs when HPr of the PTS is phosphorylated at a serine site (P-Ser-HPr) by HPr kinase which forms a complex (P-Ser-HPr-CcpA) that binds to a specific DNA site in the presence of FBP (Figure 1.3) (Deutscher *et al.*, 2008). Moreover, *B. subtilis* contains a homologous protein to HPr termed Crh which can be involved in CCR (Figure 1.3). All these elements of CCR are explained below in greater detail.

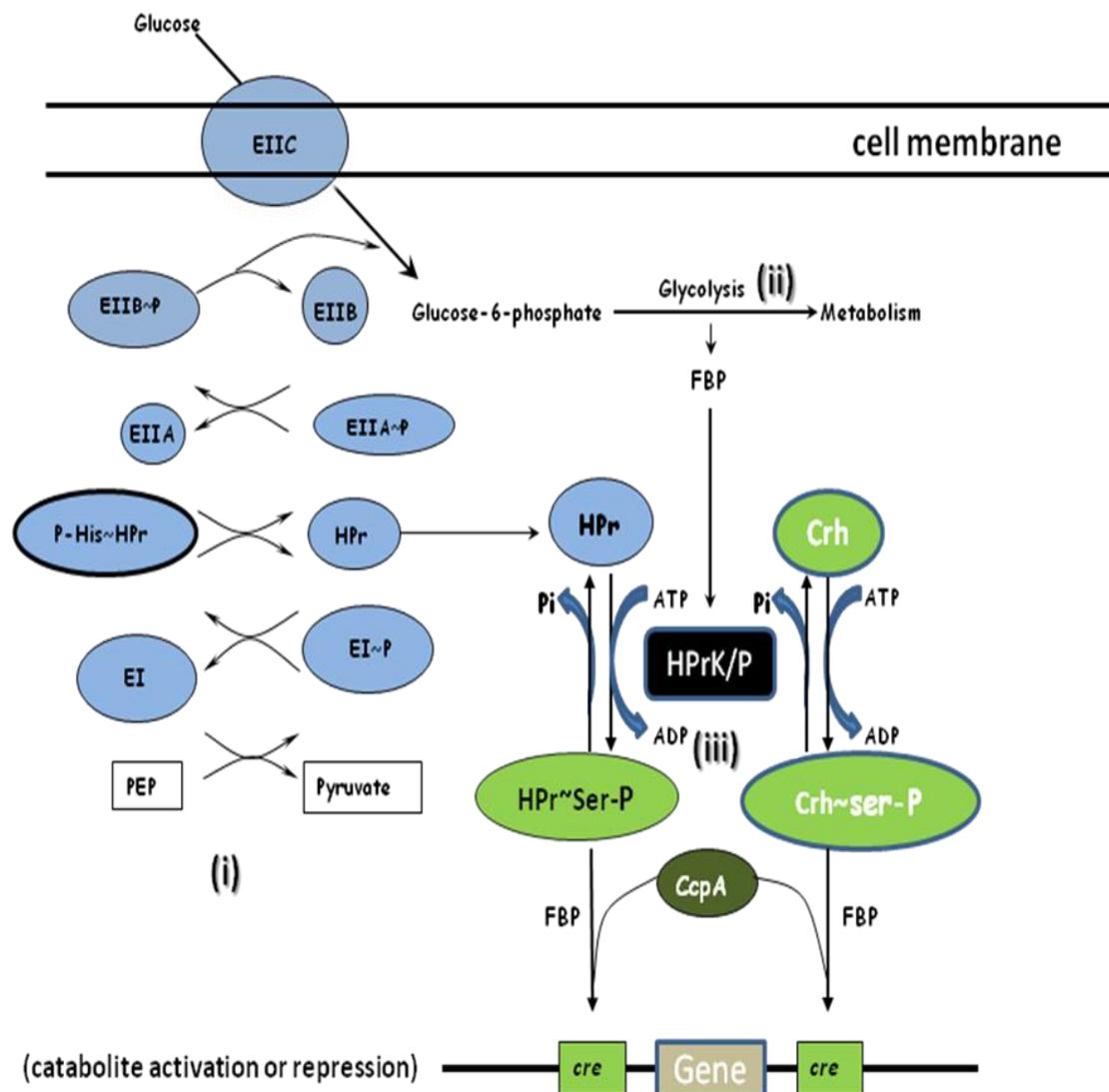


Figure 1.3: Model of catabolite activation or repression in *B. subtilis*.

Figure adapted from Lorca *et al.*, 2005. Proposed sensory transduction pathway by which exogenous glucose is believed to activate the CcpA transcription factor to promote catabolite repression (CR) or catabolite activation (CA) by binding to a catabolite responsive element (*cre*) in the control region of a target gene. (i) Proteins primarily involved are the glucose phosphotransferase system, including enzyme I, HPr, and IIBCA^{Glc}, (ii) glycolytic enzymes, (iii) the ATP-dependent fructose 1,6-bisphosphate (FBP)-dependent HPr/Crh kinase/phosphatase HPrK, and the two HPr and Crh protein targets of HPrK that independently bind to CcpA to activate it for binding to *cre* sites.

1.9.1 HPr and Crh

The general PTS protein HPr is the central processing unit in the regulation of carbon metabolism in the low GC Gram positive bacteria. HPr is a small protein consisting of about 90 amino acids only (9 to 10 kDa) (Postma *et al.*, 1993). HPr can be phosphorylated at either His-15 by enzyme I and PEP or Ser-46 by ATP. As a result, four different forms exist; HPr (dephosphorylated form), HPr phosphorylated at either His-15 (P-His-HPr) or Ser-46 (P-Ser-HPr), and the double phosphorylated HPr (Deutscher *et al.*, 2006). The active-site histidyl residue (His-15) in HPr which participates in the PTS is located in the N-terminal region, and the region around the histidyl residue is conserved in most bacteria (Deutscher *et al.*, 1986; Deutscher *et al.*, 2006). In the *B. subtilis* sequencing project, a gene was identified that encodes a protein that shares 45% sequence identity to HPr. It was therefore termed catabolite repression HPr-like protein (Crh) (Martin-Verstraeter *et al.*, 1999; Galinier and Haiech *et al.*, 1997). To date, Crh has been found only in bacillus spp. has the same regulatory site (Ser-46) as HPr which is involved in carbon catabolite repression. However, His-15 that is involved in sugar transport is replaced by glutamine in Crh, thus it cannot participate in sugar transport (Galinier *et al.*, 1997; Schumacher *et al.*, 2006).

Martin-Verstraete *et al.*, (1999) reported that HPr can completely substitute for Crh deletion in carbon catabolite repression of the *lev* operon which encodes for a fructose-specific PTS and levanase that catalyse metabolism of fructose polymers. However, Crh can only partly substitute for HPr deletion. However, specific function for Crh in *B. subtilis* has been described by Warner *et al.*, (2000). They reported that the Mg-citrate transporter (CitM) of *B. subtilis*, which is the main citrate uptake transporter during growth on citrate as the sole carbon source, is exclusively regulated by Crh during growth on gluconeogenic substrates such as succinate and glutamate (Warner *et al.*, 2000). CitM transports citrate or isocitrate into the cell in complex with divalent cations such as Mg, Zn, and Mn, hence it is termed the Mg-citrate transporter (Warner *et al.*, 2000).

1.9.2 HPrK/P

The metabolism of several carbon sources affects the activity of HPrK/P which is a bifunctional protein kinase/ phosphorylase, that reacts to changes of the concentration of FBP, ATP, pyrophosphate (PP_i), and inorganic phosphate (P_i) (Hogema, *et al.*, 1998). The phosphorylation of HPr at Ser-46 requires the HPrK/P protein kinase and ATP (Deutscher and Saier Jr, 1983). The substitution of Ser-46 with alanine, tyrosine, or aspartate in *B. subtilis* HPr prevented the FBP-stimulated ATP-dependent phosphorylation of HPr (Eisermann *et al.*, 1988). Also, mutation at this site affected the PEP-dependent phosphorylation at His-15 (Reizer *et al.*, 1989). FBP is needed to stimulate HPrK/P to phosphorylate either HPr or Crh in *B. subtilis* (Jault *et al.*, 2000). HPrK/P kinase activity was identified in several Gram positive bacteria more than 28 years ago. However, the *hprK* gene was discovered only during the *B. subtilis* sequencing project (Jault *et al.*, 2000; Deutscher and Saier Jr, 1983). When the concentrations of FBP and ATP increase, the kinase activity of HPrK/P is predominant; on the other hand, the phosphorylase activity becomes predominant when the concentration of P_i increases (Kravanja *et al.*, 1999; Jault *et al.*, 2000). HPrK/P can dephosphorylate P-Ser-HPr in presence of inorganic phosphate producing PP_i . The dephosphorylation is stimulated by P_i , and due to P-Ser-HPr dephosphorylation being a reversible mechanism, hence PP_i can substitute for ATP as the phosphoryl donor for HPr at high concentration (Figure 1.3) (Deutscher *et al.*, 1985; Nessler *et al.*, 2002; Kravanja *et al.*, 1999). Kravanja *et al.*, (1999) reported that the kinase activity of HPrK/P was inhibited at high concentration of P_i (50 mM) present under starvation conditions, however, a drop in the concentration of P_i (4 mM) was detected after adding glucose to the medium which led to restoration of the kinase activity of HPrK/P. Therefore, the HPrK/P protein plays a vital role in adjusting the level of P-His-HPr and P-Ser-HPr according to the nutritional state of the cell (Kravanja *et al.* 1999) (Figure 1.3).

Jault *et al.*, (2000) showed that when the FBP concentration increases, HPrK/P was strongly stimulated to phosphorylate HPr at Ser-46. Therefore, there is a correlation between FBP concentration and HPrK/P stimulation. Also, it has been shown that FBP (20 mM) stimulates the kinase activity of *B. subtilis* and *C. acetobutylicum* HPrK/P and the kinase activity is stronger when low concentrations of ATP (0.4 mM) are used (Jault *et al.*, 2000; Tangney *et al.*, 2003).

The intracellular concentrations of ATP, P_i, PP_i and FBP are determined by whether the bacteria are utilizing a favoured sugar or not (Deutscher *et al.*, 2006). *B. subtilis* cells grown on glucose were found to contain more FBP (14.1 mM) and PP_i (6 mM) than cells grown on less favourable substrate such as succinate which was found to contain only 1.4 mM for FBP and 1.2 mM for PP_i (Blencke *et al.*, 2003; Poncet *et al.*, 2004; Singh *et al.*, 2008). This explains why about 65% of the HPr in *B. subtilis* grown on glucose is present as P-Ser-HPr and the rest is present as HPr, P-His-HPr, and the doubly phosphorylated HPr (Singh *et al.*, 2008). However, when succinate is the sole carbon source, the majority of the protein is in the form of HPr, together with a small amount of P-His-HPr. Therefore, the two antagonistic activities of HPrK/P and the percentage of various form of HPr are regulated mainly by these metabolites (Moreno *et al.*, 2001; Singh *et al.*, 2008). HPrK/P kinase activity in the presence of less than 1 mM FBP was very low, however, higher FBP concentrations in the assay reaction mix led to an increase in the HPrK/P kinase activity (Jault *et al.*, 2000; Singh *et al.*, 2008). A few mM of FBP can overcome the inhibition of the kinase activity of HPrK/P exerted by P_i, hence, the major physiological function of FBP may be to stop the inhibition of HPrK/P by P_i (Kravanja *et al.* 1999; Brochu and Vadeboncoeur, 1999).

In the firmicutes, the *hprK* gene, which encodes the bifunctional enzyme HPr K/P, is frequently the first within an operon, and the second gene encodes a prolipoprotein diacylglycerol transferase (*lgt*) which may be involved in metabolism and membrane anchoring. This association of these two genes may indicate that the functions of the two enzymes are related but until now there is no experimental evidence to support the assumption (Deutscher *et al.*, 2006). However, in clostridia the *lgt* and *hprK* genes are located in different places in the genome. In the case of *C. acetobutylicum*, *hprK* is preceded by the *glpX* gene which appears to encode a fructose 1.6 bisphosphatase (Tangney *et al.*, 2003).

No Gram negative bacteria have been identified which possess the HPrK/P kinase, even though their HPr proteins possess Ser-46 and share 45% overall sequence identity to HPr's from the Gram positive bacteria (Deutscher and Engelmann, 1984).

However, the difference around the Ser-46 region is responsible for the failure of HPrK/P kinase from Gram positive organisms to phosphorylate the *E. coli* HPr (Deutscher and Saier, 1983; Reizer *et al.*, 1989). The replacement of the amino acids next to Ser-46 in the *Mycoplasma capricolum* HPr with the equivalent amino acids of *E. coli*, resulted in poor phosphorylation with ATP (Zhu *et al.*, 1998).

1.9.3 CcpA

CcpA is the catabolite control protein A which belongs to the GalR–LacI regulatory protein family as mentioned before (Deutscher *et al.*, 2006). It has a helix-turn-helix at the N-terminal which is responsible for DNA recognition and binding to a specific site called *cre* (catabolite responsive element) which is found in the promoter region of catabolic operons. The C-terminal of CcpA is engaged in the recognition of the phosphoserine and binding to P-Ser-HPr complex in the presences of FBP (Jones *et al.*, 1997 and Loll *et al.*, 2007). CcpA either by activation or repression is needed in order to maximize the efficiency of energy and carbon metabolism in the cell (Fujita, 2009). CcpA regulates the expression of around 400 genes which is about 10% of the *B. subtilis* genome (Nessler *et al.*, 2003). In *B. subtilis*, CcpA (global regulator) and XylR (specific regulator) are required for full glucose repression of the *xyl* operon, however, only CcpA is needed for full fructose repression (Dahl and Hillen, 1995). CcpA can act as a repressor or activator depending on the location of the *cis* acting catabolite responsive element (*cre*) which is a 14 bp palindromic sequence, reported by Weickert and Chambliss (1990). The *cre* consensus sequence is: TGWAANC*GNTNWCA, where the star signs (*) indicates the centre of the palindrome, W is adenine or thymine, and N is any base (Weickert and Chambliss, 1990). Therefore, repression occurs when the *cre* site is located within or downstream of the promoter, but, gene activation occurs when the *cre* site is located upstream of the promoter (Henkin, 1996). In *B. megaterium* and *B. subtilis*, the *ccpA* gene is expressed in a constitutive manner and this does not depend on the growth conditions. Instead, CcpA is controlled at a post translational level (Hueck *et al.*, 1995; Henkin, 1996). Also, CcpA controls the efficiency and optimizes the catabolism of glucose through repression of alternative metabolism pathways (Van der Voort *et al.*, 2008).

Moreno *et al.*, (2001) reported that CcpA can mediate gene repression or activation in a glucose independent manner. Computer analysis can identify the putative *cre* sites that are involved in catabolite repression but fails to identify *cre* sites that are involved in catabolite activation. CcpA binds to *cre* sites as a dimer in a tetrameric protein complex in the presence of either co-factors HPr-P-Ser or Crh-P-Ser (Moreno *et al.*, 2001).

The amounts of CCR elements such as HPrK/P, CcpA, HPr, and Crh in the cell are not affected by the presence of different carbon sources (Singh *et al.*, 2008). Fructose 1,6 bisphosphate is needed to stimulate the interaction between CcpA and HPr-P-Ser and is needed to enhance binding of the CcpA-HPr-P-Ser complex to *cre* sites (Henkin, 1996; Presecan-Siedel *et al.*, 1999; Antunes *et al.*, 2010). Kim *et al.*, (2005) reported that interaction between CcpA and RNA polymerase is needed in order to inhibit the transcription of the *amyE* gene. Actually, a *cre* site (centered at +4.5) is overlapping with the *amyE* promoter area, but CcpA would not prevent RNA polymerase from binding to the promoter (Kim *et al.*, 2005). Also CcpA cannot bind to *cre* sites in the absence of HPr-P-Ser (Kim *et al.*, 1998). Some operons have more than one *cre* site, for example, the *xyl* operon has 3 *cre* sites. The major one is centered at +130.5, and two supplementary sites are located at -35.5 and +219.5 (Kim *et al.*, 2005). When CcpA binds to the major *cre* site that is located downstream of the *xyl* operon transcriptional start site, the CcpA would not behave as roadblock (inhibit the elongation) instead, it inhibits the initiation of transcription. These results suggest that CcpA has a regulatory function other than simply binding to *cre* sites (Blencke *et al.*, 2003; Lorca *et al.*, 2005). Not all putative *cre* sites that have been identified have been shown to be functional (Van der Voort *et al.*, 2008; Miwa *et al.*, 2000). Actually, most of the genes that are activated by CcpA appear to lack a *cre* site and this indicates that the catabolite repression mechanism is unable to elicit catabolite activation by CcpA due to the absence of the *cre* sites (Blencke *et al.*, 2003; Lorca, *et al.*, 2005).

1.10 Elements of carbon catabolite repression in clostridia

The *ptsH* gene which encodes for HPr has been identified in the *C. acetobutylicum* genome sequence, and when cloned into an *E. coli ptsH* mutant, the clostridial gene complemented the mutation. The alignment of the HPr amino acid sequence of low GC Gram positive bacteria with HPr of *C. acetobutylicum* showed that the sequence around the His-15 site which participates in PTS transport was conserved (Tangney *et al.*, 2003), whereas the region around the Ser-46 site, which is involved in CCR and is well conserved in HPr proteins from other Gram-positive bacteria, is not conserved in *C. acetobutylicum* HPr (Tangney *et al.*, 2003). This protein can, however, be phosphorylated at Ser-46 by HPrK/P kinase in an ATP-dependent manner, and this activity is stimulated by FBP (Figure 1.4). Moreover, the clostridial HPrK/P kinase can phosphorylate the purified HPr from *B. subtilis* in the presence of FBP. Thus, it is suggested that the carbon catabolite repression model of low GC Gram positive bacteria also applies to *C. acetobutylicum* (Tangney *et al.*, 2003; Tangney and Mitchell, 2005).

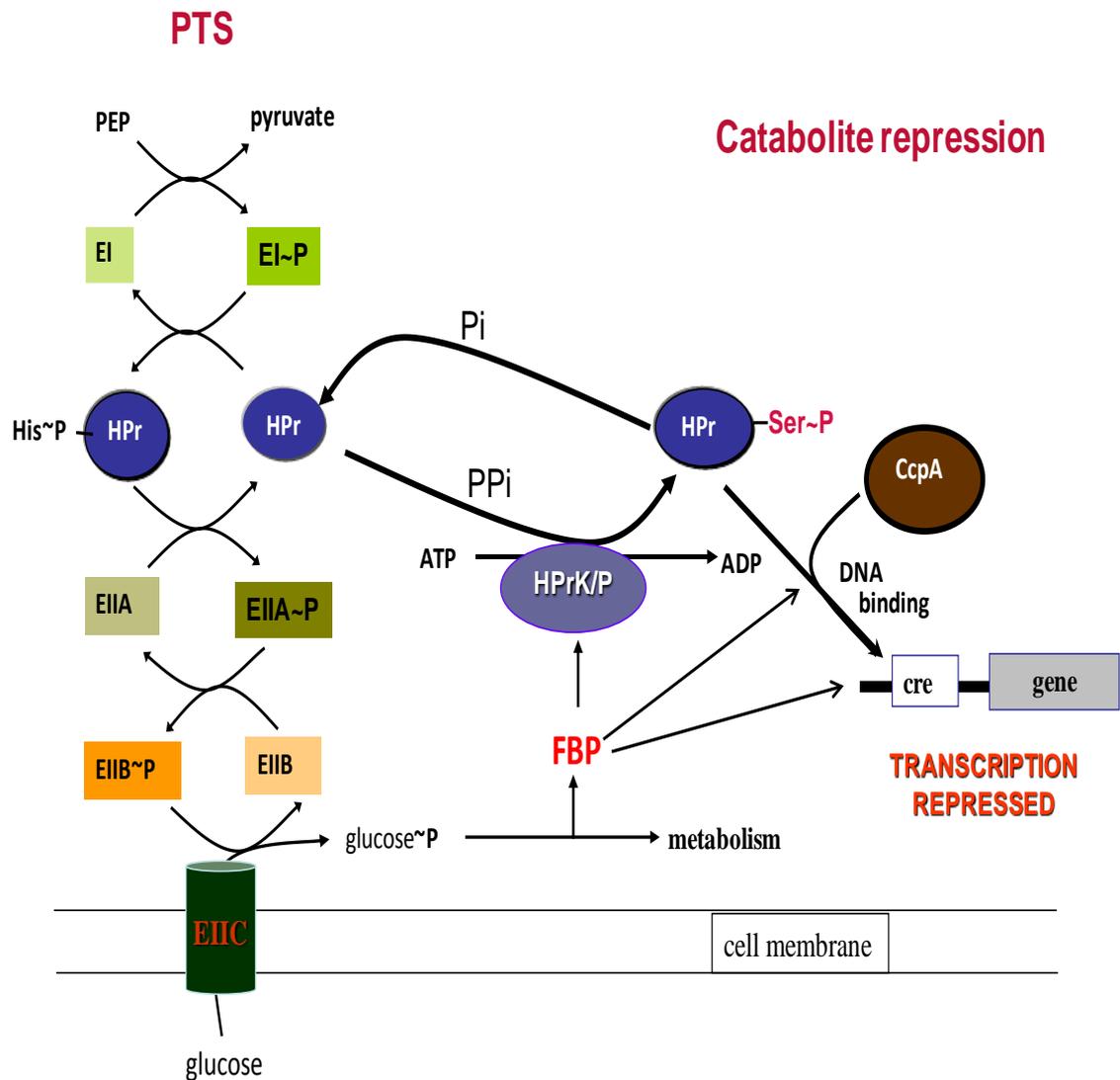


Figure 1.4: The model of the phosphotransferase system (PTS) mediated by the catabolite repression in *C. acetobutylicum* based on *B. subtilis* model.

The PTS is a transport system that has a vital role in CCR in low GC Gram positive microorganisms. When glucose is present, a high intracellular level of FBP will stimulate HPrK/P to phosphorylate the HPr at serine-46 site. This modified HPr forms a complex with the catabolite control protein A (CcpA). The complex binds to specific sequences called catabolite responsive elements (*cre*) and the transcription of a target gene is repressed. However, under unpreferred nutritional conditions such as starvation or grown on gluconeogenic substrate, the concentration of inorganic phosphate increases, and phosphorylase activity of HPrK/P becomes predominant. Figure adapted from Durre (2005).

Until now, HPrK/P kinase activity has been observed only in one species of *Clostridium* which is *C. acetobutylicum*, whereas, PTS activity has been reported in several species of *Clostridium* (Tangney *et al.*, 2003). A putative *ptsH* gene has been identified in the genome of three species of *Clostridium*, *C. perfringens*, *C. tetani*, and *C. beijerinckii*. Surprisingly, the *ptsH* gene is not followed by *ptsI* in the three species nor in *C. acetobutylicum* which means that *ptsH* and *ptsI* are not in an operon, which is unusual (Tangney and Mitchell, 2005). However, this situation has also been reported in some other bacteria such as *Mycoplasma capricolum* and *Streptomyces coelicolor* (Zhu *et al.*, 1993). As for *C. acetobutylicum*, analysis of the completed genomes of *C. perfringens*, *C. tetani*, and *C. beijerinckii* reveals that there is only one gene encoding for HPr and that there is no HPr-like protein Crh found in these bacteria (Tangney and Mitchell, 2005).

Bioinformatic analysis has shown a putative *C. acetobutylicum* HPrK/P and this protein has a signature sequence that is involved in the interaction with HPr and also a Walker A box which is involved in ATP binding. The *hprK* gene is located downstream of an open reading frame encoding a putative fructose 1,6-bisphosphatase (GlpX) (Tangney *et al.*, 2003). To date, this gene arrangement has not been observed in any other low GC Gram positive bacterium, since as mentioned previously the *hprK* gene is usually linked to the *lgt* gene encoding prolipoprotein diacylglyceryl transferase (Tangney *et al.*, 2003).

A putative gene, termed *regA*, encoding a CcpA-like protein in *C. saccharobutylicum* was cloned into a *B. subtilis* *ccpA* mutant and was found to complement the mutation (Davison *et al.*, 1995). This suggested that the RegA protein is actually CcpA. Genes encoding putative CcpA proteins have been identified in three other species of *Clostridium*; *C. acetobutylicum*, *C. tetani*, and *C. perfringens* and all these proteins contain an N-terminal DNA recognition binding domain (Tangney and Mitchell, 2005). Putative *cre* sites have been identified in the promoter region of nearly all the *pts* operons in species of *Clostridium* that have been studied. As an example, a *cre* site has been identified in both the sucrose (*scr*) and lactose (*lac*) operons of *C. acetobutylicum*. These operons are induced by sucrose and lactose respectively, and both are repressed by glucose (Tangney and Mitchell, 2000; Yu *et al.*, 2007). Ren *et al.* (2010) have recently characterized and then disrupted the *ccpA* gene in *C. acetobutylicum* ATCC 824 in order to generate a strain that is able to utilize a mixture of sugars such as xylose and glucose simultaneously.

However, the growth of the mutant strain (*ccpA*) was impaired with very low consumption of glucose and xylose compared to the wild type. Also, acid reassimilation by the mutant was inhibited which resulted in high acid accumulation. To solve the problem of the resulting low pH, calcium carbonate was added to the medium. As a result the mutant strain was able to grow on glucose and xylose. Transcriptional analysis revealed that the expression level of the *ctfAB* operon was very low in the *ccpA* mutant strain compare to the wild type (Ren *et al.*, 2010). This indicates that CcpA could be acting as a catabolite activator in *C. acetobutylicum* ATCC 824.

1.11 Fructose 1,6-bisphosphatase

Gluconeogenesis is a vital metabolic pathway that is undertaken to produce glucose when bacteria grow on non-carbohydrate substrates such as glycerol, amino acids, organic acids or fatty acids (Figure 1.5) (Brown *et al.*, 2009). Under gluconeogenic growth conditions, fructose 1,6-bisphosphatase (FBPase) is a key enzyme which produces fructose 6-phosphate and orthophosphate from the hydrolysis of fructose 1,6-bisphosphate. However, under glycolytic conditions, the reverse of the reaction is catalyzed by 6-phosphofructokinase to produce fructose 1,6-bisphosphate from fructose 6-phosphate at the expense of ATP (Brown *et al.*, 2009; Rittmann *et al.*, 2003; Jules *et al.*, 2009).

Five different classes of FBPase enzyme have been identified and classified based on their primary structure. The classes I, II, and III have been reported in bacteria, class IV in archaea, and class V in both domains of the thermophilic prokaryotes. However, only FBPase class I enzymes have been found in eukaryotes (Brown *et al.*, 2009; Hines *et al.*, 2006). These five classes of FBPases require divalent cations for activity (Mn^{2+} , Zn^{2+} , Mg^{2+}) (Brown, *et al.*, 2009; Donahue *et al.*, 2000). The classes I, III, IV, and V are all termed Fbpase, However, class II has been termed GlpX (Brown, *et al.*, 2009; Donahue *et al.*, 2000). Most microorganisms appear to have more than one FBPase enzyme, to date, no bacteria have the combination of class I and class III FBPases, however, the combinations of classes I and II and classes II and III were reported (Donahue *et al.*, 2000; Kuznetsova *et al.*, 2011). All the resolved crystal structures of class I, II, IV and V have the same α - β - α - β - α secondary structure which explains why they have the same function despite differences in primary structure (Brown, *et al.*, 2009; Hines *et al.*, 2006; Nishimasu *et al.*, 2004; Johnson *et al.*, 2001).

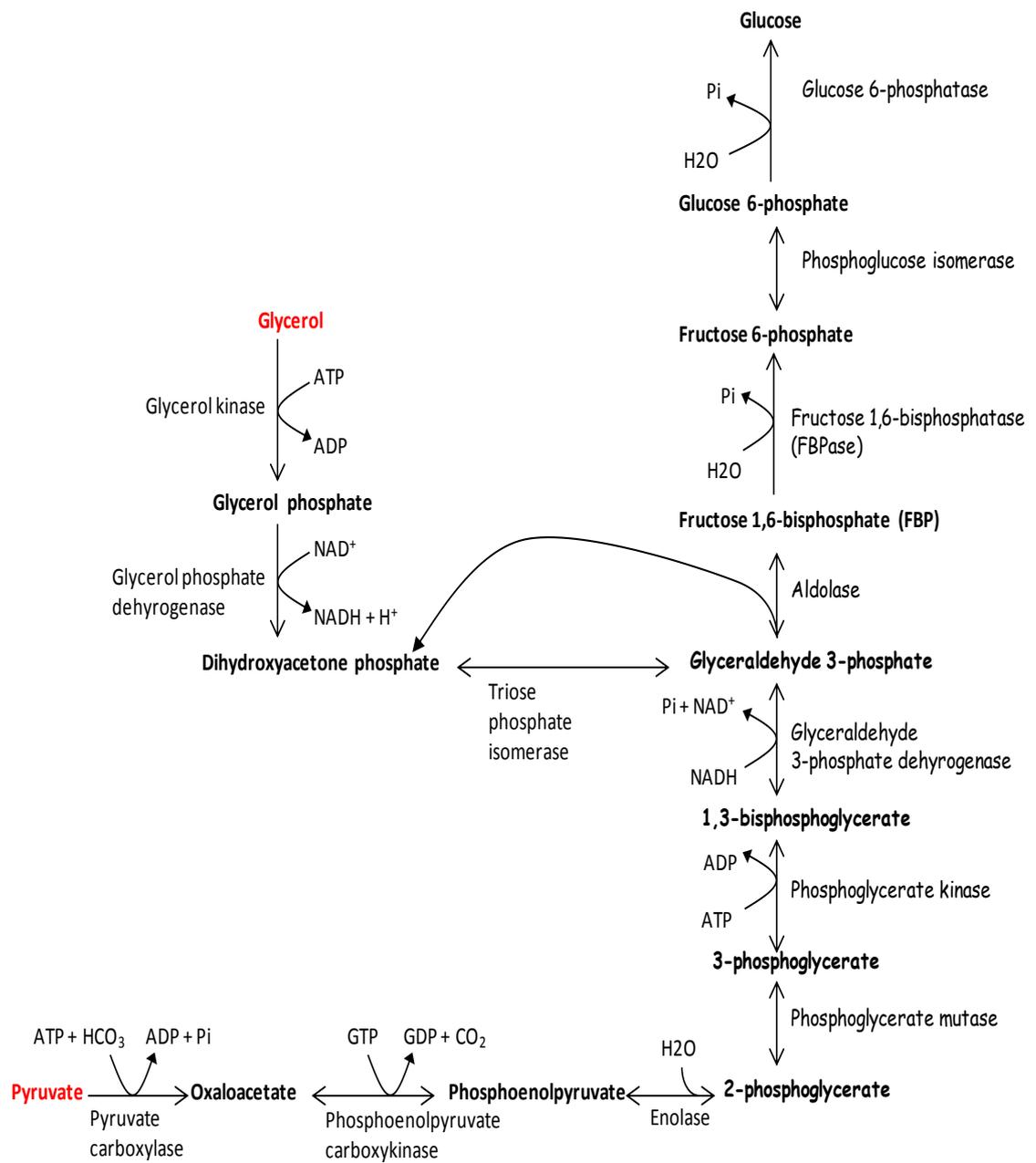


Figure 1.5: Pathway of gluconeogenesis. Adapted from Berg et al., (2002).

1.11.1 Fbpase class I

The Fbp class I has been identified in both eukaryotes and some prokaryotes. In mammals, the FBPase enzyme is a tetramer and can be inhibited by AMP and fructose 2,6-bisphosphate. In *E. coli*, the class I FBPase is strongly inhibited by AMP, however, no inhibition was detected after the addition of fructose 2,6-bisphosphate, and the enzyme is activated at least 3-fold by PEP and sulphate. The enzyme requires Mg^{2+} for activity (Brown *et al.*, 2009; Hines *et al.*, 2006). The FBPase protein in *E. coli* is the main FBPase which is involved in gluconeogenesis and this explains why a Δfbp strain cannot grow on gluconeogenesis substrates such as glycerol (Donahue *et al.*, 2000).

1.11.2 Fbpase class II (GlpX)

1.11.2.1 In *E. coli*

GlpX was first identified in *E. coli* as being encoded by one of the genes in the *glpFKX* operon involved in glycerol utilization. The *glpFKX* operon encodes for a glycerol facilitator (*glpF*), glycerol kinase (*glpK*), and an enzyme of unidentified function (*glpX*) (Donahue *et al.*, 2000). Brown *et al.*, (2009) reported that another gene also encodes for a FBPase class II enzyme (*yggF*). The activity of both GlpX and YggF enzymes were strongly inhibited by inorganic phosphate and the resolved 3 dimensional structure of the GlpX complex with phosphate confirmed that the inorganic phosphate molecule binds to the active site of the enzyme which explains why the enzyme was strongly inhibited. However, other effectors such as ADP, AMP and PEP each at 1 mM had no significant effect on the activity of either enzyme (Brown *et al.*, 2009). GlpX and YggF proteins are homodimers of subunits of 36 and 34.3 kDa respectively. Both GlpX and YggF enzymes show maximal activity at pH 7.5-8.0 and Mn^{2+} was the required divalent cation for the activity of both enzymes. However, GlpX has higher activity and affinity for FBP which leads to three times higher catalytic efficiency compared to YggF. In the GlpX structure, the active site is located in the large cavity between the two protein domains where many conserved residues are located (Brown *et al.*, 2009).

1.11.2.2 In *Corynebacterium glutamicum*

C. glutamicum is a Gram positive, high GC content bacterium which plays an important role in industrial amino acid production (Rittmann *et al.*, 2003; Kalinowski *et al.*, 2003). This bacterium has only a class II FBPase which is a homotetramer of subunits of 35.5 kDa and shares 44% identity with *E. coli* GlpX (Rittmann *et al.*, 2003). The optimal pH to assay the activity was 7.7 and Mn^{2+} was required for activity. FBPase activity was completely inhibited by 90 μ M AMP, and activity was reduced to 50% after adding 360 μ M PEP, 140 μ M LiCl, 1.6 mM phosphate and 1.2 mM fructose 1-phosphate. This gene was named *fbp*, even though the enzyme belongs to class II FBPases (Rittmann *et al.*, 2003). The *C. glutamicum* FBPase class II shows similar behaviour to the mammalian FBPase in terms of inhibition by AMP (Hines *et al.*, 2006; Rittmann *et al.*, 2003). This enzyme is an abundant protein and its synthesis does not depend on the growth conditions. An *fbp* mutant could not grow on gluconeogenic substrates which indicate that this is the major FBPase in *C. glutamicum* (Rittmann *et al.*, 2003). PTS EI, HPr and EII permeases have been identified for uptake of sugars such as fructose and glucose/mannose, but no HPrK/P kinase was identified and also the HPr lacks the conserved Ser-46 which indicates that the carbon catabolite repression is different in this high GC content Gram positive bacterium from that in the low GC content Gram positive bacteria such as *B. subtilis* and *C. acetobutylicum* (Rittmann *et al.*, 2003).

1.11.2.3 In *Mycobacterium tuberculosis*

There is only one major FBPase class II encoded by the *Rv1099c* gene in this bacterium, and the enzyme shares 43% identity to *E. coli* GlpX (Movahedzadeh *et al.*, 2004; Gutka *et al.*, 2011). The optimal pH for enzyme activity is 7.3 and Mg^{2+} was required for activity. AMP, citrate, Li^+ , PEP and ADP were tested as potential effectors. No inhibition was seen in the presence of any of these metabolites up to 1 mM except for Li^+ which inhibited the activity by 50% at 0.2 mM, and by more than 90% at 2.5 mM indicating that this enzyme belongs to the Li^+ -sensitive phosphatases (Movahedzadeh *et al.*, 2004; Gutka *et al.*, 2011).

1.11.2.4 In *Bacillus* species

GlpX in *B. subtilis* is encoded by the *ywjI* gene and the enzyme shares 54% identity with GlpX from *C. glutamicum* and requires Mn^{2+} for activity (Jules *et al.*, 2009). The GlpX activity was completely inhibited by addition of 1 mM PEP. A $\Delta ywjI$ strain was able to grow on gluconeogenic substrates such as glycerol and malate. GlpX in *B. subtilis* is a constitutive enzyme with expression which is not dependent on gluconeogenic or glycolytic conditions (Jules *et al.*, 2009). Van der Voort *et al.*, 2008 reported that *B. cereus* has two *glpX* genes. One of them (*ywjI*) was affected by the deletion of *ccpA* which led to relief of glucose repression and this might indicate that *ywjI* is involved in gluconeogenesis. However, the second *glpX* was not affected by deletion of *ccpA* and the presence of glucose (Van der Voort *et al.*, 2008).

1.11.3 Fbpase class III

The first and only FBPase belonging to class III has been reported in *B. subtilis* (Fujita *et al.*, 1998). The enzyme is encoded by a monocistronic *yjdE* gene and the expression of this gene depends on the growth phase rather than growth conditions (Fujita *et al.*, 1998). Like the *glpX* mutant, an *fbp* mutant was able to grow on gluconeogenic substrates such as glycerol indicating that each enzyme can compensate for the absence of the other (Jules *et al.*, 2009). The GlpX activity was completely inhibited by additional of 1 mM PEP, however Fbp activity was increased by 30-fold after adding 1 mM PEP which indicates that GlpX and Fbp are both regulated by PEP (Jules *et al.*, 2009). Like *glpX*, *fbp* was expressed at a constant level independent of the growth conditions (Jules *et al.*, 2009; Fujita *et al.*, 1998).

1.11.4 Fbpase class IV and V

The Fbpase class IV was first identified in *Methanocaldococcus jannaschii* which is anaerobic archaeon that produces methane as a by-product (Bult *et al.*, 1996). The MJ0109 protein can act as both inositol monophosphatase (IMPase) and FBPase (Stec *et al.*, 2000). The first class V FBPase was reported in the sulphur-reducing hyperthermophilic euryarchaeon *Thermococcus kodakaraensis* to be encoded by the *fbpTK* gene. The enzyme showed substrate specificity for FBP only, indicating that this is the true FBPase, unlike the bifunctional class IV IMPase/FBPase enzyme.

The expression of the *fbpTK* gene was strongly repressed under glycolytic conditions and a *fbpTK* mutant was unable to grow under gluconeogenic conditions even though *impTK* was constitutively expressed (Rashid *et al.*, 2002). The activity of FbpTK was detected without any divalent cations, but was enhanced 5-fold by the addition of 20 mM Mg^{2+} , and 4-fold by the addition of 1 mM Mn^{2+} . Also, Zn^{2+} increased the activity at 1 mM but Ca^{2+} and Ni^{2+} had no effect on the activity and the optimal pH was 8.0 (Rashid *et al.*, 2002). A recent study reported that class V FBPase is a bifunctional fructose 1,6 bisphosphate aldolase/phosphatase (FBP A/P) enzyme (Say and Fuchs, 2010; Fushinobu *et al.*, 2011). Fructose 1,6-bisphosphate aldolase (FBP aldolase) is an enzyme that catalyzes FBP under glycolytic growth conditions into dihydroxyacetone phosphate and glyceraldehyde 3-phosphate (aldo cleavage step). However, under gluconeogenic growth conditions the same enzyme catalyzes the reverse step which is aldol condensation of dihydroxyacetone phosphate and glyceraldehyde 3-phosphate to produce FBP (Siebers *et al.*, 2001; Berry and Marshall, 1993). Two different classes of FBP aldolase have been identified and classified based on their primary structure. The aldolase activity of FBP A/P only has irreversible step of aldol condensation of dihydroxyacetone phosphate and glyceraldehyde 3-phosphate to produce FBP. Bioinformatics have shown that analysis for all the archaeal genomes sequenced encode a bifunctional FBP A/P, although the mechanism by which FBP A/P exhibits its dual activity remains unclear (Say and Fuchs, 2010; Fushinobu *et al.*, 2011).

1.12 Project aim

Interest in the putative GlpX of *C. acetobutylicum* started when Tangney *et al.*, (2003) stated that the putative *hprK* gene, which is a vital gene in carbon catabolite repression, is followed by an open reading frame that might encode for a class II GlpX enzyme. This unique gene arrangement has not been reported in any other low GC Gram positive up to date, which might indicate that the putative GlpX enzyme plays a vital and unique role in regulating the FBP-dependent HPrK/P kinase activity (Figure 1.4) and thus in regulating carbon catabolite repression in *C. acetobutylicum* ATCC 824. For these reasons, GlpX needs to be characterized at the biochemical and genetic levels. In addition, the unique profile of the gene arrangement should be characterised in terms of gene expression. Moreover, bioinformatic analysis revealed a putative class III *fbp* gene in *C. acetobutylicum* which is an orthologue to the class III *fbp* gene in *B. subtilis*. Therefore, this gene and the encoded enzyme should be characterized and compared to the putative class II GlpX in order to know the physiological roles of both proteins. The overall aim of this project is to isolate by cloning, characterise and compare the physiological function of *glpX* and *fbp* genes in the solventogenic *C. acetobutylicum* ATCC 824, and also to investigate the significance of the *hprK-glpX* unique gene arrangement in the bacterium.

Chapter 2:

2 General materials and methods

2.1 Bacterial strains, growth conditions, primer and vector details

C. acetobutylicum ATCC 824 strain was obtained from Dr P. Soucaille of the Institut National des Sciences Appliquées-DGBA, Toulouse, France, and was originally obtained from American Type Culture Collection, Rockville, Maryland, USA. The *E. coli* strains JB103, JB108, JLD2402, JLD2403, JLD2404, and JLD2405 were obtained from Dr T. J. Larson of the Virginia Polytechnic Institute and State University, Blacksburg, USA. For TOPO cloning, the *E. coli* strain Mach1[™]-T1[®] was acquired from Invitrogen, Carlsbad, CA. For Gateway cloning the *E. coli* strains One Shot[®] OmniMAX[™] 2-T1R (Invitrogen, Carlsbad, CA) and Rosetta[™] 2(DE3) (Novagen Madison, WI) were used. The *E. coli* strains for Ligation Independent Cloning (LIC) were acquired from Novagen. Bacterial strains were stored at -70°C in the appropriate glycerol media. Details of the *E. coli* strains used in this study are listed in Table 2.1. Plasmids were stored at -20°C in nuclease free water (Qiagen, Valencia, CA). Details of the plasmids used in this study are listed in Table 2.2.

Table 2.1: Details of <i>E. coli</i> strains used in this study and their genotypes		
<i>E. coli</i> strain	Genotype	Source
TOP10F ⁻	F ⁻ { <i>lacIq</i> Tn10(TetR)} <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) Φ80 <i>lacZ</i> ΔM15 Δ <i>lacX74 recA1 araD139 Δ(ara-leu)7697</i> <i>galU galK rpsL endA1 nupG</i>	Invitrogen
Mach1 TM -T1 ^R	F ⁻ φ80(<i>lacZ</i>)ΔM15 Δ <i>lacX74 hsdR</i> (r _K ⁻ m _K ⁺) Δ <i>recA1398 endA1</i> <i>tonA</i>	Invitrogen
JB108	BL21 (DE3) <i>zjg920::Tn10 Δfbp287</i>	Donahue <i>et al.</i> , 2000
JLD2402	TL524 <i>glpX::spcr Δfbp287 zjg-920::Tn10</i>	Donahue <i>et al.</i> , 2000
JLD2403	TL524 <i>glpX::spcr fbp+ zjg-920::Tn10</i>	Donahue <i>et al.</i> , 2000
JLD2404	TL524 <i>glpX+ Δfbp287 zjg-920::Tn10</i>	Donahue <i>et al.</i> , 2000
JLD2405	TL524 <i>glpX+ fbp+ zjg-920::Tn10</i>	Donahue <i>et al.</i> , 2000
JB103	F ⁻ <i>araD139 Δ(lacZYA-argF)U169 rpsL150 deoC1 relA1</i> <i>ptsF25 flbB5301 thiA1zjg-920::Tn10 Δfbp287(λ lacIq tetR+</i> <i>spcR)</i>	Lutz & Bujard, 1997
One Shot® OmniMAX TM 2-T1R	F ⁻ [<i>proAB+ lacIq lacZ</i> ΔM15 <i>Tn10</i> (TetR) Δ(<i>ccdAB</i>)] <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>)φ80(<i>lacZ</i>)ΔM15Δ(<i>lacZYA-argF</i>) <i>U169 endA1 recA1 supE44 thi-1 gyrA96 relA1 tonA pand</i>	Invitrogen
Rosetta TM 2(DE3)	F ⁻ <i>ompT hsdSB</i> (rB ⁻ mB ⁻) <i>gal dcm</i> (DE3) pRARE23 (CamR)	Novagen
NovaBlue GigaSingles TM	F ⁻ [<i>proA⁺ B⁺ lacI^aZ</i> ΔM15::Tn10 (Tc ^R)] <i>endA1 hsdR17</i> (r _{K12} ⁻ m _{K12} ⁺) <i>supE44 thi-1 recA1 gyrA96 relA1 lac</i>	Novagen
BL21(DE3) pLysS	F ⁻ <i>ompT hsdSB</i> (rB ⁻ mB ⁻) <i>gal dcm</i> (DE3) pLysS (CamR)	Novagen

Plasmid + cloned gene	Code assigned to clones after transformation
pCR2.1-TOPO + <i>glpX</i>	pM ₁ B, pM ₁ W, and pM ₂ W
pDONR TM 221 + <i>glpX</i>	pMAX- <i>glpX</i> -1, pMAX- <i>glpX</i> -8, and pMAX- <i>glpX</i> -10
pDONR TM 221 + <i>fbp</i>	pMAX- <i>fbp</i> -1, pMAX- <i>fbp</i> -2
pET-60-DEST + <i>glpX</i>	pRosetta- <i>glpX</i> -6 and pRosetta- <i>glpX</i> -7
Pet-41 Ek/LIC + <i>glpX</i>	pNova- <i>glpX</i> -3 and pNova- <i>glpX</i> -4
Pet-41 Ek/LIC + <i>fbp</i>	pNova- <i>fbp</i> -3 and pNova- <i>fbp</i> -4

2.1.1 *E. coli* strain growth conditions

Luria-Bertani (LB) and M9 glycerol or glucose minimal medium were used to grow these strains. The compositions of these media are shown in Table 2.3. Media were autoclaved at 121°C and ampicillin or tetracycline or kanamycin was added to the media as required after cooling to 55°C (Table 2.4). All the strains were grown at 37°C.

M9 minimal medium: 0.4 % (w/v) glycerol (per litre)	
M9 salts (see below)	100 ml
Glycerol	4 g
0.1M MgSO ₄	10 ml
0.01M CaCl ₂	10 ml
dH ₂ O	to 1000 ml
M9 Minimal medium: 0.2 % (w/v) glucose (per litre)	
M9 salts	100 ml
Glucose	2 g
0.1M MgSO ₄	10 ml
0.01M CaCl ₂	10 ml
dH ₂ O	to 1000 ml
M9 Salts (per litre)	
Na ₂ HPO ₄	60 g
KH ₂ PO ₄	30 g
NaCl	5 g
NH ₄ Cl	10 g
dH ₂ O	to 1000 ml

Luria-Bertani (LB) medium (per litre)	
Tryptone	10 g
Yeast extract	5 g
NaCl	5 g

Table 2.4: List of antibiotics		
Full name	Abbreviation	Final concentration
Ampicillin	+Amp	(100 µg/ml)
Tetracycline	+Tet	(10 µg/ml)
Kanamycin	+Kan	(30 µg/ml)

2.1.2 Primer sequences

The sequences of the forward and reverse primers used in this study to amplify specific fragments of DNA are presented in Table 2.5. All primers were obtained from Eurofins MWG Operon UK, except for RT-PCR control primers which were obtained from Novagen. The extra overhang sequences for Gateway and LIC cloning primers are underlined.

Table 2.5: Primer details	
Designated name	Sequence (5' – 3')
TOPOglpX	Fwd: ACTTTAGGACACTATAGCGC
	Rev: TACCTGCATTTACTATCCCC
TOPOfbp	Fwd: ACTTTAGGACACTATAGCGC
	Rev: TACCTGCATTTACTATCCCC
GATEglpX	Fwd: <u>GGGGACAAGTTTGTACAAAAAAGCAGGCTTC</u> ATGTTTGATAATGATATATATCC
	Rev: <u>GGGGACCACTTTGTACAAGAAAGCTGGGTTT</u> TCTACCACTAATTTACTTTT
GATEfbp	Fwd: <u>GGGGACAAGTTTGTACAAAAAAGCAGGCTTC</u> ATGCTATTAGAAAGTAAC
	Rev: <u>GGGGACCACTTTGTACAAGAAAGCTGGGTA</u> ACTTTTTACAAATTCCTT
LICfbp	Fwd: <u>GACGACGACAAGATGCTATTAGAAAGTAACACCA</u>
	Rev: <u>GAGGAGAAGCCCGGTACTTTTTACAAATTCCTTAATAAG</u>
LICglpX	Fwd: <u>GACGACGACAAGATGTTT</u> GATAATGATATATCC
	Rev: <u>GAGGAGAAGCCCGGTTT</u> TCTACCACTAATTTACTTTT
RNAhprK	Fwd: TATACGTCCAGGGAGGAATG
	Rev: CCTTGGAATAACCTTCGGC
RNAglpX	Fwd: TGCAGCTACAGCAATAACAG
	Rev: CAAACTGTAGTCCTGGTCTG
RT-PCR control	Fwd: TCCACCACCCTGTTGCTGTA
	Rev: ACCACAGTCCATGCCATCAC
T7 promoter	Fwd: TAATACGACTCACTATAGGG
	-

2.1.3 Growth conditions for *C. acetobutylicum* ATCC 824

C. acetobutylicum was maintained as a spore suspension in sterile water at 4°C. Two ml of spore suspension was transferred into a sterile test tube and heated-shocked at 80°C for 10 minutes and transferred into 20 ml of Reinforced Clostridial Medium (RCM) which was prepared according to the manufacturer's instructions (OXOID LTD, England). To prepare the starter culture, the inoculated RCM was then incubated in an anaerobic cabinet at 37°C overnight under N₂:H₂:CO₂ (80:10:10) mixed gas atmosphere. Five ml of starter culture was then transferred into 100 ml of Clostridial Basal Medium (CBM) [casein hydrolysate, 4 g; MgSO₄·7H₂O, 0.2 g; FeSO₄·7H₂O, 10 mg; MnSO₄·4H₂O, 10 mg; α-aminobenzoic acid, 1 mg; thiamine-HCl, 1 mg; d-biotin, 2 µg; KH₂PO₄, 0.5 g; 5 mM of glucose in 1 L of dH₂O] and incubated at 37°C under the same anaerobic conditions. The cells were harvested by transferring them into 1.5 ml microcentrifuge tube, centrifuged at 18000 xg in room temperature for 10 minutes and the supernatant was discarded and the cell pellet was kept at -70°C.

2.2 Chromosomal DNA extraction from *C. acetobutylicum* ATCC 824

The chromosomal DNA was extracted by using a DNA Isolation Kit for Gram-positive bacteria (Gentra[®], Cat# D-6000A) with some modifications in the protocol. Cell Suspension Solution (300 µl) was added to the *C. acetobutylicum* cell pellet that was prepared as described above, and then the cell pellet was pipetted up and down until the cells were suspended. Lytic Enzyme Solution (1.5 µl) was added to the mix and vortexed for 30 seconds, and then incubated on the shaking incubator at 100 rpm for 30 minutes at 37°C. The reaction mix was centrifuged again at 18000 xg for 1 minute and the supernatant was discarded.

Cell Lysis Solution (300 µl) was added to the cell pellet, to break open the cells and release the DNA, and then gently pipetted up and down to lyse the cells, 1.5 µl of RNase A Solution was added to the lysed cells, and incubated in a shaking incubator at 100 rpm for 1 hour at 37°C. Protein Precipitation Solution (100 µl) was added to the reaction mix and vortexed for 30 seconds and centrifuged at 18000 xg for 3 minutes, a white pellet containing the precipitated proteins was then formed. The supernatant containing the DNA was transferred into a sterile 1.5 ml microcentrifuge tube containing 300 µl of 100% isopropanol, vortexed for 30 seconds, and then centrifuged at 18000 xg for 1 minute.

The supernatant was discarded and the pellet which contained the genomic DNA was resuspended again in 300 µl of 70% (v/v) ethanol, vortexed for 30 seconds, and then centrifuged at 18000 xg for 1 minute. The supernatant was discarded and the pellet was hydrated in 60 µl of DNA Hydration solution, then incubated for 1 hour at 65°C and the purified genomic DNA was stored at -20°C.

2.3 Polymerase Chain Reaction (PCR) for TOPO™ TA cloning

C. acetobutylicum ATCC 824 chromosomal DNA was purified to be used as DNA template in the PCR reaction. The total volume of the PCR reaction was 50 µl which consisted of; 1 µl of chromosomal DNA, 1 µl of each of the TOPO forward and reverse primers for both *glpX* and *fbp* genes, 10 µl of 5 × PCR buffer [100 Mm of dNTPs, 20 µl; 10 × Buffer (Stratagene®, Cat# 600400), 230 µl; dH₂O, 250 µl] and 36.5 µl of deionised water. After the first denaturation stage of 95°C for 5 minutes, 0.5 µl of Easy-A high fidelity PCR cloning enzyme (Stratagene®, Cat# 600400) was added to the PCR mixture. The second stage was 30 cycles of 1 minute at 95°C, 1 minute at the appropriate annealing temperature and 3 minutes at 72°C. The final stage was 10 minutes at 72°C to complete primer extension.

2.4 Gel electrophoresis

To enable visualisation of nucleic acids, 5 µl of PCR product were mixed with 2 µl of DNA loading buffer (Bioline, Cat# BIO-37045). Then 7 µl of the resultant mixture was loaded onto a 1% (w/v) agarose gel (containing 1 µl of Ethidium Bromide at 10 mg/ml) along with Hyperladder I (Bioline), as size standards and run at 60 V for 1 hour, and the DNA in the gels were visualised by UV light.

2.5 TOPO™ cloning procedure

A TOPO TA Cloning® Kit (Invitrogen, Cat# K4500-01) was used to perform the cloning of both *glpX* and *fbp* genes. *Taq* polymerase has a non-template dependent terminal transferase activity that adds a single deoxyadenosine (A) to the 3' ends of the PCR product. The pCR2.1-TOPO vector possesses single overhanging 3' deoxythymidine (T) residues. This enables cloning of the PCR product into the pCR2.1-TOPO vector (appendix 7.4).

Fresh PCR products (see section 2.3) (3 μ l) were mixed with 1 μ l water, 1 μ l Salt Solution and 1 μ l pCR2.1-TOPO vector (all provided from the TOPO kit). The TOPO cloning reaction was incubated for 15 minutes at room temperature with the protocol carried out as per manufacturer's instructions.

2.6 Transformation of chemically competent *E. coli*

A vial of Mach1TM-T1^R competent cells (Table 2.1) stored at -70°C was thawed on ice prior to use. After thawing, 2 μ l of TOPO cloning reaction was added into the vial and incubated for 10 minutes on ice and then the cells were heat-shocked for 30 seconds at 42°C. Immediately the vial was placed on ice for 2 minutes and then a 250 μ l of S.O.C medium (provided from the TOPO kit) was added at room temperature. The vial was incubated horizontally at 37°C on a shaking incubator at 170 rpm for 1 hour. LB plates (Table 2.3) supplemented with 40 μ l/plate of 40 μ g/ml X-gal and (+Amp) were incubated for 1 hour at 37°C to pre-warm the plates. Aliquots from the transformation mixture (50 μ l or 100 μ l) were then spread onto the surface of the agar. The plates were incubated overnight at 37°C.

2.7 Screening of clones

2.7.1 PCR analysis

White colonies contain an inserted gene which disrupts the *lacZ* gene fragment on the pCR2.1-TOPO vector. Therefore the white colonies were tested for harbouring a plasmid containing the desired PCR insert. This was done by transferring a colony, using a sterile toothpick, to 20 μ l of dH₂O in a 1.5 ml microcentrifuge tube and heating to 100°C in a heat block for 10 minutes. The mixture was transferred to ice for 10 minutes to cool down and then centrifuged at 18000 \times g for 2 minutes. After the centrifugation, the supernatant was used as the DNA template for a PCR reaction. The PCR screening was carried out using the same primers that were used in the PCR amplification with 1 μ l of the supernatant. The reaction mix and conditions were the same as previously described in Section 2.3. For determination of the orientation of the insert, the T7 promoter primer (Table 2.5) was used in conjunction with one of the TOPOglpX primers. PCR products were visualised following electrophoresis in 1% (w/v) agarose gels in 1 \times Tris-acetate-EDTA (TAE) buffer at 60 V.

2.7.2 Plasmid purification

Plasmids were purified using the Qiagen Plasmid Midi Kit (Cat# 12145) as described in the manufacturer's protocol. A single colony of the cells containing the plasmid of interest was transferred to 10 ml of LB Broth supplemented with appropriate antibiotic and incubated overnight with shaking at 37°C. Cell suspension (1 ml) was transferred into 100 ml of LB Broth supplemented with appropriate antibiotic and incubated again overnight with shaking at 37°C. The cell suspension (50 ml) was then pelleted by centrifuging at 2700 x g for 10 minutes at 4°C. The pellet was resuspended in 4 ml of Buffer P1 and vigorously shaken for 1 minute. Lysis Buffer (4 ml, Buffer P2) was added to the resuspended pellet and mixed by inverting 6 times and incubated at room temperature for 5 minutes. Then, 4 ml of chilled Buffer P3 was added to the mixture to stop the lysis reaction and incubated on ice for 15 minutes. The reaction mixture was centrifuged at 20000 xg for 30 minutes at 4°C, the supernatant was transferred to another tube and centrifuged at 2700 xg for 10 minutes at 4°C and the supernatant was kept and the pellet discarded. The purification column (QIAGEN-tip 100) was equilibrated by adding 4 ml of Buffer QBT to the purification column and allowing it to empty by gravity flow.

The supernatant was applied to the purification column and allowed to move through the column by gravity flow and then the column was washed twice by applying 10 ml of Buffer QC. The plasmid was eluted by adding 5 ml of Buffer QF and collected in a 15 ml tube. Isopropanol (3.5 ml) was added and centrifuged at 2700 xg for 1 hour at 4°C to precipitate the plasmid. The plasmid pellet was washed by the addition of 2 ml of 70% (v/v) ethanol and centrifuged at 2700 xg for 30 minutes, after which the supernatant was discarded. The pellet was resuspended in 50 µl of deionized water. The concentration of plasmid DNA was quantified by comparing it to concentration of Hyperladder I (Bioline).

2.7.3 DNA Sequencing

All DNA sequencing was conducted commercially by Beckman Coulter Genomics. Plasmid DNA was sent in 20 µl of nuclease free dH₂O at a concentration of 50 ng/µl.

2.7.4 Restriction analysis

Colonies thought to possess a PCR insert, were grown in 100 ml LB Broth (+Amp). The Qiagen Plasmid Midi Kit was used to purify the recombinant plasmid. After purification, 5 µl of the sample was electrophoresed in a 1% (w/v) agarose gel in 1× TAE buffer at 60 V for 1 hour, to check whether the plasmid of interest was purified or not. Separate digestion reactions using *EcoRI* (Fermentas, Cat# ER0271, cutting sites: 5'...G↓A A T T C...3' and 3'...C T T A A↓G...5') and *HindIII* (Fermentas, Cat# ER0501, cutting sites: 5'...A↓A G C T T...3' and 3'...T T C G A↓A...5') restriction enzymes were performed to determine the orientation of the insert. The reaction mixture (Table 2.6) was incubated at 37°C overnight before being visualised by electrophoresis. The restriction digest sites of *HindIII* and *EcoRI* in the vector are shown in (Appendix 7.4).

Table 2.6: Composition of the restriction digest mixture	
Substance	Volume (µl)
10 × buffer of <i>EcoRI</i> or <i>HindIII</i>	2
Plasmid DNA	5
Restriction digest enzyme	1
dH ₂ O	12
Total volume	20

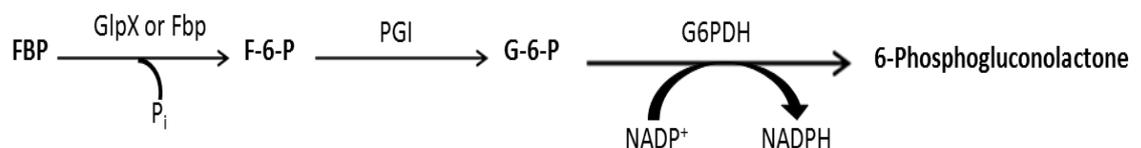
2.7.5 Transformation of *E. coli* mutant strains

The *E. coli* strains JB103, JB108, JLD2402, JLD2403, JLD2404, and JLD2405 were grown on Nutrient Agar. Single colonies of these strains were transferred to 10 ml of LB supplemented with tetracycline (+Tet), and incubated on the shaking incubator at 170 rpm overnight at 37°C. After that, 1 ml of overnight culture was transferred into 100 ml of LB (+Tet) and incubated at 37°C (170 rpm) for 4 hours until the OD₆₀₀ reached 0.6. The culture (50 ml) was transferred into a 50 ml polypropylene tube and then incubated on ice for 10 minutes. The cells were centrifuged at 1700 xg at 4°C for 10 minutes. The supernatant was discarded and the pellet was resuspended in 10 ml of ice-cold 0.1 M CaCl₂ and again incubated on ice for 10 minutes. The suspension was centrifuged again at 1700 xg at 4°C for 10 minutes and then the cells were resuspended in 2 ml of the same ice-cold 0.1M CaCl₂.

Purified recombinant of interest plasmid (2.5-5 μ l) was added to 200 μ l of cell suspension using a chilled pipette tip and then mixed gently and incubated on ice for 10 minutes. After that, the reaction mixture was heat shocked for 2 minutes at 42°C in a water bath and placed on ice again for 5 minutes. An 800 μ l aliquot of LB broth was added to the mixture and incubated on the shaking incubator (100 rpm) at 37°C for 1 hour. After the incubation, 100 μ l of the mixture were spread onto LB plates (+Amp and +Tet) and incubated again at 37°C overnight. The colonies which formed were given a reference number and transferred onto M9 glucose (+Amp and +Tet), LB (+Amp and +Tet) and M9 glycerol (+Amp and +Tet) plates, incubated overnight at 37°C and then PCR screened the same way as described in section 2.7.1.

2.7.6 Fructose 1,6-bisphosphatase assay

A colony that contains a gene of interest was transferred using a sterile toothpick into 10 ml LB broth (+Kan) and incubated overnight with shaking at 200 rpm at 37°C. The overnight culture (1 ml) was transferred into 500 ml of Overnight Express™ Instant LB medium and incubated overnight with shaking at 200 rpm at 37°C [glycerol, 10 g; Overnight Express™ Instant LB medium, 45 g in 1 L of dH₂O] (Novagen, Cat# 71757). The 500 ml culture was centrifuged at 1700 xg for 10 minutes at 4°C. The supernatant was discarded and the cell pellet was suspended in ice-cold 20 mM Tris-Cl (8 pH) and centrifuged again similar to the previous conditions. The suspension was transferred into a 50 ml polypropylene tube and centrifuged at 1700 xg at 4°C for 10 minutes. The weight of the cell pellet was recorded before it was stored at -20°C. The pellet was thawed on ice and resuspended in 50 mM Tris-Cl (pH8) containing 1 mM dithiothreitol (DTT) and 5 mM MgCl₂ to reach a concentration of 3 ml/g pellet. The cells were broken twice using a French Pressure Cell (SLM instruments, INC.) then centrifuged at 2700 xg for 10 minutes at 4°C. Finally, the supernatant was transferred into 15 ml polypropylene tube and stored at -20°C. The cell extracts were assayed spectrophotometrically, by using a Beckman Coulter DU 800 spectrophotometer at 340 nm at 37°C, for FBPase activity which involved measuring the production of fructose 6-phosphate coupled to the reduction of NADP⁺ in the presence of glucose 6-phosphate dehydrogenase (G6PDH) (Sigma, 4.6 mg protein/ml; 715 units/mg protein, Cat# G8404-1KU) and phosphoglucose isomerase (PGI) (Sigma, 10.1 mg protein/ml; 547 units/ mg protein, Cat# P5381-1KU) as follows:



The formation of NADPH in the final step of the reaction was measured by the increase of the absorbance at 340 nm and FBPase specific activity is calculated using the molar extinction coefficient of NADPH which is 6.22×10^3 nmoles (Jules *et al.*, 2009; Fraenkel *et al.*, 1966; Babul *et al.*, 1983) as follows:

$$\text{Specific activity (units/mg protein)} = \frac{\Delta\text{OD}}{\text{time (min)} \times (6.22 \times 10^3) \times \text{protein conc (mg/ml)}}$$

The activity of the coupling enzymes glucose 6-phosphate dehydrogenase and phosphoglucose isomerase were assayed in same way but instead of the FBP, 20 μ l of F-6-P (500 mM) was added to the reaction mixture to check whether the enzymes were able to form NADPH. The composition of the assay reaction mixture is shown in Table 2.7.

Substance	Volume (μl)
1 M Tris-Cl pH8	50
1 M MgCl ₂	10
200 mM DTT	5
0.1 mM EDTA	4
50 mM Fbp	4
50 mM NADP	4
Coupling enzymes	1 of each
Cell extract	50
dH ₂ O	871
Total reaction volume	1000

2.7.6.1 Microbuiuret protien assay

A microbuiuret assay was used to measure the concentration of the protein present in the crude extract. A serial standard was made using BSA as shown in Table 2.8, in 1 ml tube with the addition of 240 μ l of 40% (w/v) NaOH. After that, 680 μ l of water and 63 μ l of 1% (w/v) CuSO₄ .5 H₂O were added to the mixture and then these were incubated in the 37°C water bath for 15 minutes. The standard samples were read at 310 and 390 nm to calculate the concentration of the standards for creating a calibration curve. The concentration of protein in 10 μ l of each crude extract sample was measured in same way but instead of BSA, which was used as standard, crude extract was used. The protein concentration of crude extract was calculated by interpolation from the standards calibration curve (appendix 7.9).

Table 2.8: Different amounts of BSA and dH₂O that were needed in the 1 ml assay mixture.

BSA (10 mg/ml) μl	dH₂O μl
20	0
15	5
10	10
5	15
0	20

2.8 Cloning of the *glpX* and *fbp* genes in the Gateway expression system

The Gateway expression system (Invitrogen, Cat# 12535-029) is based on two cycles of bacteriophage lambda in *E. coli*. The bacteriophage lambda can either enter, a lytic cycle which results in viral replication and destruction of the infected cell, or a lysogenic cycle in which it integrates itself into the host genomic DNA at a specific site. Also in the lysogenic cycle, the phage genome can either be transmitted into daughter cells or it can be induced to go back again to the lytic cycle from which the phage genome can be excised from the host genomic DNA. This system consists of two reactions. The first reaction is BP reaction, which is the integration process (lysogenic pathway) that facilitates recombination of *attB*-flanked PCR product with pDONRTM vector to produce an *attL*-containing entry clone, and this reaction or pathway is catalyzed by bacteriophage lambda Integrase (Int) and *E. coli* Integration Host Factor (IHF) proteins (BP ClonaseTM II enzyme mix). The second reaction is the LR reaction which is the excision process (lytic pathway) that facilitates recombination of an entry clone with a destination vector, and this reaction or pathway is catalyzed by bacteriophage lambda Excisionase (Xis), Integrase (Int) and *E. coli* Integration Host Factor (IHF) proteins.

2.8.1 BP cloning reaction

Primers were designed to amplify *C. acetobutylicum glpX* and *fbp* genes flanked by *attB1* and *attB2* overhangs (Table 2.5), that can recombine with the donor vector (pDONRTM221) at *attP1* and *attP2* sites to produce entry clones with two sites, *attL1* and *attL2*. Both *glpX* and *fbp* genes were amplified to produce *attB*-PCR products for the BP reaction as previously described in section 2.3. The *attB*-PCR product (1 µl) was mixed with 1 µl of the donor vector (pDONRTM221) and 6 µl Tris-EDTA (TE) buffer (pH 8.0) was added to the mixture, both provided with the kit. The reaction was performed in a 1.5 ml microcentrifuge tube in room temperature. The BP ClonaseTM II enzyme mix was removed from the freezer and left on ice 2 minutes for thawing, and then mixed by vortexing twice for 2 seconds, 2 µl of BP ClonaseTM II enzyme mix was added to the reaction mixture and mixed by vortexing twice for 2 seconds. The reaction mixture was incubated for 1 hour at room temperature, and after that 1 µl of proteinase K solution was added and incubated again for 10 minutes at 37°C.

2.8.2 Transformation of chemically competent cells

One vial of One Shot[®] OmniMAX[™] 2-T1R Chemically Competent *E. coli* (Invitrogen, Cat# 12535-029) was left on ice to thaw for 5 minutes and then 1 μ l of BP reaction mixture was added to the One Shot[®] OmniMAX[™] 2-T1R cells and mixed gently. The vial was incubated on ice for 30 minutes and then heat shocked in a water bath for 30 seconds at 42°C, and placed again on ice for 2 minutes. S.O.C medium (250 μ l) was added to the vial at room temperature, and incubated for 1 hour at 37°C, shaking horizontally at 200 rpm. After that, 20 μ l of the transformation mixture was spread onto LB plates (+Kan) and left in the incubator overnight at 37°C.

2.8.3 LR reaction

Plasmids of *fbp* and *glpX* were purified as described in section 2.7.2 using the Qiagen plasmid purification midi kit, and sent for sequencing to Eurofins MWG Operon. The following reaction was performed in a 1.5 ml microcentrifuge tube at room temperature, 1 μ l of 150 ng purified plasmid (entry clone) was mixed with 1 μ l of pET-60-DEST vector, and then 6 μ l of TE buffer, pH 8.0 was added to the mixture. The LR Clonase[™] II enzyme mix was removed from the freezer and left on ice for 2 minutes for thawing, and then mixed by vortexing twice for 2 seconds, 2 μ l of LR Clonase[™] II enzyme mix was added to the reaction mixture and mixed by vortexing twice for 2 seconds. The reaction mixture was incubated for 1 hour at room temperature, after that 1 μ l of proteinase K solution was added and then incubated again for 10 minutes at 37°C.

2.8.4 Transformation into expression host

One vial of Rosetta[™] 2(DE3) Competent Cells (Novagen, Cat# 71397) was left on ice to thaw for 5 minutes and then 1 μ l of LR reaction mixture was added to the Rosetta[™] 2(DE3) Competent Cells and mixed gently. The transformation was performed as described in section 2.8.2.

2.9 Cloning of *glpX* and *fbp* by Ligation-Independent Cloning (LIC)

Novagen ligation-independent cloning (LIC) allows directional cloning of the insert with no need for ligation or restriction digest reaction. This system uses the activity of the bacteriophage T4 DNA polymerase enzyme to make specific 15-base single stranded overhangs in the Ek/LIC vector. Complementary overhangs in the insert must be present by adding suitable 5' extension into the primers (Table 2.5). To create specific vector compatible overhangs, purified insert DNA must be treated with T4 DNA polymerase in the presence dATP. Transformation into competent *E. coli* was carried out after cloning.

2.9.1 LIC cloning reaction

Primers LIC*glpX* and LIC*fbp* were designed for amplification of *glpX* and *fbp* genes flanked by specific 13-bases at the ends which were complementary to the overhangs in the Ek/LIC vector (Table 2.5). Hot Start DNA Polymerase (KOD) (Novagen, Cat# 71086-5) was used for amplification of both *glpX* and *fbp* genes from the purified *C. acetobutylicum* chromosomal DNA. For reaction set up and cycle conditions see Table 2.9 and Table 2.10 respectively.

Component	Volume (µl)
10X Buffer for KOD Hot Start DNA Polymerase	5
25 mM MgSO ₄	3
dNTPs (2 mM each)	5
PCR Grade Water	33
10 pmol/µl Sense Primer	1
10 pmol/µl Anti-Sense Primer	1
Template DNA	1
KOD Hot Start DNA Polymerase (1 U/µl)	1
Total reaction volume	50

Table 2.10: Temperatures and times that were used for amplification by KOD Hot Start DNA Polymerase.

Cycle	Temperature (°C)	Time
1-Polymerase activation	95	2 min
2-Denature	95	20 s
3-Annealing	51 (<i>fbp</i>)	10 s (<i>fbp</i>)
	56 (<i>glpX</i>)	10 s (<i>glpX</i>)
4-Extension	70 (<i>glpX</i>)	20 s (<i>glpX</i>)
	70 (<i>fbp</i>)	40 s (<i>fbp</i>)
Repeat steps 2–4	30 times	
5-Final extension	70	3 min

2.9.2 Purification and precipitation of PCR product

After amplification, PCR products were purified to remove any contaminating DNA. Therefore, DNA extraction from agarose gel with Ultrafree[®]-DA kit (Millipore, Cat# 42600) was used for purification of the PCR product. PCR products (60 µl) were run in a modified TAE electrophoresis gel [40 mM Tris-acetate, pH 8.0, 0.1 mM Na₂EDTA] for 1 hour alongside hyperladder I, and then the gel was placed on a UV box to locate the bands. The bands were cut out using a blade and transferred into a Gel Nebulizer and the device was sealed with the cap attached to the vial, centrifuged at 6800 xg for 10 minutes, and finally the supernatant which contained the DNA was captured in the vial and the Gel Nebulizer which contained the agarose gel was discarded.

To inactivate the polymerase enzyme and remove dNTPs, isopropanol precipitation was used, 80 µl of gel extracted DNA was added to 80 µl of phenol-chloroform isoamyl alcohol, 24:1), mixed by vortexing for 1 minute, and centrifuged at 18000 xg for 1 minute and then the supernatant mixed with an equal volume of chloroform to remove any residue of phenol, and centrifuged at 18000 xg for 5 minutes. The top phase of the supernatant which contained the DNA was transferred into a new 1.5 ml tube, and 0.1 volume of 3 M sodium acetate pH 5.2 and 1 volume isopropanol were added, mixed by vortexing and incubated at room temperature for 5 minutes to precipitate DNA.

After centrifugation at 18000 xg for 5 minutes, the supernatant was removed and the DNA pellet was washed with 70% (v/v) ethanol, spin at 18000 xg for 5 minutes, and then the DNA pellet was resuspended in 30 μ l TE buffer [10 mM Tris-HCl, 0.1 mM EDTA pH 8.0] which was provided with the kit.

2.9.3 T4 DNA Polymerase treatment of the purified PCR product

To generate compatible overhangs on the purified PCR product, T4 DNA Polymerase (Novagen, Cat# 71086-5) was added to the reaction mix as shown in Table 2.11. To start the reaction, the enzymes was added to the reaction mix, stirred with a pipette tip and incubated at room temperature for 30 minutes. The enzyme was inactivated by incubating at 75°C for 20 minutes.

Table 2.11: T4 DNA polymerase reaction setup.	
Component	Volume (μl)
100 ng/ μ l Purified PCR product	4
10X T4 DNA Polymerase Buffer	2
25 mM dATP	2
100 mM DTT	1
Nuclease-free Water	10.6
2.5 U/ml T4 DNA Polymerase	0.4
Total reaction volume	20

The treated PCR product (1 μ l) was added to the pET-41 Ek/LIC vector (2 μ l) in order to perform the annealing reaction. The reaction mix was incubated at room temperature for 5 minutes and finally, 1 μ l of 25 mM EDTA was added to stop the reaction and preserve recombinant vector.

2.9.4 Transformation into the cloning strain

The vector that contained the cloned PCR product was transformed into NovaBlue GigaSingles™ Competent Cells which harbor *recA* and *endA* mutations to increase the transformation efficiency and yield of recombinant plasmid DNA (Novagen, Cat# 71086-5) as described in section 2.8.2. The recombinant vector was purified as described in section 2.7.2.

2.9.5 Transformation into the expression strain

The BL21 (DE3) pLysS Competent Cells (Novagen, Cat# 69451) were used to overexpress GlpX and Fbp proteins. The transformation was carried out as described in the instructions of the competent cell kit. The tube that contained the competent cells was taken from the -70°C freezer and immediately placed on ice for 5 minutes to thaw. Then 1 µl of purified plasmid (10 ng/µl) was added to the cells, mixed gently and placed on ice. The tube was incubated on ice for 5 minutes, and then heated in a water bath at 42°C for 30 seconds, returned to ice for 2 minutes and then 250 µl of S.O.C medium was added. The tube was incubated at 37°C for 1 hour at 200 rpm in the shaking incubator, and then 100 µl of the mixture was plated on a selective medium which contained (+Kan), and incubated overnight at 37°C.

2.9.6 PCR Screening

PCR screening was conducted as per Section 2.7.1 to verify whether the plasmid was transferred to the BL21 (DE3) pLysS Competent Cells.

2.9.7 Expression of the Fbp and GlpX proteins

After the *fbp* and *glpX* genes were identified as being cloned correctly in the Pet-41 Ek/LIC plasmid, the Fbp and GlpX proteins were expressed as His tagged proteins using Overnight Express™ Instant LB media (Novagen, Cat# 71757). The procedure for the Fbp and GlpX proteins expression are summarized as follows.

A colony of recombinant BL21 cells containing either *fbp* or *glpX* gene was transferred using a sterile toothpick into 10 ml LB Broth (+Kan) and was incubated overnight with shaking at 200 rpm at 37°C. The overnight culture (1 ml) was transferred into 500 ml of Overnight Express™ Instant LB media, and incubated overnight with shaking at 200 rpm at 37°C. After induction, cells were centrifuged at 20000 xg at 4°C for 10 minutes. The supernatants were discarded and the pellets were washed by resuspending in distilled water, and a second centrifugation was carried out at 2700 xg at 4°C for 20 minutes and finally the supernatants were discarded and the cell pellets were resuspended in Ni-Nitrilotriacetic acid (NTA) His•Bind Matrix buffer (3 ml for each 1 g of cell pellet) [50 mM NaH₂PO₄, pH 8.0; 300 mM NaCl; 10 mM imidazole].

The suspended cells were lysed using a French Pressure Cell Press[®] under 1000 PSI twice and then the lysates were centrifuged at 2700 xg at 4°C for 20 minutes to separate the soluble fraction from insoluble material.

2.9.8 Purification of protein under native conditions using immobilizing metal-affinity chromatography (IMAC)

This purification system is based on a separation technique which uses chelating compounds with immobilized metal ions that covalently bind to amino acid residues exposed on the surface of the target protein (Gaberc-Porekar and Menart, 2001). This purification system uses nitrilotriacetic acid (NTA) as the chelating compound, Ni²⁺ as the immobilized metal ion and the hexahistidine tags as amino acid residues exposed on the surface of the target protein (Figure 2.1).

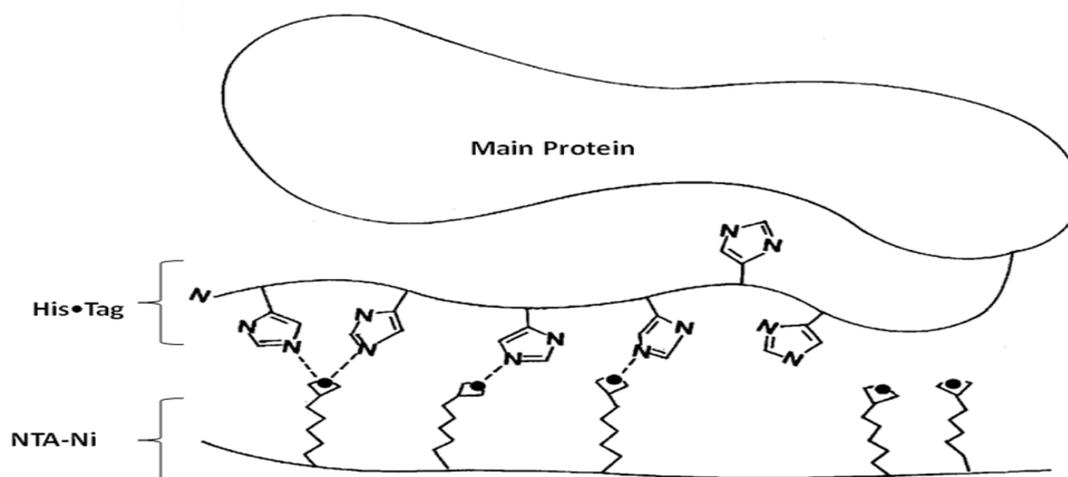


Figure 2.1: Schematic illustration of a protein with a His-Tag attached at the N-terminal binding to a metal-chelated affinity support (NTA-Ni).

Strong binding of a protein onto the IMAC matrix is achieved predominantly by multi-point attachment of the engineered surface histidine tag which is located at the N- or and C-terminus of the protein. Figure adapted from Gaberc-Porekar and Menart, (2001).

The His tagged Fbp and GlpX proteins were purified using BugBuster[®] Ni-NTA His-Bind Purification kit (Novagen, Cat# 70751-3) under native conditions by following the manufacturer's protocol with the modification of using 1 M imidazole in the third and fourth elutions and being left for 1 hour to collect the elution.

The 50% Ni-NTA His•Bind slurry that was provided with the kit (1 ml) was added to 4 ml of 1X Ni-NTA Bind Buffer and mixed gently and left for 20 minutes to settle. A 4 ml of soluble lysate was added to the Ni-NTA His•Bind slurry and mixed gently and incubated 1 hour at 4°C.

After incubation, the mixture was loaded into a column with the outlet capped and left for 20 minutes at room temperature and then the cap was removed and the flow-through was collected in a 15 ml tube. The washing step was carried out twice by added 4 ml of 1X Ni-NTA Wash Buffer [50 mM NaH₂PO₄, pH 8.0; 300 mM NaCl; 20 mM imidazole] and the washing fractions were collected. Finally, 4 × 0.5 ml of Ni-NTA Elution Buffer [50 mM NaH₂PO₄, pH 8.0; 300 mM NaCl; 250 mM imidazole] was added to elute the His tagged protein and collected in four tubes.

2.9.9 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

A 10 µl of both purified proteins were added to 10 µl of 2 × SDS protein samples loading buffer and heated in a heating block at 100°C for 10 minutes, then the mixtures were loaded on the gel. Also 10 µl of insoluble fractions of the cell extract were removed and added to the separate volumes of 2 × SDS protein samples loading buffer and heated in a heating block at 100°C for 10 minutes, then the mixtures were also loaded on the gel. The AE-6450 Dual Mini Slab Kit (ATTO, Cat# 2322222) cassette was assembled for mini gel as described in the manufacturer`s manual and the SDS-PAGE gel was prepared with the stacking gel at the top and the running gel at the bottom. A 7% and 4% (v/v) of acrylamide/bis-acrylamide (30%, Sigma) in buffer were used in making the stacking and running gel, respectively. First, the running gel was prepared by loading around 4/5 of the cassette and then a layer of 98% (v/v) butanol was added above the running gel and left for 1 hour. After the running gel had polymerized, the butanol was discarded and replaced by the stacking gel and left for 45 minutes to polymerise.

The 7% (v/v) running gel contained the following (30 ml): 15.3 ml of H₂O, 7.5 ml of 1.5 M Tris-HCl (pH 6.8), 140 µl of 20% (w/v) SDS, 6.9 ml of acrylamide/bis-acrylamide (30%, Sigma), 140 µl 10% (w/v) ammonium persulfate, 20 µl of Tetramethylethylenediamine (TEMED).

The 4% (v/v) stacking gel contained the following (5.05 ml): 3.075 ml of H₂O, 1.25 ml of 1.5 M Tris-HCl (pH 6.8), 25 µl of 20% (w/v) SDS, 670 µl of acrylamide/ bis-acrylamide (30 %, Sigma), 25 µl 10% (w/v) ammonium persulfate and 5 µl of TEMED. After the 20 µl samples were loaded alongside 10 µl of HyperPage Prestained Protein Marker (Cat# 33066, Bioline), the gel was run in the SDS-PAGE tank containing a reservoir buffer [30.3 g Tris base; 144 g Glycine; 10 g SDS with 1L dH₂O] at 200 V and left for 2 hours. The gel was then removed from the cassette and placed in a box containing Coomassie blue staining buffer [45% (v/v) methanol; 10% (v/v) acetic acid; 0.2% (w/v) Coomassie Brilliant Blue; 44.8% (v/v) water] to stain the gel and left for 1 hour at 37°C, followed by overnight destaining at room temperature in the destaining buffer [60% (v/v) methanol; 10% (v/v) acetic acid; 30% (v/v) water].

2.9.10 Dialyzing of the recombinant proteins

The recombinant proteins were dialyzed by using Dialysis Cassette Slide-A-Lyzers (3 ml) (Fisher Scientific, Cat# 66382). Dialysis cassette is a single use, self-floating syringe-loadable unit. The dialysis cassette membrane is made from regenerated cellulose that allows a complete retention of proteins and macromolecules that are larger than 10 kDa, at the same time, allowing the removal of unwanted buffer. These dialysis cassettes can be loaded with up to 3 ml of the sample needs to be dialyzed.

The recombinant proteins were dialyzed as described in the manufacturer`s protocol. A 5 ml syringe was filled with the recombinant protein up to 3 ml and injected through the gasket of a needle port at the cassette`s corner by the needle. After that, the needle was removed and the gasket resealed and the cassette corner was marked to note which needle port had been used. The cassette was placed vertically with foam attached in 1 litre of dialysis buffer and stirred gently, and was left to dialyze for overnight at 4°C. The dialysis cassette was removed from the dialysis buffer, and then the needle was inserted into an unused needle port to remove the dialyzed recombinant protein.

2.9.11 Cleaving the fusion protein with recombinant enterokinase

The Enterokinase Cleavage Capture Kit from Novagen (Cat# 69067) can be used in order to remove tags from proteins in the N-terminal (GST.Tag, His.Tag and S.Tag). Recombinant enterokinase (rEK) is a catalytic subunit of bovine enterokinase which cleaves after the C-terminal end of a lysine residue in the sequence Asp-Asp-Asp-Asp-Lys↓-insert. After the cleavage of the target protein, rEK is removed from the reaction by using affinity capture on EKapture™ Agarose and finally the agarose is removed using a spin-filtration tube. First, small scale optimization is needed to estimate the appropriate amount of target protein Table 2.12.

Table 2.12: Reaction components for small scale optimization incubated in room or 37°C temperature.	
Component	Volume (µl)
10X rEK Cleavage/Capture Buffer	5
Target protein	20, 10
Diluted rEK (1 unit/µl)	1, 2
Deionized water	Until the total volume reached 50

2.10 Growth of *C. acetobutylicum* ATCC 842 for RNA extraction

C. acetobutylicum ATCC 842 was kept as spores in the cold room 4°C in sterile water. Spore suspension (2 ml) was transferred into an empty sterile tube and heat shocked at 80°C for 10 minutes, and then transferred into 20 ml of Reinforced Clostridial Medium, (RCM, Oxoid) and incubated at 37°C for 2 days in the anaerobic cabinet (Forma Scientific, Marietta, Ohio) to prepare starter cultures. Starter culture (3 ml) was transferred into 100 ml of Clostridial Basal Medium (CBM) and incubated at 37°C in the anaerobic cabinet overnight. Overnight culture (1.5 ml) was transferred into a microcentrifuge tube, 1 ml of RNA Protect Bacteria Reagent (Qiagen, Cat# 76506) was added and incubated at room temperature for 10 minutes, and then centrifuged for 10 minutes at 10600 x g. The supernatant was discarded and the pellet was immediately frozen in liquid nitrogen and then stored at -70°C.

2.10.1 RNA extraction for *C. acetobutylicum* ATCC 824

The RNA was extracted by using RNeasy Mini Kit as described in the manufacturer's protocol (Qiagen, Cat# 76506). A tube of cell pellet was transferred from the -70°C freezer into ice to thaw for 5 minutes, and then 200 μl of QIAGEN Proteinase K was added and the pellet was resuspended by pipetting up and down several times. The mixture was mixed by vortexing for 10 seconds and then incubated at room temperature for 10 minutes, during this period of incubation, the sample was vortexed for 10 seconds every 2 minutes. The RLT Buffer (700 μl) was added to the mixture and mixed by vortexing then, 500 μl of ethanol (96%) was added to the sample and mixed by pipetting. Lysate (700 μl) was transferred into an RNeasy Mini spin column and placed into a 2 ml collection tube, and centrifuged for 15 seconds at 10600 $\times g$, and the flow-through was discarded. The Buffer RW1 (700 μl) was added to the RNeasy Mini spin column, and centrifuged for 15 seconds at 10600 $\times g$ to wash the membrane of the spin column and the flow-through and the collection tube were discarded. The RNeasy Mini spin column was placed in the top of a new 2 ml collection tube and 500 μl of Buffer RPE was added and centrifuged for 15 seconds at 10600 $\times g$ to wash the spin column, and again the flow-through was discarded. Again 500 μl of Buffer RPE was added to the RNeasy Mini spin column and centrifuged for 15 seconds at 10600 $\times g$ to wash the spin column, and again the flow-through was discarded. The RNeasy Mini spin column was placed in a new 1.5 ml collection tube and 50 μl of RNeasy-free water was added, and centrifuged for 1 minute at 10600 $\times g$ to elute the RNA.

2.10.2 Analysis of purified RNA

The total purified RNA was separated using a 1% (w/v) agarose gel in $1 \times$ MOPS (morpholinopropanesulfonic acid) buffer. The MOPS stock buffer ($10 \times$ MOPS) contained: 50 mM sodium acetate, 200 mM MOPS, and EDTA at pH 7.0. The RNA loading buffer consisting of 50% (v/v) formamide, 16% (v/v) formaldehyde, 10% (v/v) of $10 \times$ MOPS buffer, 0.1 mg/ml ethidium bromide and 0.01% (w/v) bromophenol blue, was mixed with the purified RNA, denatured at 70°C for 10 minutes, placed on ice for 2 minutes, and then the total RNA was separated by electrophoresis at 80 V for 1 hour.

2.10.3 Reverse transcriptase PCR

RNA was prepared as described in section 2.10.1 and used for cDNA synthesis and then PCR amplification of two fragments, first one between *glpX* and *hprK* genes and second between a gene that encodes for 5-formyltetrahydrofolate cyclo-ligase and *hprK* gene using a QIAGEN[®] One-Step RT-PCR Kit (Cat# 210210). This kit contains a special blended enzyme mix for both reverse transcription and PCR amplification. Omniscript and Sensiscript Reverse Transcriptases are responsible for generating cDNA from the RNA. Omniscript Reverse Transcriptase is specially engineered for reverse transcription of RNA amounts greater than 50 ng. However, Sensiscript Reverse Transcriptase is designed for use with very small amounts of RNA, less than 50 ng. HotStarTaq DNA Polymerase is needed to generate the PCR products from the cDNA and this enzyme does not interfere with the reverse-transcriptase reaction due to it only being activated when the enzyme is heated to 95°C for 15 minutes. Details of the PCR reaction mix and PCR cycles are presented in Table 2.13 and Table 2.14 respectively.

Component	Volume (µl)
5x QIAGEN OneStep RT-PCR Buffer	10
dNTP Mix (containing 10 mM of each dNTP)	2
100 pmol/µl RNAglpX or RNAhprK forward primer	1
100 pmol/µl RNAglpX or RNAhprK reverse primer	1
QIAGEN OneStep RT-PCR Enzyme Mix	2
(4 U/µl) RNase inhibitor	2
Template RNA	Variable
RNase-free water	Variable
Total volume	50

Cycling step	Time (min)	Temperature (°C)
1-Reverse transcription step	30	50
2-Initial PCR activation step	15	95
3-Denaturation step	1	94
4-Annealing step	1	Variable
5-Primers extension step	1	72
Repeats steps 3 to 5 for 40 cycles		
6-Final extension step	10	72

Also, the One Step RT-PCR Master Mix Kit from Novagen® was used to generate cDNA and then PCR product. This kit utilizes an engineered *Thermus thermophilus* (*rTth*) DNA polymerase, which acts both as reverse transcriptase and DNA polymerase. Details of the PCR reaction mix and PCR cycles are presented in Table 2.15 and Table 2.16 respectively.

Component	Volume (µl)
2X One Step RT-PCR Master Mix	25
50 mM Mn(OAc) ₂	2.5
10 pmol/µl RNAgpX or RNAhprK forward primer	1
10 pmol/µl RNAgpX or RNAhprK reverse primer	1
Template RNA	Variable
RNAse-free water	Variable
Total volume	50

Cycling step	Time	Temperature (°C)
1-Enzyme Activation step	30 s	90
2-Reverse transcription step	30 min	60
3-Denaturation step (enzyme)	1 min	94
4-Denaturation step (DNA)	30 s	94
5-Annealing step	30 s	variable
6-Extension step	1 min	72
Repeat steps 4–6 for 40 cycles		
7-Final Extension	7 min	60

2.11 Proteomic analysis

A sample of crude extract of *C. acetobutylicum* grown on glucose minimal medium was obtained from Dr. Wilfrid J Mitchell and sent to Moredun Research Institute by Dr. Douglas Fraser-Pitt. The proteomic analysis was carried out by using liquid chromatography electrospray ionization tandem mass spectrometry (LC/ESI-MS/MS).

2.12 Bioinformatics analysis

The National Centre for Biotechnology Information (NCBI) database was used to obtain the DNA and amino acid sequence of each gene and protein respectively:

<http://www.ncbi.nlm.nih.gov/>.

Also, the Basic Local Alignment Search Tool (BLAST), which is one of the search tools in NCBI database, was used to determine the percent homology in sequence alignments:

http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastHome.

ClustalW2 on The European Molecular Biology Laboratory- European Bioinformatics Institute (EMBL-EBI) database was used to align multiple DNA or protein sequences and visualised via the programmes GeneDoc or TreeView:

<http://www.ebi.ac.uk/Tools/msa/clustalw2/>.

Chapter 3

3 Characterisation of *glpX* gene

3.1 Introduction

FBPase is one of the main irreversible gluconeogenic enzymes and is involved in converting fructose 1,6-bisphosphate to fructose 6-phosphate and orthophosphate. This protein has been characterised in many Gram positive and Gram negative bacteria such as *B. subtilis* and *E. coli* (Brown *et al.*, 2009; Jules *et al.*, 2009). To date, there is no study reporting FBPase activity in the genus *Clostridium*. Tangney *et al.*, (2003) reported that there is an open reading frame that might encode for a FBPase enzyme in *C. acetobutylicum* ATCC 824. Therefore, the aim of this chapter was to examine whether the open reading frame that has been reported does encode for FBPase and also to examine whether there is any other gene that might encode for FBPase enzyme. This was investigated by bioinformatic analysis, genetic complementation and biochemical analysis (enzyme assay).

3.2 Results

3.2.1 Bioinformatic analysis

The whole genome of *C. acetobutylicum* ATCC 824 has been sequenced and is accessible for analysis (Nölling *et al.*, 2001; GenBank accession number AE001437 <http://www.ncbi.nlm.nih.gov/nucleotide/AE001437>). *C. acetobutylicum* like all other bacteria requires fructose-1,6-bisphosphatase activity to enable the organism to grow under gluconeogenic conditions. The genome sequence of *C. acetobutylicum* was searched for potential homologues of fructose-1,6-bisphosphatase encoding genes using the NCIB database. Two potential homologues were identified, cac1088 and cac1572. Phylogenetic analysis of cac1088 and cac1572 with both experimentally proven and putative FBPase proteins from diverse species including human, yeast, plants and bacteria revealed that cac1088 clusters with FBPase class II proteins (Figure 3.1). It appears that cac1088 is evolutionarily divergent from *E. coli* FBPase class II. Interestingly, cac1572 clusters with FBPase class III enzymes which including *B. subtilis* FBPase enzyme (Fujita *et al.*, 1998).

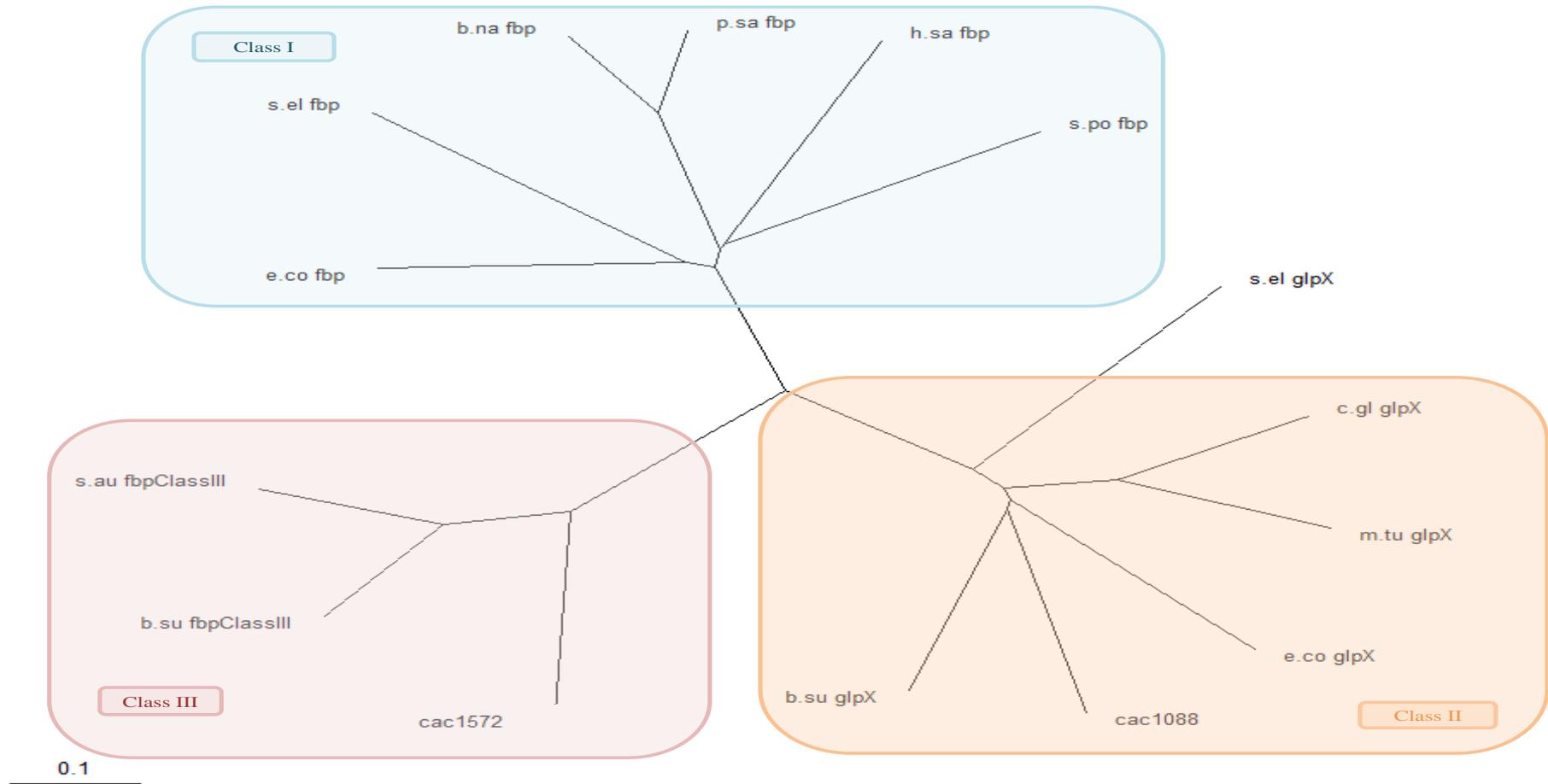


Figure 3.1: Unrooted phylogenetic tree of characterised FBPs and the two putative clostridial FBPs

Continuation of Figure 3.1. Abbreviations for FBPsases Class I: b.na fbp; *Brassica napus* (Rape), s.po fbp; *Schizosaccharomyces pombe* (Fission yeast), h.sa fbp; *Homo sapiens* (Human), e.co fbp; *Escherichia coli fbp*, s.el_fbp; *Synechococcus elongatus* PCC 7942 and p.sa fbp: *Pisum sativum* (Garden pea). For FBPsases Class II: s.el glpX; *Synechococcus elongatus* PCC 7942, c.gl glpX; *Corynebacterium glutamicum*, m.tu glpX; *Mycobacterium tuberculosis*, e.co glpX; *Escherichia coli*, cac1088; *C. acetobutylicum glpX*, b.su glpX; *Bacillus subtilis*. For FBPsases Class III: cac1572; *C. acetobutylicum fbp*, b.su fbp; *Bacillus subtilis*, s.au fbp; *Staphylococcus aureus*.

3.2.2 The cac1088 gene

The first gene identified cac1088, covers 975 bp and putatively encodes a class II GlpX-like protein of 324 aa. The estimated molecular weight of the protein is 34.81 kDa. A BLAST search was conducted on the NCIB protein database using the deduced amino acid sequence of the putative GlpX protein; the results are presented in Table 3.1. All of the top ten hits from the BLAST search were from the genus *Clostridium*. The putative FBPsase of *C. botulinum* A str. ATCC 3502 shares 78% identity with the putative *C. acetobutylicum* FBPsase protein (cac1088). The cac1088 also shares similarity with putative FBPsase of other *Clostridium* spp.

The protein sequences of the top ten BLAST hits were then aligned in ClustalW2 (www.ebi.ac.uk/Tools/msa/clustalw2/) together with the experimentally proven type II FBPsase protein sequences such as *E. coli* and *M. tuberculosis* and then visualised in GeneDoc (www.nrbsc.org/gfx/genedoc/). There are large areas of conservation throughout the length of the proteins (Figure 3.2). There are four blocks of conserved motifs (Brown *et al.*, 2009) consisting of VIGEGE, APML, AVDPIDGT and DGDV in sequence block I, II, III and IV respectively (Figure 3.2). The amino acid residues AVDPIDGT conserved in motif III have been shown to be involved in metal ion binding and phosphatase catalytic activities by crystallographic and mutational analysis (Brown *et al.*, 2009). It is interesting to note that the presence of the conserved block III in proteins is not sufficient for identification of FBPsase members (Brown *et al.*, 2009; York *et al.*, 1995).

Table 3.1: BLAST homology results for the deduced amino acid sequence of the putative fructose 1,6-bisphosphatase II (cac1088).

Species	Putative protein	% Identity	% Similarity	Predicted molecular mass (Da)	Accession number
<i>C. carboxidivorans</i> P7	Fructose 1,6-bisphosphatase , class II	78	89	34364	ZP_05392875
<i>C. sporogenes</i> ATCC 15579	Fructose 1,6-bisphosphatase , class II	78	89	34306	ZP_02995269
<i>C. botulinum</i> A str. ATCC 3502	Fructose 1,6-bisphosphatase , class II	78	89	34304	YP_001255113
<i>C. botulinum</i> A3 str. Loch Maree	Fructose 1,6-bisphosphatase II	77	89	34304	YP_001787935
<i>C. botulinum</i> B1 str. Okra	Fructose 1,6-bisphosphatase II	77	89	34286	YP_001782231
<i>C. novyi</i> NT	Fructose 1,6-bisphosphatase II	77	89	34227	YP_877736
<i>C. botulinum</i> C str. Eklund	Fructose 1,6-bisphosphatase , class II	76	90	34281	ZP_02621295
<i>C. botulinum</i> D str. 1873	Fructose 1,6-bisphosphatase , class II	75	89	34216	ZP_04862266
<i>C. botulinum</i> BKT015925	Fructose 1,6-bisphosphatase II	75	89	34255	YP_004395515
<i>C. tetani</i> E88	Fructose 1,6-bisphosphatase II	74	87	34339	NP_781657

```

                *           20           *           40           *
C.ace       : -----MFDNDISMSLVRVTEAAAALQSSKYMGRGDKIGADQAAVDG : 40
C.boA       : -----MLNTDIAMGLARVTEAAAALSASKFMGRGDKNAADQAAVDG : 40
C.bo_BKT    : -----MKNLDVSMGLVRVTEAAAALNGAKLMGRGDKNAADQAAVDG : 40
C.bo_B1     : -----MLNTDIAMGLARVTEAAAALSASKFMGRGDKNAADQAAVDG : 40
C.bo3       : -----MLNTDIAMGLARVTEAAAALSASKFMGRGDKNAADQAAVDG : 40
C.sp        : -----MLNTDIAMGLARVTEAAAALSASKFMGRGDKNAADQAAVDG : 40
C.nov       : -----MNNLDVSMGLVRVTEAAAALNGAKLMGRGDKNAADQAAVDG : 40
C.bo_C      : -----MNNLDVSMGLIRVTEAAAALNGSKLMGRGDKNAADQAAVDG : 40
C.bo_D      : -----MKNLDVSMGLVRVTEAAAALNGAKLMGRGDKNAADQAAVDG : 40
C.car       : -----MLNTDIAMGLARVTEAAAALCSSKFMGRGDKIAADQAAVDG : 40
C.tet       : -----MIKQDIAMGLVRVTEAAAALCSSKYMGRGDKIAADQAAVDG : 40
Mtb         : MASHDPSHTRPSRREAPDRNLAMELVRVTEAGAMAAGRWWVGRGDKEGGGAAVDA : 55
Eco         : -----MRRELATIEFSRVTESAALAGYKWLGRGDKNTADGAAVNA : 39
    
```

```

                60           * I           II           *           100           III           *
C.ace       : MEKAFSFMPPVRGQVVIIGEGELDEAPMLYIIGQKLGMGKDYMPEMDIIVDPLDGTII : 95
C.boA       : MHKAFQIMPIRGKVVIGEGELDEAPMLYIGEEVGI GAEDMVEMDIIVDVPDGTIKL : 95
C.bo_BKT    : MEKAFNMMPIRGKVVIGEGEMDEAPMLYIGQSVGVGEEKMPEMDIIVDPI DGTIVL : 95
C.bo_B1     : MHKAFQIMPIRGKVVIGEGELDEAPMLYIGEEVGI GAEDMVEMDIIVDVPDGTIKL : 95
C.bo3       : MHKAFQIMPIRGKVVIGEGELDEAPMLYIGEEVGI GAEDMLEMDIIVDVPDGTIKL : 95
C.sp        : MHKAFQIMPVRGKVVIGEGELDEAPMLYIGEEVGI GAEDMVEMDIIVDVPDGTIKL : 95
C.nov       : MEKAFNMMPIRGNVVIGEGEMDEAPMLYIIGQKVGVGQEDMPEMDIIVDPI DGTIVL : 95
C.bo_C      : MEKAFNMMPIRGNVVIGEGEMDEAPMLYIIGQKVGIGREDMPEMDIIVDPI DGTIVL : 95
C.bo_D      : MEKAFNMMPIRGRRVIGEGEMDEAPMLYIIGQNVGVGVEENMPEMDIIVDPI DGTIVL : 95
C.car       : MEKAFAMMPVRGTVVIGEGELDNAPMLYIGQSVGVGSADMPEMDIIVDPLDGTII : 95
C.tet       : MEKAFSMLPIKGTVVIGEGEMDEAPMLYIIGQKLGIGTDGMEEMDIIVDPLDGTIVL : 95
Mtb         : MRELVNSVSMRGVVVIGEGEKDHPAPMLYNGEEVGN--DGPECFIVDVPIDGTTL : 108
Eco         : MRIMLNQVNI DGTIVIGEGELDEAPMLYIIGEKVGTG--RGDAVDIIVDPIEGTRM : 92
    
```

```

                120           *           140           *           160
C.ace       : ISKGLPNAIAVAVIAMPKGSLLH-APDMYMKKIVVGPAGKGAIDINKSPEENILNV : 149
C.boA       : IAKGLENAAIAVVAMGPKGSLLH-APDMYMKKIIVGAGAKGAIDINKSPKENILNV : 149
C.bo_BKT    : IAKGLPNSIAVVAMGPKGSLLH-APDMYMQKIAVGPARGSIDINKSPRENIINV : 149
C.bo_B1     : IAKGLENAAIAVVAMGPKGSLLH-APDMYMKKIIVGAGAKGAIDINKSPKENILNV : 149
C.bo3       : IAKGLENAAIAVVAMGPKGSLLH-APDMYMKKIIVGAGAKGAIDINKSPKENILNV : 149
C.sp        : IAKGLENAAIAVVAMGPKGSLLH-APDMYMKKIIVGSGAKGAIDINKSPKENILNV : 149
C.nov       : IAKGLPNSIAVVAMGPKGSLLH-APDMYMKKIIVGPGAVGAIDINKSPKENIINV : 149
C.bo_C      : IAKGLPNSIAVVAMGPKGSLLH-APDMYMKKIIVGPGAKGAIDINKSPKENIINV : 149
C.bo_D      : IAKGLPNSIAVVAMGPKGSLLH-APDMYMQKIAVGPAGKGAIDINKSPKENIINV : 149
C.car       : IAKGLPNAIAVVAMGPKESLH-APDMYMKKIIVGPGAKGAIDINKSPKENIINV : 149
C.tet       : IAKGLPNAIVSMAMGPKGSLLN-APDMYMKKIIVGASAKGVIDINKTNKENIINV : 149
Mtb         : MSKGMTNAISVLAADRGTMFDPSPAFYMNKIAVGPDAAHVLDITAPISENIRAV : 163
Eco         : TAMGQANALAVLAVGDKGCFLN-APDMYMEKLIIVGPGAKGTIDLNLPLADNLRNV : 146
    
```

```

                *           180           *           200           IV           *           220
C.ace       : AKALNKDISELTVIVQERERHDIYVKAATIEVGARVKLFGEGDVAAILACGFEDTG : 204
C.boA       : ARALNKDVTELTIVIVQERDRHDIYVKDAREVGARVKLFGEGDVAAILACGFEDTG : 204
C.bo_BKT    : AKALNKDIEDLTVIVQERERHNYIVEQAREVGARVKLFAEGDVAAILACGFENTG : 204
C.bo_B1     : ARALNKDVTELTIVIVQERDRHDIYVKDAREVGARVKLFGEGDVAAILACGFEDTG : 204
C.bo3       : ARALNKDVTELTIVIVQERDRHDIYVKDAREVGARVKLFGEGDVAAILACGFEDTG : 204
C.sp        : ARALNKDVTELTIVIVQERDRHDIYVKDAREVGARVKLFGEGDVAAILACGFEDTG : 204
C.nov       : AKALNKDIEDLTIIVQERERHDIYVEQAREVGARVKLFGEGDVAAILACGFENTG : 204
C.bo_C      : AKALSKEIEDLTVIVQERERHNYIVEQAREVGARVKLFGEGDVAAILACGFENTG : 204
C.bo_D      : AKALNKDIEDLTVIVQERERHNYIVEQAREVGARVKLFAEGDVAAILACGFENTG : 204
C.car       : SKALNKDITEMTVIVQERERHNYIVQAREVGARVKLFGEGDVAAILACGFENTG : 204
C.tet       : AMALNKDVTELTIMVQERERHNDIYIKDAREVGARVKLFGEGDVAAILACGFESTG : 204
Mtb         : AKVKDLSVRDMTVCILDRPRHAQLIHDVIRATGARIRLITTDGDVAGATISACRPHSG : 218
Eco         : AAALGKPLSELTVTILAKPRHDAVIAEMQQLGVRVFAIPDGDVAASLLTCMPDSE : 201
    
```

Continuation of Figure 3.2

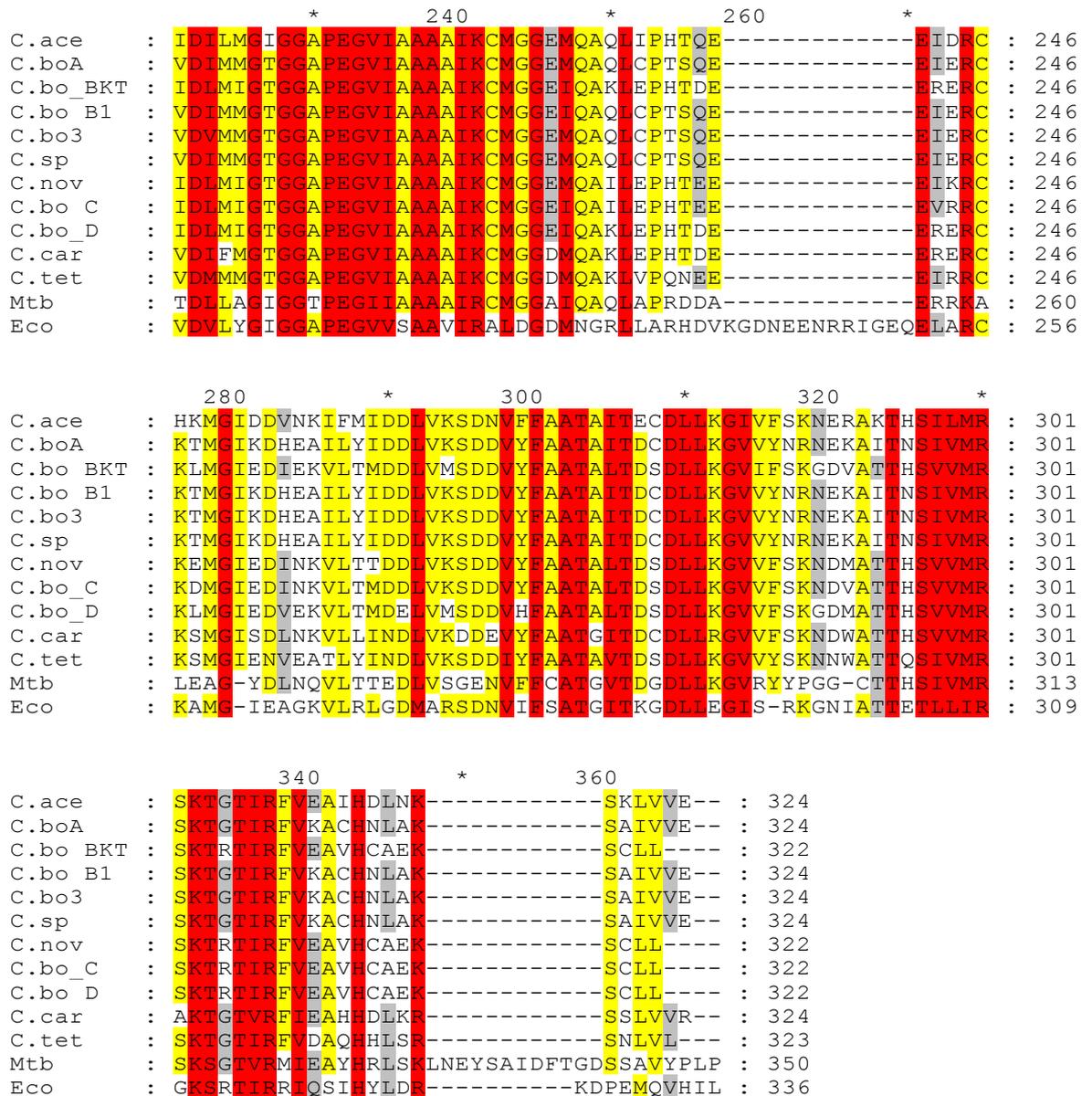


Figure 3.2: Alignment of the putative *C. acetobutylicum* fructose 1,6-bisphosphatase II (cac1088) sequence with the closest homologues.

The deduced amino acid sequence of the putative *C. acetobutylicum* fructose 1,6-bisphosphatase II protein is aligned with the fructose 1,6-bisphosphatase protein sequences with the highest percentage identities from the BLAST output and experimentally verified FBPase II enzymes in *M. tuberculosis* (Mtb) and *E. coli* (Eco). Dashes within the sequences indicate gaps giving optimal alignment. The amino acids are highlighted according to 90% or more (red), 80% (yellow), 60% (grey) and non-highlighted less than 60% conservation. Abbreviations: C.ace, *C. acetobutylicum* ATCC 824;

C.car, *C. carboxidivorans* P7; *C.sp*, *C. sporogenes* ATCC 15579; *C.boA*, *C. botulinum* A str. ATCC 3502; *C.bo3*, *C. botulinum* A3 str. Loch Maree; *C.bo B1*, *Clostridium botulinum* B1 str. Okra; *C.nov*, *C. novyi* NT; *C.bo C*, *C. botulinum* C str. Eklund; *C.bo D*, *C. botulinum* D str. 1873; *C.bo BKT*, *C. botulinum* BKT015925; *C.tet*, *C. tetani* E88; *Mtb*, *M. tuberculosis* and *Eco*, *E. coli*.

3.2.3 The *cac1572* gene

Bioinformatic analysis of the locus tag *cac1572* of the *C. acetobutylicum* genome revealed the presence of a second putative fructose 1,6-bisphosphatase-like class III protein (Fbp) of 665 aa. A BLAST search using the deduced amino acid sequence of the putative *C. acetobutylicum* Fbp protein revealed that *cac1572* is closely related to FBPase_2 family of proteins (Pfam06874). This family consists of several bacterial fructose-1,6-bisphosphatase proteins (EC:3.1.3.11) belonging to the Firmicutes phylum. The Pfam06874 appears to be not structurally related to class II FBPase_GlpX enzymes (Pfam003320). Also, the BLAST search results are presented in Table 3.2. All of the top ten hits from the BLAST search are from the genus *Clostridium*. The putative FBPase of *C. cellulovorans* 743B shares 73% identity with the putative *C. acetobutylicum* FBPase protein (*cac1572*). The *cac1572* also shares similarity with putative FBPases of other *Clostridium* spp. The protein sequences of the top ten BLAST hits were then aligned in ClustalW2 together with the experimentally proven class III FBPases protein sequences of *B. subtilis* (Fujita *et al.*, 1998) and then visualised in GeneDoc. Several conserved motifs can be seen through the alignment result together with the *B. subtilis* enzyme indicating that putative *C. acetobutylicum* *fbp* gene (*cac1572*) is most likely to encode for a class III FBPase protein.

Table 3.2 BLAST homology results for the deduced amino acid sequence of the putative fructose 1,6-bisphosphatase III (cac1572).

Species	Putative protein	% Identity	% Similarity	Predicted molecular mass (Da)	Accession number
<i>C. cellulovorans</i> 743B	Fructose 1,6-bisphosphatase	73	87	76707	YP_003842980
<i>C. beijerinckii</i> NCIMB 8052	Fructose 1,6-bisphosphatase	73	87	76726	YP_001309581
<i>C. botulinum</i> D str. 1873	Fructose 1,6-bisphosphatase	71	83	76535	ZP_04863125
<i>C. botulinum</i> C str. Eklund	Fructose 1,6-bisphosphatase	70	84	76294	ZP_02620031
<i>C. botulinum</i> BKT015925	Fructose 1,6-bisphosphatase	70	83	76412	YP_004395366
<i>C. novyi</i> NT	Fructose 1,6-bisphosphatase	68	83	76286	YP_877593
<i>C. ljungdahlii</i> DSM 13528	Fructose 1,6-bisphosphatase	68	83	77329	YP_003781055
<i>C. botulinum</i> A2 str. Kyoto	Fructose 1,6-bisphosphatase	69	85	77256	YP_002802849
<i>C. botulinum</i> Bf	Fructose 1,6-bisphosphatase	69	85	77257	ZP_02616118
<i>C. sporogenes</i> ATCC 15579	hypothetical protein	68	85	77298	ZP_02994084

```

                *           20           *           40           *
C.acet      : -----MLLESNTKNEEIKD-----NLKYIVLLSKQYPTINEAAFTIINLQA : 41
C.bot       : -----MTFCDENNLDLIKR-----DLRYINLLSKQYPDISSASTFIIVNLQA : 41
C.bot BKT   : -----MTFCDENNLDLIKR-----DLRYLKLAKQYYPDISSASTFIIVNLQA : 41
C.bot C     : -----MTFCDENNLDLIKK-----DLRYLKLLANQYYPNISSASTFIIVNLQA : 41
C.novy      : -----MTFCDENNLDLIKK-----DLRYLKLLANQYYPNISSASTFIIVNLEA : 41
C.ljun      : -----MIFYDENNLDIIEK-----DLRYLQLLSREYPTISSASTFIINLQA : 41
C.botA2     : -----MTLYDENNLHTIKD-----NLRYLKLKLSKQYPSISSASSETIINLQA : 41
C.botBf     : -----MTLYDENNLHTIKD-----NLRYLKLKLSKQYPSISSASSETIINLQA : 41
C.cell      : -----MSIHKELNELDVEK-----DLRYINLLSKEYYSTINSVCTEFINLQS : 41
C.beij      : -----MKKYDLSTNEISD-----NLRYLELLSKQYPTINEASTFIINLQA : 40
B.sub       : MFKNNVILLNSPYHAHAHKEGFILKRGWTVLESKYLDLLAQKYDCEEKVVTEIINLKA : 58
    
```

```

                60           *           80           *           100           *
C.acet      : ILNLPKGTTEHFLSDVHGGEYEQFIHVLKNASGVIKRKIDDIFGNRLMQSEKSLATLIY : 99
C.bot       : ILNLPKSTTEHFLSDIHGEYESFTHVLKNASGVIKRKIDDVFGNSLRDSEKLTLATVIY : 99
C.bot BKT   : ILNLPKSTTEHFLSDIHGEYESFTHVLKNASGVIKRKIDDVFGNSLRDSEKLTLATVIY : 99
C.bot C     : ILNLPKSTTEHFLSDIHGEYESFTHVLKNASGVIKRKIDDVFGNSLRDSEKLTLATVIY : 99
C.novy      : ILNLPKSTTEHFLSDIHGEYESFTHVLKNASGVIKRKIDDVFGNSLRDNEKLTLATVIY : 99
C.ljun      : ILNLPKATEHFISDIHGGEYESFTHMLRNASGVIKRKIDDVFGNSLREEQKASLATLVY : 99
C.botA2     : ILNLPKGTTEHFISDVHGGEYESFTHMLKNASGVIKRKIDDVFGTSLRECDKKNLATLIY : 99
C.botBf     : ILNLPKGTTEHFISDVHGGEYESFTHMLKNASGVIKRKIDDVFGTSLRECDKKNLATLIY : 99
C.cell      : ILNLPKGTTEHFLTDIHGEYEQFNHVLKNASGVIKRKIEDIFGNTLRQNEKKSATLIY : 99
C.beij      : ILNLPKGTTEHFLTDIHGEYEPFIHVLKNASGVIKRKIEDLFGNSLMQSEKSLATLIY : 98
B.sub       : ILNLPKGTTEHFVSDLHGGEYQAFQHVLRN GSRVKEKIRDFISGVLYDREIDEALALVY : 116
    
```

```

                120           *           140           *           160           *
C.acet      : YPEQKLDIILK---QEKNIDDWYKITLYRLIEVCRNVSSKYTRSKVRKALPKDFSYII : 154
C.bot       : YPEQKLELIKQ---SEKDLSDWYKITLYRLIELCRVVS SKYTRSKVRKALPHDFAYII : 154
C.bot_BKT   : YPEQKLELIKQ---CEKDLSDWYKITLYRLIELCRAVSSKYTRSKVRKALPHDFAYII : 154
C.bot_C     : YPEQKLELIKQ---SEKDLSDWYKITLYRLIELCRVVS SKYTRSKVRKALPHDFAYII : 154
C.novy      : YPEQKLELIKQ---SEKDLSDWYKITLYRLIELCRVVS SKYTRSKVRKALPHDFAYII : 154
C.ljun      : YPERKIEELIKE---KEHNLEEWYRISLYQLIELCKNVSSKYTRSKVRKALPSDFSYII : 154
C.botA2     : YPEQKLDLIKK---SEKNLEDWYKITLYRLIRLCQIVSSKYTRSKVRKSLPSDFAYII : 154
C.botBf     : YPEQKLDLIKK---SEKNLEDWYKITLYRLIRLCQIVSSKYTRSKVRKSLPSDFAYII : 154
C.cell      : YPEQKLDIVLK---EENINDWYTITLYRLIELCRHVSSKYTRSKVRKALPKDFAYII : 154
C.beij      : YPEQKLEIVLK---QEENIDDWYKINLYRLIELCRYVSSKYTRSKVRKALPKDFTYII : 153
B.sub       : YPEDKLLKLIKHFDAKEALNEWYKETIHRMIKLVSYCSSKYTRSKLRKALPAQDFAYIT : 174
    
```

```

                180           *           200           *           220           *
C.acet      : EELLHEQPKGVDKQEYDETIKTIISIDRAKEFTATSKLIQRLVDRLHIIGDIFDR : 212
C.bot       : EELLHEHDGAFNKQEYNGIVSTIIDIDRAPEFTATSKVIQRLVDRLHIIGDIYDR : 212
C.bot_BKT   : EELLHEHDGTINKQEYNGIISTIIDIDRAPEFTATSKVIQRLVDRLHIIGDIYDR : 212
C.bot_C     : EELLHEHDGTINKQEYNGIISTIIDIDRAPEFTATSKVIQRLVDRLHIIGDIYDR : 212
C.novy      : EELLHEHDGTINKQEYNGIVSTIIDIDRAPEFTATSKVIQRLVDRLHIIGDIYDR : 212
C.ljun      : EELLHDQDDKVDKQAYNGIIKTIIDINRAPEFTATSNVIQRLVDRLHIIGDIYDR : 212
C.botA2     : EELLNEQGDRVDKQEYNSIETIIDIDRAPEFTATSNVIQRLVDRKLIIGDIYDR : 212
C.botBf     : EELLNEQGDRVDKQEYNSIETIIDIDRAPEFTATSNVIQRLVDRKLIIGDIYDR : 212
C.cell      : EELLHEYPESLDKHGYGEIVKTIIVADRAREFTMALSNLIQRLVIDRHLIIGDIFDR : 212
C.beij      : EELLHEQPKGIDKYEYEQIIRTIIDTDRSKEFTVALSKLIQRLVIDRHLIIGDIFDR : 211
B.sub       : EELLYKTEQAGNKEQYSEIIDQIIELGQADKLITGIAYSVQRLVDRHLHVVGDIYDR : 232
    
```

```

                240           *           260           *           280           *
C.acet      : GFRADIIMDKLEEYFAVDIQWGNHDILWMGAASCSVCMANVIRISARYANLSTIEDG : 270
C.bot       : GFGAEIILEALMKHHSVDIQWGNHDILWMGAAAGSEACICNVLRISLRYANLNTIEDG : 270
C.bot BKT   : GFGAEIILEALMKHHSVDIQWGNHDILWMGAAAGSEACICNVLRISLRYANLNTIEDG : 270
C.bot C     : GFGAEIIMEELMKHHSVDIQWGNHDILWMGAAAGSEACICNVLRISLRYANLNTIEDG : 270
C.novy      : GFGAEIIMEELMKHHSVDIQWGNHDILWMGAAAGSEACICNVLRISLRYANLNTIEDG : 270
C.ljun      : GFGAEIIMDALMKHHSVDIQWGNHDMVWMGAAAGCEACIANVLRISLRYANLHTIEDG : 270
C.botA2     : GFGAEIILEALSKHHSVDIQWGNHDIVWMGAAAGCEACIANVIRISLRYANLSTLEDG : 270
C.botBf     : GFGAEIILEALSKHHSVDIQWGNHDIVWMGAAAGCEACIANVIRISLRYANLSTLEDG : 270
C.cell      : GFGAEIILDTLKNYHSVDIQWGNHDVLMGACAGSKACIANVIRISARVSNLDTIEDG : 270
C.beij      : GFGADIIMDTLVEYHSVDIQWGNHDILWMGAACSSDVCIANVIKNSLKYANLDTLENG : 269
B.sub       : GFGPDRIMEELINYSVDIQWGNHDVWLWIGAYSSKVCLANIIRICARVNDLDTIEDV : 290
    
```

Continuation of Figure 3.3

```

          300          *          320          *          340
C.acet   : YGINLLPLATFAMDFYGNDKCKNFEPKIESDKSYTVKEIELIGKMHKAIIIQFKLEG : 328
C.bot    : YGINLLPLATFAMDIYENDPCNSFIPKT-INKELTQNEINLISKMHKAIIIQFKLQG : 327
C.bot BKT : YGINLLPLATFAMDIYENDPCNSFIPKT-INKELTQNEINLISKMHKAIISIIIQFKLQG : 327
C.bot C  : YGINLLPLATFAMDVYENDPCNSFIPKT-INKELTQNELNLISKMHKAIIAIIIQFKLQG : 327
C.novy   : YGINLLPLATFAMDVYENDPCNSFIPKT-INKELTQNELNLISKMHKAIIAIIIQFKLQG : 327
C.ljun   : YGINLLPLATFALEFYKDDPCKNFIPKT-LNKDISKNELNLLAKMHKAIIAIMQFKLEA : 327
C.botA2  : YGINLLPLATFAMDFYKEDNCENFKPRT-IDKNLNETDIKLLSKMHKAIISIIIQFKLEG : 327
C.botBf  : YGINLLPLATFAMDFYKEDNCENFKPRT-IDKNLNETDIKLLSKMHKAIISIIIQFKLEG : 327
C.cell   : YGINLLPLATFALDFYKEDPCTGFFPKFDAEKNYSVTEIELIAQMHKAIIAIIIQFKLEG : 328
C.beij   : YGINLLPLATFSMDFYKDHPCNIFLPKMDCDKKYSINEINLIAQMHKAIIAIILFKLEG : 327
B.sub    : YGINLRPLLNLAEKYVDNNP--AFRPKA--DENRPEDEIKQITKIHQAIAIMIQFKLES : 344
    
```

```

          *          360          *          380          *          400
C.acet   : EAIKRHPEFKMEHRMLLNKINFEDSTIELDGKKYKLNDTSFPTIDKNDFYKILIDERE : 386
C.bot    : QIIKNHPEFKMDDQLLLDKINYEKGTINLDGNVYKLNDTIFFTINPKDFYTISPREND : 385
C.bot BKT : QIIKNHPEFKMDDQLLLDKINYEKGTISLDGNIYKLNDTVFPTIDPKNPYTTPGEDD : 385
C.bot C  : QIIQNHSEFKMDDQLLLDKIDYEKGTINLSGKIYKLNDTYFPTIDPKNPYKLTESED : 385
C.novy   : QIIKNHPEFKMDDQLLLDKINYEKGTIDLDGHIYKLNDTFFFPTVDPKDFYKLTENED : 385
C.ljun   : EIIKRHPEFKMDDRLLLDKIDIEKEWIDTYGHNYKLNDSKFPTIDWNDFYKLTDREQE : 385
C.botA2  : KIIKRRPEFKMEERLLLDKINIEKGTINLNEKIYKLNDTNFPTLDKENPYELNERER : 385
C.botBf  : KIIKRRPEFKMEERLLLDKINIEKGTINLNEKIYKLNDTNFPTLDKENPYELNERER : 385
C.cell   : EIINRRPGFNMDDRLLLNKINFEAGTIELDGKTYKLNDTYFPTIDPKNPYKLLKEEQE : 386
C.beij   : QVILRHPEFNMNRLLLLNKINYAEGTINLNGKTHKLKSFFFPTIDPKNPYELTYDEKE : 385
B.sub    : PIIKRRPENFMEERLLLEKIDYDKNEITLNGKTYQENTCFATINEQDQLLEEEEAE : 402
    
```

```

          *          420          *          440          *          460
C.acet   : VEEKLRSSSVNSEKLNRHVRFLFSHGNLYLKFNSNLLYHGCIPLNEDGTFKEVLIGSH : 444
C.bot    : LIRKITKSFVNSEKLQRHIRFLYSKGSMYLIYNSNLLYHGCIPLNEDGTFKEVTIDGK : 443
C.bot BKT : LIKKITKSFVNSEKLQRHIRFLYSKGSMYLIYNSNLLYHGCIPLNEDGTFKEVTIDGK : 443
C.bot C  : LIKKLARSFVNSEKLQRHIRFLYSKGSMYLVYNSNLLYHGCIPLNEDGTFREISINGT : 443
C.novy   : LIKKLTRSFVNSEKLQRHIRFMYSKGSMYLVYNSNLLYHGCVPLNEDGTFKEITIDGV : 443
C.ljun   : LIEKLTSSFVNSEKLQRHIKFLYGKSIYLVYNSNLLYHGCIPLNKDGSFREVKIGKI : 443
C.botA2  : LVEKLTNSFINSEKLQRHIKFLYSNGSLYLKYNSNLLYHGCIPLNEDGSLKEVTLCKE : 443
C.botBf  : LVEKLTNSFINSEKLQRHIKFLYSNGSLYLKYNSNLLYHGCIPLNEDGSLKEVTLCKE : 443
C.cell   : LVEKLRSSFMNSDKLNKHVRELFFANGSIYLKYNSNLLYHGCIPMNENGFFKVRLRDK : 444
C.beij   : LIDKLKTSFINSDKYNKHVRFLYSNGSLYLKFNSNLLYHGFIPLNEDGSFKVKIADK : 443
B.sub    : VIDKLLFSVQHSEKLGRHMNFMMKGSLYLKYNGNLIHGCIPVDENGNMTMIEDK : 460
    
```

```

          *          480          *          500          *          520
C.acet   : KYKGKALLDKLDVLARKSFFYEENSKNSKYENDMIWYLWSGPFSPLFGREKMTTFERY : 502
C.bot    : KYWGKSLLDKFDLAREAFFKEDSRLKFFAMDVWYLWCGPNSPEFGRKYMATFERY : 501
C.bot BKT : KYFGKSLLDKFDLAREAFFKEDSRLKFFAMDVWYLWCGPNSPEFGRKYMATFERY : 501
C.bot C  : RYKGKSLLDKFDSLARESFFKKGSKAKFFALDMWYLWCGPNSPEFGRKLRMTTLERY : 501
C.novy   : RYSGKSLLDKFDSLAREAFFKNGSSAKRFALDMMWYLWCGPNSPEFGRKLRMTTLERY : 501
C.ljun   : SYSGRELLDRYDRLARDAFFKNNTNSKKYGMDMMWYMWCGEDSPVFGRVRMTTFERY : 501
C.botA2  : TLKGKSLLDKLDRLAREAYFFKKDPESKLYGMDMMWYLWCGPNSPLFGRKKKMTTFERY : 501
C.botBf  : TLKGKSLLDKLDRLAREAYFFKKDPESKLYGMDMMWYLWCGPNSPLFGRKKKMTTFERY : 501
C.cell   : EYNGKSLLDRLEILAREGYFHEDNPAAKLFGTDMWYLWTGPYSPLFGREKMTTFERY : 502
C.beij   : EYKGKELLDKLMLAREAYFSKDKDDSDNK-KDIMWYLWCGASSPLFGRDMTIFEQY : 500
B.sub    : PYAGRELLDVFERFLAREAFAHPEET--DDLATDMAWYLWTCEYSSLFGRAMTTTFERY : 516
    
```

```

          *          540          *          560          *          580
C.acet   : FIDDKKTHEYKKDPYYHYRDDEDICINILREFGLDSEQAHIINGHVPESKNGENPIK : 560
C.bot    : FIDNEGTHIEYRNPYYKYRSDEKIVINILKEFGLDPNCSHIINGHIPVKTKAGENPIK : 559
C.bot_BKT : FIDNEGTHIEYRNPYYKYRSNEKVVINILKEFGLDPNCSHIINGHIPVKTKAGENPIK : 559
C.bot C  : FIDDKGTHIEYRNPYYKYRSNEDVVINILKEFGLDPDCSHIINGHIPVKTKAGENPIK : 559
C.novy   : FIDDKGTHIEYRNPYYKYRSNEKVVTNINILKEFGLDPDCSHVINGHIPVKTKAGENPIK : 559
C.ljun   : FIDDKIAHYEKKNYYQYRDDENVCKNILAEFNLPPESSHIINGHIPVKTKEGESPIK : 559
C.botA2  : FIDDKNTHKEQKNPYYKYRNDEKMCTMIFEEFELDADNSHIINGHIPVKTKEGENPIK : 559
C.botBf  : FIDDKNTHKEEKNPYYKYRNDEKMCTMIFEEFELDADNSHIINGHIPVKTKEGENPIK : 559
C.cell   : FIDDKKTHIEKKDPYNFRDKEEVYDNIFSEFGLDPNSHIVNGHVPVEKKGESPIK : 560
C.beij   : FIEEKETHIEKKDPYESLDNEDICKKILKEFGLSSPESHINGHMPVEEKNGESPIK : 558
B.sub    : FIKEKETHKEEKNPYYYLFEDEATCRNILAEFGLNPDHGHIINGHTPVKEIEGEDPIK : 574
    
```

Continuation of Figure 3.3

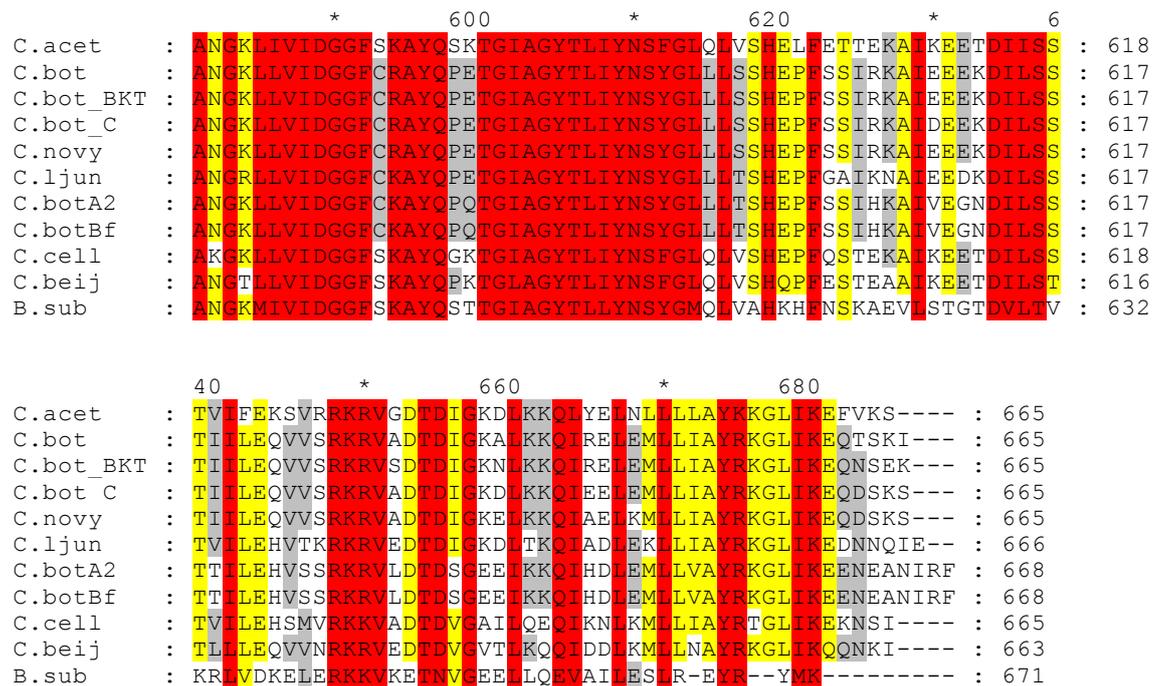


Figure 3.3: Multiple alignment of the putative *C. acetobutylicum* fructose 1,6-bisphosphatase (cac1572) sequence with the top 10 results from the BLAST search.

The deduced amino acid sequence of the putative *C. acetobutylicum* fructose 1,6-bisphosphatase protein is aligned with the fructose 1,6-bisphosphatase protein sequences with the highest percentage identities from the BLAST output. Dashes within the sequences indicate gaps giving optimal alignment. The amino acids are highlighted according to 90% or more (red), 80% (yellow), 60% (grey) and non-highlighted less than 60% conservation. Abbreviations: C.acet, *C. acetobutylicum* ATCC 824; C.cell, *C. cellulovorans* 743B; C.beij, *C. beijerinckii* NCIMB 8052; C.bot, *C. botulinum* D str. 1873; C.bot C, *C. botulinum* C str. Eklund; C.bot BKT, *C. botulinum* BKT015925; C.novy, *Clostridium novyi* NT; C.botA2, *C. botulinum* A2 str. Kyoto; C.botB, *C. botulinum* Bf; C.ljun, *C. ljungdahlii* DSM 13528 and B.sub, *B. subtilis* str. 168.

3.3 Cloning *glpX* and *fbp* genes using TOPO cloning system

3.3.1 Cloning of the *C. acetobutylicum* class II *glpX* gene

The putative class II *glpX*-like gene (*cac1088*) was amplified by PCR using the TOPO*glpX* reverse and forward primers (Table 2.5: Primer details with annealing temperature 52°C and the purified genomic DNA of *C. acetobutylicum* as template, to produce a PCR product of about 1.1 kbp (Figure 3.1). This fragment which contains a single deoxyadenosine (A) at the 3' ends was then cloned into the pCR2.1-TOPO vector (Appendix 7.4) that contains single 3'-thymidine (T) overhangs allowing the PCR product to ligate easily with the vector. Also, this vector contained the *lacZα* gene encoding for α part of beta-galactosidase, an enzyme that converts the colourless X-gal (5-bromo-4-chloro-3-indolyl-[beta]-D-galactopyranoside) modified galactose sugar into a coloured product (blue). In the pCR2.1-TOPO vector, the cloning site is located within the *lacZα* gene (Appendix 7.4). When the foreign gene is cloned into that site, this results in disturbing the *lacZα* gene which in turn cannot produce the beta-galactosidase peptide. Therefore, the colonies that contain the foreign gene would be white and this is the foundation of the blue/white screening technique. The recombinant plasmids were transformed into two chemically competent strains of *E. coli* Mach1TM-T1R and TOP10F^ϕ. The Mach1TM-T1R is a genetically modified *E. coli* strain which harbours a gene (*tonA*) conferring resistance to T1 phage. However, TOP10F^ϕ is a genetically modified *E. coli* strain which harbours the *lacIq* repressor gene, hence, reduces the background expression of the cloned gene. Different amounts of both transformed cells (25, 50, 75, 100, and 125 µl) were plated on five different LB plates (+Amp) for plasmid selection and X-gal for white/blue screening. The efficiency of transformation was poor and the plates had 1, 6, 14, 49, and 74 white colonies, respectively. These colonies were transferred onto fresh plates and given reference numbers. Six colonies were selected (five white and one blue) for PCR screening to check whether these colonies contained the insert or not. PCR screening was carried out using the same TOPO*glpX* reverse and forward primers and the DNA template which was the recombinant plasmid extracted using the boiling method. Four colonies out of 6 were positive (contained the cloned *glpX* gene), 3 were white colonies and 1 was blue (Figure 3.1). These colonies were named and numbered based on two factors; competent cells in which the gene was transformed (M for Mach1TM-T1R and T for TOP10F^ϕ) and colony colour (W for white and B for blue).

Table 3.3: Locations of TOPOglpX primers on the sequence of *cac1088* (*glpX* gene).

Bases underlined represent the location of TOPOglpX primers. Bases highlighted in green, red and gray represent the start codon, stop codon and the *hind*III cutting site respectively. Obtained from Genome Information Broker.

http://gib.genes.nig.ac.jp/single/index.php?spid=Cace_ATCC824

TOPOglpX forward primer →

ACTTTAGGACACTATAGCGCATATATAAATTTTTATGGTTGCAATGTGTTTATAAACAAATGTTTAATAT
 AATCAAGGAGGGAAATTCATGTTTGGATAATGATATATCCATGAGTTTAGTAAGAGTAACAGAAGCAGCAG
 CACTACAATCTTCAAAGTATATGGGAAGAGGAGATAAAAATTGGAGCTGATCAAGCAGCAGTAGATGGAAT
 GGAGAAGGCATTTAGTTTTATGCCAGTAAGAGGCCAGGTTGTAATAGGAGAGGGGAGAACTTGATGAAGCT
 CCTATGCTTTATATAGGTCAAAAGCTTGGTATGGGAAAAGACTATATGCCTGAAATGGATATAGCAGTAG
 ATCCTTTAGATGGAACGATTTAATTTCTAAGGGACTACCTAATGCAATAGCAGTAATAGCAATGGGACC
 AAAAGGAAGTTTACTTCATGCCCCAGATATGTATATGAAGAAAATTGTTGTGGGACCTGGAGCAAAGGT
 GCTATAGATATAAATAAATCTCCTGAAGAGAATATTTTAAATGTAGCAAAGGCATTAACAAGGACATAT
 CTGAATTAACAGTTATAGTTCAAGAAAAGAGAAAAGACATGACTACATAGTAAAAGCAGCTATAGAAGTTGG
 AGCAAGAGTTAAGCTATTTGGTGAGGGCGATGTTGCAGCTGCACCTTGCTTGTGGTTTTGAAGATACTGGA
 ATAGACATACTTATGGGAATTGGAGGAGCTCCAGAAGGAGTTATAGCCGCAGCAGCTATCAAGTGCATGG
 GCGGAGAAATGCAGGCTCAGCTTATACCTCATACTCAGGAAGAAAATAGATAGATGTCACAAAATGGGAAT
 AGATGATGTAAATAAAATTTTCATGATAGATGATTTAGTTAAAAGTGATAATGTGTTTTTGCAGCTACA
 GCAATAACAGAATGTGATCTTCTTAAGGGCATAGTATTTTCTAAAAATGAACGTGCAAAAACCCATTCCA
 TATTAATGAGATCTAAAAGTGGTACAATAAGATTTGTTGAAGCTATTCATGACTTGAATAAAAAGTAAATT
 AGTGGTAGAAATAAAA

← TOPOglpX reverse primer

TTAAAATGTTTCGATTTACGGCGAAGGGGATAGTAAATGCAGGTAA

Therefore, for example the number one blue Mach1™-T1R colony on the plate is referenced as M₁B. White and blue colonies were tested for the presence of the insert using the same TOPOglpX primers which were used to amplify the gene from chromosomal DNA of *C. acetobutylicum*. Also the blue colony, considered to be lacking an insert, was tested as a control. Surprisingly, the blue colony was positive (contained an insert) and was termed M₁B. Moreover, the 3 white colonies, T₁W, M₁W, and M₂W were positive, while the other two white colonies, M₃W and T₂W were negative (no insert).

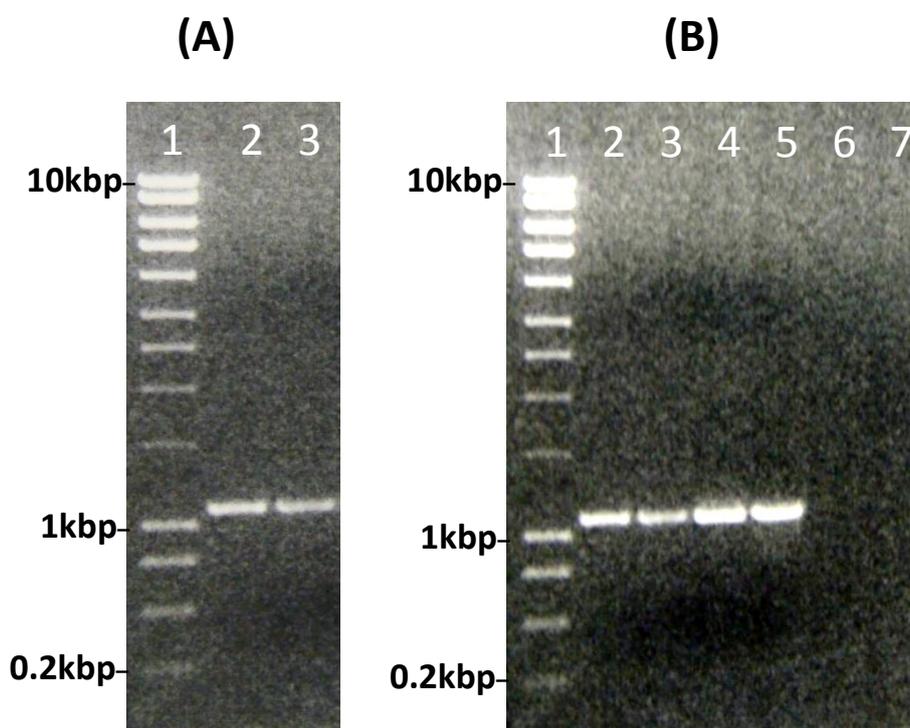


Figure 3.4: Two agarose gels of PCR amplification and screening of *glpX* gene.

(A) Amplification of *C. acetobutylicum* putative class II FBPase gene (*glpX*) using TOPOglpX primers forward and reverse as indicated by lanes 2 and 3. (B) PCR screening of 6 colonies as indicated by lane 2, T₁W strain, lane 3, M₁W strain, lane 4, M₂W, lane 5, M₁B strain, lane 6, T₂W strain, and lane 7, M₃W strain. Lane 1, represents molecular weight makers (hyperladder I) from Bioline.

3.3.2 Determination of the orientation of the insertion using PCR analysis

In the pCR2.1-TOPO vector, the cloned gene can be placed under the control of the *lac* or T7 promoter (appendix 7.4). Therefore, the orientation of the insert needs to be determined using PCR and restriction digest analysis. After testing for the presence of the *glpX* gene, the orientation of the insert (*glpX*) was determined by PCR using one of the TOPOglpX primers (forward or reverse) along with the T7 promoter primer. For a gene under the control of the *lac* promoter, the result will be a fragment of about 1.2 kbp generated by the TOPOglpX forward and T7 promoter primers and there will be no product from the TOPOglpX reverse and T7 promoter primers. However, for a gene under the control of the T7 promoter, the result will be a fragment of about 1.2 kbp generated by the TOPOglpX reverse and T7 promoter primers, with no product for the TOPOglpX forward and T7 promoter primers (Figure 3.5). Only one from the 6 colonies tested gave a product by using the TOPOglpX forward and T7 promoter primers and this colony was the blue one (M₁B) suggesting that the *glpX* gene was under the control of *lac* promoter, and hence, it was selected for further analysis. Strains T₁W, M₁W, and M₂W contained inserts under the control of the T7 promoter. Surprisingly, even though they were white colonies in the plate; M₃W and T₂W did not give any product using either combination of the TOPOglpX forward and T7 promoter primers or the TOPOglpX reverse and T7 promoter primers (Figure 3.6).

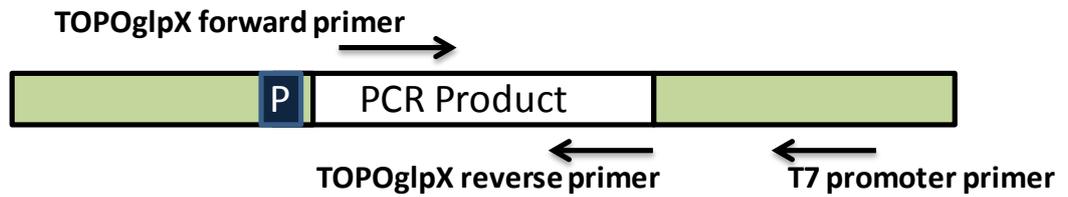
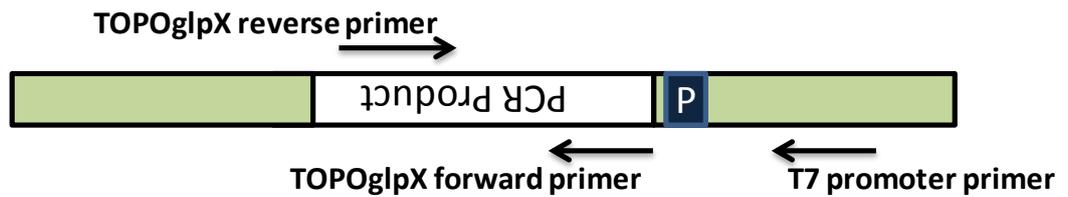
(A)**(B)**

Figure 3.5: Illustration of the strategy to determine the orientation of the insert in the pCR2.1-TOPO vector.

The TOPOglpX reverse, TOPOglpX forward and T7 promoter primers were used in this strategy. (A) When the insert is under the control of *lac* promoter and (B) when the insert is under the control of T7 promoter. P in blue indicates the location of the promoter that controls expression of the PCR product.

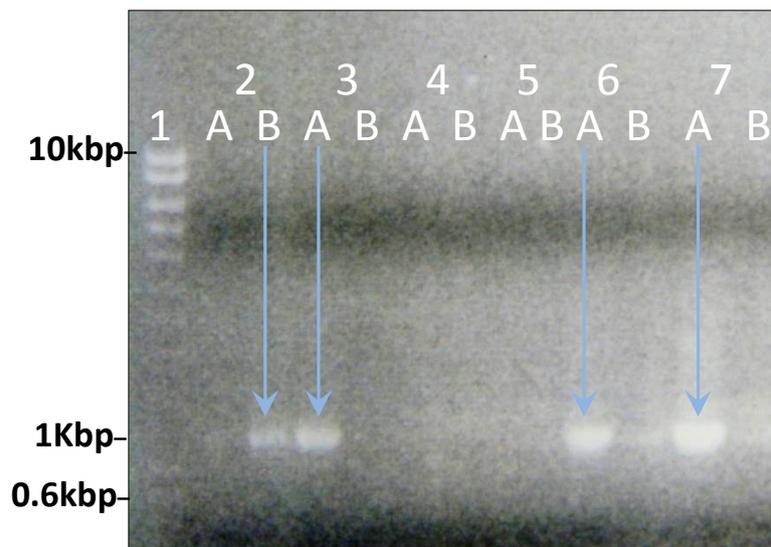


Figure 3.6: Agarose gel of PCR analysis to determine the orientation of the *glpX* gene in the vector.

The lanes A show amplification of the insert under the control of T7 promoter using the TOPO*glpX* reverse and T7 promoter primers. The lanes B show amplification of the insert under the control of *lac* promoter using the TOPO*glpX* forward and T7 promoter primers. The lanes 2A and 2B for strain M₁B, lanes 3A and 3B for strain T₁W, lanes 4A and 4B for strain M₃W (no insert), lanes 5A and 5B for strain T₃W (no insert), lanes 6A and 6B for strain M₁W and lanes 7A and 7B for strain M₂W. Lane 1, represents molecular weight makers (hyperladder I) from Bionline.

Recombinant vectors from strains M₁B, M₁W, and M₂W were purified using the Qiagen Plasmid Midi Kit as described in materials and methods section 2.7.2, and then again screened by PCR to check whether plasmids that harbour *glpX* were isolated or not. All the purified plasmids gave positive results for inserts. Restriction digestions were performed in order to reconfirm the orientation of the inserts. These purified plasmids were used in the restriction digest reaction mixture as described in materials and methods section 2.7.4. The size of the pCR2.1-TOPO vector is 3931 bp and PCR fragment is 1111 bp, hence, the total size of the recombinant plasmid containing *glpX* gene is 5042 bp.

The *EcoRI* enzyme has two restriction sites in the recombinant vector which can be found at 11 bp upstream and at 6 bp downstream of the PCR product (Figure 3.7). Therefore, after *EcoRI* digestion two fragments 1128 bp and 3914 bp should be formed. In the case of pM₁B, pM₁W, and pM₂W plasmids, *EcoRI* gave two products, one was around 1000 bp which represents the insert and the second product was around 4000 bp and that represents the vector (Figure 3.8). These results reconfirm that *glpX* was successfully cloned in the plasmid which is consistent with the PCR screening results

However, *HindIII* restriction sites can be found 60 bp upstream from the insert and 302 bp from the 5' end of the PCR product (Figure 3.7). For the insert under control of the T7 promoter, products of 869 bp and 4173 bp should be formed, however, products of 367 bp and 4675 bp should be formed for insert under the control of the *lac* promoter (Figure 3.7). *HindIII* digestion of the plasmids pM₁W and pM₂W gave products of around 800 bp and 4,000 bp (Figure 3.8), which indicates that the *glpX* gene is under the control of the T7 promoter. Whereas for pM₁B plasmid, *HindIII* produces two products, one was around 5,000 bp and the second one around 400 bp which indicates that the *glpX* gene is under the control of the *lac* promoter (Figure 3.8). These results were consistent with the PCR screening.

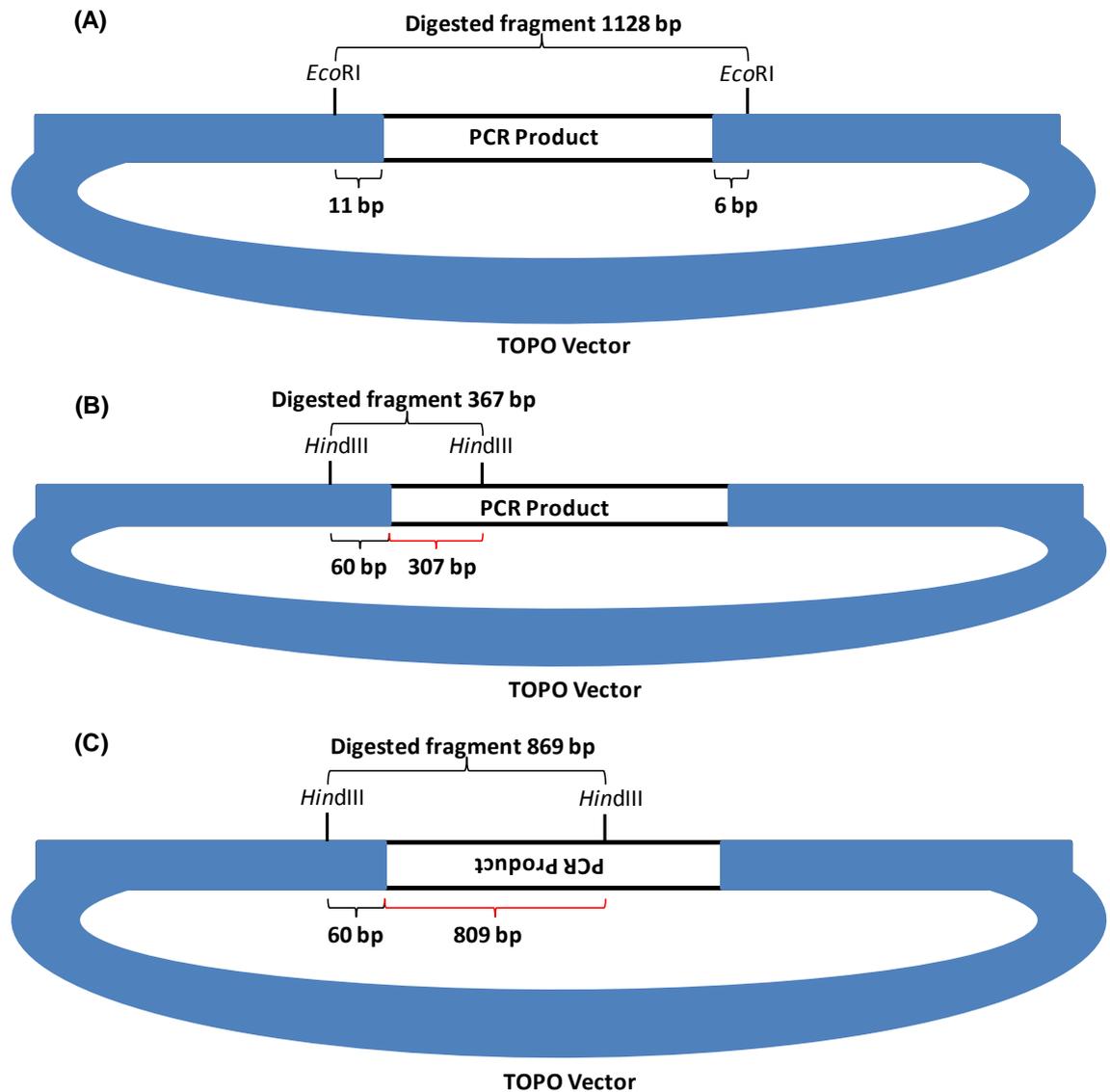


Figure 3.7: Illustration of the strategy to determine the orientation of the insert in the vector by restriction enzymes.

The blue circle represents the TOPO vector (3931 bp) and the white oblong represents the cloned PCR product (1111 bp). (A) *EcoRI* cuts the TOPO vector 11 bp from the 5' end and 6 bp from 3' end of the PCR product. (B) When the insert is under the control of the *lac* promoter, *HindIII* cuts at 60 bp upstream of the 5' end and 307 bp downstream from the 5' end of the PCR product. (C) when the insert is under the control of T7 promoter, *HindIII* cuts at 60 bp upstream from the 5' end and 809 bp downstream from 5' end of the PCR product.

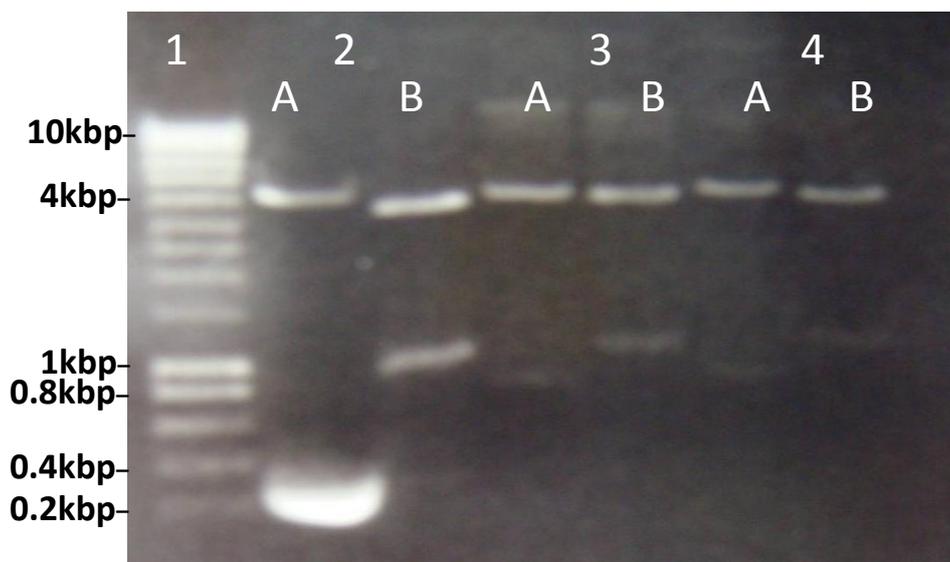


Figure 3.8: Agarose gel showing *Hind*III (A) and *Eco*RI (B) digests of purified plasmid.

In the case of plasmid pM₁B which, *Hind*III produced two products in lane 2A; one was around 5,000 bp and the second one around 400 bp. The *Hind*III digests of the plasmids pM₁W (lane 3A), and pM₂W (lane 4A) gave products of around 800 bp and 4 kbp, whereas, plasmid pM₁B, pM₁W, and pM₂W, *Eco*RI digests produced two fragments, one around 4 kbp and the other around 1 kbp in lanes B2, B3 and B4 respectively. Lane 1, represents molecular weight makers (hyperladder I) from Bionline.

3.3.3 Growth of *E. coli* strains on different selective media for phenotype studies.

The phenotypes of all the *E. coli* mutant strains needed to be checked before transferring the plasmids that harbour the *glpX* gene. LB Agar plates (+Tet) were used to grow the *E. coli* strains JB103, JB108, JLD2402, JLD2403, and JLD2405. All these strains are tetracycline resistant because they carry the *Tn10* transposon (Donahue *et al.*, 2000). Also, the *E. coli* strains JB103, JB108 and JLD2402 all lack a functional *fbp* gene and therefore cannot grow on gluconeogenic substrates such as glycerol. The JB103 strain has multiple copies of the *lac* repressor gene which allows tight regulation of the recombinant gene expression (Lutz & Bujard, 1997). Whereas, the JB108 strain, which also lacks *fbp*, can express T7 RNA polymerase when induced with Isopropyl β -D-1-thiogalactopyranoside (IPTG) (Donahue *et al.*, 2000), and therefore, genes under the control of the T7 promoter can be indirectly induced by the addition of IPTG to the medium. Moreover, JB108 still has a functional *glpX* class II gene, but the expression level of this gene is very low, hence, it is not sufficient for growth on glycerol (Donahue *et al.*, 2000). However, the JLD2402 strain is a double mutant (Δfbp and $\Delta glpX$) (Donahue *et al.*, 2000). The strains were streaked onto M9 minimal medium that contained 0.2% glucose or 0.4% glycerol (+Tet). All mutants grew on the M9 glucose medium as expected. Whereas, only two out of five strains were able to grow on glycerol, JLD2403 (*fbp*⁺, $\Delta glpX$) and JLD2405 (*glpX*⁺, *fbp*⁺) (Figure 3.9), due to both strains harbouring a functional *fbp* gene which enables them to grow on gluconeogenic substrates such as glycerol (Table 3.4). The strains JB103, JB108 and JLD2402 were unable to grow on glycerol due to the Δfbp mutation (Table 3.4). However, all the *E. coli* strains JB103, JB108, JLD2402, JLD2403, and JLD2405 were able to grow on LB (Table 3.4). These results indicate that JB103, JB108 and JLD2402 can be used for genetic complementation of the *fbp* mutation.

Table 3.4: Growth of the <i>E. coli</i> mutants on LB, glycerol and their phenotype.			
Plus sign (+) represents growth and minus sign represents no growth.			
Strain	Genotype	LB	Glycerol
JB103	Δfbp , $glpX^+$ and $lacIq$	+	-
JB108	Δfbp , $glpX^+$ and T7 RNA polymerase	+	-
JLD2402	Δfbp and $\Delta glpX$	+	-
JLD2403	fbp^+ and $\Delta glpX$	+	+
JLD2405	$glpX^+$ and fbp^+	+	+

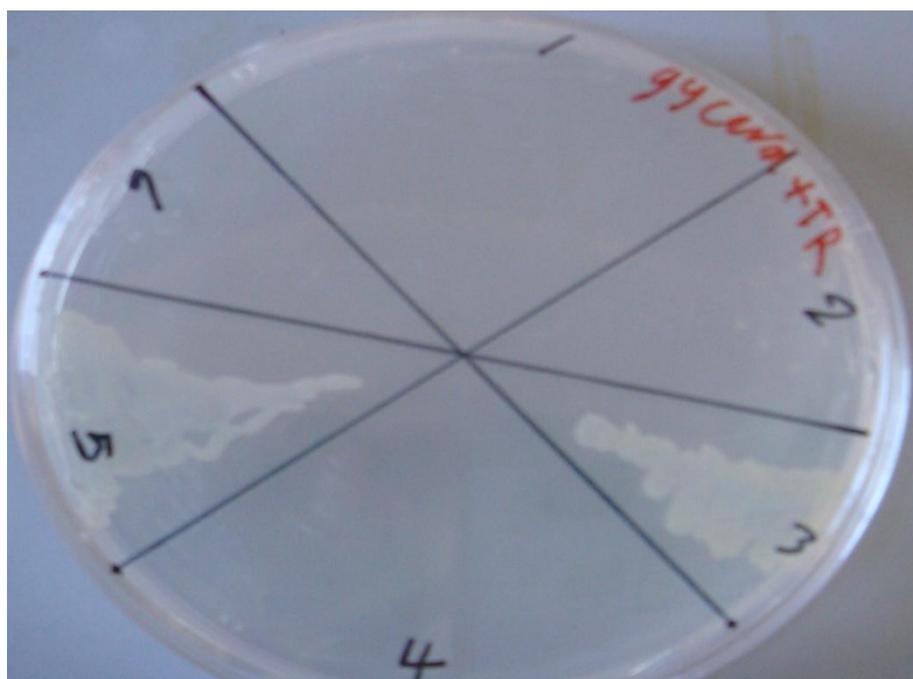


Figure 3.9: Growth of *E. coli* strains on M9 minimal medium containing 0.4% glycerol.

Each section represents, 1: JB103, 2: JB108, 3: JLD2403, 4: JLD2402, 5: JLD2405, 6: JB108.

3.3.4 Chemical transformation and complementation of different *E. coli* mutants.

Since the PCR analysis of plasmids pM₁B and pM₁W matched the restriction analysis in term of insert orientation, the next stage was to transform selected *E. coli* mutant strains with these plasmids that contain *glpX* gene. The aim of this step was to investigate whether the *C. acetobutylicum glpX* gene does indeed encode for FBPase or not by genetic complementation. Therefore, different *E. coli* Δfbp mutants (JB103, JB108 and JLD2402) were used to study various phenotypes exerted by the *C. acetobutylicum glpX* gene and then to select the best transformant for enzyme assay. Purified plasmid pM₁B containing the insert under the control of *lac* promoter was transformed into *E. coli* JB103 and JLD2402 mutants all of which lack a functional *fbp* gene. However, the JB108 strain was transformed as described in section 2.6, using the pM₁W plasmid containing the insert under the control of T7 promoter because it can express T7 RNA polymerase when induced with IPTG.

Due to the vector pCR2.1-TOPO harbouring a gene that encodes for ampicillin resistance, hence, transformants were spread out on LB plates containing ampicillin and tetracycline (all the *E. coli* mutant strains are resistant to tetracycline) to confirm the presence of plasmid in the transformants. After that, each colony was given a name such as JB108/pM₁W strain which indicates the host strain name and transformed plasmid. Then colonies were transferred onto fresh LB plates (+Amp and +Tet), and to M9 glycerol plates (+Amp and +Tet) to observe the differences in the phenotype and whether the *C. acetobutylicum glpX* gene can complement the mutation or not.

The JB108/pM₁W strain was also transferred onto M9 glycerol plates (+Amp and +Tet) supplemented with IPTG to induce the expression of *glpX*. Surprisingly, the JB108/pM₁W strain showed no growth on the M9 glycerol medium containing IPTG (Figure 3.10), whereas, in the absence of IPTG, a lot of colonies were formed. This may be due to the cloned *glpX* gene harbouring an extra sequence upstream of the transcriptional start site that RNA polymerase might recognise as a promoter. Therefore, the cloned *glpX* gene would be expressed from that promoter rather than from T7 promoter which means that IPTG is not needed (Table 3.5).

Another possibility is that addition of IPTG will potentially result in stimulation of expression of the cloned DNA in both directions; and therefore a nonsense mRNA expressed from the lac promoter might interfere with translation of the mRNA derived from the T7 promoter.

Alternatively, the reason why the IPTG inhibits growth on glycerol may be that the overproduction of the foreign clostridial GlpX protein is toxic to the *E. coli* JB108 strain grow on glycerol minimal medium (a poor medium). Further investigations were carried out by streaking JB108/pM₁W on LB agar (+Amp and +Tet) supplemented with IPTG (Figure 3.11) to test whether the effect of the overexpression of the GlpX protein was still present in rich media. The growth on LB agar was not affected by the addition of IPTG and this might indicate that cells can tolerate a high level of GlpX protein when grown on rich medium such as LB (Figure 3.11 and Table 3.5)

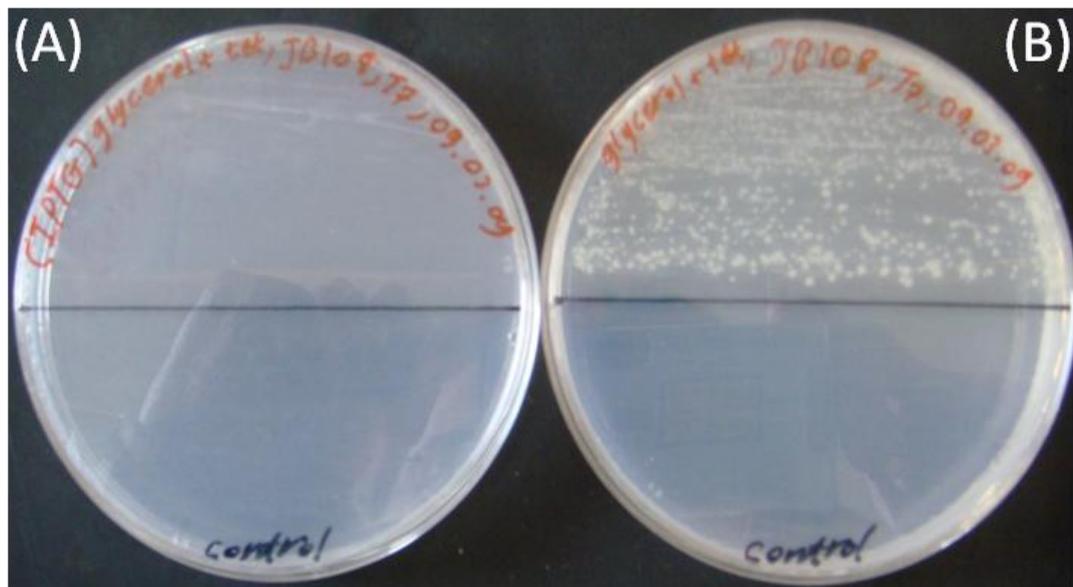


Figure 3.10: Growth of JB108/pM₁W strains on two M9 glycerol plates.

Plate (A) supplemented with IPTG (+Tet), and plate (B) without IPTG (+Tet). Control is the mutant JB108 strain without the recombinant plasmid which harbours *glpX*.

Table 3.5: Growth of different *E. coli* mutant strains in LB and glycerol medium supplemented with or without IPTG.

Plus sign (+) represents growth and minus sign represents no growth

Strains/plasmids	LB	LB (IPTG)	Glycerol	Glycerol (IPTG)
JLD2402/pM ₁ B (<i>lac</i>)	+	Not tested	+	+
JB108/pM ₁ W(T7)	+	+	+	-
JB103/pM ₁ B (<i>lac</i>)	+	Not tested	+	Not tested

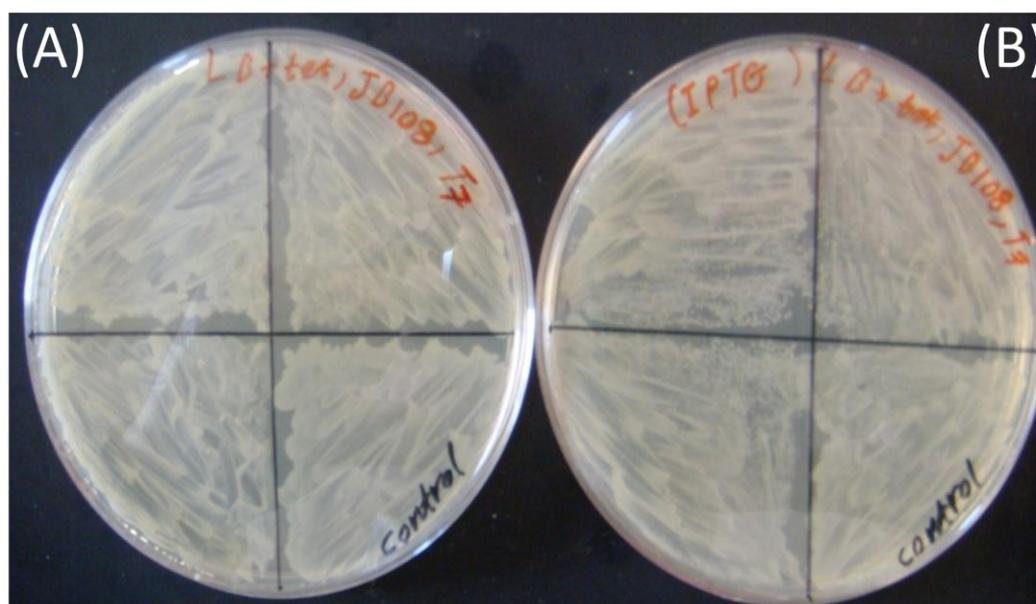
**Figure 3.11: Growth of JB108/pM₁W strains on two LB plates.**

Plate (A) supplemented with IPTG (+Tet), and plate (B) without IPTG (+Tet). Control is the JB108 strain without the recombinant plasmid which harbours *glpX*.

The JB103/pM₁B and JLD2402/pM₁B strains in which the cloned *glpX* gene is under the control of the *lac* promoter, were tested on M9 glycerol medium (+Tet) together with their controls, the JB103 and JLD2402 mutants. Only 15 colonies of JB103/pM₁B and only one colony of JLD2402/pM₁B grew (Table 3.5 and Figure 3.12). Four out of 15 colonies of JB103/pM₁B were transferred onto another M9 glycerol plate (+Tet) for further confirmation (Figure 3.13). All the four colonies were able to thrive on the glycerol plate. This could be due to that the JB103 strain contains a *lac* repressor gene, which would lower the expression of the clostridial *glpX* gene that is under the control of the *lac* promoter, thus reducing toxicity and this would allow the cells to thrive and grow in poor media such as M9 glycerol. These results are consistent with the previous finding that a high expression level of the clostridial *glpX* gene is toxic to the tested JB108 mutant strain under glycerol growth conditions. However, poor growth was still observed after transferring the JLD2402/pM₁B strain onto another M9 glycerol plate (+Tet) (Table 3.5).

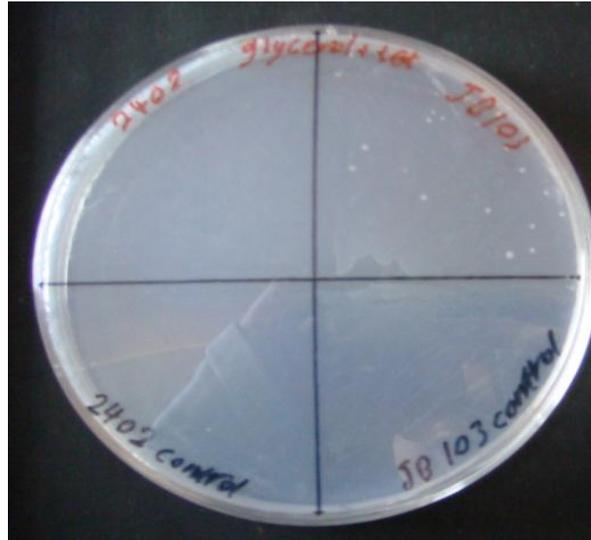


Figure 3.12: Growth of JB103/pM₁W and JLD2402/pM₁B on M9 glycerol plate.

Controls represent the JB103 and JLD2402 strain without the recombinant plasmid which harbours *glpX*. Only one colony of the JLD2402 strain grew and 15 colonies of JB103 grew on glycerol. The controls, JB103 and JLD2402, were unable to grow on glycerol.



Figure 3.13: Growth of JB103/ M₁B on M9 glycerol plate (+Tet).

3.3.5 Fructose biphosphatase assay

Based on genetic complementation results, the JB108/pM₁W strain was selected for assay analysis due to the fact that the strain can tolerate overexpressed GlpX protein and also was the only mutant available which harbours the T7 RNA polymerase to express the *glpX* gene under the control of T7 promoter. A colony of the JB108/pM₁W strain was inoculated into 10 ml of LB broth (+Amp and +Tet) and incubated overnight at 37°C and then 5 ml was transferred into 500 ml of Overnight Express™ Instant LB medium and then incubated overnight at 37°C. The cells were broken using a French Pressure Cell as described in the materials and methods section to produce a crude extract and then the activity was assayed using by Beckman Coulter DU 800 spectrophotometer at 37°C. The protein concentration in crude extract (12.3 mg/ml) was determined by the microbuiuret assay, using bovine serum albumin as a standard.

The crude extract was used in a spectrophotometric coupled enzyme assay to measure the FBPase activity by detecting the reduction of NADP⁺ to NADPH spectrophotometrically at 37°C and wavelength of 340 nm. The activity of the coupling enzymes, glucose-6-phosphatase dehydrogenase and phosphoglucose isomerase were confirmed using fructose 6-phosphate as substrate.

The assay was initiated by adding FBP to the 1 ml mixture. F-6-P will be formed as a result of the FBPase reaction and this was converted to glucose 6-phosphate and then to 6-phosphogluconate by coupling to phosphoglucoisomerase and glucose-6-dehydrogenase. As a result the NADP⁺ will be converted to NADPH which was followed at 340 nm. All the reported classes of FBPase required divalent cations such as Mg²⁺, Mn²⁺, Fe²⁺, Cu²⁺ or Zn²⁺ to maximize the activity (Brown *et al.*, 2009). Therefore, all these divalent cations were tested. Maximum activity was obtained with 10 mM Mn²⁺, whereas no activity was detected using Fe²⁺, Cu²⁺, Ca²⁺ or Zn²⁺. No activity was detected in the absence of a cation. Replacing Mn²⁺ with 10 mM Mg²⁺ reduced the activity about 75%. The effects of 1 mM of ATP, ADP, AMP, PEP, and KH₂PO₄ on the FBPase activity were tested (Table 3.6). The optimal activity was achieved using fructose 1,6-bisphosphate as substrate at pH 8.0. The FBPase activity of GlpX was completely inhibited by the addition of 1 mM inorganic phosphate. Moreover, 1 mM PEP reduced the activity by 16.5%, whereas, an increase in the activity about 20.5% was detected after the addition of 1 mM ADP.

Other metabolites that were tested, AMP and ATP at 1 mM, produced no significant effect on the FBPase activity of GlpX (Table 3.6). These results are comparable to those obtained for GlpX proteins from *M. tuberculosis* and *E. coli*, except for the effect of ADP and PEP. Although strain JB108 has a functional *glpX* gene in the chromosome, it is not expressed at a sufficient level to enable growth on minimal medium supplemented with a gluconeogenic substrate such as glycerol. Therefore, the *glpX* gene in the chromosome is unable to compensate the loss of *fbp* gene (Donahue *et al.*, 2000). Crude extracts from JB108 mutant strain (control) were tested to detect any FBPase activity of the chromosomal *glpX* in the extract and as expected no activity was detected. Brown *et al.*, 2009 reported that the *E. coli* GlpX enzyme is structurally related to the Lithium-sensitive phosphatases which are strongly inhibited by lower than 0.3 mM Li⁺. Also they found that GlpX and YggF (another class II) were not sensitive to lithium even at 10 mM Li⁺, hence, the class II FBPases in *E. coli* belong to the more resistant Li⁺-sensitive phosphatases such as the class IV bifunction IMPase/FBPase enzymes (Brown *et al.*, 2009). This encouraged the testing of Li⁺ as a potential inhibitor. No effect was detected after the addition of 1 mM or even 10 mM Li⁺, which indicates that the *C. acetobutylicum* GlpX belongs to the more resistant Li⁺-sensitive phosphatases. This assay again showed that *glpX* gene encodes for FBPase enzyme.

Table 3.6: The effect of different metabolites (effectors) on the reaction rate of the crude GlpX enzyme compare to the control (no effector).			
Effectors(1 mM)	Reaction rate %	(-)Inhibition or (+)activation rate %	FBPase specific activity (nmol/min/mg)
None	100	0	18.51 ± 0.88
ATP	93.7	-6.3	17.34 ± 0.04
ADP	120.5	+20.5	22.30 ± 0.79
AMP	93.1	-6.9	17.24 ± 0.11
PEP	83.5	-16.5	15.45 ± 0.13
Inorganic phosphate	No activity	No activity	No activity

3.3.6 Analysis of sequenced recombinant plasmid

The recombinant plasmid from the JB108/pM₁W strain that contains the insert under the control of the T7 promoter was purified using a QIAGEN[®] plasmid midi kit. The region that contains the insert was sequenced by Beckman Coulter Genomics using the T7 promoter and M13 reverse primers. No mutations or mismatches were detected in the sequence analysis on comparison with the original putative *C. acetobutylicum* chromosomal *glpX* gene.

3.4 Cloning *fbp* gene using TOPO cloning system.

Based on the BLAST search, the *C. acetobutylicum fbp* gene is an orthologue of the *B. subtilis fbp* gene which belongs to the class III FB Pase enzymes. The *fbp* gene was amplified using the TOPOfbp primers (Table 2.5 and Table 3.7) and *C. acetobutylicum* genomic DNA was used as a template with 49°C as annealing temperature. Two PCR products were seen on an agarose gel, the first fragment was *fbp* with a size of around 2 kbp and the second fragment was an unwanted band of a size around 800 bp (Figure 3.14). Therefore, two different approaches were used to attempt to purify the *fbp* gene. The first approach was using different annealing temperatures (48, 50 and 51°C) to try produce a PCR product around 2 kbp without the unwanted fragment (Figure 3.14). However, the unwanted fragment was still present even after using various annealing temperatures. Therefore, to overcome the problem, the target fragment was purified using a gel purification kit as described in materials and methods section 2.9.2. After gel purification only the target gene of around 2 kbp was seen on agarose gel (Figure 3.14). The purified PCR product was cloned into the pCR2.1-TOPO vector and transformed into *E. coli* Mach1TM-T1^R and TOP10F⁺ which as mentioned before has the *lacI^q* repressor to reduce the background expression of the cloned gene. After transformation, PCR screening was carried out using the same primers, however, after plating the transformants on LB agar supplemented with (+Amp and +Tet) and X-gal for blue white screening, all the colonies were blue which indicated that these transformants did not contain the *fbp* gene. Further confirmation was carried out by PCR screening, however, no PCR product was detected on the agarose gel. These steps were repeated several times with no success. After the *fbp* gene was cloned and expressed using the Ligation Independent Cloning system (LIC) which will be explained later on in the next chapter, attempts to clone it using the TOPO system were discontinued.

Table 3.7: Locations of TOPOfbp primers on the sequence of *cac1572* (*fbp* gene).

Bases underlined represent the location of TOPOfbp primers. Bases highlighted in green and red represent the start codon and stop codon respectively. Obtained from Genome Information Broker.

http://gib.genes.nig.ac.jp/single/index.php?spid=Cace_ATCC824

TOPOfbp forward primer →

CCTTGAAAATGGCTATAAAGTTAAATTTGTTGACGGTAAAAGCATAATAATATAAAAATACCATTAAGACA

CAAACGACACACTGTATACGGAGGGTTAAAATTATTATGCTATTAGAAAAGTAACACCAAAAACGAAGAAA
 TTAAGGACAATTTAAAGTACTTAGTTCTTCTTTTCGAAACAGTACCCAACAATTAACGAAGCAGCTACGGA
 AATAATCAATCTACAGGCTATTTTAAACCTACCTAAAGGAACGGAACACTTTTTATCAGATGTTTCATGGG
 GAATATGAACAATTCATACATGTAAGAACGCCTCTGGAGTAATAAAAAAGAAAAATAGACGATATTT
 TCGGAAATAGACTTATGCAGAGTGAAAAGAAAAGTCTTGCTACGTTGATTTATTATCCGGAGCAGAAGCT
 GGATATAATATTTAAAGCAGGAAAAAATATTGATGACTGGTATAAAAATAACACTGTATAGGCTTATAGAG
 GTTTGCAGGAATGTCTCCTCAAAGTATACTCGTTCTAAAGTAAGAAAAAGCTCTTCCTAAAGAAATTTTCGT
 ATATAATTGAGGAGCTTTTGCATGAACAACCCAAGGGAGTAGATAAGCAGGAATATTATGACGAGATAAT
 AAAGACTATTATAAGCATAGATAGGGCTAAGGAGTTTATAACTGCAATATCAAAGCTTATACAGAGGCTT
 GTAGTAGATAGACTTCACATAATAGGTGATATCTTTGATAGAGGTCCAAGGGCGGATATTATAATGGATA
 AGCTTGAAGAGTATCATGCGGTAGATATTCAATGGGGAAATCATGATATTTTATGGATGGGAGCTGCATC
 CGGTTCTTCAGTATGCATGGCAAATGTAATAAGAATTTCTGCAAGATATGCAAATCTGTCAACTATAGAA
 GATGGTTACGGAATTAATTTGTTGCCATTAGCCACTTTTGTCTATGGACTTTTACGAAAACGATAAGTGCA
 AGAATTTTGAACCTAAGATAGAATCTGATAAAAAGTTATACAGTCAAGGAAATGAACTTATAGGTAAAAT
 GCATAAGGCTATTGCGATAATACAGTTTAAAGCTTGAAGGAGAAGCTATAAAAAAGACATCCTGAGTTTAAAG
 ATGGAACATAGGATGCTTCTAAACAAAATCAACTTTGAAGACAGTACCATAGAATTTGGATGGTAAAAAGT
 ATAAATTAATGATACAAGCTTTCCAACCTATAGATAAAAAATGATCCATATAAACTTATAGATGAGGAGAG
 GGAGGTAGTTGAAAAATTAAGATCTTCATTTGTAAATAGTGAGAACTAACAGGCATGTAAGGTTTTTA
 TTTTCTCATGGAAATCTGTACCTAAAATTTAACTCTAATTTGTTATACCATGGCTGTATACCTTTAAATG
 AGGACGGAACCTTTAAAGAGGTGTTAATAGGAAGTCATAAGTATAAGGGAAAGGCACTTTTAGATAAACT
 TGATGTTTTAGCTAGAAAAAGCTTTTTTTATGAAGAAAATTTCTAAAAACAGTAAGTATGAGAATGATATG
 ATATGGTATCTTTGGTCAGGCCCTTTTTTCGCCACTTTTTGGTAAAGAAAAAATGACAACATTTGAAAGAT
 ATTTTATTGACGATAAGAAGACACATTATGAAAAGAAAGACCCATATTATCATTACAGAGATGATGAAGA
 TATATGTATCAATATTTTAAAGGGAATTTGGACTTGATTCTGAGCAAGCTCATATTATAAATGGGCATGTT
 CCTGTAGAAAAGTAAAAATGGAGAAAATCCAATAAAGGCCAAATGGGAAATTAATAGTTATTGATGGAGGGT
 TTTCAAAGGCATATCAAAGTAAAACCTGGAATCGCTGGATATACCTTAATATATAACTCTTTTGGACTTCA
 ACTTGATCACATGAACTGTTTGAACAACCTGAAAAAGCAATAAAAAGAAACAGATATAATATCTTCT

ACTGTTATATTTGAAAAATCGGTAAAAGAAAAAGAGTAGGAGATACTGATATAGGAAAGATTTGAAAA
GCAGCTCTATGAGTTGAACCTATTGTTTTAGCTTATAAGAAGGGGCTTATTAAGGAATTTGAAAAAGT
AAGATTATGTACAAAAAA

← TOPOfbp reverse primer

GAGAGATTTTATGTCCTCT

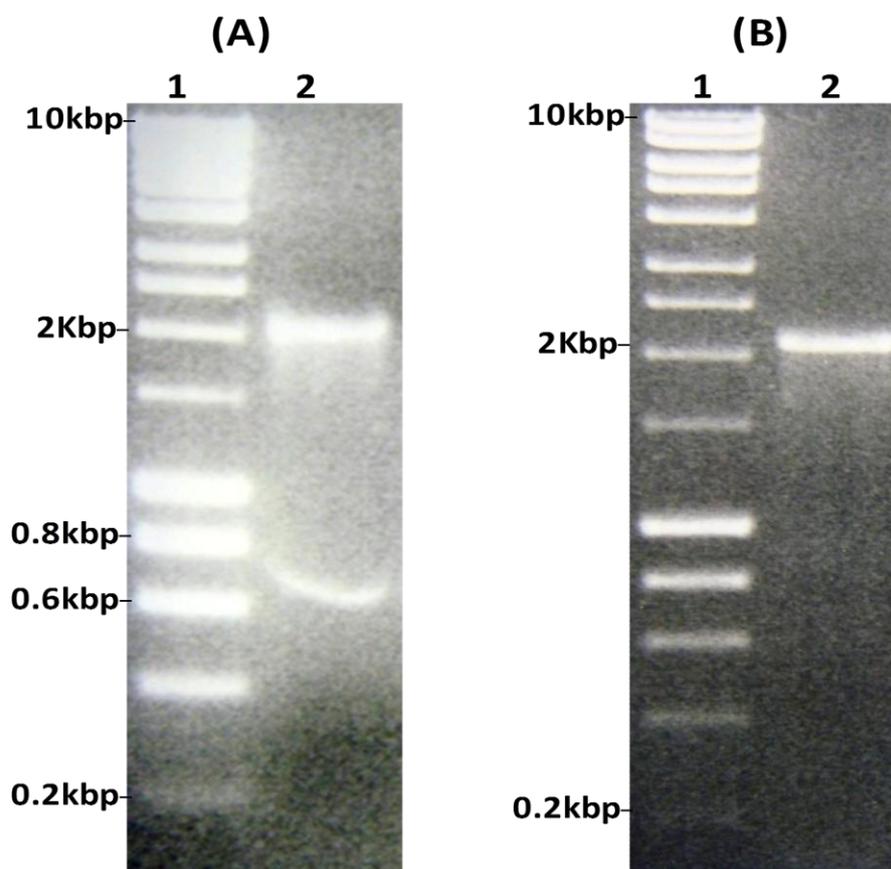


Figure 3.14: Two agarose gels of PCR amplification and purification of *fbp* gene.

(A) amplification of *C. acetobutylicum* putative class III FBPase gene (*fbp*) without purification using TOPOfbp primers forward and reverse and the unwanted fragment as indicated by Lane 2. Lane 1, molecular weight makers (hyperladder I) from Bioline. (B) Purified PCR product (*fbp*) indicated by lane 2. Lane 1, molecular weight makers (hyperladder I) from Bioline.

3.5 Discussion

3.5.1 Cloning of *glpX* gene using TOPO cloning system

The bifunctional HPr kinase/phosphorylase is the most important element of carbon catabolite repression in low GC Gram positive bacteria. It requires fructose 1,6-bisphosphate for activation of kinase activity, however, inorganic phosphate is required for activation of phosphorylase activity (Kravanja *et al.*, 1999; Jault *et al.*, 2000). Genome sequence analysis and experimentation on *C. acetobutylicum* ATCC 824 have revealed the presence of a HPr kinase/phosphorylase gene, *hprK*, which encodes a protein of 304 amino acids and a molecular mass of around 34.5 kDa. Moreover, this solventogenic organism shows a unique gene arrangement in which the HPr kinase/phosphorylase gene (*hprK*) is followed by a gene encoding a putative class II FBPase (*glpX*), hence, this gene might have a role in carbon catabolite repression (Tangney *et al.*, 2003). Until now, five classes of FBPase have been reported, and that includes the class I, encoded by *fbp* which plays an essential role in the gluconeogenesis pathway and is found in almost all organisms (Sato *et al.*, 2004). A class II enzyme was first described in *E. coli* (Hines *et al.*, 2006), whereas, class III was firstly reported in *B. subtilis* (Fujita *et al.*, 1998). The class V enzyme was found in archaea, and the class IV enzyme was discovered in thermophilic prokaryotes of both domains (Brown *et al.*, 2009).

In *E. coli*, the class I FBPase is the main FBPase which is involved in gluconeogenesis. Hence, a Δfbp strain cannot grow on gluconeogenic substrates such as glycerol. The *glpFKX* operon was found in *E. coli* and *Shigella flexneri*, encoding a glycerol kinase (*glpF*), glycerol facilitator (*glpK*) and a gene of unknown function (*glpX*) (Donahue *et al.*, 2000; Truniger *et al.*, 1992). The protein encoded by *glpX* was almost identical in these two organisms (Truniger *et al.*, 1992). Donahue *et al.*, (2000) reported that the GlpX protein had FBPase activity by overexpressing the *glpX* gene in a *fbp* mutant to complement growth on glycerol. This protein has a molecular weight around 40 kDa and a dimeric structure (Donahue *et al.*, 2000). Brown *et al.*, (2009) reported that another gene encodes for a FBPase class II enzyme (*yggF*). Therefore, there are three genes (*fbp*, *glpX* and *yggF*) which encode for FBPase activity in *E. coli*; however, the functions of *glpX* and *yggF* are still unknown (Donahue *et al.*, 2000; Brown *et al.*, 2009).

Several studies reported that *M. tuberculosis* can use fatty acids as a carbon source, however, there was no gene assigned to encode for class I FBPase activity in the genome sequence of *M. tuberculosis*. The *Rv 1099c* gene encodes a GlpX-like class II FBPase that shares 43% identity to the *E. coli* class II FBPase (Movahedzadeh *et al.*, 2004). This gene complemented *E. coli fbp* mutants lacking FBPase activity indicating that *Rv 1099c* is the main and only gene in *M. tuberculosis* that encodes for the missing FBPase which is the key enzyme in the gluconeogenesis pathway (Movahedzadeh *et al.*, 2004). Also, another bacterium *C. glutamicum*, has only one main class II FBPase enzyme, hence a *fbp* mutant cannot grow on gluconeogenic substrates (Rittmann *et al.*, 2003). The clostridial *glpX* gene encodes a protein showing 45% identity to the *E. coli* class II GlpX and also has a molecular weight of 40 kDa. It was the aim of this study to show that the clostridial GlpX protein has FBPase activity, which may be involved in regulating the FBP levels in the cell, hence, regulating the important element in CCR which is the bifunctional HPr kinase/phosphorylase activity and thus the phosphorylated state of HPr on the serine-46 site.

E. coli mutant strains that lack the class I FBP were grown on Nutrient Agar and minimal media to ensure the stable phenotype which was described by Donahue *et al.*, (2000). As expected, all the strains matched the phenotype, being unable to grow on glycerol due to lacking a functional class I *fbp* gene. It is important to test these mutants on a gluconeogenic substrate, as a mutation that restored the growth on a gluconeogenic substrate such as glycerol would give false results when cells are transformed with the putative clostridial *glpX* gene.

This clostridial *glpX* gene was amplified from genomic DNA of the *C. acetobutylicum* ATTC 824. Using a TOPO TA Cloning[®] Kit from Invitrogen, the PCR product was ligated into the pCR2.1-TOPO vector and then transformed into *E. coli* Mach1[™]-T1[®] and TOP10F[™] cells. The transformation efficiency was poor, with a total of 14 white colonies that may have the insert, many fewer than expected. However, this was a sufficient number of clones to undertake the next stage of screening. The PCR fragments can be inserted into the pCR2.1-TOPO vector in one of two orientations, under the control of the T7 or *lac* promoter.

The desirable orientation is that the *glpX* gene would be placed in the pCR2.1-TOPO vector under control of the *lac* promoter, as this would allow direct control of the *glpX* expression in strains such as JLD2402 and JB103 that lacked both class I and II FBPsases. However, for the JB108 strain the *glpX* gene may be under the control of the T7 promoter, due to this strain having T7 RNA polymerase which can be induced by IPTG to express the *glpX* gene. Therefore, the JB108 strain was transformed with the pCR2.1-TOPO vector containing the *glpX* gene under the control of the T7 promoter.

PCR screening was carried out to verify the orientation of the inserts using the TOPO*glpX* forward and reverse primers along with the T7 promoter primer. Only one PCR product should be produced, when the combination of TOPO*glpX* forward and T7 promoter primers was used for the inserts under the control of *lac* promoter. However, when a combination of TOPO*glpX* reverse and T7 promoter primers was used, one product which is under the T7 promoter would be given and no product for TOPO*glpX* forward and T7 promoter primers. All the strains (T₁W, M₁W and M₂W) produced fragments from the combination of TOPO*glpX* reverse and T7 promoter primers indicating that the insert was under the control of T7 promoter. Except one strain termed M₁B which produced a fragment from the combination of TOPO*glpX* forward and T7 promoter primers, indicating that the insert was under the control of *lac* promoter. This indicated that something was powerfully discriminating against an insert under the control of the *lac* promoter. Surprisingly, the strain M₁B produced blue colonies when plated on a medium containing X-gal, even though an insert was present in the pCR2.1-TOPO vector.

Restriction analysis of the recombinant plasmids from the strains M₁B, M₁W, and M₂W also indicated that the inserts in both M₁W, and M₂W were under the control of the T7 promoter. However, the *glpX* gene in strain M₁B was under the control of the *lac* promoter. Therefore, these results were consistent with PCR screening. Two *fbp* mutant strains JB103 and JLD2402 were transformed with the M₁B plasmid in which the insert was under the control of *lac* promoter. Also the JB108 strain was transformed with the M₁W plasmid in which the insert was under the control of T7 promoter. The main purpose of using different *E. coli* mutant strains to test genetic complementation was to investigate the phenotypes that were excreted by GlpX in order to find out what is the best *E. coli* mutant that could be used in FBPase assay analysis.

All the JB103, JB108 and JLD2402 mutants were successfully transformed and able to grow on LB (+Amp and +Tet). The mutant strain JLD2402 carries deletions in both the *fbp* and *glpX* genes, which allows testing of the putative clostridial *glpX* FBPase activity. Therefore, the transformed JLD2402 strain was streaked onto M9 glycerol plates (+Tet). Only one colony grew on the plate, but this was enough to prove that *glpX* encodes for an enzyme that has the FBPase activity. Also the JB103 strain which contains a *lac* repressor allowing tight regulation of the insert under the control of the *lac* promoter was tested on M9 glycerol plates (+Tet). Only 15 colonies were able to grow on the glycerol plate but growth of these colonies showed that glycerol could be used as a carbon source. The JB108 strain contained the T7 polymerase which can be induced by IPTG to overexpress a cloned gene. Transformants of JB108 were tested on two different glycerol plates, one with IPTG and the other one without IPTG. The results were, no growth on the IPTG plate and a high level of growth on the plate without IPTG, which indicates that the *E. coli* cells cannot cope with the over expression of the foreign gene under poor growth conditions, which leads to inhibition of growth.

Also, this strain was tested on LB plates one supplemented with IPTG and the other one without IPTG. A high level of growth was seen on both plates, and a possible explanation of this was that on a rich medium such as LB, the cells may tolerate the high level of clostridial GlpX protein but, when transferred to poor environment such as glycerol minimal medium, transformants cannot withstand the overexpression, of the foreign protein and are hence unable to grow. The JB108/pM₁W strain was selected for FBPase assay analysis due to the fact that the GlpX protein can be overexpressed by Overnight ExpressTM Instant LB media with inhibiting the growth.

To assay the activity, the strain JB108/pM₁W was grown on Overnight ExpressTM Instant LB medium, to make crude extracts for testing the FBPase activity. The enzyme activity was demonstrated using a coupled enzyme assay technique which linked the production of fructose 6-phosphate to the reduction of NADP⁺ to NADPH (Jules *et al.*, 2009; Brown, *et al.*, 2009; Donahue *et al.*, 2000). The addition of cell extracts to the reaction mixture resulted in formation of NADPH which can be detected by an increase of the absorbance at 340 nm. Crude cell extract from the mutant JB108 strain was also made to be tested as a control, and as expected no activity was detected for this mutant.

Although as mentioned before, this strain does contain the *glpX* gene, the low level of expression of this gene cannot support growth on gluconeogenic conditions and enzyme activity cannot be detected by the coupled enzyme reaction (Donahue *et al.*, 2000).

The GlpX activity was optimized using Mn^{2+} as other reported FBPase enzymes such as class I, II, and III need divalent cations as such Mg^{2+} , Zn^{2+} , and Mn^{2+} for maximum activity (Brown *et al.*, 2009; Donahue *et al.*, 2000). Also other divalent cations including Fe^{2+} , Cu^{2+} , Ca^{2+} and Ni^{2+} were tested but no activity was detected. The effects of AMP, ADP, ATP, and PEP (1 mM) on GlpX activity were tested. ATP and AMP (1 mM) had almost no effect on the activity of *C. acetobutylicum* GlpX enzyme (Table 3.8), whereas, in *E. coli*, the GlpX enzyme was not effected by AMP and ADP (1 mM). Moreover, ATP inhibited the activity only by 12%, but inorganic phosphate inhibited the *E. coli* GlpX activity completely (Table 3.8) (Brown, *et al.*, 2009; Donahue *et al.*, 2000). On the other hand, ATP and AMP (1 mM) had no effect on the *C. glutamicum* GlpX enzyme (Rittmann *et al.*, 2003). Moreover, inorganic phosphate inhibited the activity completely, suggesting that phosphate may have a regulatory role on GlpX activity in *C. acetobutylicum* ATCC 824 (Table 3.8).

PEP (1 mM) stimulated the GlpX activity in *E. coli* by 1.7-fold, however, no effect was detected in *M. tuberculosis* GlpX (Table 3.8). But in *B. subtilis*, PEP (1 mM) inhibited the GlpX activity completely (Jules *et al.*, 2009). In this study, however, PEP inhibited the GlpX activity only by 16.5%. However, ADP (1 mM) does not exert any effect on GlpX activity for any of the bacteria reported (Table 3.8). Instead, in this study, GlpX activity was stimulated by 20.5% (Table 3.8). The complete inhibition via inorganic phosphate was explained by solving the 3D structure of *E. coli* in which inorganic phosphate was shown to block the active site of the enzyme (Brown, *et al.*, 2009). This indicates that same mechanism might be responsible for complete inhibition by inorganic phosphate in *C. acetobutylicum*. Also, *C. acetobutylicum* GlpX was shown to belong to resistant Li^+ -sensitive phosphatases such as *E. coli* GlpX due to the fact it was not completely inhibited by even 10 mM Li^+ . Unlike the *M. tuberculosis* GlpX which is inhibited by more than 90% at 2.5 mM.

microorganism	Effectors (inhibit or stimulate FBPase activity of GlpX)					Source
	ATP	ADP	AMP	PEP	Phosphate	
<i>E. coli</i>	Inhibited by 12% (1 mM)	No effect (1 mM)	No effect (1 mM)	Stimulate by 1.7-fold (1mM)	Inhibited completely (1 mM)	Brown <i>et al.</i> , 2009; Donahue <i>et al.</i> , 2000
<i>M. tuberculosis</i>	Not tested	No effect (1 mM)	No effect (1 mM)	No effect (1 mM)	Not tested	Movahedzadeh <i>et al.</i> , 2004
<i>C. glutamicum</i>	No effect (1 mM)	No effect (1 mM)	Inhibited completely (90 µM)	Inhibited only by 50% (360 µM)	Inhibited only by 50% (1.6 mM)	Rittmann <i>et al.</i> , 2003
<i>C. acetobutylicum</i>	Inhibited by 6.3% (1 mM)	Stimulated by 20.5% (1 mM)	Inhibited by 6.9% (1 mM)	Inhibited by 16.5% (1 mM)	Inhibited completely (1 mM)	This study
<i>B. subtilis</i>	Not tested	Not tested	Not tested	Inhibited completely (1 mM)	Not tested	Jules <i>et al.</i> , 2009

3.5.2 Cloning of *fbp* gene using TOPO cloning system

The putative *fbp* gene in *C. acetobutylicum* encodes for a class III Fbp enzyme. To date, this class has only reported in *B. subtilis* (Fujita *et al.*, 1998; Jules *et al.*, 2009). Based on bioinformatic analysis the size of in *C. acetobutylicum* enzyme was 77.23 kDa which is almost twice the size of the GlpX class II enzyme (34.81 kDa). Therefore, this enzyme needs to be investigated as a second potential FBPase protein in *C. acetobutylicum*. The assay results of *B. subtilis* class III FBPase enzyme indicate that PEP is needed for full activity and adding PEP (1 mM) to the assay reaction mix resulted in 30-fold stimulation compared to a control with no effector added to the assay reaction mix (Jules *et al.*, 2009). Also in *B. subtilis*, this gene enables the cell to grow on a gluconeogenic substrate such as glycerol even after disruption of the *glpX* gene which might imply that the *fbp* gene is the main FBPase enzyme in *B. subtilis* that is involved in the gluconeogenesis pathway. However, recent results revealed that *glpX* gene can substitute for the FBPase after *fbp* has been knocked out (Jules *et al.*, 2009).

Significant problems were encountered in cloning the *C. acetobutylicum fbp* gene using the TOPO cloning system. After amplification of *fbp*, two PCR products were detected, the first fragment was the target which was around 2 kbp and the second fragment was small unwanted band of around 800 bp. It was attempted to solve this problem by two different techniques, the first was changing the annealing temperature from 49 to 48, 50 and 51°C with no success. Secondly, the amplified *fbp* was purified using gel purification kit to ensure that only the target fragment which was *fbp* gene would be cloned into the vector. After the transformation, no clones were detected that had the *fbp* gene which might indicate that the *fbp* gene was toxic to the *E. coli* host strain which would prevent the growth of the transformed cells. Problems in cloning genes encoding an FBPase enzyme have been reported by Col, (2004). He could not clone the *yggF* which is the third gene in *E. coli* that encodes FBPase enzyme, but he had no explanations for this problem. However, this gene was cloned later (after 5 years) by Brown *et al.*, (2009) using an expression vector obtained from the Genobase collection.

3.6 Conclusion

In this chapter, the open reading frame *cac1088* was investigated using bioinformatic analysis which showed that this open reading frame might encode for a class II FBPase enzyme (GlpX). The putative *glpX* gene was cloned using TOPO cloning technology, and transformed into an *E. coli fbp* mutant that cannot grow on glycerol. As a result, the transformed cells were able to grow on glycerol which indicates that *glpX* does indeed encode for FBPase activity. Different phenotypes were detected by growing the transformants which carry *glpX* on glycerol and LB agar with or without IPTG. The JB108/pM₁W strain could grow on LB agar with or without IPTG, but no growth was detected when the JB108/pM₁W strain was streaked on glycerol agar supplemented with IPTG. Instead, cells were able to grow on glycerol without IPTG. This might imply that overexpression of the *C. acetobutylicum glpX* gene is toxic to the *E. coli* mutant (JB108) on this medium. FBPase activity was detected by assaying crude extract of the JB108/pM₁W strain in the presence of 10 mM Mn²⁺. The GlpX enzyme was completely inhibited by 1 mM inorganic phosphate and stimulated about 20.5% by 1 mM ADP. To date, this is the first study to report FBPase activity in clostridia. Attempts to clone the putative *fbp* gene, which might encode for Fbp class III were unsuccessful using TOPO cloning technology.

Chapter 4:

4 GlpX and Fbp overexpression and purification

4.1 Gateway cloning system

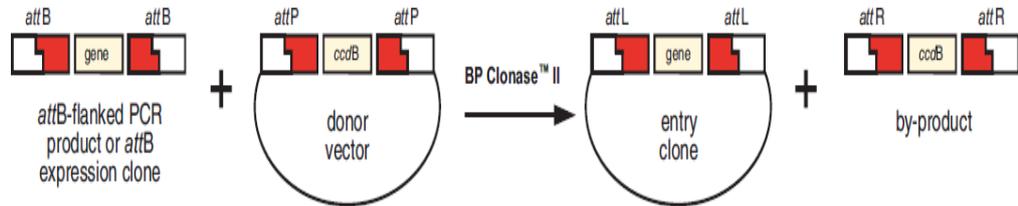
4.1.1 Introduction

To assay FBPase activity of purified GlpX and Fbp enzymes, cloning, overexpression of both proteins and then purification needed to be carried out. Different overexpression strategies were explored in order to obtain the possible highest expression and purity level without aggregation. The first expression system that was explored was the Gateway expression system (Invitrogen®). The Gateway® Technology is based on two infectious stages of the bacteriophage lambda in *E. coli*. Bacteriophage lambda can enter the lysogenic pathway, in which it integrates itself into the genomic DNA of *E. coli* at a specific site or alternatively it releases active progeny viruses in a process called lytic pathway. In the lysogenic pathway, the integrated phage can either be excised from the host genome and go back to lytic behavior or can be transferred to daughter cells. There are four *att* binding sites, termed *attP*, *attB*, *attL*, and *attR* respectively, that govern the integration and excision process of the phage lambda DNA catalysed by integrase (Int) and excisionase (Xis) enzymes encoded by the phage lambda genome, and the *E.coli* integration host factor (IHF). The integration reaction, called the BP reaction (*attB*×*attP*) and the excision reaction, called the LR reaction (*attL*×*attR*). The gene of interest can be integrated or excised *in vitro* by the BP or LR reactions, and this is the foundation of the Gateway® cloning Technology.

Attempts were made to express the *fbp* and *glpX* of *C. acetobutylicum* genes by the Gateway system, two reactions were applied by creating an entry clone for each gene with the Donor vector pDONR™221 (BP reaction) and then transferring each gene to the expression vector pET-60-DEST (LR reaction). PCR primers were designed to amplify each gene flanked by two *attB* overhangs (*attB1* and *attB2*) to enable the recombination with two *attP* regions (*attP1* and *attP2*) in the donor vector pDONR™221 (BP reaction) to form an entry clone with two *attL* sites (*attL1* and *attL2*) (Figure 4.1).

After the integration reaction, the entry clone was mixed in the LR reaction with the expression vector (pET-60-DEST) (Appendix) that contains *attR* sites (*attR1* and *attR2*), so that each gene was transferred from the entry clone (pDONR™221) (Appendix) to the expression vector (pET-60-DEST).

(A)



(B)

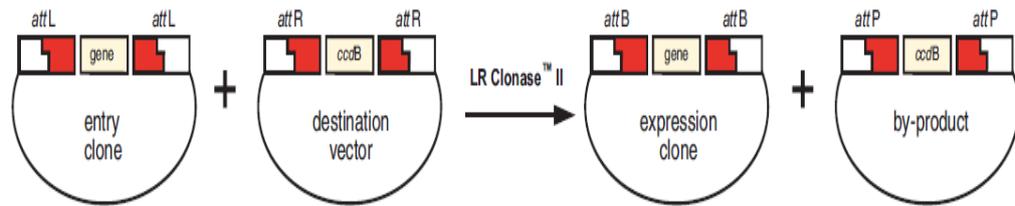


Figure 4.1: Illustration of the two BP and LR reactions of Gateway[®] cloning Technology.

(A) BP Reaction: Facilitates recombination of an *attB* substrate (*attB*-PCR product) with an *attP* substrate (donor vector) to create *attL* sites in the entry clone and this reaction is catalyzed by BP Clonase[™] II enzyme mix. (B) LR reaction: Facilitates recombination of an *attL* substrate (entry vector) to create *attB* sites in the expression clone and this reaction is catalyzed by LR Clonase[™] II enzyme mix. Figure taken from Invitrogen Gateway manual, 2009.

4.1.3 Results

4.1.3.1 Cloning of the *glpX* gene into the donor vector pDONRTM221

To clone the *glpX* gene into the donor vector pDONRTM 221 of the Gateway system, the GATE_{glpX} primers (Table 2.5 and Table 4.1) were designed to amplify the *glpX* gene with the *attB* overhangs, with *C. acetobutalicum* genomic DNA used as template. The amplified PCR product consisted of the 975 bp *glpX* gene, flanked by a 31 bp *attB* site at the 5` terminus, and a 30 bp *attB* site at the 3` terminus.

In the integration reaction (BP reaction), the two *attB* sites of the PCR products recognize the two *attP* sites in the pDONRTM 221 vector and are recombined by integrase enzyme to form the entry clone, in which the *glpX* gene was integrated into pDONRTM 221 vector. The reaction mix was transformed into One Shot[®] OmniMAXTM 2-T1R chemically competent *E. coli* and then plated on LB (+Kan) for screening (Section 2.8.2). Cells that contain plasmid without the integrated *glpX* gene could not grow due to the presence of *ccdB* gene (Figure 4.1) which is lethal to *E. coli* by inhibiting gyrase which has the ability to relax the negative supercoiling DNA to allow replication (Bernard, 1996). Therefore, makes the screening for transformants that contain the *glpX* gene straightforward. Twelve Colonies out of over 300 resistant to (+Kan) were selected, and then PCR screening (Section 2.7) was carried out with the same primers that were used to amplify the *glpX* gene. A product of around 1 kbp was observed (Figure 4.2). Thus, indicated that the plasmid containing *glpX* has successfully transformed into *E. coli*. The plasmids from three different selected isolates MAX-*glpX*-1, MAX- *glpX*-8, and MAX *glpX*-10 were purified (Figure 4.3) as described section 2.7.2. These purified plasmids gave different sizes on the gel even though their size around 6 kbp due to the fact that the plasmid can exist in two forms; supercoiled and relaxed (Figure 4.3).

Table 4.1: Locations of GATEglpX primers on the sequence of *cac1088* (*glpX* gene).

Bases underlined represent the locations of GATEglpX primers. Bases highlighted in green and red represent the start codon and stop codon respectively. Obtained from Genome Information Broker.

http://gib.genes.nig.ac.jp/single/index.php?spid=Cace_ATCC824

GATEglpX forward primer →

ATCAAGGAGGGAAATTCATGTTTGATAATGATATATCCATGAGTTTAGTAAGAGTAACAGAAGCAGCAGC
 ACTACAATCTTCAAAGTATATGGGAAGAGGAGATAAAAATTGGAGCTGATCAAGCAGCAGTAGATGGAATG
 GAGAAGGCATTTAGTTTTATGCCAGTAAGAGGCCAGGTTGTAATAGGAGAGGGAGAACTTGATGAAGCTC
 CTATGCTTTATATAGGTCAAAGCTTGGTATGGGAAAAGACTATATGCCTGAAATGGATATAGCAGTAGA
 TCCTTTAGATGGAACGATTTTAATTTCTAAGGGACTACCTAATGCAATAGCAGTAATAGCAATGGGACCA
 AAAGGAAGTTTACTTTCATGCCCCAGATATGTATATGAAGAAAATTGTTGTGGGACCTGGAGCAAAAGGTG
 CTATAGATATAAATAAATCTCCTGAAGAGAATATTTTAAATGTAGCAAAGGCATTAAACAAGGACATATC
 TGAATTAACAGTTATAGTTCAAGAAAAGAGAAAAGACATGACTACATAGTAAAAGCAGCTATAGAAGTTGGA
 GCAAGAGTTAAGCTATTTGGTGAGGGCGATGTTGCAGCTGCACTTGCTTGTGGTTTTGAAGATACTGGAA
 TAGACATACTTATGGGAATTGGAGGAGCTCCAGAAGGAGTTATAGCCGAGCAGCTATCAAGTGCATGGG
 CGGAGAAATGCAGGCTCAGCTTATACCTCATACTCAGGAAGAAAATAGATAGATGTCACAAAATGGGAATA
 GATGATGTAAATAAAAATTTTCATGATAGATGATTTAGTTAAAAGTGATAATGTGTTTTTTGCAGCTACAG
 CAATAACAGAATGTGATCTTCTTAAGGGCATAGTATTTTCTAAAAATGAACGTGCAAAAACCCATTCCAT
 ATTAATGAGATCT

← GATEglpX reverse primer

AAAAGTGGTACAATAAGATTTGTTGAAGCTATTCATGACTTGAATAAAAAGTAAATTAGTGGTAGAAATAA

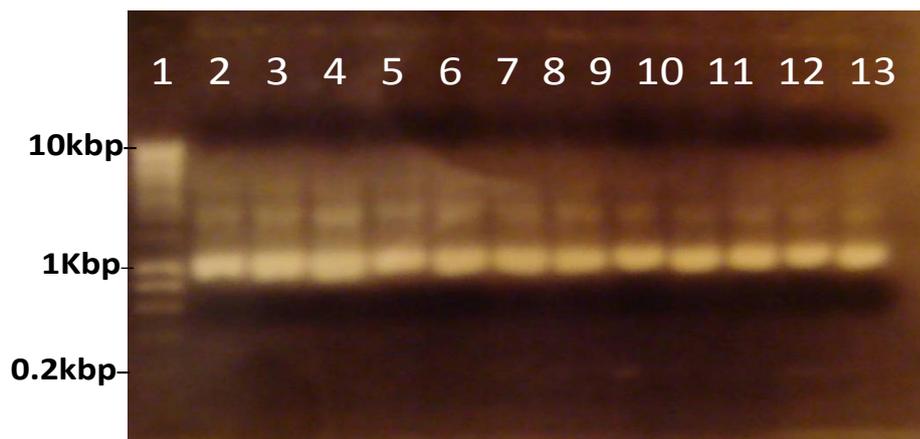


Figure 4.2: Agarose gel of PCR screening of the *glpX* gene in MAX-*glpX* isolates.

Twelve different isolates (MAX-*glpX*) were used for PCR screening as indicated by lane 2 to 13. Lane 1 represents molecular weight makers (hyperladder I) from Bioline.

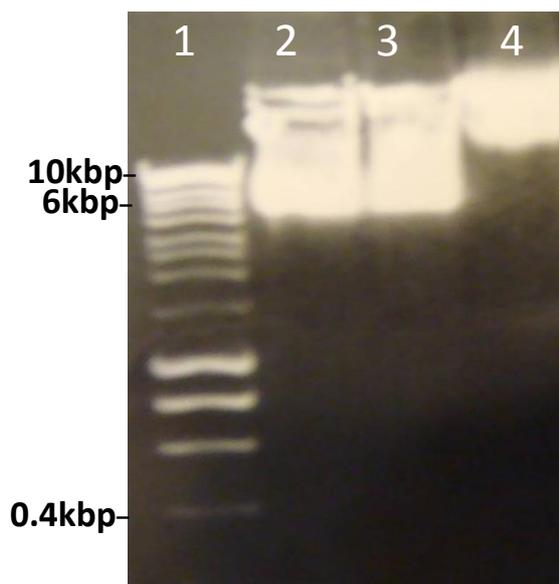


Figure 4.3: Agarose gel of the purified plasmids (pMAX-*glpX*).

Three plasmids (recombinant pDONRTM221 carrying *glpX*) were purified from different colonies. Lane 2, 3, and 4 represent pMAX-*glpX*-1, pMAX-*glpX*-8, and pMAX-*glpX*-10 plasmids respectively. Lane 1, represents molecular weight makers (hyperladder I) from Bioline.

4.1.3.2 Transferring the *glpX* gene into the expression vector pET-60-DEST

The *glpX* gene was transferred from the donor vector to the expression vector pET-60-DEST by the LR reaction (excision reaction), in which the *attL* sites at each end of the *glpX* gene in pDONRTM 221 and the *attR* sites in the expression vector pET-60-DEST were recombined. Thus, the lethal *ccdB* gene was replaced in the pET-60-DEST vector by the *glpX* gene (Figure 4.1). The reaction mixture was transformed into the expression host, *E. coli* RosettaTM 2(DE3), and transformants were selected on LB plated agar (+Amp). Twelve colonies out of over 300 were screened by PCR using the same cloning primers. The PCR products from 12 (termed Rosetta-*glpX*-1-12) different isolates were around 1 kbp as expected (Figure 4.4).

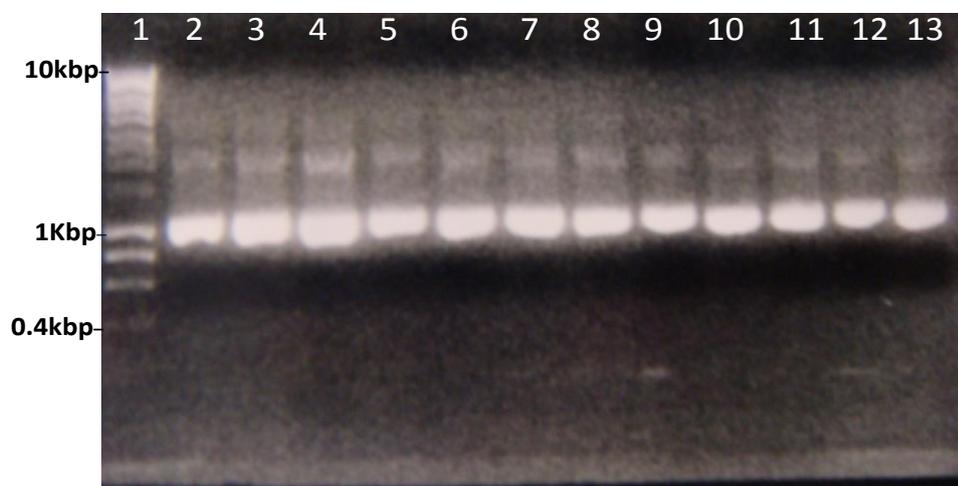


Figure 4.4: Agarose gel of PCR screening of the *glpX* gene in Rosetta-*glpX* isolates. Twelve Rosetta-*glpX* isolates were used for PCR screening as indicated by lane 2 to 13. Lane 1, represents molecular weight makers (hyperladder I) from Bioline.

Two colonies (Rosetta-*glpX*-6 and Rosetta-*glpX*-7) were selected for plasmid purification, and carried out as described in material and methods section 2.7.2. After purification, the two plasmids pRosetta-*glpX*-6 and pRosetta-*glpX*-7 (recombinants pET-60-DEST carrying *glpX*) were examined via sequencing (Section 2.7.3) with GATE_{glpX} primers by Beckman Coulter Genomics. The sequencing result was a disappointment, with a lot of mismatching and unidentified bases, even although Easy-A High-Fidelity PCR Cloning Enzyme had been used to amplify *glpX* for the cloning in the donor vector. The same problem was found also with other students using this system to express different genes.

4.1.3.3 Cloning of the *fbp* gene into the donor vector pDONR™ 221

To clone the *fbp* gene into the donor vector pDONR™ 221 of the Gateway system, the GATEfbp primers were designed to amplify the *fbp* gene with *attB* overhangs (Table 2.5 and Table 4.2), and the *C. acetobutylicum* genomic DNA was used as template. The amplified PCR products consisted of the 2000 bp *fbp* gene, flanked by a 31 bp *attB* site at the 5` terminus, and a 30 bp *attB* site at the 3` terminus.

The *fbp* gene was cloned in the same way as *glpX* (Section 0). After transformation of the recombinant pDONR™ 221 into One Shot® OmniMAX™ 2-T1R chemically competent *E. coli*, transformants were plated on LB agar (+Kan) for screening. Four colonies out of over 250 that were resistant to (+Kan) were isolated for potential insertion and then PCR screening was carried out with the same primers that used in clone *fbp* gene and the product of this pair was the same size as that one used in the cloning, around 2 kbp (Figure 4.5). This Thus indicated the plasmid that contains *fbp* has successfully transformed into One Shot® OmniMAX™ 2-T1R Chemically Competent *E. coli*. The plasmids from two different clones MAX-*fbp*-1 and MAX-*fbp*-2 were purified as described in section 2.7.2 and their size was around 9 kbp (Figure 4.6). After purification, the two plasmids (pMAX-*fbp*-1, pMAX-*fbp*-2) were examined via sequencing (Section 2.7.3) with same cloning primers (GATEfbp primers) by Beckman Coulter Genomics. Again the sequencing result was a disappointment, with a lot of mismatching and unidentified bases.

Table 4.2: Locations of GATEfbp primers on the sequence of *cac1572* (*fbp* gene).

Bases underlined represent the locations of GATEfbp primers. Bases highlighted in green and red represent the start codon and stop codon respectively. Obtained from Genome Information Broker.

http://gib.genes.nig.ac.jp/single/index.php?spid=Cace_ATCC824

GATEfbp forward primer →

CAAACGACACACTGTATACGGAGGGTTAAAATTATTATGCTATTAGAAAGTAACACCAAAAACGAAGAAA
 TTAAGGACAATTTAAAGTACTTAGTTCTTCTTTTCGAAACAGTACCCAACAATTAACGAAGCAGCTACGGA
 AATAATCAATCTACAGGCTATTTTAAACCTACCTAAAGGAACGGAACTTTTTATCAGATGTTTCATGGG
 GAATATGAACAATTCATACATGTACTTAAGAACGCCTCTGGAGTAATAAAAAGAAAAATAGACGATATTT
 TCGGAAATAGACTTATGCAGAGTGAAAAGAAAAGTCTTGCTACGTTGATTTATTTATCCGGAGCAGAAGCT
 GGATATAATATTAAGCAGGAAAAAATATTGATGACTGGTATAAAAATAACTGTATAGGCTTATAGAG
 GTTTGCAGGAATGTCTCCTCAAAGTATACTCGTTCTAAAGTAAGAAAAGCTCTTCCTAAAGAATTTTCGT
 ATATAATTGAGGAGCTTTTGCATGAACAACCCAAGGGAGTAGATAAGCAGGAATATTATGACGAGATAAT
 AAAGACTATTATAAGCATAGATAGGGCTAAGGAGTTTATAACTGCAATATCAAAGCTTATACAGAGGCTT
 GTAGTAGATAGACTTCACATAATAGGTGATATCTTTGATAGAGGTCCAAGGGCGGATATTATAATGGATA
 AGCTTGAAGAGTATCATGCGGTAGATATTCAATGGGGAAATCATGATATTTTATGGATGGGAGCTGCATC
 CGGTTCTTCAGTATGCATGGCAAATGTAATAAGAATTTCTGCAAGATATGCAAATCTGTCAACTATAGAA
 GATGGTTACGGAATTAATTTGTTGCCATTAGCCACTTTTGCTATGGACTTTTACGGAAACGATAAGTGCA
 AGAATTTTGAACCTAAGATAGAATCTGATAAAAAGTTATACAGTCAAGGAAATTGAACCTTATAGGTAAAAT
 GCATAAGGCTATTGCGATAATACAGTTTAAAGCTTGAAGGAGAAGCTATAAAAAGACATCCTGAGTTTAAAG
 ATGGAACATAGGATGCTTCTAAACAAAATCAACTTTGAAGACAGTACCATAGAAATTTGGATGGTAAAAAGT
 ATAAATTAATGATACAAGCTTTCCAACTATAGATAAAAATGATCCATATAAACTTATAGATGAGGAGAG
 GGAGGTAGTTGAAAAATTAAGATCTTCATTTGTAAATAGTGAGAACTAAAACAGGCATGTAAGGTTTTTA
 TTTTCTCATGGAAATCTGTACCTAAAATTTAACTCTAATTTGTTATACCATGGCTGTATACCTTTAAATG
 AGGACGGAACCTTTAAAGAGGTGTTAATAGGAAGTCATAAGTATAAGGGAAAGGCACTTTTAGATAAACT
 TGATGTTTTAGCTAGAAAAAGCTTTTTTTATGAAGAAAATTTCTAAAAACAGTAAGTATGAGAATGATATG
 ATATGGTATCTTTGGTCAGGCCCTTTTTTCGCCACTTTTTTGGTAAAGAAAAAATGACAACATTTGAAAGAT
 ATTTTATTGACGATAAGAAGACACATTATGAAAAGAAAAGACCCATATTATCATTACAGAGATGATGAAGA
 TATATGTATCAATATTTTAAGGGAATTTGGACTTGATTTCTGAGCAAGCTCATATTATAAATGGGCATGTT
 CCTGTAGAAAGTAAAAATGGAGAAAAATCCAATAAAGGCAAAATGGGAAATTAATAGTTATTGATGGAGGGT
 TTTCAAAGGCATATCAAAGTAAAAGTGAATCGCTGGATATACCTTAATATATAACTCTTTTGGACTTCA
 ACTTGTATCACATGAACTGTTTGAACAACCTGAAAAAGCAATAAAAAGAAAGAACAGATATAATATCTTCT
 ACTGTTATATTTGAAAAATCGGTAAGAAAAGAAAAGAGTAGGAGATACTGATATAGGGAAAAGATTTGAAAAA
 GCAGCTCTATGAGTTGAACCTATTGT

← GATEfbp reverse primer

TTTAGCTTATAAGAAGGGGCTTATTAAGGAATTTGTAAAAAGTTAAGATTATGTACAAAAAA

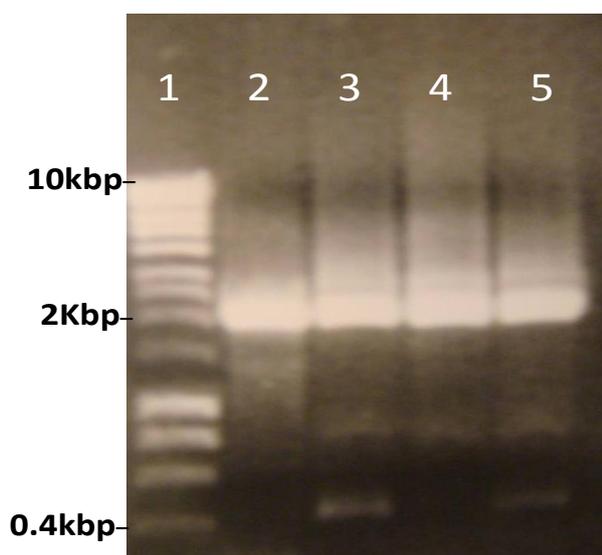


Figure 4.5: Agarose gel of PCR screening of the *fbp* gene in MAX-*fbp* isolates.

Three isolates (MAX-*fbp*-1, 2, 3 and 4) were used for PCR screening as indicated by lanes 2, 3, 4 and 5. Lane 1, represents molecular weight makers (hyperladder I) from Bioline.

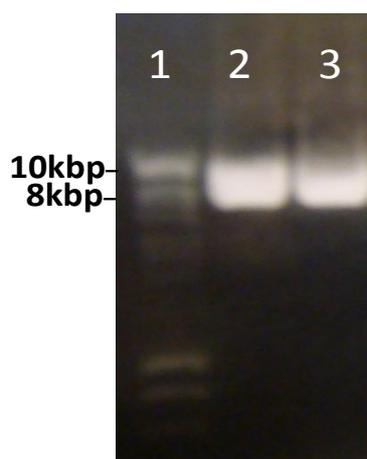


Figure 4.6: Agarose gel of the purified plasmids (pMAX-*fbp*).

Two plasmids (recombinant pDONRTM221 carrying *fbp*) were purified from different colonies (pMAX-*fbp*-1, pMAX-*fbp*-2) as indicated by lanes 2 and 3 respectively. Lane 1, represents molecular weight makers (hyperladder I) from Bioline.

4.2 Cloning and expression of the *glpX* and *fbp* genes using Ligation Independent Cloning system (LIC)

4.2.1 Introduction

To overcome the problem of unsatisfying cloning in the Gateway cloning system, another overexpression system was investigated which is the Ligation Independent Cloning system (LIC) from Novagen. This system was developed for directional cloning of amplified PCR products with no need for restriction enzymes, DNA ligase or alkaline phosphatase (Aslanidis and de Jong 1990). The system uses compatible 15 base single stranded overhangs, which are created in the amplified PCR product by the 3' \rightarrow 5' exonuclease activity of T4 DNA polymerase and anneal to the same complementary 15 base single stranded overhangs in the vector (Figure 4.7).

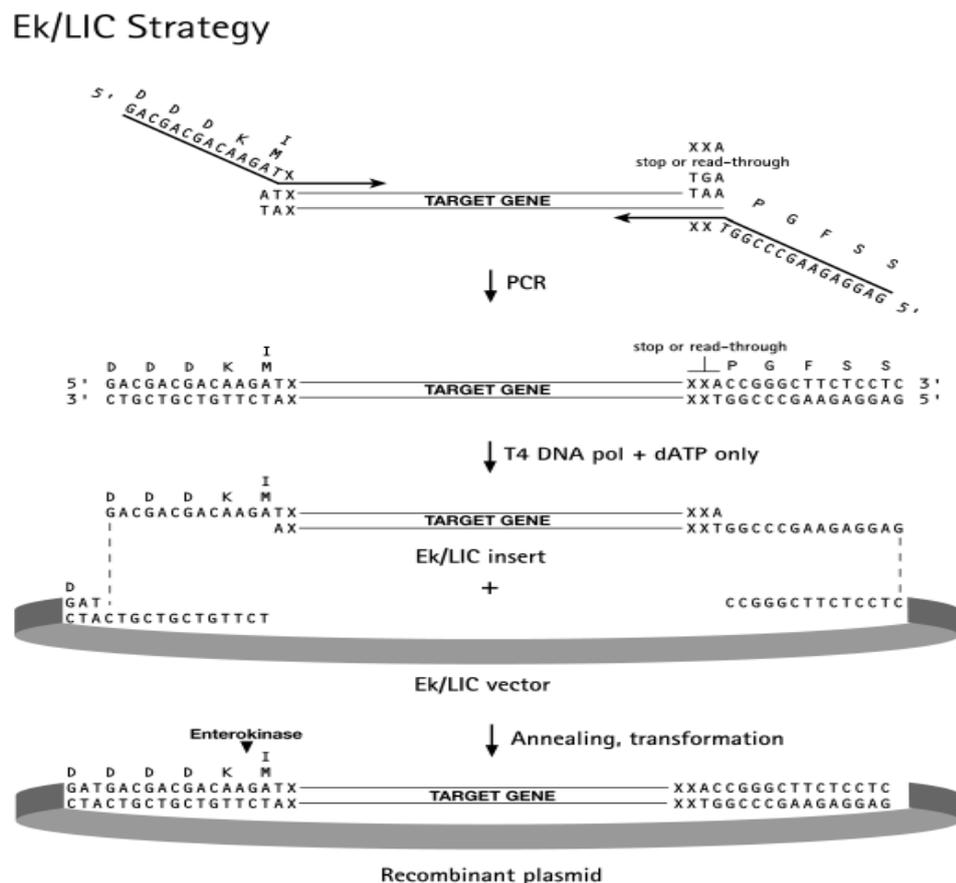


Figure 4.7: Diagram of the Ek/LIC strategy.

After amplification with primers that include the indicated 5' LIC extensions, the PCR insert is treated with LIC-qualified T4 DNA Polymerase (+dATP), annealed to the Ek/LIC vector, and the resultant nicked, circular plasmid is transformed into competent *E. coli*. Figure taken from Novagen[®] LIC manual, 2010.

To create PCR products with the required overhangs, 15 extra bases is added to the primers at the 5' end. The PCR products after the purification are treated by T4 DNA polymerase in the presence of dATP to create specific vector-compatible overhangs and then the treated PCR product is annealed to the pET-41 Ek/LIC vector. To express *glpX* and *fbp* genes by the LIC system, the recombinant plasmid was transformed into the expression host BL21 (DE3) pLysS.

4.2.2 Results

4.2.2.1 Amplification and purification of the *glpX* and *fbp* PCR products

To clone the *glpX* and *fbp* genes into the pET-41 Ek/LIC vector of the LIC system, the LIC $glpX$ primers (Table 4.3 and Table 2.5) were designed to amplify the *glpX* gene and also the LIC fbp primers (Table 4.4 and Table 2.5) were designed (Section 2.1.2) to amplify the *fbp* gene with the extra 15 bases overhangs at each terminus, using the *C. acetobutylicum* genomic DNA as template. For the *glpX* gene, the amplified PCR products consist of the 975 bp of the cloning sequence, flanked by the 15 extra bases at each terminus. For the *fbp* gene, the amplified PCR products consist of the 1998 bp of the cloning sequence, flanked by the 15 extra bases at each terminus (Figure 4.8).

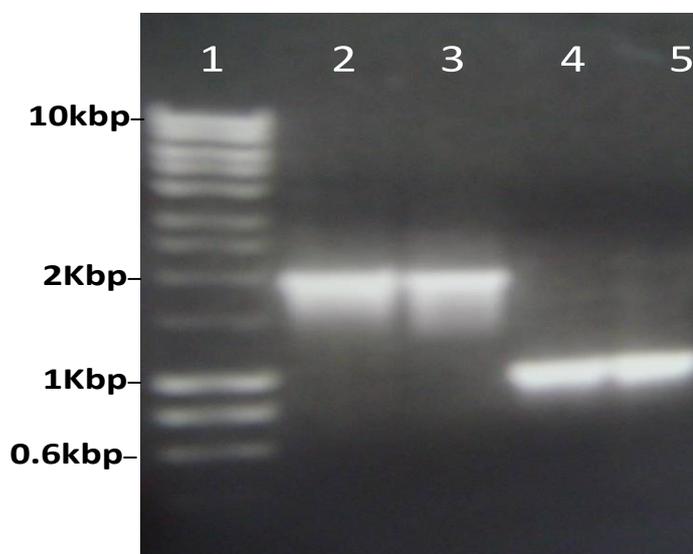


Figure 4.8: Agarose gel of PCR amplification of the *glpX* and *fbp* genes.

Lanes 2 and 3, represent amplified *fbp* and lanes 4 and 5, represent amplified *glpX*. Lane 1, represents molecular weight makers (hyperladder I) from Bioline.

Table 4.3: Locations of LIC $glpX$ primers on the sequence of *cac1088* (*glpX* gene).

Bases underlined represent the locations of LIC $glpX$ primers. Bases highlighted in green and red represent the start codon and stop codon respectively. Obtained from Genome Information Broker.

http://gib.genes.nig.ac.jp/single/index.php?spid=Cace_ATCC824

LIC $glpX$ forward primer →

ATCAAGGAGGGAAATTCATGTTTGATAATGATATATCCATGAGTTTAGTAAGAGTAACAGAAGCAGCAGC
 ACTACAATCTTCAAAGTATATGGGAAGAGGAGATAAAAATTGGAGCTGATCAAGCAGCAGTAGATGGAATG
 GAGAAGGCATTTAGTTTTATGCCAGTAAGAGGCCAGGTTGTAATAGGAGAGGGAGAACTTGATGAAGCTC
 CTATGCTTTATATAGGTCAAAAGCTTGGTATGGGAAAAGACTATATGCCTGAAATGGATATAGCAGTAGA
 TCCTTTAGATGGAACGATTTTAATTTCTAAGGGACTACCTAATGCAATAGCAGTAATAGCAATGGGACCA
 AAAGGAAGTTTACTTTCATGCCCCAGATATGTATATGAAGAAAATTGTTGTGGGACCTGGAGCAAAGGTG
 CTATAGATATAAAATAAATCTCCTGAAGAGAATATTTTAAATGTAGCAAAGGCATTAAACAAGGACATATC
 TGAATTAACAGTTATAGTTCAAGAAAGAGAAAGACATGACTACATAGTAAAAGCAGCTATAGAAGTTGGA
 GCAAGAGTTAAGCTATTTGGTGAGGGCGATGTTGCAGCTGCACTTGCTTGTGGTTTTGAAGATACTGGAA
 TAGACATACTTATGGGAATTGGAGGAGCTCCAGAAGGAGTTATAGCCGCAGCAGCTATCAAGTGCATGGG
 CGGAGAAATGCAGGCTCAGCTTATACCTCATACTCAGGAAGAAATAGATAGATGTCACAAAATGGGAATA
 GATGATGTAAATAAAATTTTCATGATAGATGATTTAGTTAAAAGTGATAATGTGTTTTTTGCAGCTACAG
 CAATAACAGAATGTGATCTTCTTAAGGGCATAGTATTTTCTAAAAATGAACGTGCAAAAACCCATTCCAT
 ATTAATGAGATCT

← LIC $glpX$ reverse primer

AAAAGTGGTACAATAAGATTTGTTGAAGCTATTCATGACTTGAATAAAAAGTAAATTAGTGGTAGAAATAA
 A

Table 4.4: Locations of LICfbp primers on the sequence of *cac1572* (*fbp* gene).

Bases underlined represent the locations of LICfbp primers. Bases highlighted in green and red represent the start codon and stop codon respectively. Obtained from Genome Information Broker.

http://gib.genes.nig.ac.jp/single/index.php?spid=Cace_ATCC824

LICfbp forward primer →

CAAACGACACACTGTATACGGAGGGTAAAATTATTATGCTATTAGAAAAGTAACACCAAAAAACGAAGAAA
 TTAAGGACAATTTAAAGTACTTAGTTCTTCTTTTCGAAACAGTACCCAACAATTAACGAAGCAGCTACGGA
 AATAATCAATCTACAGGCTATTTTAAACCTACCTAAAGGAACGGAACACTTTTTATCAGATGTTTCATGGG
 GAATATGAACAATTCATACATGTACTTAAGAACGCCTCTGGAGTAATAAAAAAGAAAAATAGACGATATTT
 TCGGAAATAGACTTATGCAGAGTGAAAAGAAAAGTCTTGCTACGTTGATTTATATCCGGAGCAGAAGCT
 GGATATAATATTTAAAGCAGGAAAAAAATATTGATGACTGGTATAAAAATAACACTGTATAGGCTTATAGAG
 GTTTGCAGGAATGTCTCCTCAAAGTATACTCGTTCTAAAGTAAGAAAAGCTCTTCCTAAAGAAATTTTCGT
 ATATAATTGAGGAGCTTTTGCATGAACAACCCAAGGGAGTAGATAAGCAGGAATATTATGACGAGATAAT
 AAAGACTATTATAAGCATAGATAGGGCTAAGGAGTTTATAACTGCAATATCAAAGCTTATACAGAGGCTT
 GTAGTAGATAGACTTCACATAATAGGTGATATCTTTGATAGAGGTCCAAGGGCGGATATTATAATGGATA
 AGCTTGAAGAGTATCATGCGGTAGATATTCAATGGGGAAATCATGATATTTTATGGATGGGAGCTGCATC
 CGGTTCTTCAGTATGCATGGCAAATGTAATAAGAATTTCTGCAAGATATGCAAATCTGTCAACTATAGAA
 GATGGTTACGGAATTAATTTGTTGCCATTAGCCACTTTTGCTATGGACTTTTACGAAAACGATAAAGTGCA
 AGAATTTTGAACCTAAGATAGAATCTGATAAAAAGTTATACAGTCAAGGAAATTGAACCTTATAGGTAAAAT
 GCATAAGGCTATTGCGATAATACAGTTTAAAGCTTGAAGGAGAAGCTATAAAAAAGACATCCTGAGTTTAAAG
 ATGGAACATAGGATGCTTCTAAACAAAATCAACTTTGAAGACAGTACCATAGAATTTGGATGGTAAAAAGT
 ATAAATTAATGATACAAGCTTTCCAACTATAGATAAAAAATGATCCATATAAACTTATAGATGAGGAGAG
 GGAGGTAGTTGAAAAATTAAGATCTTCATTTGTAAATAGTGAGAACTAAACAGGCATGTAAGGTTTTTA
 TTTTCTCATGGAAATCTGTACCTAAAATTTAACTCTAATTTGTTTATACCATGGCTGTATACCTTTAAATG
 AGGACGGAACCTTTAAAGAGGTGTTAATAGGAAGTCATAAGTATAAGGGAAAGGCACTTTTAGATAAACT
 TGATGTTTTAGCTAGAAAAAGCTTTTTTTATGAAGAAAATCTAAAAACAGTAAGTATGAGAATGATATG
 ATATGGTATCTTTGGTCAGGCCCTTTTTCGCCACTTTTGGTAAAAGAAAAATGACAACATTTGAAAGAT
 ATTTTATTGACGATAAGAAGACACATTATGAAAAGAAAGACCCATATTATCATTACAGAGATGATGAAGA
 TATATGTATCAATATTTTAAAGGAATTTGGACTTGATTCTGAGCAAGCTCATATTATAAATGGGCATGTT
 CCTGTAGAAAAGTAAAAATGGAGAAAAATCCAATAAAGGCAAAATGGGAAATTAATAGTTATTGATGGAGGGT
 TTTCAAAGGCATATCAAAGTAAAACCTGGAATCGCTGGATATACCTTAATATATAACTCTTTTGGACTTCA
 ACTTGTATCACATGAACTGTTTGAACAACCTGAAAAAGCAATAAAAAGAAACAGATATAATATCTTCT
 ACTGTTATATTTGAAAAATCGGTAAAAGAAAAAGAGTAGGAGATACTGATATAGGGAAAGATTTGAAAAA
 GCAGCTCTATGAGTTGAACCTATTGT

← LICfbp reverse primer

TTTAGCTTATAAGAAGGGGCTTATTAAGGAATTTGTAAAAAGTTAAAGATTATGTACAAAAAAA

In order to clone the *glpX* and *fbp* genes in the LIC system a large amount of ultra pure PCR products was needed. Therefore, PCR amplification was carried out again to allow DNA purification from an agarose gel (Section 2.3). After amplification, 60 μ l of PCR products of both genes instead of 5 μ l was loaded onto the agarose gel for DNA agarose gel purification procedure and purified as described in materials and methods section 2.9.2. This step is to ensure that only the band of interest is purified (Figure 4.9).

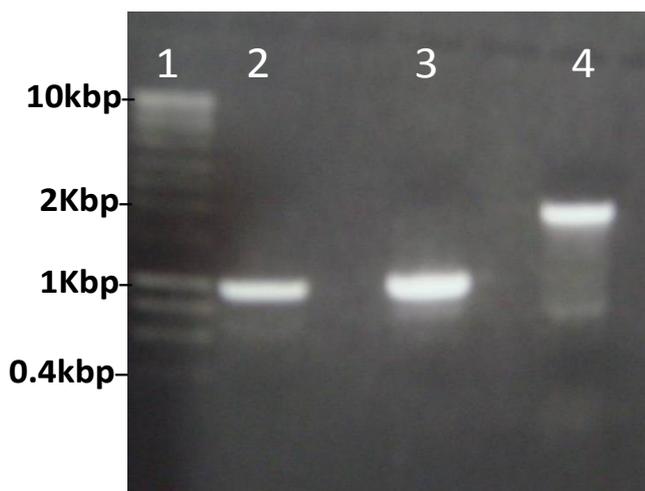


Figure 4.9: Agarose gel of purified PCR products of the *glpX* and *fbp* genes using agarose gel purification (Ultrafree-DA).

Lanes 2 and 3, represent *glpX* and lane 4, *fbp*. Lane 1, represents molecular weight makers (hyperladder I) from Bioline.

Another purification step is also needed to remove dNTPs and residual enzyme and this was done by chloroform extraction and ethanol precipitation method. The samples from the agarose gel purification were extracted with chloroform and DNA was precipitated with isopropanol as described in section 2.9.2 and again run in duplicate on an agarose gel to show purity (Figure 4.10).

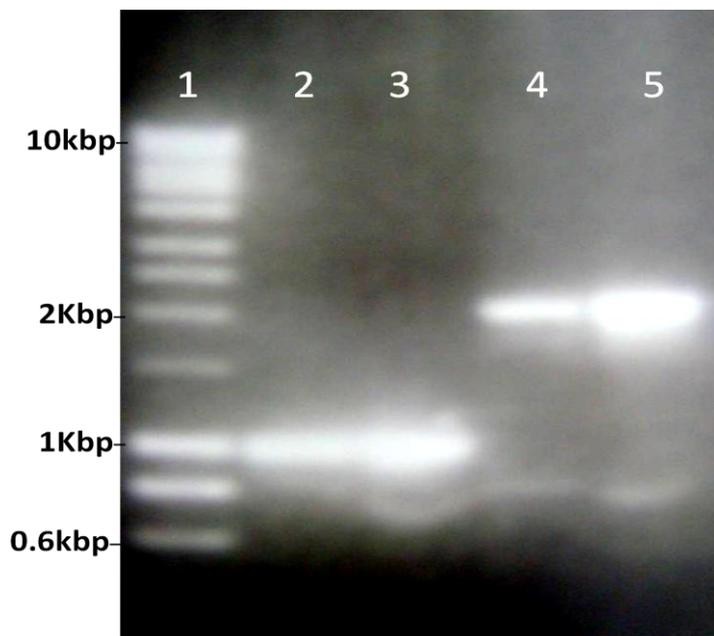


Figure 4.10: Agarose gel of purified PCR products of the *glpX* and *fbp* genes following chloroform extraction and ethanol precipitation.

Lanes 2 and 3, represent *glpX* and lanes 4 and 5, represent *fbp*. Lane 1, represents molecular weight makers (hyperladder I) from Bionline.

4.2.2.2 Cloning the *glpX* into the pET-41 Ek/LIC vector

After purification of the PCR products, T4 DNA polymerase treatment was carried out to create compatible overhangs on the purified insert as described in section 2.9.3. The treated insert was used in the annealing reaction with the pET-41 Ek/LIC vector. The recombinant vector that contained *glpX* insert was transformed into NovaBlue GigaSingles™ Competent Cells (cloning host). The transformants (Nova-*glpX*) were plated on LB (+Kan). Seven colonies out more than 200 were selected for PCR screening using the same primers that were used in cloning *glpX*. All the selected screened colonies contained the plasmid pET-41 Ek/LIC vector that harbour the *glpX* insert (Figure 4.11).

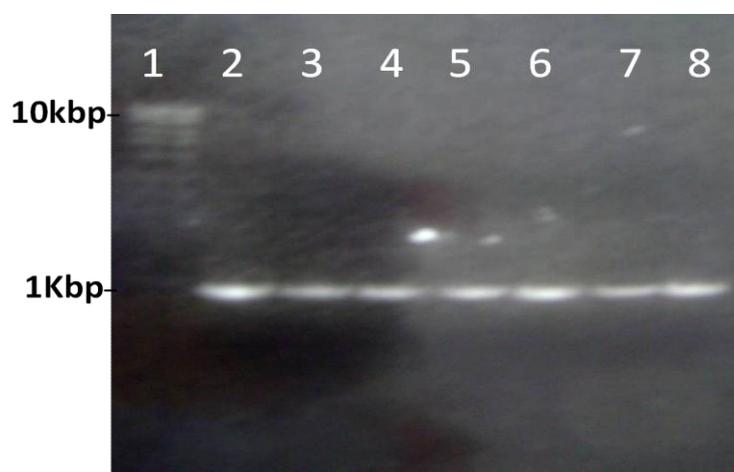


Figure 4.11: Agarose gel of PCR screening of the *glpX* gene in Nova-*glpX* isolates.

Seven isolates (Nova-*glpX*-1 to Nova-*glpX*-7) were PCR screened as indicated by lanes 2 to 8. Lane 1, represents molecular weight makers (hyperladder I) from Bioline.

Two (Nova-*glpX*-3 and Nova-*glpX*-4) from the seven isolates that contain the insert were used for plasmid purification. The recombinant pET-41 Ek/LIC vector was purified as described in section 2.7.2 and their sizes were around 7 kbp (Figure 4.12). The two purified plasmids, pNova-*glpX*-3 and pNova-*glpX*-4, were again screened by PCR using the cloning primers, and as expected both plasmids contained the *glpX* gene (Figure 4.13).

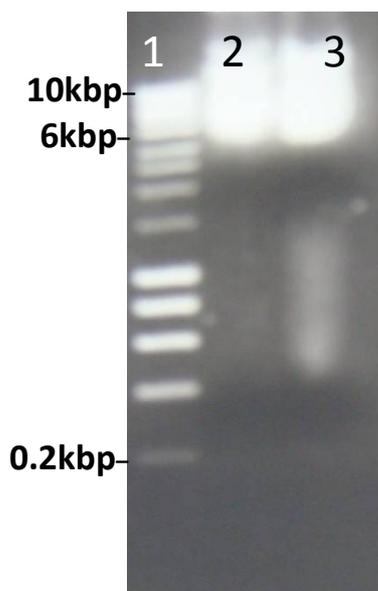


Figure 4.12: Agarose gel of the purified pNova-*glpX* plasmids.

Two plasmids (pNova-*glpX*-3 and pNova-*glpX*-4) were purified from two different colonies (Nova-*glpX*-3 and Nova-*glpX*-4) as indicated by lanes 2 and 3. Lane 1, represents molecular weight makers (hyperladder I) from Bioline.

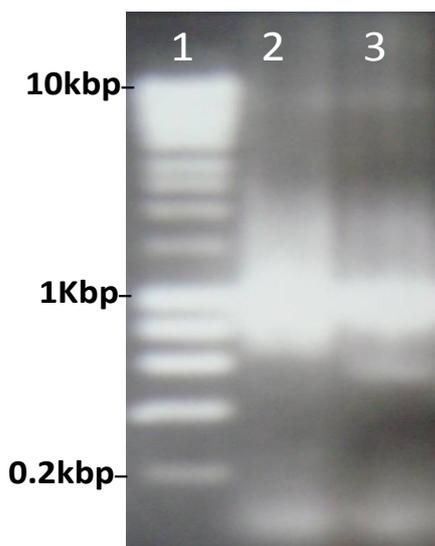


Figure 4.13: Agarose gel of PCR screening of the *glpX* after purification from Nova- *glpX*-3 and Nova- *glpX*-4 isolates.

Two purified plasmid (pNova-*glpX*-3 and pNova-*glpX*-4) were PCR screening as indicated by Lanes 2 and 3. Lane 1, represents molecular weight makers (hyperladder I) from Bioline.

The two purified plasmids pET-41 Ek/LIC containing the *glpX* gene (pNova-*glpX*-3 and pNova-*glpX*-4) were examined via sequencing with same cloning primers. The sequencing results indicate that the *glpX* gene was cloned into the pET-41 Ek/LIC vector.

4.2.2.3 Expression of the *glpX* gene in the recombinant pET-41 Ek/LIC vector

The purified plasmid pNova-*glpX*-3 was selected to be transformed into the expression host, BL21 (DE3) pLysS. This strain contains an IPTG-inducible T7 RNA polymerase, and is designed for protein expression from a pET vector. pLysS means that the strain carries a pACYC184-derived plasmid that encodes T7 lysozyme, which is an inhibitor of T7 RNA polymerase, and hence, represses basal expression of the target gene under the control of the T7 promoter to reduce possible toxicity. Therefore, the plasmid pNova-*glpX*-3 was transformed into the BL21 (DE3) pLysS cells as described in section 2.9.5. The transformants (BL21/pNova-*glpX*-3) were plated on LB (+Kan). To demonstrate that the pNova-*glpX*-3 we successfully transformed into the expression host, PCR screening (Section 2.9.6) was carried out for 4 selected isolates out of more than 200 using the same cloning primers (Figure 4.14).

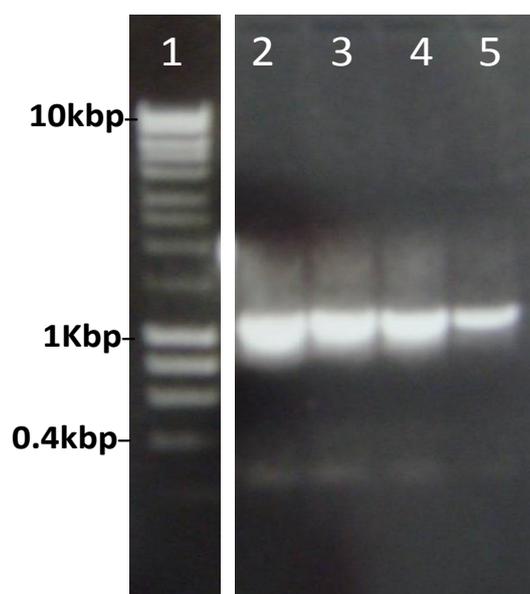


Figure 4.14: Agarose gel of PCR screening of the *glpX* gene in BL21/ pNova-*glpX*-3 isolates.

Four different isolates (BL21₁/pNova-*glpX*-3 to BL21₄/pNova-*glpX*-3) were PCR screened. Lanes 2 to 5, represent amplified *glpX*. Lane 1, represents molecular weight makers (hyperladder I) from Bioline.

To express the *glpX* gene, 500 ml of Overnight ExpressTM Instant LB medium (Section 2.9.7) was used to grow the transformant (BL21₁/pNova-*glpX*-3), and this was incubated overnight at 37°C. The cells were harvested and a crude extract prepared using a French Pressure Cell Press[®] as described in section 2.7.6. The recombinant GlpX protein was purified under native conditions from the crude extract using the Bug Buster[®] Ni-NTA His•Bind Purification kit as described in section 2.9.8, and then all the purified fractions collected were analyzed by SDS-PAGE method as described in section 2.9.9 (Figure 4.15). The recombinant fusion GlpX was purified to homogeneity as estimated by SDS-PAGE with a molecular weight of about 73 kDa (Figure 4.15) which was as expected and is bigger than the native protein which is 34.81 kDa, due to the extra two His-Tags, GST-Tag, and S-Tag (appendix 7.1).

These tags are important for the expressed recombinant protein, since they increase the solubility, the efficiency of purification and prevent aggregation of the recombinant protein. This was demonstrated by analysing the insoluble fraction of the GlpX crude extract together with the 2X diluted insoluble fraction of the GlpX crude extract in SDS-PAGE gel in order to identify any aggregation, however, no sign of aggregation was detected (Figure 4.15). The total molecular weight of all the tags in the recombinant fusion protein is about 37 kDa. Thus, the total expected molecular weight of the recombinant fusion protein is $37 + 34.81 = 71.81$ kDa, which is very close to size of the GlpX (about 73 kDa) seen on the SDS-PAGE gel (Figure 4.15).

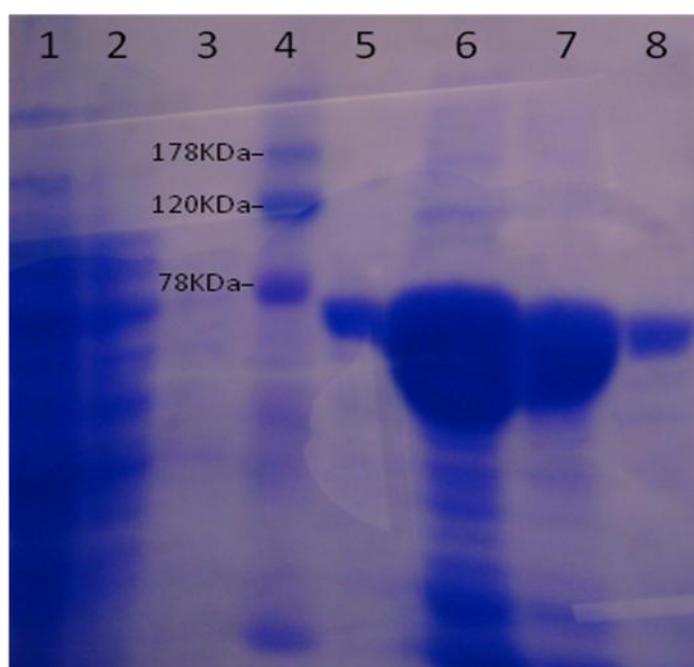


Figure 4.15: SDS-PAGE of purified GlpX proteins in the eluted fractions after His-tag purification under native conditions.

Lanes 1 and 2, insoluble fraction of the crude extract and 2X diluted insoluble fraction of the GlpX crude extract respectively. Lanes 5 to 8, eluate fractions of GlpX. Lane 4, HyperPage prestained protein marker (Bioline).

4.2.2.4 Cleavage of the recombinant fusion GlpX protein

Before assay of the recombinant GlpX enzyme, attempts were made to cut the recombinant fusion tag part of GlpX using enterokinase. This enzyme is able to remove all fused GST-Tag, His-Tag, and S-tag in the N-terminal from the recombinant fusion GlpX by cleaving at the C-terminal end of the lysine residue which is before the start codon site of the target gene (Figure 4.16).

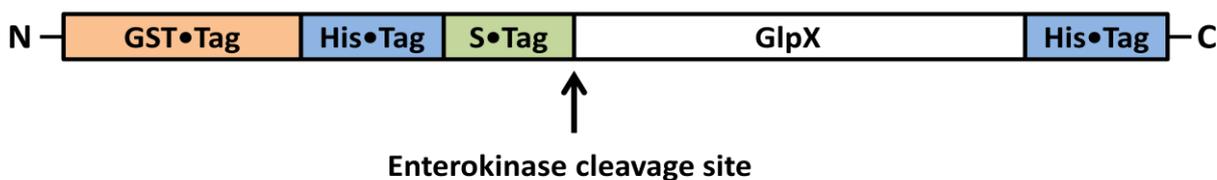


Figure 4.16: Illustration of GlpX protein attached to tagged proteins at the N- and C-terminus. Enterokinase cleavage site is shown by the horizontal arrow.

During purification, the recombinant GlpX protein was eluted using 1 M imidazole. Since the enterokinase enzyme is inhibited by imidazole, the recombinant protein was dialyzed against distilled water using a Dialysis cassette Slide-A-Lyzer as described in materials and methods section 2.9.10, and then the dialyzed recombinant GlpX was added to the reaction mix containing the enterokinase (Section 2.9.11). The enterokinase enzyme appeared however to be unable to cut the tag region from the recombinant GlpX, despite various optimizations of the reaction mix such as increasing the amount of the recombinant protein from 10 to 20 μ l, and changing the incubation temperature (room temperature to 37°C). This might have been due to the inhibition of some remaining imidazole in the reaction mix.

4.2.2.5 Assay of the fructose 1,6 biphosphatase activity of the purified recombinant GlpX protein.

The purified recombinant GlpX was dialyzed against 50 mM Tris-Cl (pH8) containing 1 mM dithiothreitol (DTT) and 5 mM MgCl₂. The recombinant GlpX was assayed as described before in materials and methods section 2.7.6. However, no activity was detected, a similar problem has also reported by Movahedzadeh *et al*, (2004) in that the purification of recombinant GlpX from *M. tuberculosis* led to the loss of FBPase activity.

4.2.2.6 Overexpression of the *glpX* gene in the mutant JB108

As an alternative means of demonstrating FBPase activity associated with the GlpX enzyme, the purified plasmid pNova-*glpX*-3 containing the *glpX* gene was transformed into the expression host, the mutant *E. coli* JB108 strain which is the same strain used for genetic complementation in chapter 3. The transformants (JB108/ pNova-*glpX*-3) were plated on LB (+Kan and +Tet) and more than 200 colonies were able to grow on the plate. To demonstrate that the recombinant plasmid was successfully transformed, two colonies (JB108₁/pNova-*glpX*-3 and JB108₂/pNova-*glpX*-3) were PCR screened using the same cloning primers and both produced a PCR product of around 1 kbp (Figure 4.17).

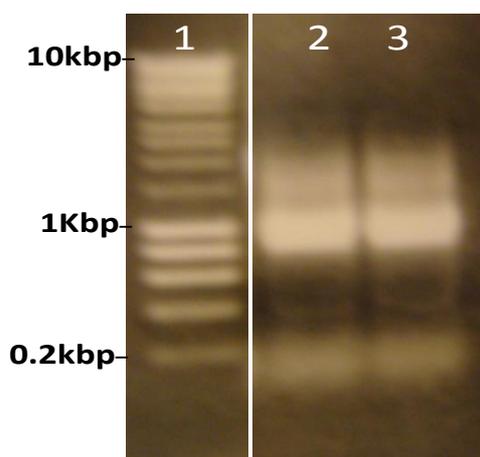


Figure 4.17: Agarose gel of PCR screening of the *glpX* gene in the JB108/ pNova-*glpX*-3 isolates.

Two different isolates JB108₁/pNova-*glpX*-3 and JB108₂/pNova-*glpX*-3 were PCR screened as indicated by lanes 2 and 3. Lane 1, represents molecular weight makers (hyperladder I) from Bionline.

The same overexpression procedure that was used with BL21 (DE3) pLysS was applied to overexpress the GlpX protein from JB108₁/pNova-*glpX*-3 strain to demonstrate whether overexpression was possible in the JB108 mutant. The recombinant GlpX protein was purified under native condition from the crude extract by Bug Buster[®] Ni-NTA His•Bind Purification kit and then all the purified fractions were run in the SDS-PAGE (Figure 4.18). As expected, all the purified proteins were around 73 kDa which was consistent with the previous purification.

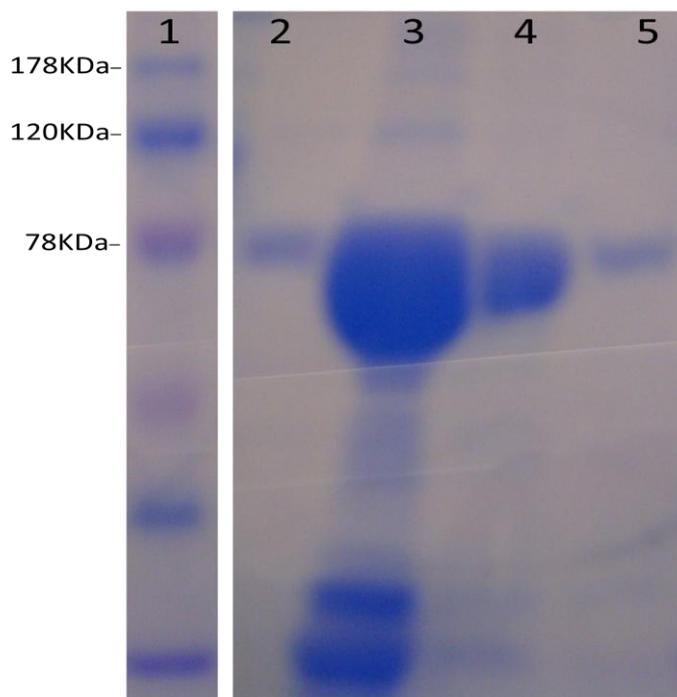


Figure 4.18: SDS-PAGE of purified GlpX proteins in the eluted fractions after His-tag purification under native conditions.

Lanes 2 to 5, eluate fractions of GlpX. Lane 1, represents the HyperPage prestained protein marker (Bioline).

4.2.2.7 Complementation of JB108 mutant

Before assaying the enzyme, the JB108₁/pNova-*glpX*-3 and JB108₂/pNova-*glpX*-3 need to be tested via genetic complementation to know whether the recombinant fusion GlpX still has FBPase activity or not. Therefore, two colonies of JB108₁/pNova-*glpX*-3 and JB108₂/pNova-*glpX*-3 grown on LB plate (+Kan and +Tet) which harbour plasmid pNova-*glpX*-3 were streaked on M9 glycerol plate (+Tet). As expected no growth was detected due to the *glpX* gene being under the control of the T7 promoter, and the JB108 strain does not express the T7 RNA polymerase without the presence of IPTG in the medium (Figure 4.19). Therefore, 50 µl of IPTG (200 mg/ml) was spread onto a M9 glycerol plate (+Tet), and the same two colonies were streaked onto it and then incubated overnight at 37°C. The two colonies were able to grow on the medium, and this indicates that the cloned *glpX* gene is expressed and has an FBPase activity (Figure 4.19).

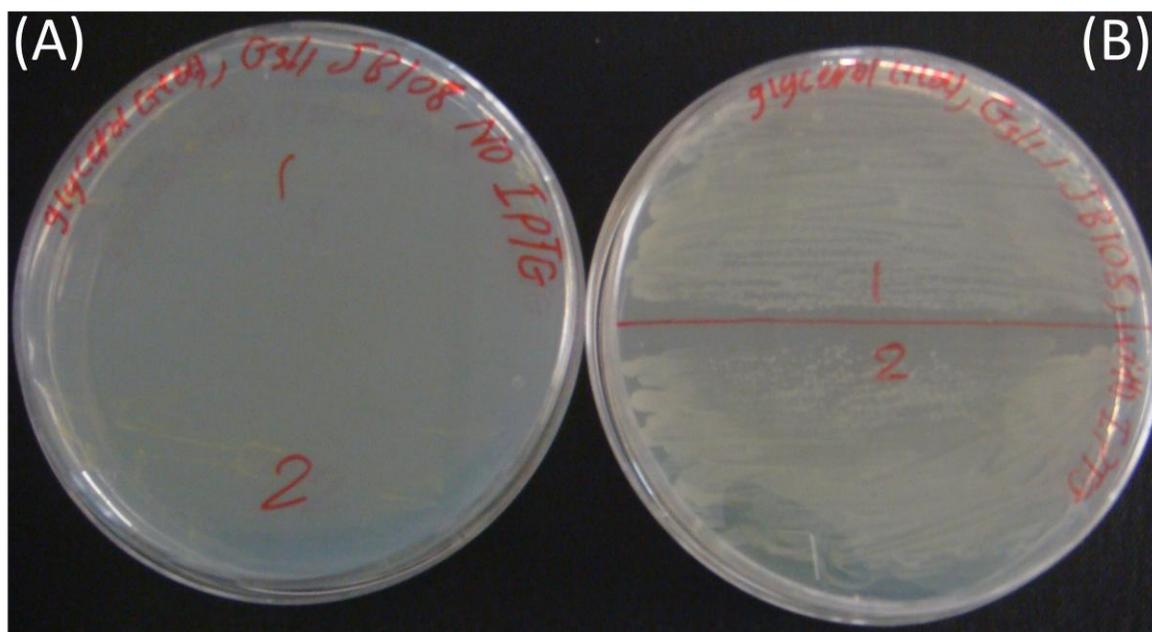


Figure 4.19: Growth of JB108₁/pNova-*glpX*-3 and JB108₂/pNova-*glpX*-3 strain on two M9 glycerol plates.

Plate (A) M9 glycerol medium (+Tet) without IPTG and plate (B) M9 glycerol medium (+Tet) supplemented with IPTG. The JB108₁/pNova-*glpX*-3 and JB108₂/pNova-*glpX*-3 that were tested are indicated by 1 and 2 respectively.

4.2.2.8 Assay the fructose 1,6-biphosphatase activity of the recombinant GlpX from JB108₁/pNova-*glpX*-3 crude extract.

The genetic complementation results shown that the recombinant fusion GlpX had FB Pase activity. This indicates that the recombinant fusion tag part did not affect the activity of GlpX enzyme. This encouraged the testing of the FB Pase activity of recombinant GlpX. Therefore, the JB108₁/pNova-*glpX*-3 strain was grown on the same overexpression media that used to grow the BL21 (DE3) pLysS strain and a crude extract was prepared as described in section 2.7.6. The protein concentration of the crude extract was 23 mg/ml and the FB Pase activity was assayed using fructose 1,6-bisphosphate as substrate at pH 8.0 (Section 2.7.6). The recombinant GlpX from JB108₁/pNova-*glpX*-3 strain had FB Pase specific activity of 29.70 ± 0.61 nmol/min/mg. Different divalent cations were tested (Mn^{2+} , Ca^{2+} , Zn^{2+} , Ni^{2+} , Cu^{2+} and Mg^{2+}), and 10 mM of Mn^{2+} was found to support the activity of GlpX as also shown in section 3.3.5. No activity was detected when Mn^{2+} was omitted from the reaction mix. However, replacing Mn^{2+} with 10 mM of Mg^{2+} led to increase the activity about 2-fold compared to the 10 mM of Mn^{2+} . In contrast, no FB Pase activity was detected when Mn^{2+} was replaced with 10 mM of Zn^{2+} or Ca^{2+} . Moreover, all divalent cations that inhibit the FB Pase activity (Ca^{2+} , Zn^{2+} , Ni^{2+} , Cu^{2+}) were tested as a potential inhibitor for the coupling enzymes that are used in the assay reaction. Ca^{2+} and Zn^{2+} did not inhibit the activity of the coupling enzymes, whereas, Ni^{2+} , Cu^{2+} in inhibited the activity of these enzymes completely.

The FB Pase activity of GlpX was completely inhibited by the addition of 1 mM inorganic phosphate. The PEP (1 mM) reduced the activity by about 26.5%. Other metabolites (effectors) tested such as AMP, ADP and ATP, each at 1 mM each, produced no major effect on the FB Pase activity of GlpX (Table 4.5). These results are compatible to that obtained for GlpX proteins from *M. tuberculosis*, except for PEP (Movahedzadeh *et al.*, 2004). Also, Li^+ was tested as a potential inhibitor. No effect was detected after the addition of 1 mM Li^+ , on the other hand, only 21% inhibition was detected after the addition of 10 mM Li^+ , which reconfirms that the *C. acetobutylicum* GlpX belongs to the more resistant Li^+ -sensitive phosphatases.

Table 4.5: The effect of different metabolites (effectors) on the reaction rate of the crude recombinant GlpX enzyme compare to the control (no effector).

Effector (1mM)	Reaction rate (%)	Inhibition rate (%)	FBPase specific activity (nmol/min/mg)
Control	100	0	29.70 ± 0.61
ATP	93.7	6.3	28.83 ± 0.68
ADP	100	0	29.70 ± 0.08
AMP	100	0	29.70 ± 0.06
PEP	73.5	26.5	21.74 ± 0.49
Inorganic phosphate	No activity	No activity	No activity

4.2.2.9 Cloning and expression of the *fbp* gene

After the second purification of the PCR product (Section 4.2.2.1), T4 DNA polymerase treatment was carried out to create compatible overhangs on the purified insert as described in section 2.9.3. The treated insert was used in the annealing reaction, by adding the treated insert that contained the compatible overhangs to the pET-41 Ek/LIC vector. The recombinant vector that contains *fbp* insert was transformed into NovaBlue GigaSingles™ Competent *E.coli* Cells (cloning host) and the transformants were plated on LB (+Kan). Five different colonies (designated from Nova-*fbp*-1 to Nova-*fbp*-5) were PCR screened (Section 2.9.6) using the same primers that were used in cloning *fbp*. All the selected screened colonies contained the insert (Figure 4.20). Two isolates (Nova-*fbp*-3 and Nova-*fbp*-4) from the five were used for plasmid purification. The recombinant pET-41 Ek/LIC vector was purified as described in section 2.7.2 and gave a band size of around 8 kbp (Figure 4.21). The two purified plasmid (pNova-*fbp*-3 and pNova-*fbp*-4) were PCR screened using the same cloning primers and as expected both plasmid contain *fbp* gene (Figure 4.22). These two purified plasmids were examined via sequencing with same cloning primers, and the sequencing results indicated that the *fbp* gene was perfectly cloned into the pET-41 Ek/LIC vector.

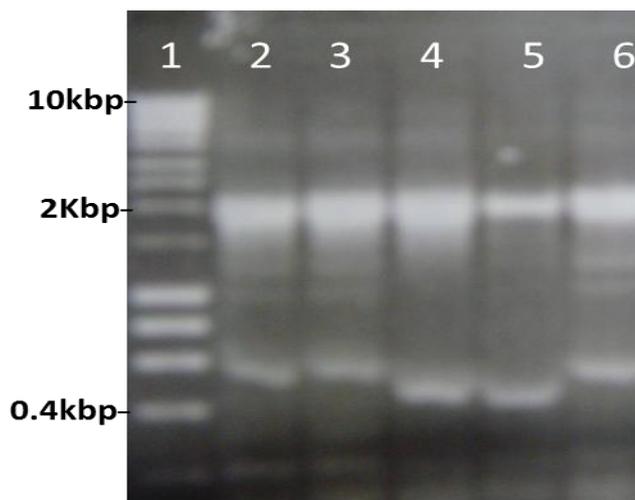


Figure 4.20: Agarose gel of PCR screening of the *fbp* gene in Nova-*fbp* isolates.

Five isolates (Nova-*fbp*) were PCR screened as indicated by lane 2 to 6. Lane 1, represents molecular weight makers (hyperladder I) from Bioline.

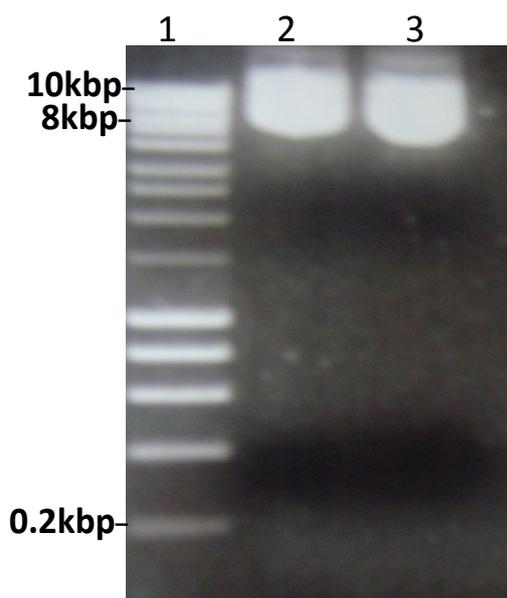


Figure 4.21: Agarose gel of the purified pNova-*fbp* plasmids.

Two plasmids (pNova-*fbp*-3 and pNova-*fbp*-4) were purified from two different colonies as indicated by lanes 2 and 3. Lane 1, represents molecular weight makers (hyperladder I) from Bioline.

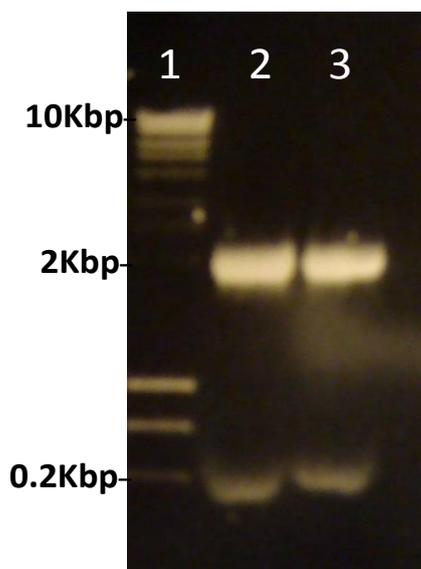


Figure 4.22: Agarose gel of PCR screening of the *fbp* gene after purification.

Two purified plasmid (pNova-*fbp*-3 and pNova-*fbp*-4) were PCR screened as indicated by lanes 2 and 3 respectively. Lane 1, represents molecular weight makers (hyperladder I) from Bioline.

4.2.2.10 Expression of the *fbp* gene in the recombinant pET-41 Ek/LIC vector

The purified plasmid pNova-*fbp*-3 was selected to be transformed into an expression host, BL21 (DE3) pLysS. The transformants (BL21/pNova-*fbp*-3) were plated on LB (+Kan). To show that the recombinant plasmid was successfully transformed, PCR screening (Section 2.9.6) was carried out using the same cloning primers and a 2 kbp band was observed on the agarose gel (Figure 4.23).

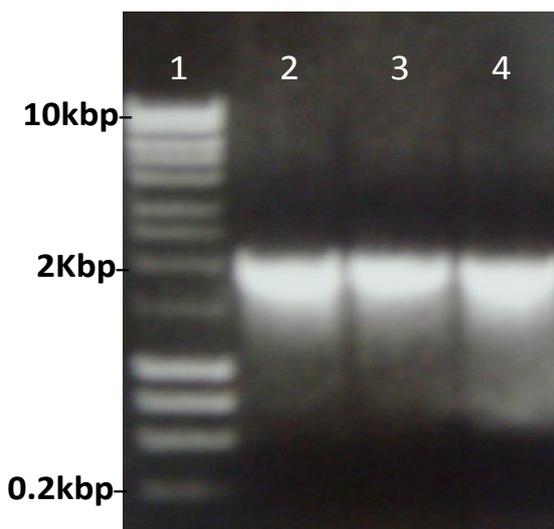


Figure 4.23: Agarose gel of PCR screening of the *fbp* gene in BL21/pNova-*fbp*-3 isolates.

Three different isolates (BL21₁/pNova-*fbp*-3, BL21₂/pNova-*fbp*-3 and BL21₃/pNova-*fbp*-3) were used for PCR screening as indicated by lane 2 to 4. Lane 1 represents molecular weight makers (hyperladder I) from Bioline.

The *fbp* gene was expressed in BL21₁/pNova-*fbp*-3 and purified in the same way as *glpX* gene (Section 4.2.2.3). The recombinant fusion Fbp was purified to homogeneity and estimated by SDS-PAGE to have a molecular weight of about 115 kDa (Figure 4.24), bigger than the native protein which is 77.23 kDa, due to the extra His-Tag, GST-Tag, and S-Tag. The total molecular weight of all the tags in the recombinant fusion protein is about 37 kDa. Thus, the total expected molecular weight of the recombinant fusion protein is $37 + 77.23 = 114.23$ kDa, which is very close to the Fbp size seen on the SDS-PAGE 115 kDa (Figure 4.24). Also, no aggregation was detected in the insoluble fraction of the Fbp crude extract together with the 2X diluted insoluble fraction of the Fbp crude (Figure 4.24).

Attempts to cut the recombinant fusion tag from the recombinant Fbp using enterokinase enzyme was carried out (Section 2.9.11). As expected, the fusion tag in the C-terminal was not removed. Also, an attempt was made to assay the purified recombinant Fbp enzyme, however, no activity was detected.

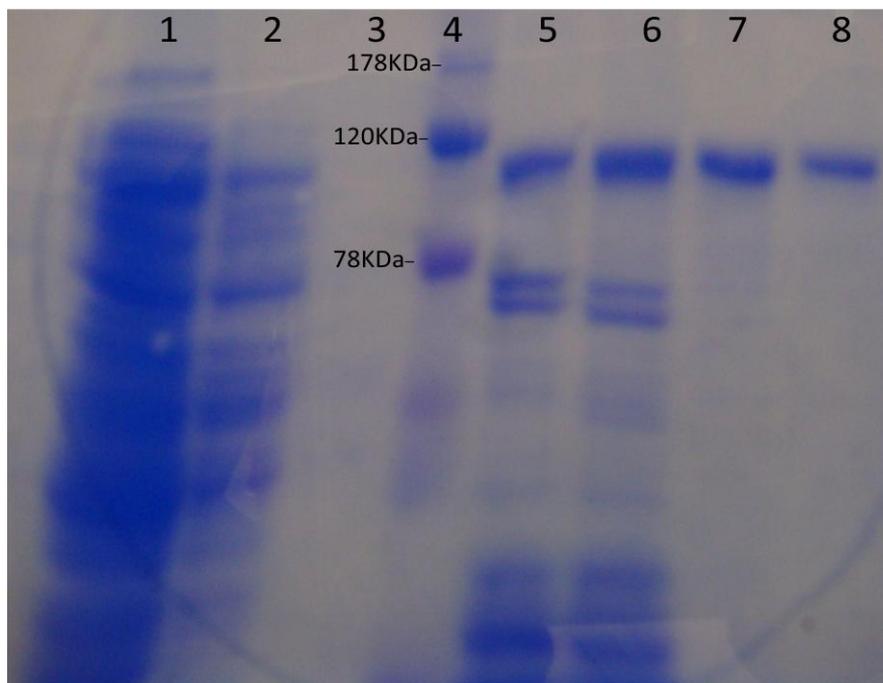


Figure 4.24: SDS-PAGE of the purified Fbp proteins in the eluate fractions after His-tag purification under native conditions.

Lanes 1 and 2, insoluble fractions of crude extract and the 2X diluted insoluble fractions respectively. The lane 3, empty and lane 4, HyperPage prestained protein marker (Bioline). Lane 5 to 8, eluate fractions of Fbp.

4.2.2.11 Expression of the *fbp* gene in the mutant JB108

The problem of unable to detect FB Pase activity was solved by transforming, the purified plasmid pNova-*fbp*-3 into the *E.coli* JB108 mutant strain, the transformants (JB108/pNova-*fbp*-3) were plated on LB supplemented with (+Kan and +Tet). To confirm that the purified plasmid was transformed into the JB108 mutant strain. PCR screening (Section 2.9.6) was carried out using the same cloning primers (Figure 4.25). The *fbp* gene was shown to be expressed in JB108₃/pNova-*fbp*-3 and purified in the same way as the *glpX* gene had been (Section 4.2.2.6) and then all the purified fractions were run in the SDS-PAGE to check whether Fbp was expressed or not (Figure 4.26).

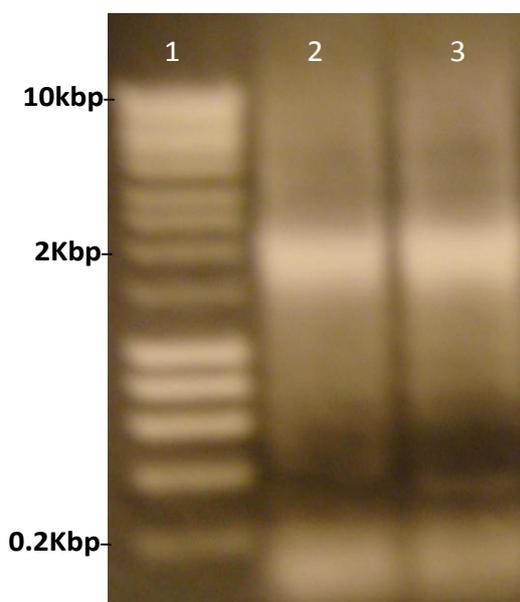


Figure 4.25: Agarose gel of PCR screening of the *fbp* gene in JB108₃/pNova-*fbp*-3 and JB108₄/pNova-*fbp*-3 isolates.

Two different isolates (JB108₃/pNova-*fbp*-3 and JB108₄/pNova-*fbp*-3) were used for PCR screening as indicated by lane 2 and 3, *fbp* fragments. Lane 1, represents molecular weight makers (hyperladder I) from Bioline.

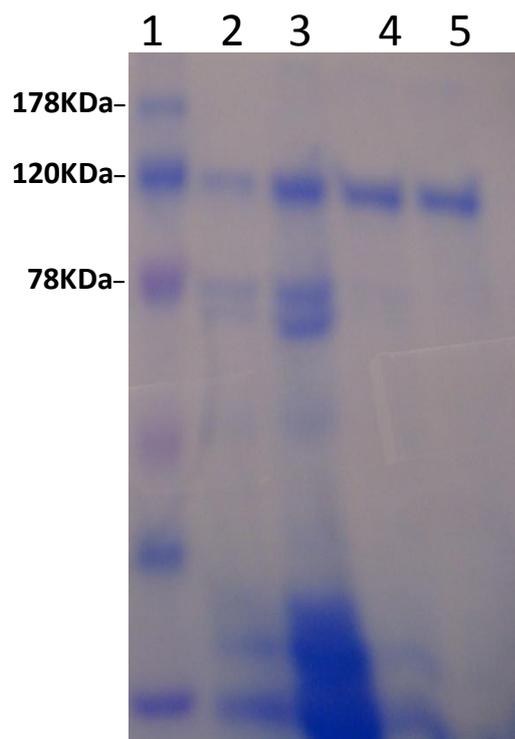


Figure 4.26: SDS-PAGE of the purified Fbp proteins in the eluate fractions after His-tag purification under native conditions.

Lanes 2 to 5, eluate fractions of Fbp. Lane 1 represents the HyperPage prestained protein marker (Bioline).

4.2.2.12 Complementation of JB108 mutant via *fbp* gene

The transformants (JB108₃/pNova-*fbp*-3 and JB108₄/pNova-*fbp*-3) cells were plated on LB plate (+Tet and +Kan). Then the two colonies were transformed onto M9 glycerol plate (+Tet). As expected no growth was detected in the palte on M9 glycerol plate (+Tet) without IPTG (Figure 4.27). However, the transformants (JB108₃/pNova-*fbp*-3 and JB108₄/pNova-*fbp*-3) were able to on the M9 glycerol plate (+Tet) supplemented with IPTG, this implies that the cloned *fbp* gene has FB Pase activity (Figure 4.27).

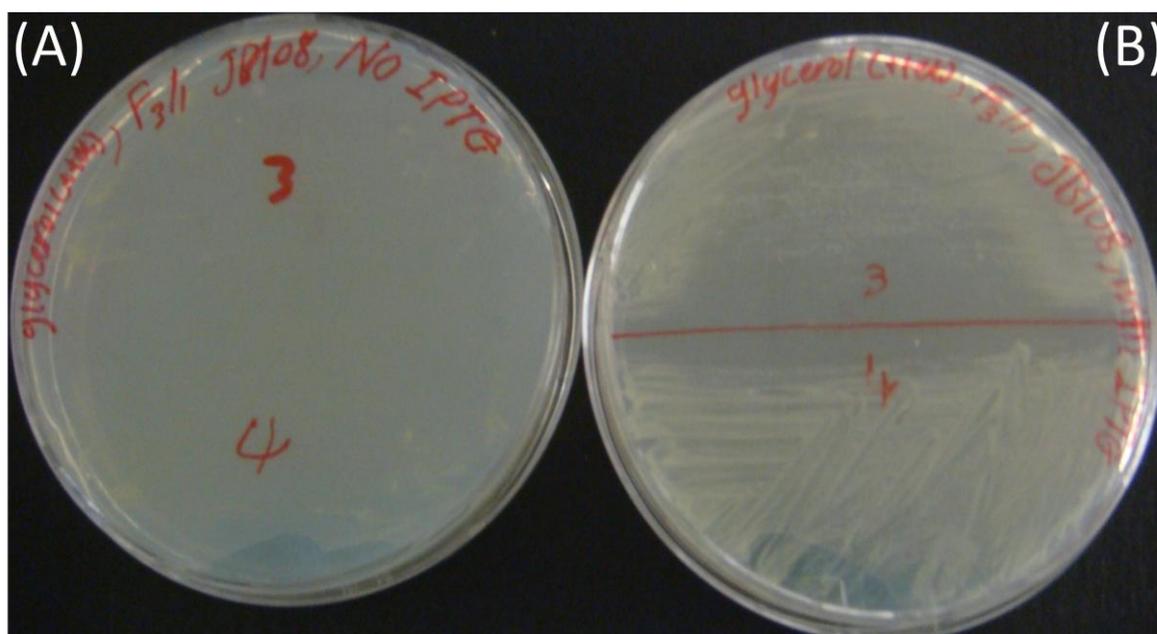


Figure 4.27: Growth of JB108₃/pNova-*fbp*-3 and JB108₄/pNova-*fbp*-3 strains on two M9 glycerol plates.

Plate (A) without IPTG (+Tet) and plate (B) with IPTG (+Tet). The JB108₃/pNova-*fbp*-3 and JB108₄/pNova-*fbp*-3 that were tested are indicated by 3 and 4 respectively.

4.2.2.13 Assay of the fructose 1,6 biphosphatase activity of the recombinant Fbp protein from a crude extract of the JB108₃/pNova-*fbp-3*

The protein concentration of the crude extract of JB108₃/pNova-*fbp-3* was 17.2 mg/ml as assessed by the Biuret method (Section 2.7.6.1) and FBPase activity in was assayed as mentioned in section 2.7.6. The recombinant Fbp from JB108₃/pNova-*fbp-3* strain had FBPase specific activity of 747.67 ± 20.97 nmol/min/mg. Different divalent cations were tested as mentioned before (Mn^{2+} , Ca^{2+} , Zn^{2+} , and Mg^{2+}). The divalent cations used that support the optimal activity of Fbp was Mn^{2+} (10 Mm) (Table 4.6). No activity was detected when Mn^{2+} was absent from the reaction mix. However, replacing Mn^{2+} with 10 Mm of Mg^{2+} or Ca^{2+} reduced the activity about 59% and 93% respectively, compares to the 10 Mm of Mn^{2+} . Instead, no FBPase activity was detected when Mn^{2+} was replaced with 10 mM of Zn^{2+} .

The FBPase activity of Fbp was completely inhibited by the addition of 1 mM inorganic phosphate. However, 1 mM of PEP and ATP reduced the activity about 10%, whereas, 1 mM of AMP reduce the activity about 16%. Also 1 mM of ADP was tested which produced no major effect on the FBPase activity of Fbp (Table 4.6). As *C. acetobutylicum* GlpX belongs to the more resistant Li^+ -sensitive phosphatases it was not completely inhibited by 1 mM or 10 mM Li^+ as mentioned before (Section 4.2.2.8), hence, Li^+ was tested as a potential inhibitor. Only 27% reduction in the activity was detected after the addition of 1 mM Li^+ , however, a 45% reduction in activity was detected after the addition of 10 mM Li^+ , which indicates that the *C. acetobutylicum* Fbp also belongs to the more resistant Li^+ -sensitive phosphatases.

Table 4.6: The effect of different metabolites on the reaction rate of the recombinant Fbp enzyme compare to the control (no effector).

Effector (1mM)	Reaction rate (%)	Inhibition rate (%)	FBPase specific activity (nmol/ min/mg)
None	100	0	747.67 ± 20.97
ATP	90.6	9.4	677.33 ± 10.56
ADP	95.3	4.7	712.21 ± 3.00
AMP	84.4	15.6	630.81 ± 34.12
PEP	90.5	9.5	680.23 ± 16.4
Inorganic phosphate	No activity	No activity	No activity

4.2.2.14 Comparison of the FBPase activities of recombinant GlpX (native and tagged) and recombinant Fbp proteins

The recombinant fusion GlpX was more active (29.70 ± 0.61 nmol/min/mg) by 1.6 fold compared to the native GlpX (18.51 ± 0.88 nmol/min/mg) due to the recombinant fusion GlpX being more concentrated than the native GlpX based on SDS-PAGE. However, the recombinant fusion Fbp had activity (747.67 ± 20.97 nmol/min/mg) which was 25 fold higher than the recombinant fusion GlpX and around 40 fold higher than the native GlpX even though the concentration of the recombinant fusion GlpX was higher than Fbp based on SDS-PAGE.

On the other hand, ADP had almost no effect on the activity of the recombinant fusion GlpX, in fact, ADP stimulated the activity of the native GlpX by 20.5%. Also, the PEP inhibited the activity for both proteins, with more inhibition 10% for the recombinant fusion GlpX (Table 4.7). Phosphate inhibited the activity completely of both enzymes which indicates that the active site is not affected by the presence of the N-terminal tags in these recombinant fusion proteins (Table 4.7). In contrast, the ADP had no dramatic effect on the FBPase activity of the recombinant fusion Fbp protein, whereas, the AMP inhibited the activity by about 15.6%, however, only 9.4% and 9.5% inhibition were detect after adding the ATP and PEP respectively (Table 4.7). Moreover, ATP exerted similar inhibition rate on both native and recombinant GlpX enzymes.

Effectors (1mM)	Percentage of Activity of FBPase enzymes compare to their control (no effector)		
	GlpX (native)	GlpX (tagged)	Fbp (tagged)
ATP	6.3% inhibition	6.3% inhibition	9.4% inhibition
ADP	20.5% stimulation	no effect	4.7% inhibition
AMP	6.9% inhibition	no effect	15.6% inhibition
PEP	16.5% inhibition	26.5% inhibition	9.5% inhibition
Phosphate	Complete inhibition	Complete inhibition	Complete inhibition

The recombinant fusion GlpX was stimulated 2-fold when the Mn^{2+} was replaced by Mg^{2+} compared to about 75% inhibition for the native GlpX which might indicate that the GST•Tag, and S•Tag have more drastic effect on the activity toward divalent cations. In contrast, Ni^{2+} , Cu^{2+} , Zn^{2+} , Ca^{2+} were tested as potential divalent cations, no activity were detected (Table 4.8). Also, the Ni^{2+} , Cu^{2+} , Zn^{2+} , Ca^{2+} were test as potential inhibitors of the coupling enzymes in the assay reaction. Ni^{2+} and Cu^{2+} were found to inhibit the coupling enzymes, indicating that the Ni^{2+} , Cu^{2+} most not used as potential divalent cations alongside with coupling enzymes in assay FBPase activity.

Divalent cations (10 mM)	Percentage of Activity of FBPase enzymes compare to control condition (10 mM of Mn^{2+})		
	GlpX (native)	GlpX (tagged)	Fbp (tagged)
Mg^{2+}	75% inhibition	200% stimulation	60% inhibition
Ni^{2+}	Inhibit the coupling enzymes	Inhibit the coupling enzymes	Inhibit the coupling enzymes
Cu^{2+}	Inhabit the coupling enzymes	Inhabit the coupling enzymes	Inhibit the coupling enzymes
Zn^{2+}	No activity	No activity	No activity
Ca^{2+}	No activity	No activity	Only 6% activity

When the Mn^{2+} was replaced by Mg^{2+} , about 60% loss of the activity of the recombinant fusion Fbp was detected, compare to 2 fold increase in the activity of the recombinant fusion GlpX and about 75% inhibition for the native GlpX. This might indicate that attached tags affect the metal bind site (Table 4.8).

Moreover, a small activity 6% compare to control was detected when Mn^{2+} was replaced by Ca^{2+} , to date this is the first study reported that Fbp class III has shown activity with Ca^{2+} as divalent cations that required for FBPases enzyme activity (Table 4.8). On the other hand, the Zn^{2+} , Ca^{2+} were test as potential divalent cations, no activity were detected.

These results shown that the N-terminal tags had only a very small effect on the activity of the target protein (GlpX) probably due to the fact that the tagged proteins was engineered on the surface of the target protein and away from the active site to reduce any effect in the activity of the target protein. These small differences in the behaviour toward the effectors compare to the native GlpX due to the GST•Tag, and S•Tag have the ability to enhance the folding of the target protein which may contribute to very small increase or decrease on the activity in the presence of effectors.

4.3 Discussion

4.3.1 Overexpression of GlpX and Fbp proteins by Gateway cloning system.

This expression system was used in order to purify the recombinant GlpX and Fbp proteins and then assay these pure proteins. The Gateway technology is based on the specific-site recombination reactions catalyzed by the bacteriophage λ Integrase (Int) protein (Hartley *et al.*, 2000). The Gateway cloning system consists of two bacteriophage pathways, the BP reaction is the first step that mediates by the bacteriophage λ Integrase (Int) and the *E. coli* Integration Host Factor (IHF) proteins and this mix of enzymes is termed BP ClonaseTM II enzyme mix and this step called lysogenic pathway. The second step of this system is the LR reaction which is the lytic pathway catalyzed by the LR ClonaseTM II enzyme mix that consists of the bacteriophage λ Integrase (Int), Excisionase (Xis), and *E. coli* Integration Host Factor (IHF) proteins. The BP ClonaseTM II enzyme mix and LR ClonaseTM II enzyme mix recognise specific-sites (*att*-sites). The GATE*glpX* and GATE*fbp* primers are designed to amplify *glpX* and *fbp* genes flanked with 31 bp *attB* sequences at both ends. These PCR products are mixed together with the Donor vector (pDONRTM221) which contains the *attP* sites, and BP ClonaseTM II enzyme mix in the first step of cloning reaction (BP reaction). As a result, the entry clone was formed containing the gene of interest and by product. The entry clone was transformed into One Shot[®] OmniMAXTM 2-T1R Chemically Competent *E. coli*.

Transformed cells grew on the LB plates (+Kan) which indicated that an entry clone had successfully transformed. The transformed cells that contain the Donor vector (pDONRTM221) without the cloned genes cannot grow due to the *ccdB* which a lethal gene which located at the cloning site. Therefore only the transformed cells that harbour the entry clones with *glpX* or *fbp* genes grew. Moreover, PCR screening reconfirmed these results by amplifying the *glpX* or *fbp* genes from the entry clones. These entry clones containing *glpX* or *fbp* genes were purified and the second step which is the lytic pathway (LR reaction) was conducted with the LR ClonaseTM II enzyme mix and the destination vector harbours the lethal gene (*ccdB*) that surrounded by *attR* sequences.

The expression clones were formed containing the *glpX* or *fbp* genes and by product vector which contains the lethal gene. These mixtures were transformed in Rosetta™ 2(DE3) Competent Cells (expression strain). Meanwhile, the purified entry clones containing the *glpX* or *fbp* genes were sent for sequencing and the results were disappointments due to a lot of mismatching and unidentified bases as mentioned before. Although the *glpX* and *fbp* genes were amplified using the Easy-A high fidelity PCR cloning enzyme which has a proofreading activity (3'- to 5'-exonuclease activity) that removes or excises the wrong base pair and replaced with a correct one (Newton and Graham, 1994). Moreover, this enzyme has a six fold lower error rate compared to *Taq* DNA polymerase and two to three fold lower than any proofreading archaeal DNA polymerases in the market (Stratagene® manual, 2009). Therefore, it is unlikely to have errors or mismatching in the PCR products. Also, the same problem was seen with another student using different genes and due to these problems, we discontinued the use of the Gateway cloning system and switched to Ligation Independent Cloning system.

4.3.2 Overexpression of GlpX and Fbp proteins by LIC system

This expression system does not require T4 ligase or restriction enzymes, however, in order to clone a gene, primers were designed with an extra 15 bp at 5' ends. The amplified PCR product as a result, would include the gene of interest flanked with the extra 15 bp at the ends. However, this system is based on the 3' → 5' exonuclease activity of T4 DNA Polymerase together with dATP which removes all the extra 13 bp (dGTP, dCTP, and dTTP) from 3' → 5' direction and stop at the 13 bp which is adenosine. Therefore, the 5' compatible overhangs on the PCR product is generated, which anneals to the compatible overhangs on the the Ek/LIC vector. The *glpX* and *fbp* genes were amplified using LIC_{glpX} and LIC_{fbp} primers respectively which contain compatible overhangs at 5' ends for easy cloning with no the needs of T4 ligase. Two PCR products were detected on the agarose gel, a fragment about 1 kbp of the amplified *glpX* and the other fragment was round 2 kbp for amplifying *fbp*. Moreover, the LIC system is sensitive to any nonspecific amplified fragments, remaining dNTPs and DNA polymerase from PCR and these reduce the efficiency of the cloning reaction.

Therefore, two purification procedures were used, the first one was the agarose gel purification to remove any unwanted PCR products and this step avoids any possibility of nonspecific cloning. The second step was Chloroform: Isoamyl Alcohol (CIAA) extraction and isopropanol precipitation to inactivate DNA polymerase and remove the remaining dNTPs which prevented interference with T4 DNA Polymerase. After purification, the *glpX* and *fbp* amplified PCR products were treated with T4 DNA Polymerase and dATP to create compatible overhangs at 5' ends and cloned into pET-41 Ek/LIC plasmid.

In this plasmid, the protein encoded by the cloned gene is fused with N-terminal GST•Tag, His•Tag, S•Tag and also another His•Tag in C-terminal, these tags are to facilitate the purification and solubilization of recombinant protein. The Glutathione *S*-transferase (GST•Tag™) is a 211 amino acid protein with a molecular mass of 26 kDa which makes it the largest fused protein. Moreover, the GST•Tag is often integrated with other tags such as His•Tag to increase the solubility of the recombinant protein and prevent aggregation. However, some recombinant proteins may aggregate even though they contain a His•Tag fusion and this is why pET-41 Ek/LIC plasmid which harbours more than one tagged fusion protein was selected and also, to increase the purity of the target protein. However, the disadvantage of using a large tag is that the fused protein has to be removed for some applications such as crystallization (Terpe, 2003). Meanwhile, the polyhistidine tag (His•Tag) is the smallest tag consisting of six histidine amino acids attached to the N-terminal of the recombinant protein and another one is attached to the C-terminal (Appendix 7.7). The advantage of using small tags is that some applications such as crystallization can be conducted without interference (Brown *et al.*, 2009; Terpe, 2003).

After cloning, the plasmids that harboured the *fbp* and *glpX* genes were transformed into NovaBlue GigaSingles™ Competent *E.coli* Cells. These were excellent cloning hosts with high transformation efficiency and this step was used to increase the yield of the plasmids containing *fbp* and *glpX* genes. PCR screening was carried out using the same primers that have been used before in the amplification of *fbp* and *glpX* genes. As expected, Two PCR products were detected on the agarose gel, a fragment about 1 kbp of the amplified *glpX* and the other fragment was round 2 kbp for amplifying *fbp*.

The plasmids were purified and then transformed into an expression host, BL21 (DE3) pLysS, a strain genetically engineered for expression vectors such as pET vectors. Also this strain contains the T7 bacteriophage gene 1(DE3), encoding T7 RNA polymerase and harbours the pLysS plasmid, which carries the gene encoding T7 lysozyme that lowers the background expression of the target gene under the control of T7 promoter under uninduced conditions. However, T7 lysozyme, under induced conditions, which was achieved by adding IPTG, does not interfere or lower with the expression level of the target gene (Davanloo *et al.*, 1984). This creates an excellent environment to increase the yield of target protein without killing the cells. After transformation to the expression host, PCR screening was conducted to confirm the presence of the plasmids that harbour the *fbp* or *glpX* genes. The cells were induced for *glpX* and *fbp* expression and the GlpX and Fbp proteins were purified using the Bug Buster[®] Ni-NTA His•Bind Purification kit.

As mentioned before, the expressed Fbp and GlpX were fused with two hexahistidine tags at N and C terminals. Therefore, after loading the crude extract into the column, covalent bonds were formed between the two His•Tags on the surface of Fbp and GlpX proteins and Ni, whereas, the rest of the proteins that not did have His•Tags were among the flow through. To elute the proteins, imidazole was added which competes with and replaces the tagged Fbp and GlpX proteins in attaching to the immobilized metal ion. As a result, the fused Fbp and GlpX proteins were purified to homogeneity estimated by SDS-PAGE, moreover, no aggregations were seen in the insoluble extracts for both proteins due to the presence of GST•Tag and S•Tag which their main functions were to prevent the aggregation and increase the solubility of Fbp and GlpX proteins.

Based on calculation, the total expected molecular weight of the fusion proteins which consist of GST•Tag, S•Tag and two His•Tags together with the extra bases between the tagged proteins was 37 kDa. The native GlpX molecular weight is around 34.81 kDa, hence, the expected size of the tagged GlpX is 71.81 kDa. After purification, the tagged GlpX was about 73 kDa which very close to the expected size. However, the native Fbp molecular weight is around 77.23 kDa, hence, the total expected size of tagged Fbp is 114.23 kDa. After purification, the tagged Fbp was 115 kDa seen in SDS-PAGE and this also was very close to the estimated size.

4.3.3 Assay of the purified Fbp and GlpX recombinant fusion proteins

Using the LIC cloning system, a specific cleavage site is engineered between the N-terminal tags and the fused proteins. Hence, all N-terminal tags proteins (GST•Tag, His•Tag, and S•Tag) should be removed by treatment with enterokinase. Several attempts to remove the tags from the dialyzed Fbp-fusion and GlpX-fusion proteins were carried out with various optimizations and conditions, but with no success. Although the reason is not clear, it might be due to remaining imidazole even after extensive dialysis.

Both JB 108 strains which harbour the plasmids that have the *fbp* or *glpX* genes, were plated onto M9 glycerol (+Tet) plates, one with IPTG and the other one without IPTG. As expected no growth occurred on the glycerol plate without IPTG due to the *fbp* or *glpX* genes were under the control of the T7 promoter which only allows for expression after induction with IPTG. However, the cells were able to thrive on the glycerol plate supplemented with IPTG, indicating that the *fbp* or *glpX* genes were able to complement the *fbp* mutation even though both the Fbp and GlpX enzymes were expressed as recombinant fusion proteins. This provided confirmation that the tags on the fusion proteins do not interfere with the FB Pase activity of the Fbp and GlpX enzymes.

Also, both purified Fbp and GlpX were assayed the same way as described before in section 2.7.6, however no activity were detected for Fbp or GlpX. A similar problem was also reported by Movahedzadeh *et al*, (2004) who found that purification of the recombinant *M. tuberculosis* GlpX led to loss of FB Pase activity. This was due to nitrilotriacetic acid (NTA) may have the ability to absorb metals and therefore purification of a protein that is dependent on a metal for activity can result in loss of the activity (Terpe, 2003). This problem was solved by transforming the expression vectors that contained the *fbp* or *glpX* genes into the strain JB108, allowing preparation of crude extracts, and assay of the enzymes as described previously in section 2.7.6. Both recombinant tagged Fbp and GlpX proteins had FB Pase activity indicating that the fusion tags did not inhibit FB Pase activity. However, there were small differences in the activity of the native and tagged GlpX toward the respond to the effectors and this was explained in section 4.2.2.14.

4.4 Conclusion

The main goal of using Gateway expression system was to express and purify GlpX class II and Fbp class III. The BP and LR reactions appeared to work successfully as mentioned before. However, the sequence results were disappointments which resulted in a decision to stop using this expression system and alternative cloning strategies were being investigated at the same time and attention turned to one of them. Although, this problem was reported by another student in the laboratory who was cloning different genes at the same time. This problem was solved using LIC cloning system and both proteins were expressed and purified, but the purified GlpX and Fbp enzymes were not active. It is presumed that purification system Bug Buster[®] Ni-NTA His•Bind is not suitable for enzymes that depend on metal for their activity. Genetic complementation results of both recombinant GlpX and Fbp enzymes confirm this hypothesis and both enzymes were able to complement the JB108 mutant. Also, it was shown that Fbp protein encodes for FB Pase and this is first study reported Fbp class III enzyme in clostridia. Moreover, the enzymes assay results were consistent with genetic complementation that both proteins indeed encode for FB Pase activity, despite the large tag protein attached in the N-terminal. The native GlpX and recombinant fusion GlpX assay results were almost alike. However, it seems that the large tag protein does affect the metal bind site. Further analysis needed to be carried out in order to determine the physiological function of *glpX* using RT-PCR and proteomics analysis.

Chapter 5:

5 Transcriptional analysis

5.1 Reverse transcriptase- Polymerase chain reaction (RT-PCR)

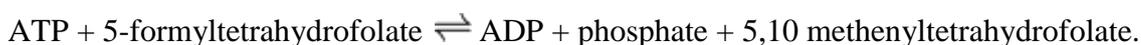
5.1.1 Introduction

The RT-PCR reaction consists of two amplification steps. The first one is reverse transcription of RNA which can be performed by a primer that anneals to the RNA template to create a complementary DNA (cDNA) by using the reverse transcriptase (RT) enzyme, hence, this reaction is termed the RT reaction. The second step is PCR amplification which can be performed by using any DNA polymerase enzyme (Newton and Graham, 1994).

5.1.2 Results

5.1.2.1 Determination the expression of *glpX* and *hprK*

The unique gene arrangement that has been proposed in *C. acetobutylicum* by Tangney *et al.*, (2003) in which *glpX* is linked to *hprK*, is not found in other low GC Gram-positive bacteria. In all the other bacteria, the identified *hprK* genes are linked to an *lgt* gene that codes for a prolipoprotein diacylglyceryl transferase enzyme (Tangney *et al.*, 2003; Boel *et al.*, 2003). Also the *hprK* gene overlaps with a gene that encodes for 5-formyltetrahydrofolate cyclo-ligase, an enzyme that forms ADP, phosphate, and 5,10 methenyltetrahydrofolate via the hydrolysis of ATP and 5-formyltetrahydrofolate as per the reaction equation (Greenberg *et al.*, 1965).



In *C. acetobutylicum* FBP is needed for the activity of HPr kinase/phosphorylase which phosphorylates HPr at the Ser-46 site and GlpX hydrolyses FBP. Therefore, GlpX might be involved in regulating the kinase activity of HPrK/P which in turn would regulate carbon catabolite repression (Tangney *et al.*, 2003). It is very important to prove this unique gene arrangement by experimental evidence. Therefore, an RT-PCR was carried out using RNA_{glpX} primers (Table 2.5 and Table 5.1) which amplify a fragment of 369 bp between the 3' end of *glpX* and the 5' end of *hprK* genes. Also, another RT-PCR was carried out using RNA_{hprK} primers (Table 2.5 and Table 5.2) which amplify the overlap fragment of 316 bp between *hprK* and the gene that encodes for 5-formyltetrahydrofolate cyclo-ligase (Figure 5.1).

Table 5.1: Locations of RNAglpX primers on the sequences of cac1088-cac1089 (*glpX-hprK* genes).

Bases underlined represent the locations of RNAglpX primers. Bases highlighted in green and red represent the start and stop codon respectively. Obtained from National Center for Biotechnology Information.

http://www.ncbi.nlm.nih.gov/nucore/NC_003030.1?report=fasta&from=1257296&to=1258210

glpX start codon

ATGTTTGATAATGATATATCCATGAGTTTAGTAAGAGTAACAGAAGCAGCAGCACTACAATCTTCAAAGT
 ATATGGGAAGAGGAGATAAAAATTGGAGCTGATCAAGCAGCAGTAGATGGAAATGGAGAAGGCATTTAGTTT
 TATGCCAGTAAGAGGCCAGGTTGTAATAGGAGAGGGAGAACTTGATGAAGCTCCTATGCTTTATATAGGT
 CAAAAGCTTGGTATGGGAAAAGACTATATGCCTGAAATGGATATAGCAGTAGATCCTTTAGATGGAACGA
 TTTTAATTTCTAAGGGACTACCTAATGCAATAGCAGTAATAGCAATGGGACCAAAAAGGAAGTTTACTTCA
 TGCCCCAGATATGTATATGAAGAAAATTGTTGTGGGACCTGGAGCAAAAAGGTGCTATAGATATAAAATAAA
 TCTCCTGAAGAGAATATTTTAAATGTAGCAAAGGCATTAACAAGGACATATCTGAATTAACAGTTATAG
 TTCAAGAAAGAGAAAGACATGACTACATAGTAAAAGCAGCTATAGAAGTTGGAGCAAGAGTTAAGCTATT
 TGGTGAGGGCGATGTTGCAGCTGCACCTTGCTTGTGGTTTTGAAGATACTGGAATAGACATACTTATGGGA
 ATTGGAGGAGCTCCAGAAGGAGTTATAGCCGCAGCAGCTATCAAGTGCATGGGCGGAGAAAATGCAGGCTC
 AGCTTATACCTCATACTCAGGAAGAAAATAGATAGATGTCACAAAATGGGAATAGATGATGTAATAAAAT

RNAglpX forward primer →

TTTCATGATAGATGATTTAGTTAAAAGTGATAATGTGTTTTTTGCAGCTACAGCAATAACAGAATGTGAT
 CTTCTTAAGGGCATAGTATTTTCTAAAAATGAACGTGCAAAAACCCATTCCATATTAATGAGATCTAAAA

glpX stop codon

CTGGTACAATAAGATTTGTTGAAGCTATTCATGACTTGAATAAAAAGTAAATTAGTGGTAGAA**TAA**ATTA

hprK start codon

AAATGTTTCGATTTACGGCGAAGGGGATAGTAA**ATG**CAGGTAAGTATTGAAGATATAATAGAAAATCTCG
 ATTTAGAGGTCTTGGTTAAGGGGAAAGATGGAATAAACTGGGGTTAAGTGATATAAA

←RNAglpX reverse primer

CAGACCAGGACTACAGTTTG

Table 5.2: Locations of RNAhprK primers on the sequences of *cac1089-cac1090* (*hprK*-5-formyltetrahydrofolate cyclo-ligase genes).

Bases underlined represent the locations of RNAhprK primers. Bases highlighted in green and red represent the start and stop codon. The overlap region is in bold.

Obtained from National Center for Biotechnology Information.

http://www.ncbi.nlm.nih.gov/nuccore/NC_003030.1?report=fasta&from=1257296&to=1258210

hprK start codon

ATGCAGGTAAGTATTGAAGATATAATAGAAAATCTCGATTTAGAGGTCTTGGTTAAGGGGAAAGATGGAA
TAAACTGGGGTTAAGTGATATAAACAGACCAGGACTACAGTTTGCCGATTCTATGATTATTTTGGCAA
TGAAAGAGTGCAGGTTATAGGAAAGGCTGAGTGGAGTTTTCTAAATGCAATGCCCTCCAGAAATAAGAGAA
AAAAGAATTAGAAAATATTTTCAATTTGAAACACCATGTATTGTTCTTGCAAGAGGATTGAAACCACAAA
AGGAACTTCTTGATTGTTCCAAAGAGTACAATAGGTGGCTTTTAAGGTCAAAGGCTCAACTACTAGATT
TATAAATAAAATAATGAACTATCTTGATGACAAGCTGGCTCCTGAAACAAGAAATTCATGGAGTGTTAGTT
GATATCTATGGATTAGGTATACTAATCACTGGAGAAAGTGGAAATCGGAAAAAGTAACTGCACTTGAGC
TCATAAAAAGAGGGCATAGATTGGTTGCAGATGATGCTGTAGATATAAAAAGAAATGAAATCAGTTCTTGT
TGGAAAATCACCATACATAACTTCTGGTATGCTTGAGGTTAGAGGAATGGGAATAATAGATGTTCCAGCA
CTTTATGGTCTTAGTTCTGTTTTGTCAGAAAAGAATATAAATCTTGTAAATATACCTTGAACAATGGAAAG
AGGGAAGAGACTACGATAGACTTGGTACAGAT

RNAhprk forward primer →

GATGAACACATAAAAAATTTTAAATATTCCTGTGAGAAAAATGACGCTGCCTATACGTCCAGGGAGGAATG
TTGCTGTTATAATAGAGGCAGCAGCTGCTAATTATAGATATAATTTAAGCAGTAAAAATATCTCCTGTAGA

5-formyltetrahydrofolate cyclo-ligase gene start codon ↘ ↙ *hprK* stop codon

TACTATTAATAAGAGAATTGAAGAGTCTACAAAT**ATGATTAA**TAGTAAAAAAGAACTTAGGAAAAATA
TGGTTTTTAAAAGGGATTCTTTAGAGCCAGATTTAAAATCAATAAAGGATAATATGATTTATAATAAAGT
CATAAATAGCATACAATATAAAAAGTGCTAATAATATATTTGTGTTTCGTTAGTTACAAAAGTGAAGTTGAT
ACTCATAATATAATAAG

← RNAhprk reverse primer

AACAGCAATTGCAGATGGAAAAAAGTTTTTGTGCCGAAGGTTATTTCCAAGG

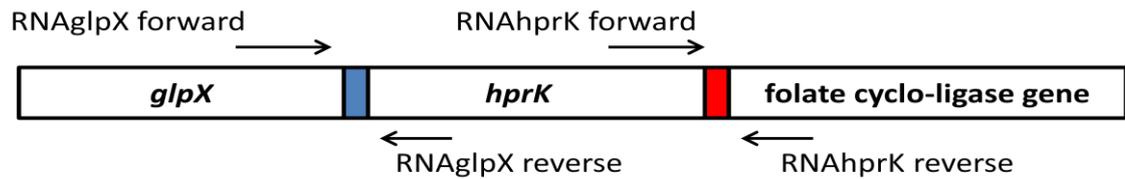


Figure 5.1: Illustration of *glpX*, *hprK* and 5-formyltetrahydrofolate cyclo-ligase gene arrangement.

The blue box is an open reading frame located between *glpX* and *hprK* genes that would be amplified by RNAglpX primers, the red box is an overlap region located between *hprK* and 5-formyltetrahydrofolate cyclo-ligase genes that would be amplified by RNAhprK primers.

First, normal PCR using a DNA template was carried out using the reverse and forward RNAglpX primers and reverse and forward RNAhprK primers to check their ability to amplify the right fragments and the best annealing temperatures (Figure 5.2). Three different annealing temperatures were tested for each gene; for the RNAglpX primers 51, 53 and 55°C were tested, only 53 and 55°C were able to anneal the primers and amplify the fragment between the the 3` end of *glpX* and the 5` end of *hprK* genes to give a product of 369 bp (Figure 5.2). However, for the RNAhprK primers 53, 55 and 57°C were tested, and only 53 and 57°C were able to anneal and amplify the overlap fragment between *hprK* and the 5-formyltetrahydrofolate cyclo-ligase gene to give a product of 316 bp (Figure 5.2). The agarose gel showed that both annealing temperatures for each fragment can be used in the cDNA amplification step of the RT-PCR (Figure 5.2). Also, the primer dimers issue needed to be solved to prevent interference with the results due to them being are very close in size to the PCR product (Figure 5.2).

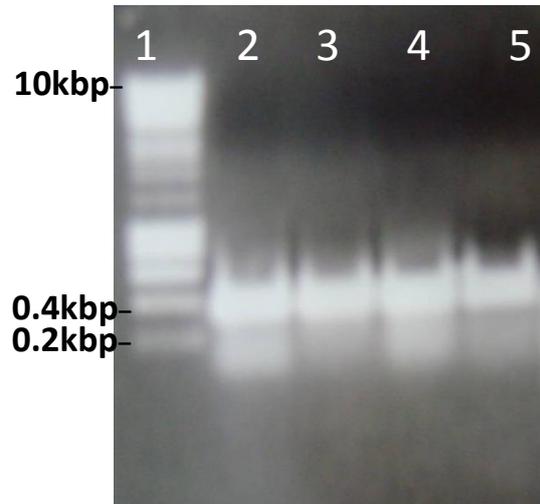


Figure 5.2: Agarose gel electrophoresis of PCR amplification of fragments between *hprK*, *glpX* and a gene that encodes for 5-formyltetrahydrofolate cyclo-ligase.

The RNA*glpX* primers amplify a fragment between the 3` end of *glpX* and the 5` end of *hprK* genes and the size of the PCR product was around 369 bp (lane 2 for 53°C and lane 3 for 55°C). The RNA*hprK* primers which amplify the overlap fragment between *hprK* and a gene that encodes for 5-formyltetrahydrofolate cyclo-ligase and the PCR product was around 316 bp (lane 4 for 55°C and lane 5 for 57°C). The lane 1 represents Hyperladder I (Bioline).

5.1.2.2 QIAGEN[®] One-Step RT-PCR Kit

The RNA template was purified as described in materials and methods section 2.10.1, from *C. acetobutylicum* grown in glucose minimal medium (Figure 5.3). Several attempts were made to amplify these fragments using the QIAGEN[®] One-Step RT-PCR Kit, such as different amount of RNA template (1, 2, and 4 μ l) (Figure 5.4), adding RNase inhibitor into the reaction mix to protect from degradation and extend the life time of the RNA template. Also the same different annealing temperatures that were used before to amplify each fragment were tested for the cDNA amplification step of the RT-PCR. Despite these various optimisations, no products were obtained from the RT-PCR reaction and only primer dimer was observed (Figure 5.4). Moreover, the QIAGEN[®] One-Step RT-PCR Kit does not come with a control reaction to test the efficiency of the kit.

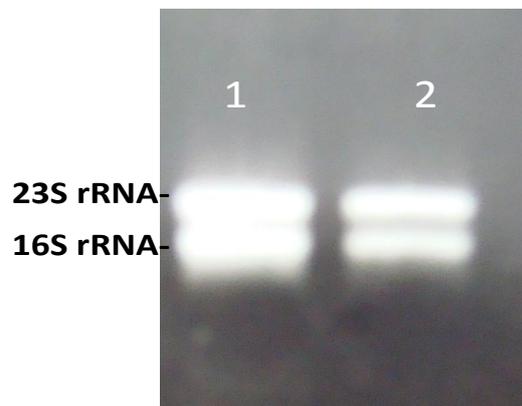


Figure 5.3: Agarose gel of purified total RNA from a culture of *C. acetobutylicum* grown in glucose minimal medium.

Lane 1 and 2 represent duplicate sample of the total purified RNA.

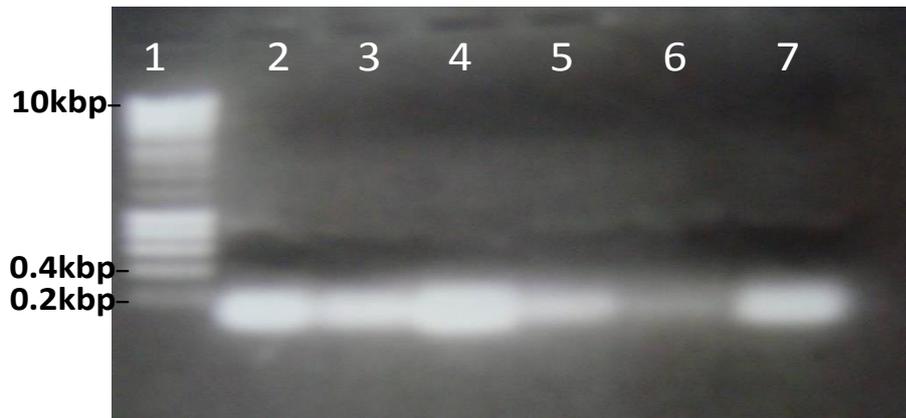


Figure 5.4: Agarose gel electrophoresis of RT-PCR attempt to amplify fragments between *hprK*, *glpX* and a gene that encodes for 5-formyltetrahydrofolate cyclo-ligase.

Attempt to amplify fragment between the downstream of *glpX* and the upstream of *hprK* genes using these different amounts of RNA template, 1 μ l for the lane 2, 2 μ l for the lane 3, and 4 μ l for the lane 4. Lanes 5 to 7, represent the attempt to amplify the overlap fragment between *hprK* and a gene that encodes for 5-formyltetrahydrofolate cyclo-ligase using the same different amounts of RNA template. The lane 1 represents the Hyperladder (Bioline).

5.1.2.3 One Step RT-PCR Master Mix Kit from Novagen®

Therefore, another kit, the One Step RT-PCR Master Mix Kit from Novagen® was used to generate cDNA and then PCR product as described in the materials and methods section 2.10.3. This kit comes with a control reaction to test the efficiency of the kit and also it can detect very small amounts of RNA, less than 2 ng, which makes it more sensitive. First, a control reaction to test the kit was carried out using the RT-PCR control primers and a positive RNA control transcribed product from the human gene *G3PDH* (glyceraldehyde 3-phosphate dehydrogenase) which is a housekeeping gene in human tissue (Barber *et al.*, 2005). The reaction mix was set up as described in materials and method section 2.10.3. A positive RNA control (2 µl) and 16.5 µl RNAs-free water to reach the reaction total volume of 50 µl. As expected a RT-PCR product of 450 bp was detected in the agarose gel (Figure 5.5).

To prove the unique gene arrangement, two RT-PCR reactions were then carried out using RNA*glpX* primers which amplify a fragment between 3` end of *glpX* and the 5` end of *hprK* genes, and RNA*hprK* primers which amplify the overlap fragment between *hprK* and the 5-formyltetrahydrofolate cyclo-ligase gene. The reaction mix was prepared as described in the materials and methods section 2.10.3, however, to overcome the primer dimer issue, primer concentrations were diluted to 10 pmol/µl (10-fold) and the amount of RNA template used was 9 µl. A PCR product of size around 300 bp was detected in the agarose gel for each of the reactions, which indicates that *glpX* and *hprK* are expressed on one polycistronic m-RNA and also the same results were seen for *hprK* and the 5-formyltetrahydrofolate cyclo-ligase gene. Therefore, all these *glpX*, *hprK* and 5-formyltetrahydrofolate cyclo-ligase genes are expressed as an operon. Also, a control reaction was carried out with normal PCR reaction (no reverse transcriptase step) using the same primers (RNA*glpX*) to detect any trace of DNA in the purified RNA. However, no amplification was detected indicating that no DNA was present in the purified RNA (Figure 5.6).

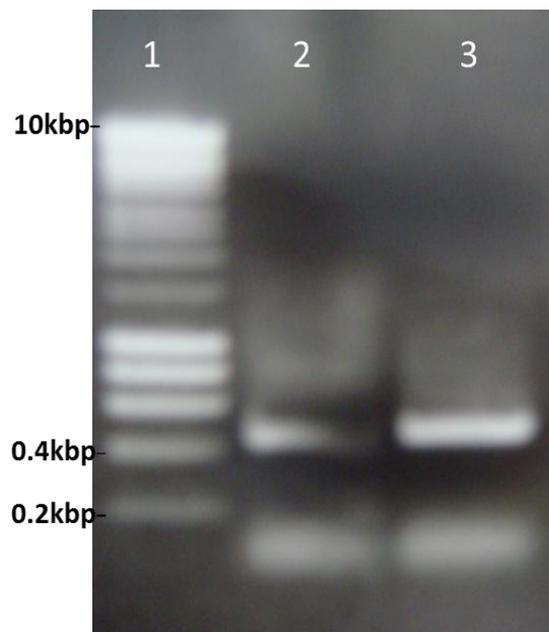


Figure 5.5: Agarose gel electrophoresis of a control RT-PCR reaction from the Novagen[®] One Step RT-PCR Master Mix Kit.

The lanes 2, 3 represent two control RT-PCR products of size 450 bp. Lane 1, represents the Hyperladder I (Bioline).

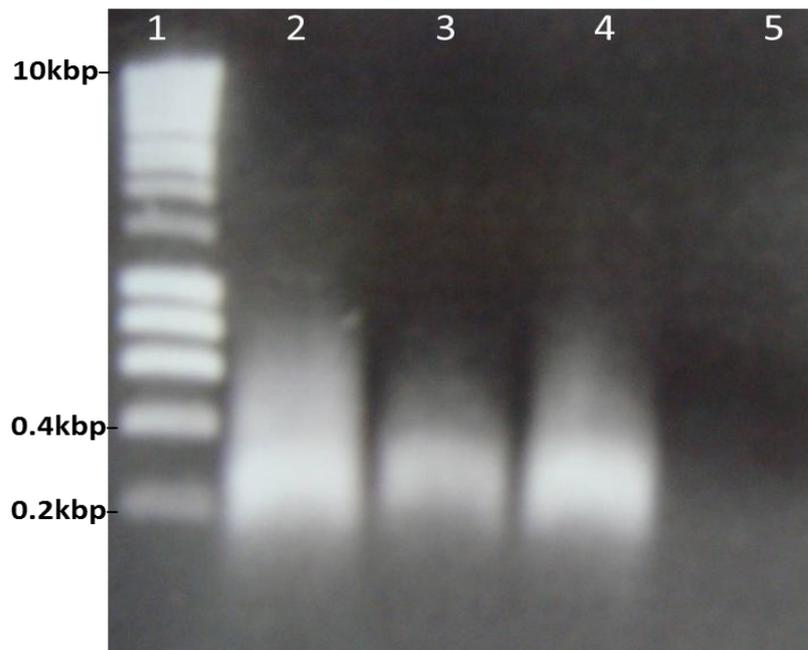


Figure 5.6: Agarose gel electrophoresis of RT-PCR reaction to amplify fragments between *hprK*, *glpX* and a gene that encodes for 5-formyltetrahydrofolate cyclo-ligase.

The lane 2, and 3 represent amplified fragments between the 3` end of *glpX* and the 5` end of *hprK* genes (300 bp). Lanes 4 and 5 represent the amplified the overlap fragment between *hprK* and a putative gene that encodes for 5-formyltetrahydrofolate cyclo-ligase and the control reaction mix (normal PCR), respectively. The lane 1 represents the Hyperladder I (Bioline).

5.2 Discussion

The kinase activity of the bifunctional HPr kinase/phosphorylase is dependent on FBP. HPrK/P is phospholyrates the HPr at serine-46 site in an ATP dependant manner to participate in carbon catabolite repression. However, the putative *hprK* gene in *C. acetobutylicum* is located downstream of a gene that encode for GlpX which is a class II FBPase enzyme (Tangney *et al.*, 2003). This gene arrangement is exclusive to *C. acetobutylicum* which might indicate the unique role of GlpX in regulating the activity of HPr kinase, and this in turn would effect carbon catabolite repression. To confirm how the expression of the genes is organised, RT-PCR analysis was carried out using specific primers as mentioned previously. The RNA was extracted and purified from *C. acetobutylicum* grown on glucose minimal medium conditions in which it would be expected that HPrK would be expressed. Several attempts to amplify these fragments using QIAGEN[®] One-Step RT-PCR, a kit that can detect amounts of RNA less than 50 ng, however no PCR products were seen on the gel. Therefore, RT-PCR analysis was done using another kit, the One Step RT-PCR Master Mix Kit from Novagen[®] which can detect very small amounts of RNA, less than 2 ng. The results showed that both regions were amplified and detected on agarose gel, indicate that GlpX, HPr kinase and the gene that encodes for 5-formyltetrahydrofolate cyclo-ligase were expressed under glucose condition.

In *B. subtilis* the *hprK* was reported to be weak constitutive expressed gene that does not depend on the growth condition (Hanson *et al.*, 2002). Therefore, Northing blotting cannot be used due to the fact that this technique is not sensitive enough to detect a weak gene expression (Rapley and Manning, 1998). This might explain why m-RNA of *hprK-glpX* cannot be detected by standard RT-PCR kit such as QIAGEN[®] One-Step RT-PCR and was only detected by very sensitive RT-PCR kit which was One Step RT-PCR Master Mix Kit from Novagen[®]. Moreover, these results shown, *glpX*, *hprK* and the 5-formyltetrahydrofolate cyclo-ligase gene were expressed together in one polycistronic m-RNA. In prokaryotes genes that are functionally related are usually expressed as a polycistronic m-RNA as seen for the *E. coli lac* operon in which all the three genes that responsible for uptake and utilization of lactose are expressed as one polycistronic m-RNA (Vilar *et al.*, 2003).

Therefore, the RT-PCR results imply that the unique gene arrangement of the *C. acetobutylicum* *glpX*, *hprK* and 5-formyltetrahydrofolate cyclo-ligase genes might indicate that they have a related function in regulating carbon catabolite repression in this industrial important strain. The physiological function of 5-formyltetrahydrofolate cyclo-ligase is still unknown in prokaryotes. However, a recent study indicates that this protein is necessary in thiamine (Vitamin B₁) metabolism (Pribat *et al.*, 2011). Also, the expression of *glpX* in cells grown in glucose minimal medium indicates that this gene is not simply an enzyme of the gluconeogenesis pathway which enables cells to grow on gluconeogenic substrate such as glycerol as reported by others (Rittmann *et al.*, 2003; Movahedzadeh *et al.*, 2004; Jules *et al.*, 2009). Instead, *glpX* in *C. acetobutylicum* can be suggested to have a specific role in regulating the activity of HPr kinase. Also, proteomic analysis was carried out in order to confirm whether *glpX* is expressed or not when the cells are grown on glucose as a sole carbon source. The results showed that liquid chromatography electrospray ionization tandem mass spectrometry was able to detect 439 proteins and GlpX was among the proteins which indicate that *glpX* gene is expressed during growth on a glycolyte substrate such as glucose which is consistent with the RT-PCR result. However no Fbp was detected under these conditions and this indicates that Fbp might have role gluconeogenic pathway.

5.3 Conclusion

C. acetobutylicum total RNA was purified from a culture grown on glucose minimal medium. Then two RT-PCR reactions were carried out using RNA_{glpX} and RNA_{fbp} primers to amplified regions between the 3` end of *glpX* and the 5` end of *hprK* genes and the overlap region between *hprK* and the 5-formyltetrahydrofolate cyclo-ligase gene. The RT-PCR results indicate that GlpX is expressed under glucose growth conditions and expressed as a polycistronic m-RNA together with *hprK* and the 5-formyltetrahydrofolate cyclo-ligase gene which also imply that these genes are expressed as an operon. Furthermore, proteomic analysis was compatible with RT-PCR results. Based on these results, GlpX seems to be expressed constitutively due to it linked to *hprK* which is constitutive expressed gene that does not depend on the growth condition and needed for CCR. Therefore, GlpX might play a vital role in CCR.

Chapter 6:

6 General discussion

6.1 The importance of intracellular FBP levels in CCR and strain improvements

In this project the *glpX* and *fbp* genes of *C. acetobutylicum* were shown by genetic complementation and enzyme assay to encode for FBPase enzymes. Both GlpX and Fbp proteins were inhibited by 1 mM inorganic phosphate and the recombinant tagged Fbp had 25 fold higher activity than the recombinant tagged GlpX. To investigate this further, proteomic and RT-PCR analyses were carried out. In proteomic analysis, the Fbp protein was not found under glucose growth conditions, but, GlpX was detected. RT-PCR results showed that *glpX* does indeed express on one polycistronic m-RNA together with *hprK* and 5-formyltetrahydrofolate cyclo-ligase genes under glucose growth conditions. These results are consistent with the GlpX would appear to have a different role, whereas, the Fbp enzyme appear to involve in utilizing gluconeogenic substrates, Given the position of the *glpX* gene in *C. acetobutylicum* genome, proteomic and RT-PCR analyses which work towards a role in regulating intracellular FBP concentration.

As mentioned before, FBP is a key metabolite effector that can enhance binding of the CcpA-HPr-P-Ser complex to a *cre* site which in turn results in repression or activation of an operon or gene (Henkin, 1996; Presecan-Siedel *et al.*, 1999; Schumacher *et al.*, 2007; Antunes *et al.*, 2010). In *B. subtilis*, disruption of the gene encoding phosphofructokinase, which prevented the production of FBP, resulted in no CCR and the growth was impaired (Nihashi and Fujita, 1984). These results indicated how FBP is important in mediating the level of CCR in Gram positive bacteria. Intracellular FBP concentration has vital role in regulating the activity of HPr kinase and in turn CCR in low GC Gram positive bacteria such as *B. subtilis* (Singh *et al.*, 2008; Deutscher *et al.*, 2006). In this microorganism the highest FBP concentration was detected in vivo when cells were grown on glucose which is the preferable sugar that exerts the strongest CCR effect (Singh *et al.*, 2008; Nihashi and Fujita, 1984).

Singh *et al.*, (2008) have tested the effects of many different substrates on CCR. These tested substrates such as ribose, sucrose, glucose, and fructose have the ability to exert CCR and the repression caused by these substrates was quantified using the activity of β -xylosidase (XynB) which is involved in xylose utilization. In the absence of xylose, a very low β -xylosidase activity was detectable (14 U/mg), however, in the presence of xylose, a high β -xylosidase was detectable (945 U/mg) demonstrating that the synthesis of the enzyme was induced.

In the quantitative analysis, cells were grown on xylose together with a repressing sugar to find out how XynB synthesis was affected, and intracellular FBP concentration was also measured. Glucose exerted the strongest repression on XynB activity which was reduced to about 7 U/mg. However, maltose, sucrose and ribose repressed the activity to about 437, 126 and 497 U/mg respectively which was very weak repression compared to glucose. These results, however, indicated that not only glucose, but also other substrates, can exert CCR (Singh *et al.*, 2008). Moreover, the amount of intracellular FBP during growth in the presence of glucose, maltose, sucrose and ribose were around 14.1, 10.7, 11.5, 6.5 mM respectively (Singh *et al.*, 2008). By comparison, under gluconeogenic growth conditions such as on malate intracellular FBP in *B. subtilis* was found to be very low (around 1.5 mM) (Kleijn *et al.*, 2009). The implication of the results is that a higher FBP concentration increases the activity of HPr kinase which phosphorylates more HPr protein at the serine 46 site to cause CCR, while substrates which generate a low level of intracellular FBP do not exert CCR (Carlos *et al.*, 2003; Singh *et al.*, 2008).

For developing a successful industrial strain, clarification is needed on metabolic and regulatory networks of this strain (Dobson *et al.*, 2011). Although several attempts to control metabolic pathway in *C. acetobutylicum* in order to increase solvent production were carried out. Recently, the *ack* gene which encodes for acetate kinase that involve in acetate formation was disrupted using the CloStron system with the aim to increase the butanol production by reducing acetate and acetone formation (Kuit *et al.*, 2012). Although, acetate production in the mutant strain was reduced by more than 80% compare to the wild type, however, the acetone production was not effected in the mutant strain. This indicates that another acetate producing pathway might be operated in *C. acetobutylicum* (Kuit *et al.*, 2012).

Therefore intracellular FBP concentration might be used as an important global regulatory signal to control CCR in Gram positive bacteria such as *C. acetobutylicum*, in which a similar regulatory mechanism appears to be present. It has been reported that *C. acetobutylicum* and *C. butyricum* AKR102a are able to grow on crude glycerol derived from biodiesel (Carlos *et al.*, 2003; Ringel *et al.*, 2011).

As mentioned before, cloning and overexpression of either the *fbp* or *glpX* gene of *C. acetobutylicum* allowed an *E. coli fbp* mutant to thrive in glycerol minimal media. Therefore, overexpression of either Fbp or GlpX in *C. acetobutylicum* might lead to an increase in growth rate and the utilization efficiency of crude glycerol and thus increase butanol production. At the same time, overexpression of either gene might also result in a drop in the intracellular FBP concentration which would most likely relieve CCR when the cells grow on multiple carbon sources and increase the ability of the cells to utilize multiple carbon sources. Therefore, this might solve the problem of CCR excreted via using fermentation substrates that contain mixture of several carbon sources concurrently.

Also, this metabolic engineered strain probably can be used in the current biodiesel industry not only to increase the profit but solve the excessive crude glycerol formed from the production of biodiesel. Although genomic DNA of the recently isolated *C. butyricum* AKR102a strain has not been sequenced yet, this strain might contain genes encoding *fbp* in order to grow on glycerol. Therefore, the same proposed metabolic engineering techniques mentioned above for *C. acetobutylicum* might be used to increase the utilization efficiency of crude glycerol and production of 1,3-PD. The availability of genomic DNA sequence may assist in confirming these hypotheses by allowing characterization of the genes that encode for FBPase activity in this strain and identifying the locations of those genes.

It has been reported that knockout of the *hprK* gene in Gram positive bacteria such as *Streptococcus mutans* and *B. subtilis* led to impaired growth even though the CCR was relieved (Deutscher *et al.*, 1994; Zeng *et al.*, 2010). Also a knockout of *ccpA* in *C. acetobutylicum* apparently had the same result (Ren *et al.*, 2010). These effects are presumably due to several genes being positively regulated by these CCR elements and any mutations introduced to these elements may lead to a defect in the growth. An alternative way to relieve CCR in Gram positive bacteria without disturbing growth rate may be to reduce the intracellular FBP concentration by overexpressing a gene that encodes for FBPase to mimic the conditions when the bacteria are grown on unpreferred substrates such as glycerol. Knockout of a gene that encodes FBPase might result in several phenotypes, depending on how many genes encode for FBPase enzyme in the bacterium.

For example, in a bacterium that has only one gene that encodes for FBPase is most likely to involve in gluconeogenesis pathway and the knockout of this gene would be expected to result in the mutant strain being unable to grow on gluconeogenic substrate such as glycerol. This phenomenon was reported in *C. glutamicum* and *M. tuberculosis* in which there is only one gene that encodes for FBPase enzyme (Movahedzadeh *et al.*, 2004; Rittmann *et al.*, 2003). However, if the bacteria contain two genes or more encoding for FBPase enzyme, one of these genes may encode for the major FBPase involved in the gluconeogenesis pathway and the others might have different physiological functions. In *E. coli* for example, knockout of the major *fbp* gene which encodes for the class I FBPase enzyme created the Δfbp strain which is unable to grow on gluconeogenic substrates such as glycerol, whereas the *glpX* mutant do not show any change in the phenotype (Donahue *et al.*, 2000). Therefore, it is most likely that knockout of the *fbp* gene in *C. acetobutylicum* would result in a strain that is unable to grow on a gluconeogenic substrate, while, disrupting of the *glpX* gene might generate a strain with very tight CCR due to high level of intracellular FBP concentration.

6.2 Conclusions and future work

This study has demonstrated that *C. acetobutylicum* contains two genes encoding for class II and III FBPase enzymes. In order to confirm their physiological functions, three mutants need to be constructed, Δfbp , $\Delta glpX$, and a double mutant (Δfbp and $\Delta glpX$) and tested with different carbon sources to detect any variations in the growth phenotype. Also the intracellular FBP concentration of these mutants together with the wild type strain need to be tested to determine whether this is affected by growth with various carbon sources or not. These experiments would shed some light on how GlpX can interfere with CCR and also confirm the proposed function of Fbp. These findings could form the basis of constructing strains that would show optimum performance in fermentations using a range of substrates and conditions.

7 Appendix

7.1 Sequence of the recombinant fusion tags

Obtained from Novagen (2010) LIC[®] Cloning User Manual.

His-tag sequence: (molecular weight = 900 Da) HHHHHH.

S-tag sequence: (molecular weight = 4500 Da) LGTAAALPGAGHMAS.

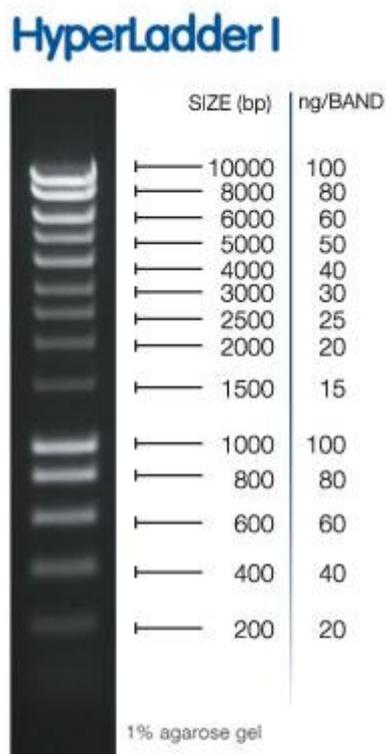
GTS-tag sequence: (molecular weight = 26 KDa)

MKLFYKPGACSLASHITLRESGKDFTLVSV DLMKKRLENGDDYFAVNPKGQVP
ALLDDGTL LTEGVAIMQYLADSV PDRQLLAPVNSISRYKTIEWLNYIATELHK
GFTPLFRPDTPEEYKSTVRAQLEKKLQYVNEALKDEHWICGQRFTIADAYLFTV
LRWAYAVKLNLEGLEHIAAFMQRMAERPEVQDALS A EGLK

7.2 Bioline Hyperladders

DNA molecular weight markers obtained from:

http://www.bioline.com/h_ladderguide.asp

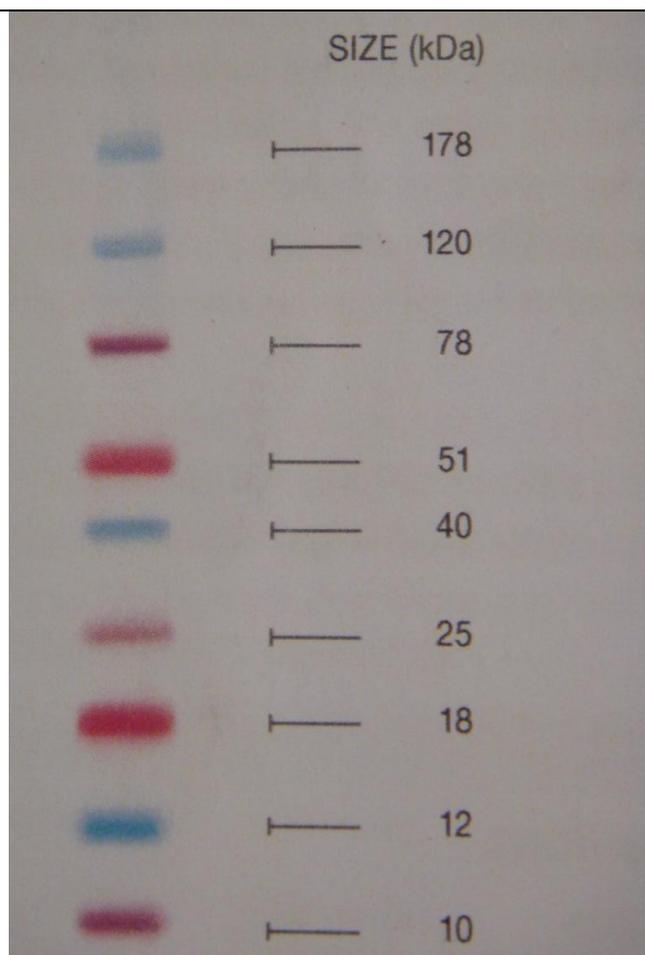


- Higher intensity bands:
1000bp and 10000bp
- Supplied in a ready-to-use format
- Each lane (5 μ l) provides
720ng of DNA

7.3 HyperPAGE Prestained Protein Marker

Prestained protein markers obtained from Bioline:

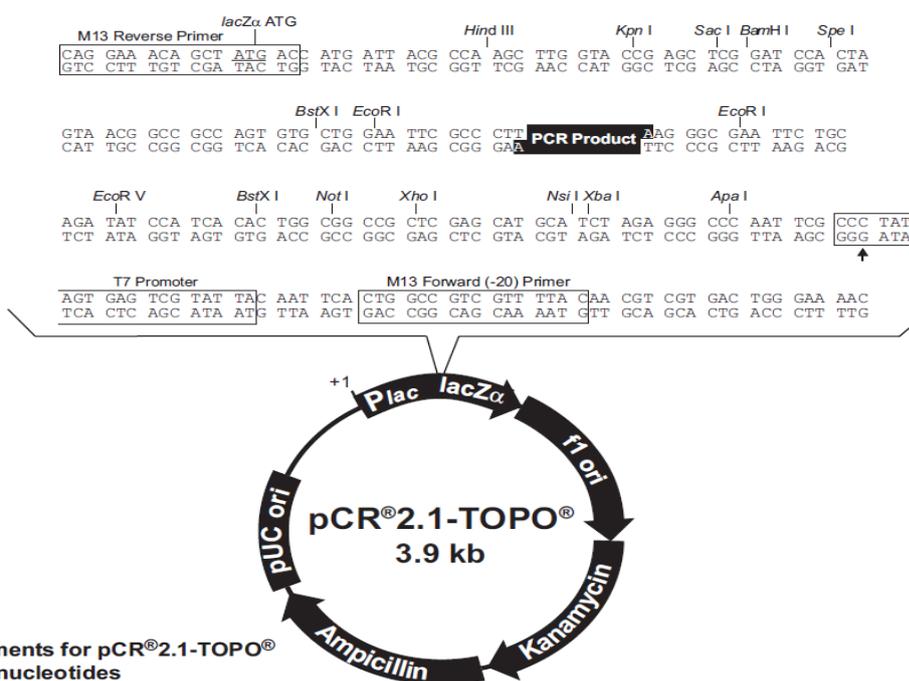
http://www.bioline.com/documents/product_inserts/HyperPAGE.pdf#zoom=130



7.4 pCR 2.1 TOPO Cloning Vector

Obtained from Invitrogen (2006) TOPO TA[®] Cloning User Manual.

http://www.invitrogen.com/content/sfs/manuals/topota_man.pdf



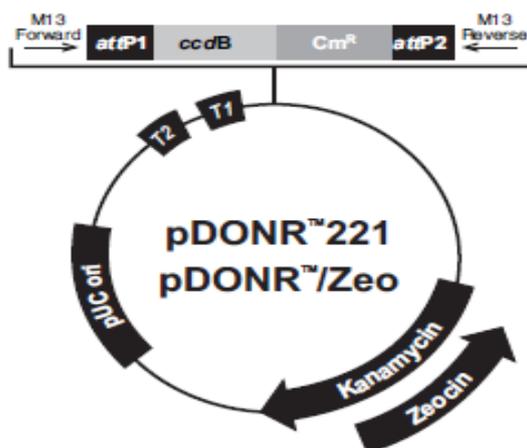
Comments for pCR[®]2.1-TOPO[®]
3931 nucleotides

LacZ α fragment: bases 1-547
M13 reverse priming site: bases 205-221
Multiple cloning site: bases 234-357
T7 promoter/priming site: bases 364-383
M13 Forward (-20) priming site: bases 391-406
f1 origin: bases 548-985
Kanamycin resistance ORF: bases 1319-2113
Ampicillin resistance ORF: bases 2131-2991
pUC origin: bases 3136-3809

7.5 The donor vector pDONRTM221.

Obtained from Invitrogen (2009) cloning user manual.

http://tools.invitrogen.com/content/sfs/manuals/gateway_pdonr_vectors.pdf



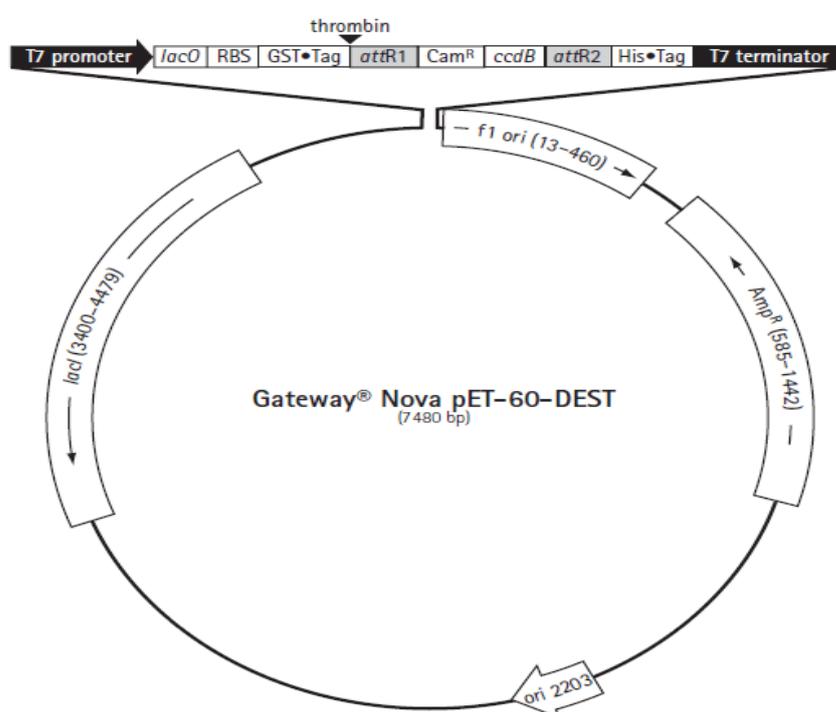
Comments for:

	pDONR TM 221 4762 nucleotides	pDONR TM /Zeo 4291 nucleotides
<i>rrbB</i> T2 transcription termination sequence (c):	268-295	268-295
<i>rrbB</i> T1 transcription termination sequence (c):	427-470	427-470
M13 Forward (-20) priming site:	537-552	537-552
attP1:	570-801	570-801
ccdB gene (c):	1197-1502	1197-1502
Chloramphenicol resistance gene (c):	1847-2506	1847-2506
attP2 (c):	2754-2985	2754-2985
T7 Promoter/priming site (c):	3000-3019	3003-3022
M13 Reverse priming site:	3027-3043	3027-3043
Kanamycin resistance gene:	3156-3965	—
EM7 promoter (c):	—	3486-3552
Zeocin resistance gene (c):	—	3111-3485
pUC origin:	4086-4759	3615-4288

7.6 The Gateway® Nova pET-60-DEST™ vector

Obtained from Novagen (2009) Gateway® Cloning User Manual.

http://www.merck-chemicals.com/united-kingdom/life-science-research/gateway-nova-pet60destexpressionsystem/EMD_BIO71863/p_C3Ob.s1OvmAAAAEjEhx9.zLX;sid=lqvX0FI_JYLd0B0mb5O6h_r_ESjFINyxuku9d2oR9QHR7kJlXIq6Hs4t1_OIJ7vpTHHzFHmESjFINW8hFvHzFHm?attachments=USP

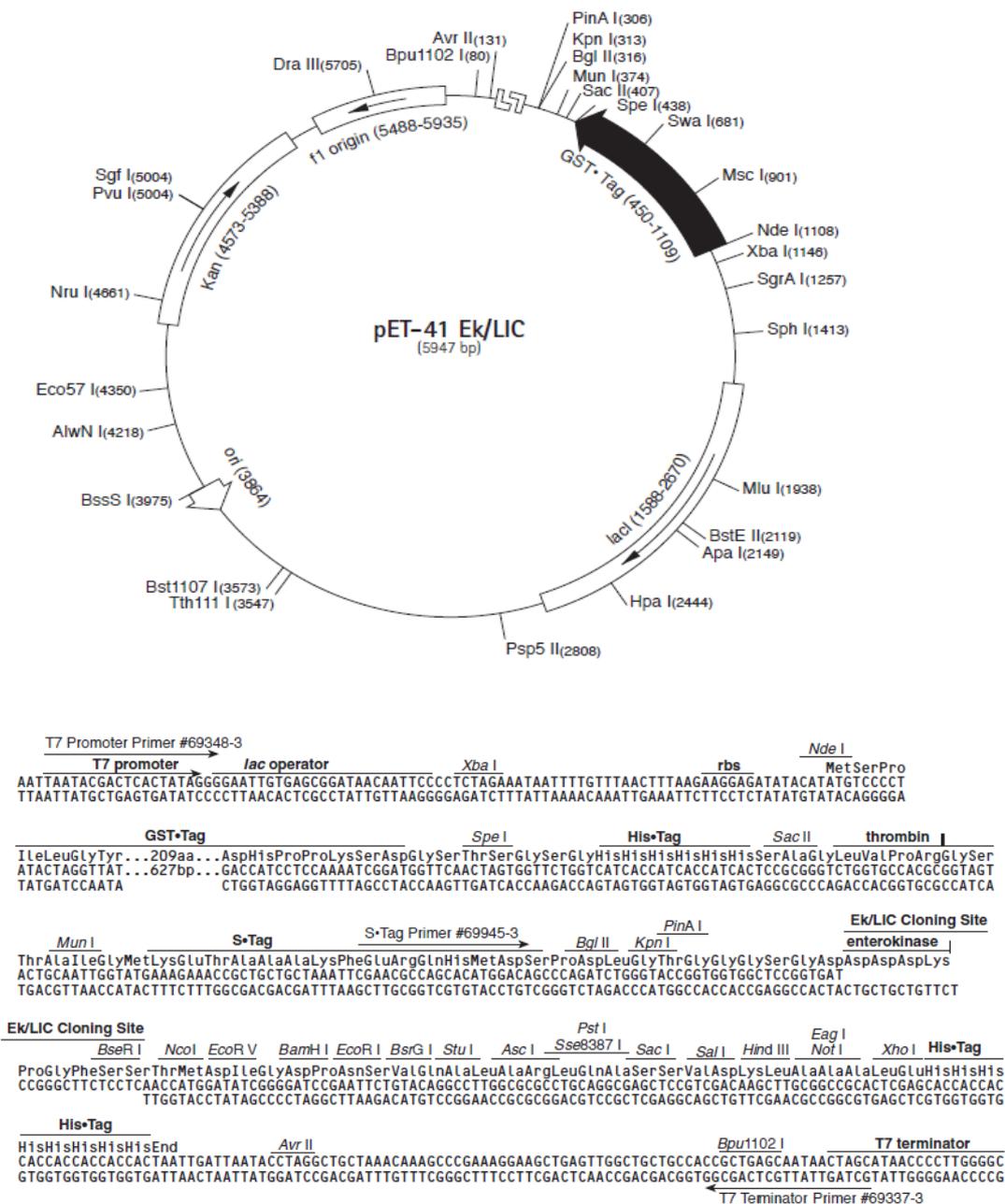


Feature	Location
f1 origin	13-460
Amp ^R	585-1442
pBR322 ori	2203
<i>lacI</i> coding sequence	3400-4479
T7 promoter	4870-4886
T7 transcription start	4887
GST•Tag™ coding sequence	4957-5616
Thrombin cleavage coding sequence	5623-5640
<i>attR1</i>	5644-5768
Cam ^R	5877-6533
<i>ccdB</i>	6878-7180
<i>attR2</i>	7224-7318
His•Tag® coding sequence	7350-7367
T7 terminator	7408-7455

7.7 The pET-41 Ek/LIC vector

Obtained from Novagen (2010) LIC[®] Cloning User Manual.

http://www.merck-chemicals.com/united-kingdom/life-science-research/pet-41-ek-lic-vector-kit/EMD_BIO-71071/p_BICb.s1O130AAAEj5Rt9.zLX?attachments=USP



7.8 Amino acids abbreviation

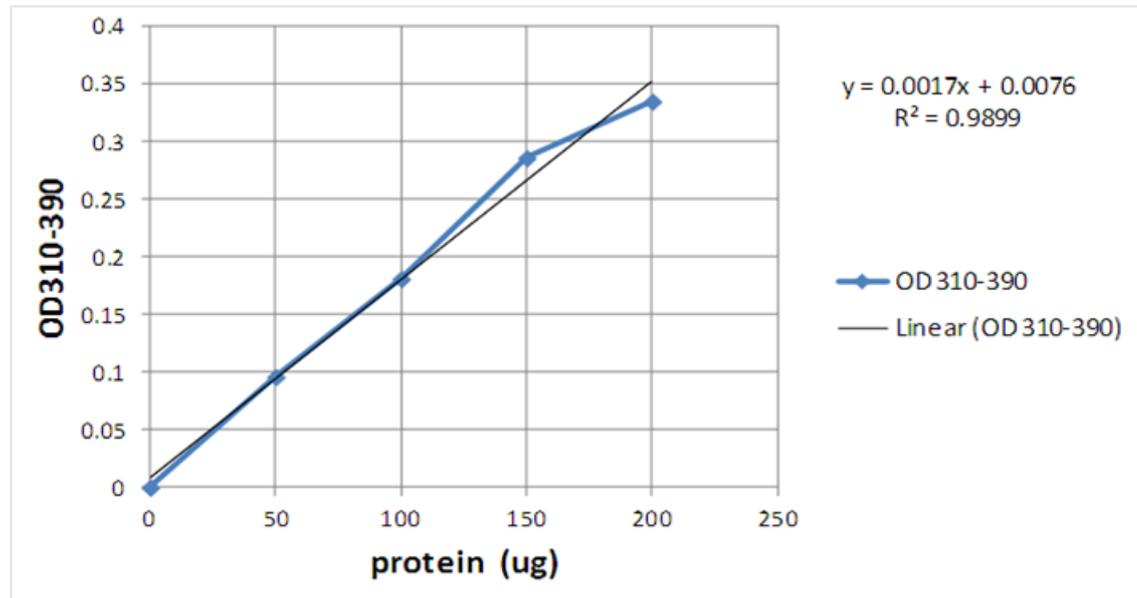
Obtained from the LabRat.com (2005).

<http://www.thelabrat.com/protocols/aminoacidtable.shtml>

Name	Abbreviation	
	3-Letter	1-Letter
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

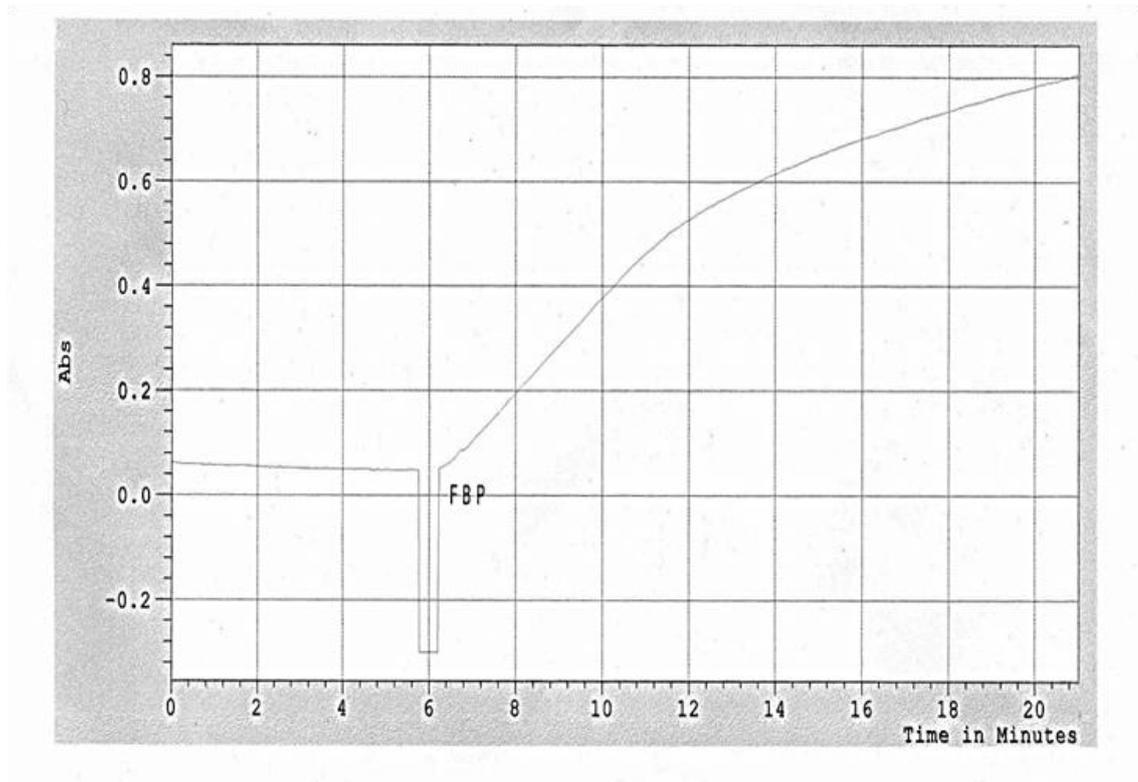
7.9 Calibration curve of protein

Typical calibration curve for estimation of concentration of protein.



7.10 FBPase enzyme assay curve

Typical FBPase enzyme assay curve for calculating the FBPase specific activity.



7.11 Proteomic analysis of FBPase class II (GlpX) protein

Typical output from liquid chromatography electrospray ionization tandem mass spectrometry (LC/ESI-MS/MS).

Match to: **gi|15894373** (*glpX*) Score: **92**

fructose 1,6-bisphosphatase II [Clostridium acetobutylicum ATCC 824]

Found in search of DATA.TXT

Nominal mass (Mr): **35015**; Calculated pI value: **5.08**

NCBI BLAST search of [gi|15894373](#) against nr

Unformatted [sequence string](#) for pasting into other applications

Fixed modifications: Carbamidomethyl (C)

Variable modifications: Oxidation (M)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

Sequence Coverage: **7%**

Matched peptides shown in **Bold Red**

```

1 MFDNDISMSL VRVTEAAALQ SSKYMGRGDK IGADQAAVDG MEKAFSEMPV
51 RGQVVIGEGE LDEAPMLYIG QKLGMGKDYM PEMDIAVDPL DGTILISKGL
101 PNAIAVIAMG PKGSL LHAPD MYMKKIVVGP GAKGAIDINK SPEENILNVA
151 KALNKDISEL TVIVQERERH DYIVKAAIEV GARVKLFGEG DVAAALACGF
201 EDTGIDILMG IGGAP EGVIA AAAIKCMGGE MQAQLIPHTQ EEIDRCHKMG
251 IDDVNKIFMI DDLVKSDNVF FAATAITECD LLKGIVFSKN ERAKTHSILM
301 RSKTGTIRFV EAIHDLNKS LVVE

```

8 References

- Antunes, A., Martin-Verstraete, I. and Dupuy, B.** (2011). CcpA-mediated repression of *Clostridium difficile* toxin gene expression. *Molecular Microbiology*, **79**(4):882-899.
- Barber, R. D., Harmer, D. W., Coleman, R. A and Clark, B. J.** (2005). GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. *Physiological Genomics*, **21**(3):389-95.
- Berg, J., Tymoczko, J. and Stryer, L.** (2002). Biochemistry (5th ed) United State of America, W. H. Freeman and Company.
- Bernard P.** (1996). Positive selection of recombinant DNA by CcdB. *BioTechniques*, **21**(2):320-3.
- Bernhauer, K., and K. Kurschner.** (1935). Butyl- und aceton- Garungen I Millerlung: uber Zwischenprodukte derbutanolaceton-Goring. *Biotechnology*. **280**:379-387.
- Berry, A. and Marshall, K. E.** (1993). Identification of zinc-binding ligands in the Class II fructose-1,6-bisphosphate aldolase of *Escherichia coli*. *FEBS Letters*, **318**(1):11-16.
- Blencke, H., Homuth, G., Ludwig, H., Mäder, U., Hecker, M. and Stülke, J.** (2003). Transcriptional profiling of gene expression in response to glucose in *Bacillus subtilis*: regulation of the central metabolic pathways. *Metabolic Engineering*, **5**(2):133-149.
- Brochu, D. and Vadeboncoeur, C.** (1999) The HPr (Ser) kinase of *Streptococcus salivarius*: purification, properties, and cloning of the *hprK* gene. *The Journal of Bacteriology* **181**(3):709-717.
- Brodeur, G., Yau, E., Badal, K., Collier, J., Ramachandran, K.B. and Ramakrishnan, S.** (2011). Chemical and physicochemical pretreatment of lignocellulosic biomass: a review. *Enzyme Research*. 10: 1-17.

- Brown, G., Singer, A., Lunin, V., Proudfoot, M., Skarina, T., Flick, R., Kochinyan, S., Sanishvili, R., Joachimiak, A., Edwards, A., Savchenko, A. and Yakunin, A.** (2009). Structural and biochemical characterization of the type II fructose-1,6-bisphosphatase GlpX from *Escherichia coli*. *Journal of Biological Chemistry*, **284**(6):3784-3792.
- Bult, C., White, O., Olsen, G., Zhou, L., Fleischmann, R., Sutton, G., Blake, J., FitzGerald, L., Clayton R., Gocayne, J., Kerlavage, A., Dougherty, B., Tomb, J., Adams, M., Reich, C., Overbeek, R., Kirkness, E., Weinstock, K., Merrick, J., Glodek, A., Scott, J., Geoghagen, N. and Venter, J.** (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**(5278):1058-73.
- Busby, S. and Ebright, R.** (1999). Transcription activation by catabolite activator protein (CAP). *Journal of Molecular Biology*, **293**(2):199-213.
- Chauvin, F., Brand, L. and Roseman, S.** (1996). Enzyme I: the first protein and potential regulator of the bacterial phosphoenolpyruvate:glycose phosphotransferase system. *Research in Microbiology*, **147**:471-479.
- Chavarría, M., Santiago, C., Platero, R., Krell, T., Casasnovas, T. M. and de Lorenzo, V.** (2011). Fructose 1-phosphate is the preferred effector of the metabolic regulator Cra of *Pseudomonas putida*. *Journal of Biological Chemistry*, **286**:9351-9359.
- Cheng, X. and Lee, J.** (1998). Interactive and dominant effects of residues 128 and 141 on cyclic nucleotide and DNA bindings in *E. coli* cAMP receptor protein. *Journal of Biological Chemistry*, **273**:705-712.
- Col, B.** (2004). Regulation of Fructose 1,6-bisphosphatase II (GlpX) Gene Expression in *Escherichia coli*, *Virginia*, Virginia Polytechnic Institute.
- Crasnier-Mednansky, M., Park, M., Studley, W. and Saier Jr, M.** (1997). Cra-mediated regulation of *E. coli* adenylate cyclase. *Microbiology*, **143**:785-792.

- Da Silva, G., Mack, M. and Contiero, J.** (2009). Glycerol: a promising and abundant carbon source for industrial microbiology. *Biotechnology Advances*, **27**(1):30-39.
- Dasari, M.** (2007). Crude glycerol potential described. *Feedstuffs*, **79**:1-3.
- Davanloo, P., Rosenberg, A. H., Dunn, J. J. and Studier, F.W.** (1984). Cloning and expression of the gene for bacteriophage T7 RNA polymerase. *The Proceedings of the National Academy of Sciences USA*, **81**: 2035-2039.
- Davanloo, P., Rosenberg, A., Dunn, J. and Studier, F.** (1984). Cloning and expression of the gene for bacteriophage T7 RNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*, **81**:2035-2039.
- Davies, R. and M. Stephenson.** (1941). Studies on the acetonebutyl alcohol fermentation. 1. Nutritional and other factors involved in the preparation of active suspensions of *C. acetobutylicum* (Weizmann). *Biochemistry Journal* **35**:1320-1331.
- Davison, S., Santangelo, J., Reid, S. and Woods, D.** (1995). A *Clostridium acetobutylicum* regulator gene (regA) affecting amylase production in *Bacillus subtilis*. *Microbiology*, **141**(4):989-996.
- Demirbas, A.** (2007). Importance of Biodiesel as Transportation Fuel. *Energy Policy* **35**: 4661-4670.
- Deutscher, J., Reizer, J., Fischer, C., Galinier, A., Saier, M. and Steinmetz M.** (1994). Loss of protein kinase-catalyzed phosphorylation of HPr, a phosphocarrier protein of the phosphotransferase system, by mutation of the *ptsH* gene confers catabolite repression resistance to several catabolic genes of *Bacillus subtilis*. *Journal of Bacteriology*, **176**(11):3336-44.

- Deutscher, J. and Saier Jr., M.** (1983). ATP-dependent protein kinase-catalyzed phosphorylation of a seryl residue in HPr, a phosphate carrier protein of the phosphotransferase system in *Streptococcus pyogenes*. *Proceedings of the National Academy of Sciences of the United States of America* **80**(22):6790-6794.
- Deutscher, D., Meilijson, I., Kupiec, M. and Ruppin, E.** (2006). Multiple knockout analysis of genetic robustness in the yeast metabolic network. *Nature Genetics* **38**(9):993-998.
- Deutscher, J.** (2008). The mechanisms of carbon catabolite repression in bacteria. *Current Opinion in Microbiology*, **11**(2):87-93.
- Deutscher, J., and Engelmann, R.** (1984). Purification and characterization of an ATP-dependent protein kinase from *Streptococcus faecalis*. *FEMS Microbiology Letters* **23**:157-162.
- Deutscher, J., Francke, C. and Postma, P.** (2006). How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiology and Molecular Biology Reviews*, **70**:939-1031.
- Deutscher, J., Pevec, B., Beyreuther, K., Kiltz, H. H. and Hengstenberg, W.** (1986). Streptococcal phosphoenolpyruvate-sugar phosphotransferase system: amino acid sequence and site of ATP-dependent phosphorylation of HPr. *Biochemistry*, **25**(21):6543-6551.
- Deutscher, J., U. Kessler, and W. Hengstenberg.** (1985). Streptococcal phosphoenolpyruvate: sugar phosphotransferase system: purification and characterization of a phosphoprotein phosphatase which hydrolyzes the phosphorylbond in seryl-phosphorylated histidine-containing protein. *The Journal of Bacteriology* **163**:1203-1209.
- Dienert, F.** (1900). Sur la fermentation du galactose et sur l'accoutumance des levures à ce sucre. *Ann. Pasteur*, **14**:139-189.

- Dobson, R., Gray, V. and Rumbold, K.** (2011). Microbial utilization of crude glycerol for the production of value-added products. *Journal of Industrial Microbiology and Biotechnology*, DOI: 10.1007/s10295-011-1038-0.
- Donahue, J., Bownas, J., Niehaus, W. and Larson, T.** (2000). Purification and characterization of glpX-encoded fructose 1,6-bisphosphatase, a new enzyme of the glycerol 3-phosphate regulon of *Escherichia coli*. *Journal of Bacteriology*, **182**(19):5624-5627.
- Du, J., Say, R. F., Lü, W., Fuchs, G. and Einsle, O.** (2011). Active-site remodelling in the bifunctional fructose-1,6-bisphosphate aldolase/phosphatase. *Nature*, **478**(7370):534-537.
- Dürre, P., Böhringer, M., Nakotte, S., Schaffer, S., Thormann, K. and Zickner, B.** (2002). Transcriptional regulation of solventogenesis in *Clostridium acetobutylicum*. *Journal of Molecular Microbiology and Biotechnology* **4**(3):295-300.
- Dürre, P.** (2007). Biobutanol: An Attractive Biofuel. *Journal of Biotechnology*. **2**: 1525–1534.
- Eisermann, R., Deutscher, J., Gonzy-Tre´boul, G. and Hengstenberg W.,** (1988). Site-directed mutagenesis with the *ptsH* gene of *Bacillus subtilis*. Isolation and characterization of heat-stable proteins altered at the ATP dependent regulatory phosphorylation site. *Journal of Biological Chemistry* **263**:17050-17054.
- Ezeji, T., Qureshi, N. and Blaschek, H. P.** (2007). Bioproduction of butanol from biomass: from genes to bioreactors. *Current Opinion in Biotechnology* **18**(3):220-227.
- Fujita, Y.** (2009). Carbon catabolite control of the metabolic network in *Bacillus subtilis*, *Bioscience, Biotechnology and Biochemistry*, **73**:245-259.

- Fushinobu, S., Nishimasu, H., Hattori, D. and Song, H. J. and Wakagi, T.** (2011). Structural basis for the bifunctionality of fructose-1,6-bisphosphate aldolase/phosphatase. *Nature*, **478**:538-541.
- Gaberc-Porekar, V. and Menart, V.** (2001). Perspectives of immobilized-metal affinity chromatography. *Journal of Biochemical and Biophysical Methods*, **49**:335-360.
- Gabriel, C. L., and Crawford, F. M.** (1930). Development of the butyl-acetonic fermentation industry. *Industrial and Engineering Chemistry*, **22**:1163-1165.
- Galinier, A., Deutscher, J. and Martin-Verstraete, I.** (1999). Phosphorylation of either Crh or HPr mediates binding of CcpA to the *Bacillus subtilis xyn cre* and catabolite repression of the *xyn* operon. *Journal of Molecular Biology*, **286**(2):307-314.
- Galinier, A., Haiech, J., Kilhoffer, M., Jaquinod, M., Stülke, J., Deutscher, J. and Martin-Verstraete I.** (1997). The *Bacillus subtilis crh* gene encodes a HPr-like protein involved in carbon catabolite repression. *Proceedings of the National Academy of Sciences*, **94**(16):8439-8444.
- Görke, B., Fraysse, L. and Galinier A.** (2004). Drastic differences in Crh and HPr synthesis levels reflect their different impacts on catabolite repression in *Bacillus subtilis*. *Journal of Bacteriology*, **186**:2992-2995.
- Greenberg, D. M., Wynston, L. K. and Nagabhushanan, A.** (1965). Further studies on N⁵-formyltetrahydrofolic acid cyclodehydrase. *Biochemistry* **4**(9):1872–1878.
- Gutka, H. J., Rukseree, K., Wheeler, P. R., Franzblau, S. G. and Movahedzadeh, F.** (2011). *glpX* gene of *Mycobacterium tuberculosis*: heterologous expression, purification, and enzymatic characterization of the encoded fructose 1,6-bisphosphatase II. *Applied Biochemistry and Biotechnology* **164**(8):1376-1389.

- Groom, M. J., Gray, E. M., Townsend, P. A.** (2008). Biofuels and biodiversity: principles for creating better policies for biofuel production. *Conservation Biology*. **3**:602-609.
- Hahn-Hägerdal, B., Galbe, M., Gorwa-Grauslund, M. F., Lidén, G. and Zacchi, G.** (2006). Bio-ethanol - the fuel of tomorrow from the residues of today. *Trends in Biotechnology* **24**(12):549-556.
- Hansen, M., Hsu, A. L., Dillin, A. and Kenyon, C.** (2005). New genes tied to endocrine, metabolic, and dietary Regulation of Longevity in *C. elegans*. *PLoS Genetics* **1**(1):e17.
- Hanson, K. G., Steinhauer, K., Reizer, J., Hillen, W. and Stülke, J.** (2002). HPr kinase/phosphatase of *Bacillus subtilis*: expression of the gene and effects of mutations on enzyme activity, growth and carbon catabolite repression. *Microbiology*, **148**(6):1805-11.
- Hartley, J. L., Temple, G. F. and Brasch M. A.** (2000). DNA cloning using in vitro site-specific recombination. *Genome Research*, **10**:1788-1795.
- Hartmanis, M. G. N., Klason, T. and Gatenbeck, S.** (1984). Uptake and activation of acetate and butyrate in *Clostridium acetobutylicum*. *Applied Microbiology and Biotechnology* **20**(1):66-71.
- Hastings, J. H. J.** (1971). Development of the fermentation industries in Great Britain, p' 1-45. In D. Perlman (ed.) *Advances in Applied Microbiology*, vol. 14. Academic Press, Inc., New York.
- Henkin, T. M.** (1996). The role of the CcpA transcriptional regulator in carbon metabolism in *Bacillus subtilis*. *FEMS Microbiology Letters* **135**: 9-15.
- Hines, J., Fromm, H. and Honzatko, R.** (2006). Novel allosteric activation site in *Escherichia coli* fructose-1,6-bisphosphatase. *Journal of Biological Chemistry* **281**(27):18386-18393.

- Hogema, B., Arents, J., Bader, R., Eijkemans, K., Yoshida, H., Takahashi, H., Aiba, H. and Postma, P.** (1998). Inducer exclusion in *Escherichia coli* by non-PTS substrates: the role of the PEP to pyruvate ratio in determining the phosphorylation state of enzyme IIAGlc. *Molecular Microbiology* **30**:487-498.
- Hueck, C. and Hillen, W.** (1995). Catabolite repression in *Bacillus subtilis*: a global regulatory mechanism for the Gram-positive bacteria? *Molecular Microbiology*, **15**(3): 395-401.
- Jault, J., Fieulaine, S., Nessler, S., Gonzalo, P., Di Pietro, A., Deutscher, J. and Galinier, A.** (2000). The HPr kinase from *Bacillus subtilis* is a homooligomeric enzyme which exhibits strong positive cooperativity for nucleotide and fructose 1,6-bisphosphate binding. *Journal of Biological Chemistry*, **275**(3):1773-1780.
- Jeanine, C. and Antonio, M. R.** (1983). Growth Inhibition Kinetics for the Acetone-Butanol Fermentation. *Foundations of Biochemical Engineering*. Washington, D.C., American Chemical Society. ISBN: 9780841207523, pp 501-512.
- Johnson, K., Chen, L., Yang, H., Roberts, M. and Stec, B.** (2001). Crystal structure and catalytic mechanism of the *MJO109* gene product: A bifunctional enzyme with inositol monophosphatase and fructose 1,6-bisphosphatase activities. *Biochemistry*, **40**(3):618-630.
- Johnson, M. J., W. H. Peterson, and Fred E. B.** (1931). Oxidation and reduction relations between substrate and products in the acetone-butyl alcohol fermentation. *Biochemistry Journal*. **91**:569-59.
- Jones, B. E., Dossonnet, V. r., Ka uster, E., Hillen, W., Deutscher, J. and Klevit, R. E.** (1997). Binding of the catabolite repressor protein CcpA to its DNA target is regulated by phosphorylation of its corepressor HPr. *Journal of Biological Chemistry* **272**(42):26530-26535.

- Jones, D. T., and Woods D. R.** (1986). Acetone-Butanol Fermentation Revisited. *Microbiological Reviews*. **50**(4):484-524.
- Jules, M., Le Chat, L., Aymerich S. and Le Coq, D.** (2009). The *Bacillus subtilis* *ywjI* (*glpX*) gene encodes a class ii fructose-1,6-bisphosphatase, functionally equivalent to the class III fbp enzyme. *Journal of Bacteriology*, **191**(9):3168-3171.
- Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bott, M., Burkovski, A., Dusch, N., Eggeling, L., Eikmanns, B., Gaigalat, L., Goesmann, A., Hartmann, M., Huthmacher, K., Krämer, R., Linke, B., McHardy, A., Meyer, F., Möckel, B., Pfefferle, W., Pühler, A., Rey, D., Rückert, C., Rupp, O., Sahm, H., Wendisch, V., Wiegräbe, I. and Tauch, A.** (2003). The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of -aspartate-derived amino acids and vitamins. *Journal of Biotechnology*, **104**(1-3):5-25.
- Killeffer, D. H.** (1927). Butanol and acetone from corn: A description of the fermentation process. *Industrial & Engineering Chemistry* **19**(1):46-50.
- Kim, J., Voskuil, M. and Chambliss, G.** (1998). NADP, corepressor for the *Bacillus* catabolite control protein CcpA. *Proceedings of the National Academy of Sciences*, **95**(16):9590-9595.
- Kim, J., Yang, Y. and Chambliss G.** (2005). Evidence that *Bacillus catabolite* control protein CcpA interacts with RNA polymerase to inhibit transcription. *Molecular Microbiology*, **56**(1):155-162.
- Kim, B. H. and Zeikus, J. G.** (1985). Importance of hydrogen metabolism in regulation of solventogenesis by *Clostridium acetobutylicum*. *Developments In Industrial Microbiology* **26**:1-14.

- Kleijn, R. J., Buescher, J. M., Chat, L. L., Jules, M., Aymerich, S. and Sauer U.** (2010). Metabolic fluxes during strong carbon catabolite repression by malate in *Bacillus subtilis*. *Journal of Biological Chemistry*, **285**:1587-1596.
- Kravanja, M., Engelmann, R., Dossonnet, V., Blüggel, M., Meyer, H., Frank, R., Galinier, A., Deutscher, J., Schnell, N. and Hengstenberg, W.** (1999). The *hprK* gene of *Enterococcus faecalis* encodes a novel bifunctional enzyme: the HPr kinase/phosphatase. *Molecular Microbiology*, **31**(1):59-66.
- Kuang-Yu, K. and Saier Jr, M.** (2002). Phylogeny of phosphoryl transfer proteins of the phosphoenolpyruvate-dependent sugar-transporting phosphotransferase system. *Research in Microbiology*, **153**:405-415.
- Kuit, W., Minton, N. P., López-Contreras, A. M., and Eggink, G.** (2012). Disruption of the acetate kinase (ack) gene of *Clostridium acetobutylicum* results in delayed acetate production. *Applied Microbiology and Biotechnology*, **94**(3):729-741.
- Kundig, W., Ghosh, S., and Roseman, S.** (1964). Phosphate bound to histidine in a protein as an intermediate in a novel phosphotransferase system. *Proceedings of the National Academy of Sciences of the United States of America*. **52**: 1067–1074.
- Lawson, C. L., Swigon, D., Murakam, K. S., Darst, S. A., Berman, H. M. and Ebright, R. H.** (2004). Catabolite activator protein (CAP): DNA binding and transcription activation. *Current Opinion in Structural Biology*, **14**(1):10-20.
- Lee, S. Y., Park, J. H., Jang, S. H., Nielsen, L. K., Kim, J. and Jung, K. S.** (2008). Fermentative butanol production by Clostridia. *Biotechnology and Bioengineering*. **101**(2):209-228.
- Leung, J. C., and Wang D. I.** (1981). Production of acetone and butanol by *Clostridium acetobutylicum* in continuous culture using free cells and immobilized cells. the Second World War Congress. *Chemical Engineering*. **1**:348-352.

- Li, J. and Lee, J. C.** (2011). Modulation of allosteric behaviour through adjustment of the differential stability of the two interacting domains in *E. coli* cAMP receptor protein *Biophysical Chemistry*, **159**(1):210-216.
- Lorca, G., Chung, Y., Barabote, R., Weyler, W., Schilling, C. and Saier Jr, M.** (2005). Catabolite repression and activation in *Bacillus subtilis*: dependency on CcpA, HPr, and HprK. *Journal of Bacteriology*, **187**(22):7826-7839.
- Martin-Verstraete, I., Deutscher, J. and Galinier, A.** (1999). Phosphorylation of HPr and Crh by HprK, early steps in the catabolite repression signalling pathway for the *Bacillus subtilis* levanase operon. *Journal of Bacteriology* **181**(9):2966-2969.
- Martin-Verstraete, I., Débarbouillé, M., Klier, A. and Rapoport, G.,** (1990). Levanase operon of *Bacillus subtilis* includes a fructose-specific phosphotransferase system regulating the expression of the operon. *Journal of Molecular Biology*, **214**(3):657-671.
- Mijakovic, I., Poncet, S., Galinier, A., Monedero, V., Fieulaine, S., Janin, J., Nessler, S., Marquez, J. A., Scheffzek, K., Hasenbein, S., Hengstenberg, W. and Deutscher, J.** (2002). Pyrophosphate-producing protein dephosphorylation by HPr kinase/ phosphorylase: A relic of early life? *Proceedings of the National Academy of Sciences of the United States of America*, **21**:13442-13447.
- Miwa, Y., Nakata, A., Ogiwara, A., Yamamoto, M. and Fujita, Y.** (2000). Evaluation and characterization of catabolite-responsive elements (*cre*) of *Bacillus subtilis*. *Nucleic Acids Research*, **28**(5):1206-1210.
- Monod, J.** (1942). Recherches sur la croissance des cultures bactériennes. *Hermann et Cie*, Paris, France.

- Moreno, M., Schneider, B., Maile, R., Weyler, W. and Saier Jr, M.** (2001). Catabolite repression mediated by the CcpA protein in *Bacillus subtilis*: novel modes of regulation revealed by whole-genome analyses. *Molecular Microbiology*, **39**(5):1366 - 1381.
- Montecucco, C. and Molgó, J.** (2005). Botulin neurotoxins: revival of an old killer. *Current Opinion in Pharmacology*, **5**(3):274-279.
- Nakas, J. P., Schaedle, M., Parkinson, C. M., Coonley, C. E. and Tanenbaum, S. W.** (1983). System development for linked-fermentation production of solvents from algal biomass. *Applied and Environmental Microbiology* **46**(5):1017-1023.
- Nessler, S., Fieulaine, S., Poncet, S., Galinier, A., Deutscher, J. and Janin, J.** (2003). HPr kinase/phosphorylase, the sensor enzyme of catabolite repression in gram-positive bacteria: structural aspects of the enzyme and the complex with its protein substrate. *Journal of Bacteriology*, **185**(14):4003-4010.
- Newton, C. R. and Graham, A.** (1994). PCR, Oxford, BIOS Scientific publishers Ltd.
- Nihashi, J. and Fujita, Y.** (1984). Catabolite repression of inositol dehydrogenase and gluconate kinase syntheses in *Bacillus subtilis*. *Biochimica et Biophysica Acta*, **798**(1):88-95.
- Nishimasu, H., Fushinobu, S., Shoun, H. and Wakagi, T.** (2004). The first crystal structure of the novel class of fructose-1,6-bisphosphatase present in thermophilic archaea, *Structure*, **12**(6):949-959.
- Poncet, S., Mijakovic, I., Nessler, S., Gueguen-Chaignon, V., Chaptal, V., Galinier, A., Boël, G., Mazé, A. and Deutscher, J.** (2004). HPr kinase/phosphorylase, a Walker motif A-containing bifunctional sensor enzyme controlling catabolite repression in Gram-positive bacteria. *Biochimica et Biophysica Acta* **1697**(1):123-135.

- Postma, P. W., Lengeler, J. W. and Jacobson, G. R.** (1993). Phosphoenolpyruvate: carbohydrate phosphotransferase systems of bacteria. *Microbiology and Molecular Biology Reviews* **57**(3):543-594.
- Presecan-Siedel, E., Galinier, A., Longin, R., Deutscher, J., Danchin, A., Glaser, P. and Martin-Verstraete, I.** (1999). Catabolite regulation of the *pta* gene as part of carbon flow pathways in *Bacillus subtilis*. *Journal of Bacteriology*, **181**(22):6889-6897.
- Pribat, A., Blaby, I. K., Lara-Núñez, A., Jeanguenin, L., Fouquet, R., Frelin, O., Gregory, J. F., Philmus, B., Begley, T. P., de Crécy-Lagard, V. and Hanson, A. D.** (2011). A 5-formyltetrahydrofolate cycloligase paralog from all domains of life: comparative genomic and experimental evidence for a cryptic role in thiamin metabolism. *Functional and Integrative Genomics*, **11**(3):467-478.
- Qureshi, N. and Maddox, I. S.** (2005). Reduction in butanol inhibition by perstraction: Utilization of concentrated lactose/whey permeate by *Clostridium acetobutylicum* to enhance butanol fermentation economics. *Food and Bioproducts Processing*, **83**(1):43-52.
- Qureshi, N. and Blaschek, H. P.** (2001). Evaluation of recent advances in butanol fermentation, upstream, and downstream processing. *Bioprocess Biosystems Engineering* **24**:219-226.
- Rapley, R. and Manning, D. L.** (1998). RNA Isolation and Characterization Protocols. United State of America, Humana Press.
- Ram, M. S. and Seenayya, G.** (2005). Ethanol production by *Clostridium thermocellum* SS8, a newly isolated thermophilic bacterium. *Biotechnology Letters* **11**:589-592.
- Ramadhas, A. S., Jayaraj, S. and Muraleedharan, C.** (2003). Use of vegetable oils as I.C. engine fuels-A review. *Renewable Energy* **29**(5):727-742.

- Ramseier, T. M., Nègre, D. Cortay, J. C., Scarabel, M., Cozzone, A. J. and Saier Jr, M.** (1993). *In Vitro* binding of the pleiotropic transcriptional regulatory protein, FruR, to the *fru*, *pps*, *ace*, *pts* and *icd* operons of *E. coli* and *S. typhimurium*. *Journal of Molecular Biology*, **234**(1):28-44.
- Rashid, N., Imanaka, H., Kanai, T., Fukui, T., Atomi, H. and Imanaka T.** (2002). A novel candidate for the true fructose-1,6-bisphosphatase in Archaea. *Journal of Biological Chemistry*, **277**(34):30649-30655.
- Reilly, J., W. J. Hickinbottom, F. R. Henley, and A. C. Thaysen.** (1920). The products of the "acetone:n-butyl alcohol" fermentation of carbohydrate material with special reference to some of the intermediate substances produced. *Biochemistry Journal*. **14**:229-251.
- Reizer, J., Sutrina, S. L., Saier, Jr., M., Stewart, G. C., Peterkofsky, A. and Reddy, P.** (1989). Mechanistic and physiological consequences of HPr (ser) phosphorylation on the activities of the phosphoenolpyruvate: sugar phosphotransferase system in gram-positive bacteria: studies with site-specific mutants of HPr. *European Molecular Biology Organization journal* **8**:2111-2120.
- Ren, C., Gu, Y., Hu, S., Wu, Y., Wang, P., Yang, Y., Yang, C., Yang, S. and Jiang, W.** (2010). Identification and inactivation of pleiotropic regulator CcpA to eliminate glucose repression of xylose utilization in *Clostridium acetobutylicum*. *Metabolic Engineering*, **12**(5):446-454.
- Ringel, A. K., Wilkens, E., Hortig, D., Willke, T. and Vorlop, K. D.** (2011). An improved screening method for microorganisms able to convert crude glycerol to 1,3-propanediol and to tolerate high product concentrations. *Applied Microbiology and Biotechnology*, DOI: 10.1007/s00253-011-3594-7.
- Rittmann, D., Schaffer, S., Wendisch, V. and Sahn, H.** (2003). Fructose-1,6-bisphosphatase from *Corynebacterium glutamicum* expression and deletion of the *fbp* gene and biochemical characterization of the enzyme. *Archives of Microbiology*, **180**(4):285-292.

- Rogers, P.** (1984). Genetics and biochemistry of *Clostridium* relevant to development of fermentation processes. *Applied. Microbiology* **31**:1-60.
- Rose, A. H.** (1961). Industrial Microbiology. Butterworths, London. ISBN 80-901087-0-9 pp.160-166.
- Saier Jr, M. and Ramseier, T. M.** (1996). The catabolite repressor/activator (Cra) protein of enteric bacteria. *Journal of Bacteriology*, **178**(12):3411-3417.
- Saier Jr, M. and Reizer, J.** (1992). Proposed uniform nomenclature for the proteins and protein domains of the bacterial phosphoenolpyruvate: sugar phosphotransferase system. *Journal of Bacteriology*, **174**:1433-1438.
- Say, R. F. and Fuchs, G.** (2010). Fructose 1,6-bisphosphate aldolase/phosphatase may be an ancestral gluconeogenic enzyme. *Nature*, **464**:1077-1081.
- Schneck, V. K., Sands, J. A. and Montenecourt, B. S.** (1984). Effect of butanol on lipid composition and fluidity of *Clostridium acetobutylicum* ATCC 824. *Applied. Environmental Microbiology* **47**:193-194.
- Schumacher, M., Seidel, G., Hillen, W. and Brennan, R.** (2006). Phosphoprotein Crh-Ser46-P displays altered binding to CcpA to effect carbon catabolite regulation. *Journal of Biological Chemistry*, **281**(10):6793-6800.
- Schumacher, M., Seidel, G., Hillen, W. and Brennan, R.** (2007). Structural mechanism for the fine-tuning of CcpA function by the small molecule effectors glucose 6-phosphate and fructose 1,6-bisphosphate. *Journal of Molecular Biology*, **368**(4):1042-1050.
- Scotcher, M. C. and Bennett, G. N.** (2005). SpoIIE regulates sporulation but does not directly affect solventogenesis in *Clostridium acetobutylicum* ATCC 824. *Journal of Bacteriology*. **6**:1930-1936.

- Shapovalov, O. and Ashkinazi, L.** (2008). Biobutanol: biofuel of second generation. *Russian Journal of Applied Chemistry* **81**(12):2232-2236.
- Shimada, T., Yamamoto, K. and Ishihama, A.** (2011). Novel members of the Cra regulon involved in carbon metabolism in *E. coli*. *Journal of Bacteriology*, **193**(3):649-659.
- Siebers, B., Brinkmann, H., Dörr, C., Tjaden, B., Lilie, H., van der Oost, J. and Verhees, C. H.** (2001). Archaeal fructose-1,6-bisphosphate aldolases constitute a new family of archaeal type class I aldolase. *Journal of Biological Chemistry*, **276**:28710-28718.
- Singh, K. D., Schmalisch, M. H. Stülke, J. and Görke B.** (2008). Carbon catabolite repression in *Bacillus subtilis*: quantitative analysis of repression exerted by different carbon sources. *Journal of Bacteriology*, **19**(21):7275-7284.
- Söhngen, N. L. and Coolhaas, C.** (1924). The fermentation of galactose by *Saccharomyces cerevisiae*. *Journal of Bacteriology*, **9**(2):131-141.
- Stec, B., Yang, H., Johnsonm, K. A., Chen, L. and Roberts, M.F.** (2000). MJ0109 is an enzyme that is both an inositol monophosphatase and the 'missing' archaeal fructose-1,6-bisphosphatase. *Nature Structural & Molecular Biology* **7**(11):1046-50.
- Tangney, M. and Mitchell W. J.** (2005). Carbohydrate uptake by the phosphotransferase system and other mechanisms. P. Dürre (Ed). CRC Press LLC, ISBN 0-8493-1618-9 Chapter 8 pp155-175.
- Tangney, M. and Mitchell W. J.** (2005). Regulation of Catabolic Gene Systems. In: *Handbook on Clostridia*. P. Dürre (Ed). CRC Press LLC, ISBN 0-8493-1618-9 Chapter 25 pp583-606.
- Tangney, M. and Mitchell W. J.** (2007). Characterisation of a glucose phosphotransferase system in *Clostridium acetobutylicum* ATCC 824. *Applied Microbiology and Biotechnology* **74**:398-405.

- Tangney, M., and Mitchell, W. J.** (2000). Analysis of a catabolic operon for sucrose transport and metabolism in *Clostridium acetobutylicum* ATCC 824. *Journal of Molecular Microbiology and Biotechnology*, **2**:71-80.
- Tangney, M., Galinier, A., Deutscher, J. and Mitchell, W. J.** (2003). Analysis of the elements of catabolite repression in *Clostridium acetobutylicum* ATCC 824. *Journal of Molecular Microbiology and Biotechnology* **6**(1):6-11.
- Tangney, M., Winters, G. T. and Mitchell, W. J.** (2001). Characterization of a maltose transport system in *Clostridium acetobutylicum* ATCC 824. *Journal of Industrial Microbiology and Biotechnology*, **27**:298-306.
- Terpe, K.** (2003). Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology*, **60**(5): 523-533.
- Thevenot, T., Brochu, D., Vadeboncoeur, C. and Hamilton, I.R.** (1995). Regulation of ATP-dependent P-(Ser)-HPr formation in *Streptococcus mutans* and *Streptococcus salivarius*. *Journal of Bacteriology*, **177**(10):2751-2759.
- Thompson, J. C. and He, B. B.** (2006). Characterization of crude glycerol from biodiesel production from multiple feedstocks. *Applied Engineering in Agriculture*, **22**(2):261-265.
- Van der voort, M., Kuipers, O. P., Buist, G., De Vos, W. M. and Abee, T.** (2008). Assessment of CcpA-mediated catabolite control of gene expression in *Bacillus cereus* ATCC 14579. *BMC Microbiology*, **8**(1):62-74.
- Varga, J. J., Stirewalt, V. L. and Melville, S. B.** (2004). The CcpA protein is necessary for efficient sporulation and enterotoxin gene (*cpe*) regulation in *Clostridium perfringens*. *Journal of Bacteriology*, **186**(16):5221-5229.

- Varga, J. J., Therit, B. and Melville, S. B.** (2008). Type IV pili and the CcpA protein are needed for maximal biofilm formation by the gram-positive anaerobic pathogen *Clostridium perfringens*. *Infection and Immunity*, **76**(11):4944-4951.
- Vasconcelos, I., Girbal, L. and Soucaille, P.** (1994). Regulation of carbon and electron flow in *Clostridium acetobutylicum* grown in chemostat culture at neutral pH on mixtures of glucose and glycerol. *Journal of Bacteriology*, **176**:1443-1450.
- Vilar, J. M., Guet, C.C. and Leibler, S.** (2003). Modeling network dynamics: the lac operon, a case study. *The Journal of Cell Biology*, **161**(3):471-6.
- Vos, P. D., G. Garrity, D., Jones, N. R., Krieg, W., Ludwig, F. A., Rainey, K., Schleifer, K. H. and W. B. Whitman** (2009). Bergys manual of systematic microbiology. Volume 3 - The Firmicutes ISBN 0-387-95041-9.
- Warner, J. B., Krom, B. P., Magni, C., Konings, W. N. and Lolkema, J. S.** (2000). Catabolite repression and induction of the Mg²⁺-citrate transporter CitM of *Bacillus subtilis*. *Journal of Bacteriology*, **182**:6099-6105.
- Weber, C., Farwick, A., Benisch, F., Brat, D., Dietz, H., Subtil, T. and Boles, E.** (2010). Trends and challenges in the microbial production of lignocellulosic bioalcohol fuels. *Applied Microbiology and Biotechnology* **4**:1303-1315.
- Weickert, M. J. and Chambliss, G. H.** (1990). Site-directed mutagenesis of a catabolite repression operator sequence in *Bacillus subtilis*. *Proceedings of the National Academy of Sciences of the United States of America*. **16**: 6238-6242.
- Weigel, N., Kukuruzinska, M. A., Nakazawa, A., Waygood, E. B. and Roseman, S.** (1982). Sugar transport by the bacterial phosphotransferase system. Phosphoryl transfer reactions catalyzed by enzyme I of *Salmonella typhimurium*. *Journal of Biological Chemistry*, **257**:14477-14491.

- Wheals, A. E., Basso, L. C., Alves, D. M. G. and Amorim, H. V.** (1999). Fuel ethanol after 25 years. *Trends in Biotechnology* **17**(12):482-487.
- Wilkins, E., Ringel, A. K., Hortig, D., Willke, T. and Vorlop, K. D.** (2011). High-level production of 1,3-propanediol from crude glycerol by *Clostridium butyricum* AKR102a. *Applied Microbiology and Biotechnology*, DOI: 10.1007/s00253-011-3595-6.
- Willke, T. and Vorlop, K.** (2008). Biotransformation of glycerol into 1,3-propanediol. *European Journal of Lipid Science and Technology*, **110**:831-840.
- Xiao, H., Gu, Y., Ning, Y., Yang, Y., Mitchell, W.J., Jiang, W. and Yang, S.** (2011). Confirmation and elimination of xylose metabolism bottlenecks in glucose phosphoenolpyruvate-dependent phosphotransferase system-deficient *Clostridium acetobutylicum* for simultaneous utilization of glucose, xylose, and arabinose. *Applied and Environmental Microbiology* **77**(22):7886-7895.
- Yazdani, S. S. and Gonzalez, R.** (2007). Anaerobic fermentation of glycerol: a path to economic viability for the biofuels industry. *Current Opinion in Biotechnology*, **18**(3):213-219.
- York, J.D., Ponder, J.W. and Majerus, P.W.** (1995). Definition of a metal-dependent/Li (+)-inhibited phosphomonoesterase protein family based upon a conserved three-dimensional core structure. *Proceedings of the National Academy* **92**(11):5149-5153.
- Yu, Y., Tangney, M., Aass, H. C. and Mitchell, W. J.** (2007). Analysis of the mechanism and regulation of lactose transport and metabolism in *Clostridium acetobutylicum* ATCC 824. *Applied and Environmental Microbiology*, **73**:1842-1850.
- Zeng, L. and Burne, R. A.** (2010). Seryl-phosphorylated HPr regulates CcpA-independent carbon catabolite repression in conjunction with PTS permeases in *Streptococcus mutans*. *Molecular Microbiology*, **75**(5):1145-1158.

- Zhu, P. P., Reizer, J., Reizer A. and Peterkofsky A.** (1993). Unique monocistronic operon (*ptsH*) in *Mycoplasma capricolum* encoding the phosphocarrier protein, HPr, of the phosphoenolpyruvate: sugar phosphotransferase system. Cloning, sequencing, and characterization of *ptsH*. *Journal of Biological Chemistry* **268**:26531-26540.
- Zhu, P. P., Herzberg, O. and Peterkofsky, A.** (1998). Topography of the interaction of HPr (Ser) kinase with HPr. *Biochemistry* **37**:11762-11770.